CONTINUOUS TIME ADAPTIVE FILTERING

A. Yashin

December 1984
WP-84-105

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
2361 Laxenburg, Austria

ABSTRACT

This paper deals with the problem of adaptive estimation in the continuous time Kalman filtration scheme. The necessary and sufficient conditions of the convergence of the parameter estimators are discussed. For systems which are characterized by constant but unknown parameters, the conditions of convergence can be checked before the observation start. The method of proof is based on the relations between singularity property of some probability measures and convergence of the Bayesian estimation algorithm.

CONTINUOUS TIME ADAPTIVE FILTERING

A. Yashin

1. INTRODUCTION

In many cases, the coordinates of partially observed plants subjected to random noise can now be successfully estimated (filtered) [1,2,3,4,5], by using good *a priori* data on the structure and numerical values of the parameters.

In real life, however, some characteristics of the plants may prove impossible to determine before the experiment or observations begin. This leads to certain difficulties in solving the filtering problem. Uncertainty in the coefficients of the equation that describe the plant subjected to random disturbances, may lead to substantial performance deterioration of coordinate estimating algorithms adjusted to certain fixed values of the parameters [1].

Fortunately, the values of unknown parameters can be updated reasonably often with the arrival of observed data. However, simultaneous estimation of both the parameters and plant coordinates may prove to be the nonlinear problem. Lack of appropriate computer methods for the solution of nonlinear problems brings to life a host of heuristic estimation algorithms, which work in many actual situations but need more thorough analytical investigation [2, 3, 4].

The temptation to apply the available coordinate estimation algorithms (which are only effective if the parameter values are known) to the solution of these nonlinear problems leads to adaptive filtering, whereby the filter equations use current parameter estimates which are obtained from processing of observations. The equations or algorithms which lead to parameter estimates are referred to as adaptation, or adjustment, algorithms [5, 6, 7, 8].

While in the case of available parameters the coordinate estimation algorithm is a kind of Kalman filter, with unknown parameters this arrangement is referred to as adaptive Kalman filtering [9, 10].

The choice of the algorithm for the parameter estimation is somewhat arbitrary [11], but one common property is frequently very important in all of these algorithms: the resultant parameter estimates should, in some sense, tend to their true values as the number of observations grows. Such estimates are referred to as consistent. The plant with unknown parameters is identifiable if the parameters' estimates are consistent.

It would be natural to investigate the property of parameter estimate consistency through studying only the properties and characteristics of the initial plant dynamics equations, the properties of noises, and the specifics of the filtering algorithms and adaptation procedure. Such an attempt for systems described by discrete time equations has been reported in [12, 13, 14].

The conditions sufficient for consistency of parameters' estimates, that take on values from a certain finite set, follow, in those papers, from the singularity of probabilistic measures associated with various values of the parameter.

This paper provides the necessary and sufficient conditions for consistency of these parameter estimates in adaptive Kalman filtering for the case of the denumerable set of the parameter values.

This proof relies on a close link between consistency and absolute continuity and singularity of a certain family of probabilistic measures. In the cases to be discussed, singularity entails consistency of estimates; the conditions for consistency may be the conditions for the family of measures to be singular. Recent research concerning absolute continuity and singularity of probabilistic measures associated with random processes [15] has made formulating the necessary and sufficient conditions for these properties possible. Representing some processes as solutions of stochastic difdifferential equations permits formulating the singularity conditions in terms of the characteristics of these equations. In adaptive Kalman filtering, the characteristics of the initial equations can be expressed as filter parameters. In this way, the condition for consistency of estimates can be tested in each specific case before the observations are made. A study of consistency has been performed for Bayesian estimates in discrete time adaptive Kalman filtering in [16, 17].

2. STATEMENT OF THE PROBLEM

The problem of parameter estimation in continuous time adaptive Kalman filtering can be investigated in the framework of the following formal description.

On probabilistic space $(\Omega, H, \mathbf{P})$ a random variable $\beta(\omega)$ is specified which takes on values from a certain denumerable set $\{\beta_i\}$, $i \in \mathbf{N}$ with *a priori* probabilities

$$p_i = \mathbf{P}(\beta = \beta_i), \ \ p_i > 0, \ \ \sum_{i \in \mathbf{N}} p_i = 1.$$

Let H^{β} be σ-algebra in Ω that is generated by the values of the parameter β. On the same space the random process $(\vartheta,\xi) = \vartheta_t,\xi_t,\ t > 0$ is specified. Denote as $H_t^{\vartheta,\xi}$ and H_t^{ξ} the σ-algebras

$$H_t^{\vartheta\xi} = \bigcap_{u>t} \sigma\{\vartheta_v,\xi_v,\ v \le u\}$$

$$H_t^{\xi} = \bigcap_{u>t} \sigma\{\xi_v,\ v \le u\}$$

that are generated by values of the processes (ϑ,ξ) and ξ up to time t, $t \ge 0$, respectively. Let $H_t^{\beta,\vartheta,\xi} = H^{\beta} \vee H_t^{\vartheta,\xi}$, $t \ge 0$ are σ-algebras generated by the union of σ-algebras H^{β} and $H_t^{\vartheta,\xi}$ while $H_t^{\beta,\xi} = H^{\beta} \vee H_t^{\xi}$, $t \ge 0$ are σ-algebras generated by the union of σ-algebras H^{β} and H_t^{ξ}, $t \ge 0$. Let $H^{\vartheta,\xi} = \vee_t H_t^{\vartheta,\xi}$, $H^{\beta,\vartheta,\xi} = \vee_t H_t^{\beta,\vartheta,\xi} = H$, $H^{\beta,\xi} = \vee_t H_t^{\beta,\xi}$. Denote by $\mathbf{H}^{\xi}$, $\mathbf{H}^{\beta,\vartheta,\xi}$ and $\mathbf{H}^{\beta,\xi}$ the respective nondecreasing right-continuous families of σ-algebras.

Assume that on probability space $(\Omega,H,\mathbf{P})$ the process (ϑ,ξ) can be represented as a system of stochastic differential equations:

$$d\vartheta_t = a(\beta,t)\vartheta_t\,dt + b_1(\beta,t)dW_{1,t} + b_2(\beta,t)dW_{2,t},\quad \vartheta_0, \tag{1}$$

$$d\xi_t = A(\beta,t)\vartheta_t\,dt + B(t)dW_{2,t},\quad \xi_0$$

where $(\vartheta_t)_{t\ge0}$ is a sequence of k-dimensional vectors; $(\xi_t)_{t\ge0}$ is a sequence of l-dimensional vectors and $W_{1,t}$, $W_{2,t}$ are independent k_1 and k_2-dimensional Wiener processes, respectively, independent of the initial values of ϑ_0,ξ_0 and the random value of β. The matrix $a(\beta,t)$ is $(k\ k)$, $A(\beta,t)$ is $(l\ k)$ and the matrices $b_1(\beta,t)$, $b_2(\beta,t)$, and $B(t)$ are $(k\ k_1)$, $(k\ k_2)$, and $(l\ k_2)$, respectively, and are the bounded functions of time for any value of β. The process ϑ describes the time variation of the unobservable coordinate of a certain dynamic plant, while the process ξ models the measurement of the coordinate ϑ with random noise.

Introduce on $(\Omega, H^{\beta,\vartheta,\xi})$ a denumerable family of probabilistic measures $\mathbf{P}^i$, $i \in \mathbf{N}$:

$$\mathbf{P}^i(A) = \frac{\mathbf{P}(A \cap \{\beta=\beta_i\})}{p_i}, \quad A \in H^{\beta,\vartheta,\xi}, \ i \in \mathbf{N}.$$

Assume that matrix $B(t)B^*(t)$ is nonsingular for any $t \geq 0$. Let $\mathbf{P}_t$, $\overline{\mathbf{P}}$ and $\overline{\mathbf{P}}_t$ denote the restrictions of the measure $\mathbf{P}$ to the σ-algebras $H_t^{\beta,\vartheta,\xi}$, H^ξ and H_t^ξ, respectively, while $\mathbf{P}_t^i$, $\overline{\mathbf{P}}^i$ and $\overline{\mathbf{P}}_t^i$, restriction of the measures $\mathbf{P}^i$ to $H_t^{\beta,\vartheta,\xi}$, H^ξ and H_t^ξ. Denote by $\mathbf{Z}_t^{i,j}$ the Radon-Nicodim derivative of the measure $\overline{\mathbf{P}}_t^i$ with respect to the measure $\overline{\mathbf{P}}_t^j$ when it exists.

Definition. H^ξ*-adapted process* $\hat{\beta}_t$, $t \geq 0$ *is considered to be a strongly consistent estimate of the parameter* β *if*

$$\lim_{t \to \infty} \hat{\beta}_t = \beta, \quad \mathbf{P}\text{-}a.s.$$

In this paper we will study the consistency conditions of the estimates $\overline{\beta}_t$, $t \geq 0$ where $\overline{\beta}_t = \mathbf{E}(\beta \mid H_t^\xi)$.

3. RESULT

Assume that the initial conditions ϑ_0 and ξ_0 are such that

$$\mathbf{E}_i(||\vartheta_0||^2 + ||\xi_0||^2) < \infty, \ i \in \mathbf{N}, \tag{2}$$

where $\mathbf{E}_i$ denotes mathematical expectation with respect to the measure $\mathbf{P}^i$. From the fact that the coefficients of the equations (1) are bounded and from the condition (2) it follows that with any $t < \infty$ and $i \in \mathbf{N}$

$$\mathbf{E}_i(||\vartheta_t||^2 + ||\xi_t||^2) < \infty. \tag{3}$$

Assume also that the joint distribution of the random variables ϑ_0 and ξ_0 is Gaussian with respect to each measure $\mathbf{P}^i$, $i \in \mathbf{N}$. Hereafter $b_1(i,t)$, $b_2(i,t)$, $\mathfrak{b}_1(i,t)$, $\mathfrak{b}_2(i,t)$, $\mathfrak{a}(i,t)$ and $A(i,t)$ will replace $b_1(\beta_i,t)$, $\mathfrak{b}_2(\beta_i,t)$, $\mathfrak{a}(\beta_i,t)$, $A(\beta_i,t)$, respectively, for convenience of the notation.

The main result of this paper is the following theorem.

Theorem 1. *The strong consistency property of the estimate* $\bar{\beta}_t$, $t \geq 0$ *takes place if and only for any* $i,j \in \mathbf{N}$

$$\int_0^\infty f_{i,j}(t)dt = \infty \tag{4}$$

where

$$f_{i,j}(t) = Sp\{\mathbf{S}_{i,j}(t)\Phi_{i,j}(t)\mathbf{S}^*_{i,j}(t)(B(t)B^*(t))^{-1}\} \tag{5}$$

$$\mathbf{S}_{i,j}(t) = (0, -A(i,t), A(j,t)) \tag{6}$$

and $\Phi_{i,j}(t)$ *are the solutions of the following linear differential equations:*

$$\frac{d\Phi_{i,j}(t)}{dt} = R_{i,j}(t)\Phi_{i,j}(t) + \Phi_{i,j}(t)R^*_{i,j}(t) + G_{i,j}(t)\ G_{i,j}(t)^* \tag{7}$$

with

$$R_{i,j}(t) = \begin{bmatrix} a(i,t) & 0 & 0 \\ K_i(t)A(i,t) & a(i,t)-K_i(t)A(i,t) & 0 \\ K_j(t)A(j,t) & 0 & a(j,t)-K_j(t)A(j,t) \end{bmatrix} \tag{8}$$

$$G_{i,j}(t) = \begin{bmatrix} b_1(j,t) & b_2(j,t) & 0 \\ 0 & K_i(t)B(t) & 0 \\ 0 & K_j(t)B(t) & 0 \end{bmatrix} \tag{9}$$

and initial conditions

$$\Phi_{i,j}(0) = \begin{pmatrix} \mathbf{E}_j(\vartheta_0\vartheta_0^*) & \mathbf{E}_j(\vartheta_0 m_0^{i*}) & \mathbf{E}_j(\vartheta_0 m_0^{j*}) \\ \mathbf{E}_j(m_0^i\vartheta_0^*) & \mathbf{E}_j(m_0^i m_0^{i*}) & \mathbf{E}_j m_0^i m_0^{j*}) \\ \mathbf{E}_j(m_0^j\vartheta_0^*) & \mathbf{E}_j(m_0^j m_0^{i*}) & \mathbf{E}_j(m_0^i m_0^{j*}) \end{pmatrix}$$

where

$$K_i(t) = (b_2(i,t)B^*(t) + \gamma_i(t)A(i,t)^*)(B(t)B^*(t))^{-1} \tag{10}$$

and $\gamma_i(t)$ *satisfy the following equations*

$$\frac{d\gamma_i(t)}{dt} = a(i,t)\gamma_i(t) + \gamma_i(t)a(i,t)^* + bb_i(t) - \tag{11}$$

$$-(b_2(i,t)B^*(t) + \gamma_i(t)A(i,t))(B(t)B^*(t))^{-1}(b_2(i,t)B^*(t) + \gamma_i(t)A(i,t)^*)^*;$$

$$bb_i(t) = b_1(i,t)b_1(i,t)^* + b_2(i,t)b_2(i,t)^*.$$

Corollary. *Assume that the coefficients of the stochastic equations (1) do not depend on time* t, *and a stationary solution of equation (7) exists. Then for the estimate* $\bar{\beta}_n$ *to be consistent it is sufficient that for any* $i,j \in \mathbf{N}$, $i = j$

$$Sp\{\mathbf{S}_{i,j}\Phi_{i,j}\mathbf{S}_{i,j}(BB^*)^{-1}\} \neq 0 \tag{12}$$

where $\Phi_{i,j}$ *is a stationary solution of equation (10).*

The Corollary follows from equalities (4) and (5), and from the stationarity assumption.

Remark. *In the case of a finite number of possible values of the parameter* β *and constant coefficients in equation (1), condition (4) can be verified before the observations are made.*

4. PROOF OF THEOREM 1

The proof will be preceded by several auxiliary lemmas.

Lemma 1. *The process* (ϑ,ξ) *on probabilistic spaces* $(OM, H^{\vartheta,\xi}, \mathbf{P}^i)$ *can be represented in the form:*

$$d\vartheta_t = a(i,t)\vartheta_t dt + b_1(i,t)dW_{1,t} + b_2(i,t)dW_{2,t}, \quad \vartheta_0 \tag{13}$$

$$d\xi_t = A(i,t)\vartheta_t dt + B(t)dW_{2,t}, \quad \xi_0$$

where $W_{1,t}, W_{2,t}$ *are independent Wiener processes independent of* ϑ_0, ξ_0.

Proof. First note that for any $i \in \mathbf{N}$ measures $\mathbf{P}^i$ are absolutely continuous with respect to measure $\mathbf{P}$ and consequenty $\mathbf{P}_t^i \ll \mathbf{P}_t$ for any $t \geq 0$. Denote $\Lambda_t^i = \dfrac{d\mathbf{P}_t^i}{d\mathbf{P}_t}$. Let us consider the transformation of the local characteristics of the process ϑ,ξ with absolutely continuous transformation of the probabilistic measure $\mathbf{P}$ into the measure $\mathbf{P}^i$, $i \in \mathbf{N}$ on $(\Omega, H^{\beta,\vartheta,\xi})$. From the definition of the measures $\mathbf{P}_t^i$, $i \in \mathbf{N}$, $t \geq 0$ it follows that for any $t \geq 0$

$$\Lambda_t^i = \frac{\mathbf{I}(\beta = \beta_i)}{p_i}$$

and consequently, Λ_t^i, $i \in \mathbf{N}$, $t \geq 0$ are independent on time. In compliance with the analog of the Girsanov theorem on transformation of the local characteristics of the processes with absolutely continuous transformation of the probabilistic measures, the local characteristics of the process ϑ,ξ, in the case of a time independent Radon-Nicodim derivative, remain unchanged and consequently the process ϑ,ξ may be represented by equation (1). Note now that following the definitions of the measures $\mathbf{P}^i$, $i \in \mathbf{N}$, the next equalities hold $\mathbf{P}^i$ -a.s.

$$a(\beta,t) = a(i,t), \quad A(\beta,t) = A(i,t), \quad b_1(\beta,t) = b_1(i,t),$$

$$b_2(\beta,t) = b_2(i,t), \quad i \in \mathbf{N},$$

while the processes $W_{1,t}, W_{2,t}$ retain their properties with respect to the measures $\mathbf{P}^i$, $i \in \mathbf{N}$. Thus on $(\Omega, H^{\beta,\vartheta,\xi})$ the process (ϑ,ξ) can be represented by stochastic differential equation (13).

Consider now the transformation of the local characteristics of the process ϑ,ξ with the restriction of σ-algebra $H^{\beta,\vartheta,\xi}$ to $H^{\vartheta,\xi}$. Since the coefficients $a(i,t)\vartheta_t$, $A(i,t)\vartheta_t$, are $\mathbf{H}^{\vartheta,\xi}$-adapted the innovation process (ϑ,ξ) which results from restricting the σ-algebras $H_t^{\beta,\vartheta,\pi i}$ to $H_t^{\vartheta,\xi}$, $t \geq 0$ can be represented on $(\Omega, H^{\vartheta,\xi}, \mathbf{P}^i)$, $i \in \mathbf{N}$ in form (13).

It is a well known fact that with the above assumptions the problem of estimation of the coordinates of the process ϑ from observations of the paths of the process ξ on each space $(\Omega, H^{\vartheta,\xi}, \mathbf{P}^i)$ can be solved in conditional Gaussian terms. Denoting by

$$m_t^i = \mathbf{E}_i(\vartheta_t \mid H_t^{\xi})$$

the mean square optimal estimate of filtering for the process ϑ and by

$$\gamma_i(t) = \mathbf{E}_i((\vartheta_t - m_t^i)(\vartheta_t - m_t^i)^* \mid H_t^{\xi})$$

the conditional variance of the estimate m_t^i, $i \in \mathbf{N}$, $t \geq 0$ we have, for these variables, the well known equations [18]:

$$dm_t^i = a(i,t)m_t^i dt + \tag{14}$$

$$+ (b_i(t)B^*(t) + \gamma_i(t)A(i,t)^*)(B(t)B^*(t))^{-1}(d\xi_t - A(i,t)m_t^i dt), \quad m_0;$$

$$\frac{d\gamma_i(t)}{dt} = a(i,t)\gamma_i(t) + \gamma_i(t)a(i,t)^* + bb_i(t\cdot) - \tag{15}$$

$$- (b_i(t)B^*(t) + \gamma_i(t)A(i,t)^*)\,(B(t)B^*(t))^{-1}\,(b_i(t)B^*(t) + \gamma_i(t)A(i,t)^*)^*,$$

$$\gamma_i(0), \quad i \in \mathbf{N}, \quad t \geq 0.$$

Denote

$$m_t^{\beta} = \mathbf{E}(\vartheta_t \mid H_t^{\beta,\xi}), \quad \gamma_t^{\beta}(t) = \mathbf{E}((\vartheta_t - m_t^{\beta})(\vartheta_t - m_t^{\beta})^* \mid H_t^{\beta,\xi})$$

The following assertion is true.

Lemma 2. *The process ξ can be represented on $(\Omega, H^{\beta,\xi}, \mathbf{P})$ in the form*

$$d\xi_t = A(\beta,t)m_t^{\beta}dt + B(t)dW_t, \quad \xi_0. \tag{16}$$

and W_t is $\mathrm{H}^{\beta,\xi}$-adapted Wiener process.

The proof of this lemma can be done using the same arguments as in [18].

Lemma 3. *On probabilistic spaces* $(\Omega, H^{\beta,\xi}, \mathbf{P}^i)$, $i \in \mathbf{N}$ *the process* ξ *can be represented as*

$$d\xi_t = A(i,t)m_t^i + B(t)dW_t, \quad \xi_0, \quad i \in \mathbf{N}, \quad \mathbf{P}^i\text{-a.s.} \tag{17}$$

Proof. Using the absolute continuity of the measure $\mathbf{P}_t^i$, $i \in \mathbf{N}$, $t \geq 0$ in relation to the measure $\mathbf{P}_t$, ≥ 0 and time independent of the derivatives $\dfrac{d\mathbf{P}_t^i}{d\mathbf{P}_t}$, $t \geq 0$, $i \in \mathbf{N}$, we have, in compliance with the analog of the Girsanov theorem on transformation of local process characteristics with absolutely continuous transformation of probabilistic measures, that the characteristics of process ξ, which is represented by equation (16), do not change with the replacement of measure $\mathbf{P}$ by measure $\mathbf{P}^i$, $i \in \mathbf{N}$. The following equalities hold $\mathbf{P}^i$ -a.s.

$$A(\beta,t) = A(i,t);$$

$$b_k(\beta,t) = b_k(i,t), \quad k = 1,2;$$

$$m_t^\beta = m_t^i;$$

whence follows presentation (17).

The properties of process ξ are found to ensure equivalence between the probabilistic measures $\overline{\mathbf{P}}_n^i$ and $\overline{\mathbf{P}}_n^j$, $i \neq j$. Let us formulate and prove this assertion.

Lemma 4. *Measures* $\overline{\mathbf{P}}_t^i$ *and* $\overline{\mathbf{P}}_t^j$ *are equivalent on measurable space* (Ω, H^ξ).

Proof. The boundedness of the coefficients $A(i,t)$, $a(i,t)$, $b(i,t)$, $B(t)$ and non-singularity of the matrix $B(t)B^*(t)$ provide for any $t < \infty$ the inequality

$$\int_0^t [\, C^*(i,u)(BB(u))^{-1}C(i,u) + C^*(j,u)(BB(u))^{-1}C(j,u))\,]du < \infty$$

where

$$C(i,t) = \begin{bmatrix} a(i,t)\vartheta_t & 0 \\ A(i,t)\vartheta_t & 0 \end{bmatrix}$$

which is true $\mathbf{P}^i$ and $\mathbf{P}^j$-a.s. In accordance with the Liptzer-Shiryaev result [18] (chapter 7) it yields the equivalence of $\mathbf{P}_t^i$ and $\mathbf{P}_t^j$ on $(OM, H^{\vartheta,\xi})$. It follows then that the restrictions of these measures to (Ω, H^{ξ}) are equivalent, that is $\overline{\mathbf{P}}_t^i \sim \overline{\mathbf{P}}_t^j$.

Lemma 5. *The processes* $(m_t^i)_{t\geq 0}$, $i \in \mathbf{N}$ *are Gaussian on the probabilistic spaces* $(\Omega, H^{\xi}, \overline{\mathbf{P}}^k)$, $k \in \mathbf{N}$; k *is not necessarily equal to* i.

Proof. It is easy to see that the processes $(m_t^i)_{t\geq 0}$, $i \in \mathbf{N}$ are Gaussian on probabilistic spaces $(\Omega, H^{\xi}, \overline{\mathbf{P}}^i)$, $i \in \mathbf{N}$. Let us prove that the processes $(m(i,t))_{t\geq 0}$, $i \in \mathbf{N}$ are Gaussian on probabilistic spaces $(\Omega, H^{\xi}, \overline{\mathbf{P}}^k)$, $k \neq i$, $k \in \mathbf{N}$. Because the simultaneous initial distribution of the variables ϑ_0, ξ_0 on any of the probabilistic spaces $(\Omega, H^{\vartheta,\xi}, \mathbf{P}^i)$, $i \in \mathbf{N}$ is Gaussian, the random variables $m_0^j = \mathbf{E}_j(\vartheta_0 | H_0^{\xi})$ are linear functions of ξ_0 with any $j \in \mathbf{N}$ and, consequently, simultaneous distributions of the variables m_0^j, ϑ_0, are also Gaussian on probabilistic spaces $(\Omega, H^{\vartheta,\xi}, \mathbf{P}^k)$ where $k \in \mathbf{N}$, $k \neq j$. Substituting into equation (14) for the values of m_t^j, ξ_t from equation (1) we have

$$dm_t^j = a(j,t)m_t^j dt + \tag{18}$$

$$+ (b_2(j,t)B^*(t) + \gamma_j(t)A(j,t)^j)(B(t)B^*(t))^{-1}(A(\beta,t)\vartheta_t - A(j,t)m_t^j) +$$

$$+ (b_2(j,t)B^*(t) + \gamma_j(t)A^*(j,t)(B(t)B^*(t))^{-1}B(t)dW_{2,t}$$

The fact that the joint distribution of $(m_0^j, \vartheta_0, \xi_0)$ is Gaussian and the formula for ϑ_n and m_t^j, $j \in \mathbf{N}$, $t \geq 0$ are linear, makes processes $(m_t^j)_{t\geq 0}$, $j \in \mathbf{N}$ and $(\xi_n)_{n\geq 0}$ Gaussian on probabilistic spaces $(\Omega, H^{\xi}, \overline{\mathbf{P}}^k)$, $k \in \mathbf{N}$, $k \neq j$ which was required. Consequently, the variables $(A(i,t)m_t^i - A(j,t)m_t^j)$, $j \in \mathbf{N}$ are Gaussian for any of the measures $\overline{\mathbf{P}}^k$, $k \in \mathbf{N}$.

Lemma 6. *Singularity conditions for measures* $\overline{\mathbf{P}}_t^i$ *and* $\overline{\mathbf{P}}_t^j$ *may be written as follows*

$$\int_0^\infty \mathbf{E}_j[A(i,u)m_u^i - A(j,u)m_u^j]^*(B(u)B^*(u))^{-1} \times \tag{19}$$

$$\times [A(i,u)m_u^i - A(j,u)m_u^j]du = \infty$$

Proof. Using the main result of paper [15] one can say that measures $\overline{\mathbf{P}}^i$ and $\overline{\mathbf{P}}^j$ are singular if and only if

$$\int_0^\infty [A(i,u)m_t^i - A(j,u)m_t^j]^*(B(u)B^*(u))^{-1}[A(i,u)m_t^i - A(j,u)m_t^j]du = \infty \tag{20}$$

with $\overline{\mathbf{P}}_j$-probability 1. Taking into account that processes m_t^j are Gaussian on any probabilistic spaces $(\Omega, H^t, \overline{\mathbf{P}}^k)$, $k \in \mathbf{N}$ and result [18], one can see that this condition is equivalent to condition (18). Condition (19) is not very convenient for checking in a general case. For calculation of the mathematical expectation in (19) perform certain additional constructions.

Let $\Phi_{i,j}(t)$ denote a block matrix

$$\begin{bmatrix} \mathbf{E}_j(\vartheta_t\vartheta_t^*) & \mathbf{E}_j(\vartheta_t m_t^{i*}) & \mathbf{E}_j(\vartheta_t m_t^{j*}) \\ \mathbf{E}_j(m_t^j\vartheta_t^*) & \mathbf{E}_j(m_t^i m_t^{j*}) & \mathbf{E}_j(m_t^i m_t^{j*}) \\ \mathbf{E}_j(m_t^j\vartheta_t^*) & \mathbf{E}_j(m_t^j m_t^{i*}) & \mathbf{E}_j(m_t^j m_t^{j*}) \end{bmatrix}$$

and $\mathbf{S}_{i,j}$, a block matrix

$$(0, -A(i,t), A(j,t)).$$

Using the well-known property of matrix multiplication [19] one can write

$$\mathbf{E}_j(A(i,t)m_t^i - A(j,t)m_t^j)^*(B(t)B^*(t))^{-1}(A(i,t)m_t^i - A(j,t)m_t^j) =$$

$$= Sp\{\mathbf{E}_j(A(i,t)m_t^i - A(j,t)m_t^j)(A(i,t)m_t^i - A(j,t)m_t^j)^*(B(t)B^*(t))^{-1}\}$$

Direct verification shows that the matrix

$$\mathbf{E}_j(A(i,t)m_t^i - A(j,t)m_t^j)(A(i,t)m_t^i - A(j,t)m_t^j)^*$$

can be represented in terms of matrices $\Phi_{i,j}(t)$ and $\mathbf{S}_{i,j}(t)$ as follows :

$$\mathbf{E}_j(A(i,t)m_t^i - A(j,t)m_t^j)(A(i,t)m_t^i - A(j,t)m_t^j)^* = \mathbf{S}_{i,j}(t)\Phi_{i,j}(t)\mathbf{S}_{i,j}^*(t)$$

Consequently

$$\mathbf{E}_j(A(i,t)m_t^i - A(j,t)m_t^j)^*(B(t)B^*(t))^{-1}(A(i,t)m_t^i - A(j,t)m_t^j) =$$

$$= Sp\, \mathbf{S}_{i,j}\Phi_{i,j}(t)\mathbf{S}_{i,j}^*(B(t)B^*(t))^{-1}$$

Let us find now the recurrence equation for the matrix $\Phi_{i,j}(t)$. Take up a block matrix $F_{i,j}(t)$ which has the form

$$F_{i,j}(t) = \begin{bmatrix} \vartheta_t \\ m_t^i \\ m_t^j \end{bmatrix} (\vartheta_t^*, m_t^{i*}, m_t^{j*}) = \begin{bmatrix} \vartheta_t\vartheta_t^* & \vartheta_t m_t^{i*} & \vartheta_t m_t^{j*} \\ m_t^i\vartheta_t^* & m_t^i m_t^{i*} & m_t^i m_t^{j*} \\ m_t^j\vartheta_t^* & m_t^j m_t^{i*} & m_t^j m_t^{j*} \end{bmatrix}$$

Using recurrence equation (14) for m_t^i and m_t^j and equation (1) for ϑ_t we have for block elements of the matrix $\Phi_{i,j}(t) = \mathbf{E}_j F_{i,j}(t)$

$$\frac{d\mathbf{E}_j(\vartheta_t\vartheta_t^*)}{dt} = a(i,t)\mathbf{E}_j(\vartheta_t\vartheta_t^*) + \mathbf{E}_j(\vartheta_t\vartheta_t^*)a^*(i,t) + bb_i(t);$$

$$\frac{d\mathbf{E}_j(m_t^j m_t^{j*})}{dt} = a(j,t)\mathbf{E}_j(m_t^j m_t^{j*}) + \mathbf{E}_j(m_t^j m_t^{j*})a^*(j,t) + K_j(t)B(t)B^*(t)K_j^*(t);$$

$$\frac{d\mathbf{E}_j(m_t^i m_t^{i*})}{dt} = a(i,t)\mathbf{E}_j(m_t^i m_t^{i*}) + \mathbf{E}_j(m_t^i m_t^{i*})a^*(j,t) +$$

$$+ K_i(t)(A(i,t)\mathbf{E}_j(m_t^i m_t^{i*}) - A(j,t)\mathbf{E}_j(m_t^j m_t^{i*}) + (\mathbf{E}_j(m_t^i m_t^{i*})A^*(i,t) -$$

$$- \mathbf{E}_j(m_t^i m_t^{j*})A^*(i,t))K_i^*(t) + K_i(t)B(t)B^*(t)K_i(t);$$

$$\frac{d\mathbf{E}_j(m_t^i m_t^{j*})}{dt} = a(i,t)\mathbf{E}_j(m_t^i m_t^{j*}) + \mathbf{E}_j(m_t^j m_t^{i*})a(j,t) + K_i(t)(A(j,t)\mathbf{E}_j(m_t^j m_t^{j*})$$

$$- A(i,t)\mathbf{E}_j(m_t^i m_t^{j*}) + K_i(t)B(t)B^*(t)K_j^*(t);$$

$$\frac{d\mathbf{E}_j(\vartheta_t m_t^{j*})}{dt} = a(j,t)\mathbf{E}_j(\vartheta_t m_t^{j*}) + \mathbf{E}_j(m_t^j\vartheta_t^*)a^*(i,t) + (\mathbf{E}_j(\vartheta_t\vartheta_t^*)A^*(j,t) -$$

$$- \mathbf{E}_j(\vartheta_t m_t^{j*})A(i,t)K_i^*(t) + b_2(i,t)B^*(t)K_i(t);$$

$$\frac{d\mathbf{E}_j(m_t^i\vartheta_t^*)}{dt} = \mathbf{E}_j(m_t^i\vartheta_t^*)a^*(i,t) + a(i,t)(m_t^i\vartheta_t^*) +$$

$$+ K_i(t)(A(j,t)\mathbf{E}_j(\vartheta_t\vartheta_t^*) - A(i,t)\mathbf{E}_j(m_t^i\vartheta_t^*)) + K_i(t)B(t)b_2^*(i,t);$$

Introducing the matrices

$$R_{ij}(t) = \begin{bmatrix} a(i,t) & 0 & 0 \\ K_i(t)A(i,t) & a(i,t)-K_i(t)A(i,t) & 0 \\ K_j(t)A(j,t) & 0 & a(j,t)-K_j(t)A(j,t) \end{bmatrix}$$

$$G_{ij}(t) = \begin{bmatrix} b_1(j) & b_2(j) & 0 \\ 0 & K_i(t)B(t) & 0 \\ 0 & K_j(t)B(t) & 0 \end{bmatrix}$$

the formula for $\Phi_{ij}(t)$ can be rearranged into a matrix form

$$\frac{d\Phi_{ij}(t)}{dt} = R_{ij}(t)\Phi_{ij}(t) + \Phi_{ij}R_{ij}(t)^* + G_{ij}(t)G_{ij}(t)^*.$$

Consequently, the conditions for singularity of the measures $\overline{\mathbf{P}}^i$ and $\overline{\mathbf{P}}^j$, $i \neq j$, $i,j \in \mathbf{N}$ are equivalent to (4) of the theorem.

REFERENCES

1. Z.A. Novoseltseva, "A priory Information in Optimal Filtering Problems," *Automatica and Telemechanica (in Russian)* **39**(6) (1968).

2. P.I. Kitsul and A.V. Lubkov, "On Design of Pseudogradient Algorithms For Adjustment of an Kalman Filter," *Automatica and Telemechanica* **37**(4) (1976).

3. P. Eykhoff , *System Identification* , *Parameter and State Estimation*, John Wiley, New York (1974).

4. J.J. Florentin, "Optimal Probing Adaptive Control of a Simple Bayesian System," *Journal Electronic Control*(11), p.571 (1962).

5. C.T. Leondes and J.O. Pearson, "Kalman Filtering of Systems with Parameter Uncertainties, a Survey.," *International Journal of Control* **17**(4) (1973).

6. K.J. Astrom and P. Eykhoff, "System Identification - a Survey," *Automatica*(7), pp.123-162 (1971).

7. K.J. Astrom and B. Wittenmark , "Problems of Identification and Control," *Journal of Mathematical Analysis and Applications* **34**, pp.90-113 (1971).

8. Y.Z. Tsypkin, *Adaptation and Optimization*, Nauka (1970).

9. C.T. Leonds and J.O. Pearson, "Kalman Filtering of Systems with Parameters Uncertainties.," *International J. Control* **17**(4) (1973).

10. P. Mehra, "On Line Identification of Linear Dynamic Systems with Applications to Kalman Filtering," *IEEE Transaction Automatic Control* **AC-16** (1), pp.12-21 (1971).

11. L. Ljung, "Convergence Analysis of Parametric Identification Methods.," *IEEE Transaction Automatic Control* **AC-23**(3) (1978).

12. Y. Baram and N.R. Sandell , "An Information Theoretic Approach To Dynamical System Modeling And Identification," *IEEE Transactions Automatic Control* **AC-23**(1), pp.61-66 (1978).

13. Y. Baram and N.Y. Sandell, "Consistent Estimation on Finite Parameter Sets with Application to Linear Systems Identification.," *IEEE Transactions Automatic Control* **AC-23**(3) (1978).

14. Y. Baram, "A Sufficient Condition for Consistent Discrimination between Stationary Gaussian Models.," *IEEE Transactions on Automatic Control* **AC-23**(5) (1978).

15. Yu.M. Kabanov, R.S. Liptzer, and A.N. Shiryaev, "Absolute Continuity and Singularity of Locally Absolutely Continuous Probability Distributions," *Math. Sbornik USSR (in Russian)* **35**(5), pp.631-680 (1979).

16. N.M. Kuznetsov and A.I. Yashin, "On the Conditions of the Identifiability of Partially Observed Systems," *Docladi Akademii Nauk SSSR (in Russian)* **259**(4), pp.790-793 (1981).

17. N.M. Kuznetsov and A.I. Yashin, "On Consistent Parameter Estimation In Adaptive Filtering," *Problems Of Control And Information Theory* **10**(5), pp.317-327 (1981).

18. R.S. Liptzer and A.N. Shiryaev, *Statistics of Random Processes*, Springer-Verlag, Berlin and New York (1978).

19. F.R. Gantmacher, *Theory of Matrices (In Russian)*, Nauka, Moscow (1960).