Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

Optimal Filtering for HMM Governed by Special Jump Processes

Andrey V. Borisov and Alexey I. Stefanovich

Abstract— **The paper presents a solution of optimal filtering problem for stochastic differential systems of random structure with switches generated by a special class of Markov jump processes. The equations for both the conditional expectation of some signal process given a noisy observation, and conditional *probability density function* (pdf) are obtained. Numerical methods for solution of corresponding Fokker-Plank and Zakai equation analogues are given and illustrated by an example.**

I. INTRODUCTION

The optimal estimation problems in systems with random structure, which can be treated as the *Hidden Markov Models* (HMM), on the basis of indirect noisy observations has been researched extensively over the last twenty years. The interest in the subject can be easily explained by presence of many practical application areas for possible investigation results: financial mathematics [7], [10], [18], [19] and [22], navigation and target tracking [1] and [2], telecommunications [16], fault detection [21], signal processing [12], automatic control [14], etc.

Meaningful number of papers related to estimation in HMMs was devoted to development of the finite-dimensional filters and extraction of hidden Markov observation systems, for which the optimal filters would be finite-dimensional ones [4], [9], [16] and [17]. Another research direction evolved identification methods for the HMM parameters [9] and [10]. However analysis of HMMs, involving infinite-dimensional objects, such as probability distributions or pdfs, was paid less attention [3].

Generally HMM is itself a dynamic system with random structure, whose transitions are generated by an unobservable Markov process. Usually this process has finite state space, that makes HMM too artificial for description of real phenomena. For example, in the asset price model with jump volatility, obtrusion for the volatility to have only *a finite set of possible values* looks as evident idealization. The same arguments can be used as critical ones towards mathematical model for description of round-trip time fluctuation in TCP/IP links. Apparently, utilization of general Markov jump processes in the HMMs gives a possibility to achieve higher level in the model adequacy. On the other hand, engagement of these processes implies multiple complication of the framework and weakening of possible results. Hence, the purpose is to propose a wider class of Markov jump processes still convenient for inferences.

In [5] it was suggested a class of special jump processes, wider than one of the finite-state Markov processes. Paper [6] introduced a class of HMMs, driven by jump processes of presented class. A system of corresponding integro-differential Fokker-Plank type equations for the transition probability was also derived.

The aim of this paper is to obtain equations of optimal filtering estimate for the signal process in HMM governed by this special type of Markov jump processes.

A. Borisov is with Institute of Informatics Problems of the Russian Academy of Sciences, 44/2 Vavilova st., 119333, Moscow, Russia, ABorisov@ipiran.ru

A. Stefanovich is with Moscow State Aviation Institute, 4, Volokolamskoye sh., A-80, Moscow 125993, Russia, AStefanovich@ipiran.ru

The structure of the paper is as follows. Section II sets basic notation and definitions. Section III contains detailed description of hidden Markov observation system, and formulation of optimal filtering problem. The main theoretical results are presented in Section IV. Equations for optimal filtering estimate of some functional of HMM state (signal process) is derived as well as one describing evolution of conditional pdf. Analogues of Zakai equations for corresponding unnormalized expectation and pdf is also obtained in this section.

Section V is directed to numerical aspects of considered nonlinear filtering problem. A numerical scheme for solution of Fokker-Plank equation analogue, stated for considered HMMs, is given as well as one for numerical solution of the Zakai equation analogue derived in the previous section.

Implementation of the numerical methods is presented in Section VI, which contains solution of the optimal filtering problem for the state of bilinear HMM given noisy observation of cubic sensor.

II. PRELIMINARIES

The following notation is used in this paper:

$(\Omega,\mathscr{F},\mathbf{P},\{\mathscr{F}_t\})$, $t\in[0,T]$ is a probability triplet with right continuous filtration;

$\theta=\{\theta_t\}_{t\in[0,T]}$ is a Markov process taking values in a finite state space $S_n=\{e_1,\dots,e_n\}$ with an initial distribution p_0 and transition intensity matrix $\Lambda(t)=\|\lambda_{ij}(t)\|_{i,j=1}^n$ having continuous components,

$\lambda(t)=(\lambda_{11}(t),\dots,\lambda_{nn}(t))^*$ is the vector collected from the diagonal elements of $\Lambda(t)$ (A^* is the transpose of any vector or matrix A); $\overline{\Lambda}(t)=\Lambda(t)-\operatorname{diag}\lambda(t)$ is an auxiliary matrix,

N_t is the counting process corresponding to transitions of θ_t,

$$\mathscr{T}_i(s,t)=\mathbf{P}\{N_t-N_s=0\mid\theta_s=e_i\}=\exp\left\{\int_s^t\lambda_{ii}(u)du\right\},$$

and $\mathscr{T}(s,t)=(\mathscr{T}_1(s,t),\dots,\mathscr{T}_n(s,t))^*$ is the distribution vector for occupation time of each state of θ_t;

$\mathbf{D}=\{\mathscr{D}_1,\dots,\mathscr{D}_n\}$ is a collection of disjoined Borel subsets of $\mathbb{R}$;

$E=\bigcup_{i=1}^n\mathscr{D}_i$ and $\mathscr{E}=\mathscr{B}(E)$ are state space and minimal σ-algebra, containing all Borel subsets of $\mathscr{D}_i$, $i=1,\dots,n$;

$\mathbf{I}_{\mathscr{D}}(x)$ is the indicator function of the set $\mathscr{D}$;

$\Theta=\Theta(x):\quad E\to S_n$ is the special indicator function: $\Theta(x)=(\mathbf{I}_{\mathscr{D}_1}(x),\dots,\mathbf{I}_{\mathscr{D}_n}(x))^*$,

$\{\pi_i(A)\}_{i=1}^n$ is a collection of probability distributions with supports $\mathscr{D}_i$, $i=1,\dots,n$:

$$\pi_i(\mathscr{D}_i)=1\quad\forall\, i=1,\dots,n;\quad\pi(B)=(\pi_1(B),\dots,\pi_n(B))^*,$$

$$\mathbf{E}_{\pi_i}\{f(y)\}=\int_{\mathscr{D}_i}f(z)\pi_i(dz),$$

$$\mathbf{E}_{\pi}\{f(y)\}=(\mathbf{E}_{\pi_1}\{f(y)\},\dots,\mathbf{E}_{\pi_n}\{f(y)\})^*,$$

$$\mathbf{E}_{\pi}^f=\operatorname{diag}\big(\mathbf{E}_{\pi}\{f(y)\}\big),$$

$\mathscr{Z} = \{\mathscr{Z}_k\}_{k\geq 0}$ is a sequence of i.i.d. random vectors $\mathscr{Z}_k = (\zeta_k^1,\dots,\zeta_k^n)^*$ with independent components ζ_k^i having distributions $\pi_i(\cdot)$.

Definition 1: The process

$$y_t = \mathscr{Z}_{N_t}^* \theta_t \tag{1}$$

is called *the special Markov jump one* generated by the process θ and random sequence $\mathscr{Z}$.

Interconnection of initial probability triplet with filtration defined above, and natural filtrations generated by θ and $\mathscr{Z}$ can be found in [5].

III. PROBLEM STATEMENT

Let us consider the following observation system defined on $(\Omega, \mathscr{F}, \mathbf{P}, \{\mathscr{F}_t\})$, $t \in [0,T]$:

$$\begin{cases} x_t = x_0 + \int_0^t a(x_{s-}, y_{s-}, s)ds + \int_0^t b(x_{s-}, y_{s-}, s)dw_s, \\ y_t = y_0 + \int_0^t \Big[y_{s-}\lambda^*(s) + \mathbf{E}_\pi^*\{y\}\overline{\Lambda}^*(s)\Big]\Theta(y_{s-})ds + M_t^y, \\ U_t = \int_0^t A_s ds + \int_0^t B_s dW_s, \\ f_t = f_0(x_0,y_0) + \int_0^t \alpha_s ds + \int_0^t \beta_s dw_s + \int_0^t \gamma_s dM_s^y, \end{cases} \tag{2}$$

where $x_t \in \mathbb{R}$ is an unobservable "switched" diffusion process, $y_t \in E \subseteq \mathbb{R}$ is corresponding unobservable special Markov jump process defined by its martingale representation [5]; $U_t \in \mathbb{R}^m$ is an observation process, and $f_t \in \mathbb{R}$ is a signal process which should be estimated.

We assume the following conditions for the state $z_t = (x_t, y_t)^*$ equations in (2) to be hold:

1) Borel functions $a = a(x,y,t)$ and $b = b(x,y,t) : \mathbb{R} \times E \times [0,T] \to \mathbb{R}$ are Lipschitz with respect to the pair (x,y):

$$\begin{aligned} &\exists\, K,\ 0<K<\infty: \quad \forall\, t \in [0,T],\ \forall\, (x_1,y_1),\,(x_2,y_2) \in \mathbb{R}\times E \\ &\quad |a(x_1,y_1,t) - a(x_2,y_2,t)|^2 \leqslant K(|x_1-x_2|^2 + |y_1-y_2|^2), \\ &\quad |b(x_1,y_1,t) - b(x_2,y_2,t)|^2 \leqslant K(|x_1-x_2|^2 + |y_1-y_2|^2), \end{aligned}$$

2) functions $a(x,y,t)$, $a'_x(x,y,t)$, $b(x,y,t)$, $b'_x(x,y,t)$ and $b''_{xx}(x,y,t)$ are continuous and bounded with respect to $(x,y,t) \in \mathbb{R}\times E \times [0,T]$, and Hölder continuous with coefficient κ ($0<\kappa<1$) with respect to x uniformly by (y,t):

$$\begin{aligned} &\exists\, C,\ 0<C<\infty: \quad \forall\, (x_1,y_1,t_1),\,(x_2,y_2,t_2) \in \mathbb{R}\times E\times[0,T] \\ &\quad |a_x^{(k)}(x_1,y_1,t_1) - a_x^{(k)}(x_2,y_2,t_2)| \leqslant C|x_1-x_2|^\kappa, \quad k=0,1; \\ &\quad |b_x^{(l)}(x_1,y_1,t_1) - b_x^{(l)}(x_2,y_2,t_2)| \leqslant C|x_1-x_2|^\kappa, \quad l=0,1,2, \end{aligned}$$

and, additionally,

$$|b(x_1,y_1,t_1) - b(x_2,y_2,t_2)| \leqslant C(|x_1-x_2|^\kappa + |t_1-t_2|^{\kappa/2}),$$

3) there exist constants $0 < \underline{\lambda} \leqslant \overline{\lambda} < \infty$ such that for any $(x,y,t) \in \mathbb{R}\times E\times[0,T]$ the equality $\underline{\lambda} \leqslant |b(x,y,t)| \leqslant \overline{\lambda}$ holds,

4) the distributions π_i, $i=1,\dots,n$ have pdfs $\frac{d\pi_i}{d\mu}(y) = \phi_i(y)$, $\phi(y) = (\phi_1(y),\dots,\phi_n(y))^*$, and there exists a constant $0<\gamma<\infty$ such that $\|E_\pi\{|y|^{2+\gamma}\}\| < \infty$,

5) $\mathscr{F}_0$-measurable initial condition y_0 has the pdf $p_y(y,0) = p_0^*\phi(y)$, where p_0 is a distribution of initial value θ_0 of the finite-state Markov process θ_t; $\varsigma_0(x)$ is the pdf of the $\mathscr{F}_0$-measurable initial condition x_0;

6) initial conditions x_0 and y_0, and $\mathscr{F}_t$-adapted Wiener process w_t are mutually independent.

Conditions 1)–6) guarantee the unique strong solution for two first equations in (2), and the following form for transition probability function of the state z_t.

Theorem 1: Let HMM (2) satisfy conditions 1)–6). Then the transition probability function is of the form

$$\begin{aligned} &P_{u,v,s}(A,B,t) = \mathbf{P}\{x_t \in A, y_t \in B | x_s = u, y_s = v\} = \\ &= \mathscr{T}^*(s,t)\Theta(v)\mathbf{I}_B(v)\int_A q_{u,v,s}(x,t)dx + \int_{A\times B} r_{u,v,s}(x,y,t)dxdy, \end{aligned} \tag{3}$$

where the pair of functions $(q_{u,v,s}(x,t), r_{u,v,s}(x,y,t))$ is a solution of the system

$$\begin{cases} (q_{u,v,s})'_t(x,t) = \mathscr{L}_v^* q_{u,v,s}(x,t), \\ (r_{u,v,s})'_t(x,y,t) = \mathscr{L}_y^* r_{u,v,s}(x,y,t) + \lambda^*(t)\Theta(y)r_{u,v,s}(x,y,t) + \\ +\phi^*(y)\overline{\Lambda}^*(t)\Big[\mathrm{diag}\big(\Theta(v)\big)\mathscr{T}(s,t)q_{u,v,s}(x,t) + \\ + \int_E \Theta(z) r_{u,v,s}(x,z,t)dz\Big], \\ 0 \leqslant s < t \leqslant T, \\ q_{u,v,s}(x,s) = \delta_u(x), \\ r_{u,v,s}(x,y,s) = 0, \end{cases} \tag{4}$$

and an operator $\mathscr{L}_h^*$ is as follows

$$\mathscr{L}_h^* f(x) = -\big(a(x,h,t)f(x)\big)'_x + \tfrac{1}{2}\big(b^2(x,h,t)f(x)\big)''_{xx}. \tag{5}$$

Proof of Theorem 1 is presented in [6] under more restrictive conditions than 1)–6).

Under conditions 1)–6) a pdf $\psi(x,y,t)$ of the state z_t also exists and satisfies the system

$$\begin{cases} \psi'_t(x,y,t) = \mathscr{L}_y^*\psi(x,y,t) + \lambda^*(t)\Theta(y)\psi(x,y,t) + \\ +\phi^*(y)\overline{\Lambda}^*(t)\int_E \Theta(z)\psi(x,z,t)dz, \\ 0 < t \leqslant T, \\ \psi(x,y,0) = p_0^*\phi(y)\varsigma_0(x). \end{cases} \tag{6}$$

Note, conditions 1)–6) are sufficient for pdf $\psi(x,y,t)$ to satisfy (6), i.e. under conditions others than 1)–6) correctness of (6) must be proved for each specific case.

We introduce additional assumptions concerning the observation U_t and estimated signal process f_t:

7) A_t, α_t, β_t and γ_t are $\mathscr{F}_t$-predictable square integrable random processes, and it being known that A_t is some measurable functional of system state of the form: $A_t = h(x_t,y_t,t)$;

8) B_t is a nonrandom process, and there exists a constant $0 < \mu < \infty$ such that $B_t B_t^* \geqslant \mu I_m$, where I_m is the $m\times m$ unit matrix;

9) initial condition $f(x_0,y_0)$ of the estimated process is a square integrable random variable: $\mathbf{E}\{f^2(x_0,y_0)\} < \infty$;

10) initial conditions x_0 and y_0, and $\mathscr{F}_t$-adapted Wiener processes $w_t \in \mathbb{R}$ and $W_t \in \mathbb{R}^m$ are mutually independent.

Conditions 7)–10) guarantee for U_t and f_t to be processes with finite moments of the second order, and the noise nondegeneracy in observation U_t.

We denote a filtration of σ-subalgebras generated by observation process U_t as $\mathscr{U}_t = \sigma\{u_s,\ 0\leqslant s\leqslant t\}$, and $\widehat{g}_t = \mathbf{E}\{g_t|\mathscr{U}_t\}$ is the optimal in the mean square sense filtering estimate of an arbitrary random process g_t ($\mathbf{E}\{g_t^2\} < \infty$) given observations $\mathscr{U}_t$.

The problem is to find equations describing evolution of optimal filtering estimate $\widehat{f}_t$ for the signal f_t in (2).

IV. OPTIMAL FILTERING EQUATIONS

Theorem 2: Let conditions 1)–10) hold for observation system (2), then

1) the process ν_t

$$\nu_t = \int_0^t (B_s B_s^*)^{-\frac{1}{2}} (dU_s - \widehat{A}_s ds) \tag{7}$$

is a standard $\mathscr{U}_t$-adapted Wiener process;

2) the optimal filtering estimate $\widehat{f}_t$ for the signal f_t is a solution of the equation

$$\widehat{f}_t = \widehat{f}_0 + \int_0^t \widehat{\alpha}_s ds + \int_0^t \left(\widehat{f_s A_s^*} - \widehat{f}_s \widehat{A}_s^* \right) (B_s B_s^*)^{-\frac{1}{2}} d\nu_s, \tag{8}$$

where $\widehat{f}_0 = \mathbf{E}\{f_0(x_0, y_0)\}$.

Proof of Theorem 2 is quite similar to ones for optimal filtering estimates brought in [8] and [20].

Let us introduce Girsanov transformant $\widetilde{\mathbf{P}}$ on the measurable space $(\Omega, \mathscr{F})$ by the following Radon-Nikodym derivative

$$\frac{d\widetilde{\mathbf{P}}}{d\mathbf{P}} = M_T,$$

where

$$M_t = 1 + \int_0^t A_s^* (B_s B_s^*)^{-\frac{1}{2}} dW_s.$$

The process $\Phi_t = \widetilde{\mathbf{E}}\left\{ \frac{d\mathbf{P}}{d\widetilde{\mathbf{P}}} \Big| \mathscr{F}_t \right\}$ (notation $\widetilde{\mathbf{E}}\{\cdot\}$ means here expectation with respect to measure $\widetilde{\mathbf{P}}$) is an $\mathscr{F}_t$-adapted martingale in the probability space $(\Omega, \mathscr{F}, \widetilde{\mathbf{P}})$ [8], which can be written as

$$\Phi_t = \exp\left\{ \int_0^t A_s^* (B_s B_s^*)^{-1} dU_s - \frac{1}{2} \int_0^t A_s^* (B_s B_s^*)^{-1} A_s ds \right\},$$

and also

$$\widetilde{\Phi}_t = \widetilde{\mathbf{E}}\{\Phi_t | \mathscr{U}_t\} = \exp\left\{ \int_0^t \widehat{A}_s^* (B_s B_s^*)^{-1} dU_s - \frac{1}{2} \int_0^t \widehat{A}_s^* (B_s B_s^*)^{-1} \widehat{A}_s ds \right\}.$$

Further, it is easy to verify that unnormalized conditional expectation $\widetilde{f}_t = \widetilde{\mathbf{E}}\{\Phi_t f_t | \mathscr{U}_t\}$ can be represented as a solution of the equation

$$\widetilde{f}_t = \widehat{f}_0 + \int_0^t \widetilde{\alpha}_s ds + \int_0^t \widetilde{f_s A_s^*} (B_s B_s^*)^{-1} dU_s, \tag{9}$$

and

$$\widehat{f}_t = \frac{\widetilde{f}_t}{\widetilde{\Phi}_t}. \tag{10}$$

By Girsanov theorem, the process $V_t = \int_0^t (B_s B_s^*)^{-1/2} dU_s$ is $\mathscr{F}_t$-adapted standard Wiener process in the probability space $(\Omega, \mathscr{F}, \widetilde{\mathbf{P}})$. The measure μ_V induced by V_t as a Wiener process, and the one $\mu_{\widehat{f}}$ induced by $\widehat{f}_t$, are equivalent [13]: $\mu_V \sim \mu_{\widehat{f}}$. Finally, from the fact $\widetilde{\mathbf{P}} \sim \mathbf{P}$, it follows that under conditions 1)–10) the state vector z_t of system (2) has a conditional pdf corresponding to the initial probability measure $\mathbf{P}$.

Theorem 3: Let conditions 1)–10) hold for observation system (2), then

1) normalized conditional pdf $\widehat{\psi}(x,y,t)$ $\big((x,y,t) \in \mathbb{R} \times E \times [0,T]\big)$ of state z_t given observations $\mathscr{U}_t$ is a solution of the stochastic equation

$$\begin{aligned} \widehat{\psi}(x,y,t) &= p_0^* \phi(y) \varsigma_0(x) + \\ &+ \int_0^t \Big(\mathscr{L}_y^* \widehat{\psi}(x,y,s) + \lambda^*(s) \Theta(y) \widehat{\psi}(x,y,s) + \\ &+ \phi^*(y) \overline{\Lambda}^*(s) \int_E \Theta(z) \widehat{\psi}(x,z,s) dz \Big) ds + \\ &+ \int_0^t \widehat{\psi}(x,y,s) \big(h(x,y,s) - \widehat{A}_s\big)^* (B_s B_s^*)^{-\frac{1}{2}} d\nu_s, \end{aligned} \tag{11}$$

where

$$\widehat{A}_t = \int_{\mathbb{R} \times E} h(x,y,t) \widehat{\psi}(x,y,t) dx dy, \tag{12}$$

2) unnormalized conditional pdf $\widetilde{\psi}(x,y,t)$ is a solution of the stochastic equation

$$\begin{aligned} \widetilde{\psi}(x,y,t) &= p_0^* \phi(y) \varsigma_0(x) + \\ &+ \int_0^t \Big(\mathscr{L}_y^* \widetilde{\psi}(x,y,s) + \lambda^*(s) \Theta(y) \widetilde{\psi}(x,y,s) + \\ &+ \phi^*(y) \overline{\Lambda}^*(s) \int_E \Theta(z) \widetilde{\psi}(x,z,s) dz \Big) ds + \\ &+ \int_0^t \widetilde{\psi}(x,y,s) h^*(x,y,s) (B_s B_s^*)^{-1} dU_s, \end{aligned} \tag{13}$$

and

$$\widehat{\psi}(x,y,t) = \frac{\widetilde{\psi}(x,y,t)}{\int_{\mathbb{R} \times E} \widetilde{\psi}(u,v,t) du dv}. \tag{14}$$

Proof: let $\xi = \xi(x,y) = f(x) g(y)$ be an arbitrary function such that $f = f(x) \in \mathbf{C}_b^2(\mathbb{R})$ has a compact support, and function $g = g(y): E \to \mathbb{R}$ satisfies the conditions $\|\mathbf{E}_\pi\{g^2(y)\}\| < \infty$ and $\int_E |g(y)| dy < \infty$. Using the Itô rule, martingale representation for special Markov jump processes, and the fact $[w, M^y]_t \equiv 0$ $\mathbf{P}$-a.s., we have for all $0 \leqslant t \leqslant T$ that

$$\begin{aligned} \Xi_t &= \xi(x_t, y_t) = \xi(x_0, y_0) + \\ &+ \int_0^t f(x_{s-}) \Big[g(y_{s-}) \lambda^*(s) + \mathbf{E}_\pi^*\{g(y)\} \overline{\Lambda}^*(s) \Big] \Theta(y_{s-}) ds + \\ &+ \int_0^t g(y_{s-}) \Big[f_x'(x_{s-}) a(x_{s-}, y_{s-}, s) + \tfrac{f_{xx}''(x_{s-})}{2} b^2(x_{s-}, y_{s-}, s) \Big] ds + \\ &+ \int_0^t f(x_{s-}) dM_s^g + \int_s^t f_x'(x_{s-}) g(y_{s-}) b(x_{s-}, y_{s-}, s) dw_s. \end{aligned}$$

Note, the last two terms in the latter equality are $\mathscr{F}_t$-adapted martingales. From Theorem 2 and the fact, the number of y_t jumps occurring in the interval $[0,T]$ is finite a.s., the normalized conditional expectation $\widehat{\Xi}_t$ takes the form

$$\begin{aligned} \widehat{\Xi}_t &= \mathbf{E}\{\xi(x_0, y_0)\} + \\ &+ \int_0^t \Big[\lambda^*(s) \widehat{\eta}_s^1 + \mathbf{E}_\pi^*\{g(y)\} \overline{\Lambda}^*(s) \widehat{\eta}_s^2 + \widehat{\eta}_s^3 + \tfrac{1}{2} \widehat{\eta}_s^4 \Big] ds + \\ &+ \int_0^t \Big[\widehat{\Xi_s A_s^*} - \widehat{\Xi}_s \widehat{A}_s^* \Big] (B_s B_s^*)^{-\frac{1}{2}} d\nu_s, \end{aligned} \tag{15}$$

where

$$\begin{aligned} \widehat{\eta}_t^1 = \mathbf{E}\{\Theta(y_t) \xi(x_t, y_t) | \mathscr{U}_t\}, \quad \widehat{\eta}_t^2 &= \mathbf{E}\{\Theta(y_t) f(x_t) | \mathscr{U}_t\}, \\ \widehat{\eta}_t^3 &= \mathbf{E}\{a(x_t, y_t, t) \xi_x'(x_t, y_t) | \mathscr{U}_t\}, \\ \widehat{\eta}_t^4 &= \mathbf{E}\{b^2(x_t, y_t, t) \xi_{xx}''(x_t, y_t) | \mathscr{U}_t\}. \end{aligned} \tag{16}$$

On the other hand, for an arbitrary process $\rho_t = \rho(x_t, y_t, t)$ such that $\mathbf{E}\{|\rho_t|\} < \infty$, its conditional expectation is defined as

$$\mathbf{E}\{\rho_t | \mathscr{U}_t\} = \mathbf{E}\{\rho(x_t, y_t, t) | \mathscr{U}_t\} = \int_{\mathbb{R} \times E} \rho(x,y,t) \widehat{\psi}(x,y,t) dx dy. \tag{17}$$

Applying Fubini theorem and integration by parts in (15–17), we obtain

$$\begin{aligned} &\int_{\mathbb{R} \times E} \widehat{\psi}(x,y,t) \xi(x,y) dx dy = \\ &= \int_{\mathbb{R} \times E} \Big\{ \int_0^t \Big[\mathscr{L}_y^* \widehat{\psi}(x,y,s) + \lambda^*(s) \Theta(y) \widehat{\psi}(x,y,s) + \\ &+ \phi^*(y) \overline{\Lambda}^*(s) \int_E \Theta(z) \widehat{\psi}(x,z,s) dz \Big] ds \Big\} \xi(x,y) dx dy + \\ &+ \int_{\mathbb{R} \times E} \Big\{ \int_0^t \widehat{\psi}(x,y,s) \big(h(x,y,s) - \widehat{A}_s\big)^* (B_s B_s^*)^{-\frac{1}{2}} d\nu_s \Big\} \xi(x,y) dx dy. \end{aligned}$$

The set of functions $\xi(x,y)$ stated above in the proof, is dense in $\mathscr{L}_1(\mathbb{R} \times E)$, hence the last equality is true iff the function $\widehat{\psi}(x,y,t)$ is a solution of (11) $\mathbf{P}$-a.s.

Equation (13) can be proved similarly, using the fact $\widetilde{\Xi}_t = \widetilde{\mathbf{E}}\{\Phi_t \Xi_t | \mathscr{U}_t\} = \widetilde{\Phi}_t \widehat{\Xi}_t$. ■

V. NUMERICAL METHODS FOR SOLUTION OF OPTIMAL FILTERING EQUATIONS

This section contains some numerical schemes for solution of both nonrandom equation (6) and stochastic equation (13).

Note that equation (6) can be rewritten in the form

$$\begin{aligned} \psi'_t(x,y,t) = Q(x,y,t)\psi'_x(x,y,t) + R(x,y,t)\psi''_{xx}(x,y,t)+ \\ +S(x,y,t)\psi(x,y,t) + \phi^*(y)\overline{\Lambda}^*(t)\int_E \Theta(z)\psi(x,z,t)dz, \end{aligned} \tag{18}$$

where

$$\begin{aligned} &Q(x,y,t) = -a(x,y,t) + 2b(x,y,t)b'_x(x,y,t),\\ &R(x,y,t) = \tfrac{1}{2}b^2(x,y,t),\\ &S(x,y,t) = -a'_x(x,y,t) + (b'_x(x,y,t))^2 + b''_{xx}(x,y,t) + \lambda^*(t)\Theta(y). \end{aligned}$$

To define a numerical scheme properly the following assumptions are made:

1) all the support sets $\mathscr{D}_i$ are intervals $[a_i, b_i]$, $i = 1, \dots, n$;
2) domain in variable x is bounded by interval $[\underline{x}, \overline{x}]$;
3) on the whole domain $[\underline{x}, \overline{x}] \times \bigcup_{i=1}^{n} [a_i, b_i] \times [0, T]$ an analytical grid G is defined with increments δ_x, δ_y and δ_t;
4) the value of numerical solution for (18) at the grid point (x_i, y_i, t_k) is denoted by ψ_{ij}^k; Q_{ij}^k, R_{ij}^k, S_{ij}^k, ϕ_j^*, $\overline{\Lambda}^k$ and Θ_j denote corresponding functions values at the grid points.

The following splitting numerical scheme is used to solve (18):

$$\begin{aligned} \frac{\psi_{ij}^{k+1/3} - \psi_{ij}^k}{\delta_t} &= L_1(\alpha\psi_{ij}^{k+1/3} + \beta\psi_{ij}^k),\\ \frac{\psi_{ij}^{k+2/3} - \psi_{ij}^{k+1/3}}{\delta_t} &= L_2(\alpha\psi_{ij}^{k+2/3} + \beta\psi_{ij}^{k+1/3}),\\ \frac{\psi_{ij}^{k+1} - \psi_{ij}^{k+2/3}}{\delta_t} &= L_3(\alpha\psi_{ij}^{k+1} + \beta\psi_{ij}^{k+2/3}), \end{aligned} \tag{19}$$

where α, $\beta \geqslant 0$ are parameters: $\alpha + \beta = 1$, and

$$\begin{aligned} L_1\psi_{ij}^k &= \frac{Q_{ij}^k}{2\delta_x}(\psi_{i+1j}^k - \psi_{i-1j}^k),\\ L_2\psi_{ij}^k &= \frac{R_{ij}^k}{\delta_x^2}(\psi_{i+1j}^k - 2\psi_{ij}^k + \psi_{i-1j}^k),\\ L_3\psi_{ij}^k &= S_{ij}^k\psi_{ij}^k + \phi_j^*(\overline{\Lambda}^k)^*\delta_y\sum_m \Theta_m\psi_{im}^k \end{aligned}$$

are difference approximations of operators $Q(x,y,t)\frac{\partial}{\partial x}(\cdot)$, $R(x,y,t)\frac{\partial^2}{\partial x^2}(\cdot)$ and $S(x,y,t)(\cdot) + \phi^*(y)\overline{\Lambda}^*(t)\int_E \Theta(z)(\cdot)dz$ respectively.

It is easy to see the local discretization error of numerical scheme (19) is of order $O(\delta_t + \delta_x^2 + \delta_y)$.

Let us consider equation (13) rewritten in the form

$$\begin{aligned} &\widetilde{\psi}'_t(x,y,t)dt = Q(x,y,t)\widetilde{\psi}'_x(x,y,t)dt+\\ &\quad+[R(x,y,t)\widetilde{\psi}''_{xx}(x,y,t) + S(x,y,t)\widetilde{\psi}(x,y,t)]dt+\\ &\quad+\phi^*(y)\overline{\Lambda}^*(t)\int_E \Theta(z)\widetilde{\psi}(x,z,t)dzdt+\\ &\quad+\widetilde{\psi}(x,y,t)h^*(x,y,s)(B_tB_t^*)^{-1/2}dV_t,\\ &\widetilde{\psi}(x,y,0) = p_0^*\phi(y)\varsigma_0(x), \end{aligned} \tag{20}$$

where the functions Q, R and S are stated above, and V_t is a standard Wiener process defined on the probability space $(\Omega, \mathscr{F}, \widetilde{\mathbf{P}})$.

The following splitting numerical scheme is used to solve this equation:

$$\begin{aligned} \frac{\widetilde{\psi}_{ij}^{k+1/4} - \widetilde{\psi}_{ij}^k}{\delta_t} &= L_1(\alpha\widetilde{\psi}_{ij}^{k+1/4} + \beta\widetilde{\psi}_{ij}^k),\\ \frac{\widetilde{\psi}_{ij}^{k+2/4} - \widetilde{\psi}_{ij}^{k+1/4}}{\delta_t} &= L_2(\alpha\widetilde{\psi}_{ij}^{k+2/4} + \beta\widetilde{\psi}_{ij}^{k+1/4}),\\ \frac{\widetilde{\psi}_{ij}^{k+3/4} - \widetilde{\psi}_{ij}^{k+2/4}}{\delta_t} &= L_3(\alpha\widetilde{\psi}_{ij}^{k+3/4} + \beta\widetilde{\psi}_{ij}^{k+2/4}),\\ \widetilde{\psi}_{ij}^{k+1} &= L_4\widetilde{\psi}_{ij}^{k+3/4}, \end{aligned} \tag{21}$$

where difference approximations L_1, L_2 and L_3 are defined above, and

$$L_4\widetilde{\psi}_{ij}^k = \widetilde{\psi}_{ij}^k \exp\left\{(h_{ij}^k)^*(B_kB_k^*)^{-1}\left[U^{k+1} - U^k - \frac{\delta_t}{2}h_{ij}^k\right]\right\}.$$

Using definition of approximation accuracy for SDE [11] it can be shown than local discretization error of numerical scheme (21) is of order $O(\delta_t^{1/2} + \delta_x^2 + \delta_y)$.

Both of schemes (19) and (21) are used in the next section with parameters $\alpha = \beta = 1/2$.

VI. COMPUTATIONAL EXAMPLE

This section demonstrates applicability of proposed numerical methods for solution both of deterministic and stochastic partial integro-differential equations (6) and (13). As an example of system (2) we consider the following observation system on the interval $[0, 100]$:

$$\begin{cases} x_t = x_0 + \int_0^t \big(y_{s-}x_{s-} + c(y_{s-})\big)ds + \int_0^t b(y_{s-})dw_s,\\ b(y) = \big(b_1^1\mathbf{I}_{\mathscr{D}_1}(y) + b_2^1\mathbf{I}_{\mathscr{D}_2}(y)\big)y + b_1^0\mathbf{I}_{\mathscr{D}_1}(y) + b_2^0\mathbf{I}_{\mathscr{D}_2}(y),\\ c(y) = \big(c_1^1\mathbf{I}_{\mathscr{D}_1}(y) + c_2^1\mathbf{I}_{\mathscr{D}_2}(y)\big)y + c_1^0\mathbf{I}_{\mathscr{D}_1}(y) + c_2^0\mathbf{I}_{\mathscr{D}_2}(y),\\ y_t = y_0 + \int_0^t \Big[y_{s-}\lambda^*(s) + \mathbf{E}_\pi^*\{y\}\overline{\Lambda}^*(s)\Big]\Theta(y_{s-})ds + M_t^y,\\ U_t = \int_0^t x_s^3 ds + \varepsilon W_t, \end{cases} \tag{22}$$

where $b_1^1 = -0.5$, $b_2^1 = -1$, $b_1^0 = 0$, $b_2^0 = 0$, $c_1^1 = 2$, $c_2^1 = 1$, $c_1^0 = -10$, $c_2^0 = 10$, $\mathscr{D}_1 = [-7, -5]$, $\mathscr{D}_2 = [-2, -1]$, $\Lambda = \begin{bmatrix} -0.01 & 0.01 \\ 0.02 & -0.02 \end{bmatrix}$, $p_0 = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$, $\varepsilon = 0.1$.

System (22) contains bilinear equation for state component x_t, and its noised cubic observation U_t. Note that conditions 1)–6) are not valid for this system, nevertheless assertions of Theorems 1 and 2 are true, and equations (6) and (13) are correct for description of pdfs $\psi(x,y,t)$ and $\widehat{\psi}(x,y,t)$. This fact can be proved, basing on the known solution of Fokker-Plank equation for linear stochastic differential systems, and the formula of total probability related to the process y_t jumps.

The problem is to obtain an optimal filtering estimate both of the state z_t and generating process $\theta_t = \Theta(y_t)$.

Calculation was implemented with the following coordinates and time increments $\delta_x = 0.1$, $\delta_y = 0.05$, $\delta_t = 0.1$.

Fig. 1 contains 3D-plot illustrating time evolution of marginal pdf $\psi^x(x,t) = \int_E \psi(x,y,t)dy$ based on the solution $\psi(x,y,t)$ of (6), and Fig. 2 contains corresponding conditional marginal pdf $\widehat{\psi}^x(x,t)$ based on the solution $\widehat{\psi}(x,y,t)$ of (13), $t \in [0, 10]$. Fig. 3 contains conditional marginal pdf $\widehat{\psi}^y(y,t)$. Fig. 4 presents both the results of component x_t optimal filtering, and its conditional accuracy characteristic $\sigma_t^x = \sqrt{\mathbf{E}\{\|x_t - \widehat{x}_t\|^2 | \mathscr{U}_t\}}$, $t \in [0, 100]$. Corresponding plots related to filtering of component y_t is drawn at Fig. 5, meanwhile the conditional probability $\widehat{\theta}_t^1 = \mathbf{P}\{y_t \in$

$\mathscr{D}_1|\mathscr{U}_t\} = \mathbf{P}\{\Theta(y_t)^* e_1 = 1|\mathscr{U}_t\}$ in comparison with its true value $\theta_t^* e_1$ is demonstrated at Fig. 6. Conditional probability $\widehat{\theta}_t^2 = \mathbf{P}\{y_t \in \mathscr{D}_2|\mathscr{U}_t\} = \mathbf{P}\{\Theta(y_t)^* e_2 = 1|\mathscr{U}_t\}$ can be easily calculated using normalization condition: $\widehat{\theta}_t^2 = 1 - \widehat{\theta}_t^1$.

As can be seen at Figs. 4 and 6, the quality of estimates $\widehat{x}_t$ and $\widehat{\theta}_t^1$ is high by contrast with one of $\widehat{y}_t$ which is rather mediocre (see Fig. 5). Apparently the reason for this feature is related in some way with identifiability of the state z_t.

VII. CONCLUSIONS

The main results of this paper are equations (11) and (13) for the conditional pdf of HMM state governed by the special Markov jump process. Numerical schemes for the solution of both the deterministic equation (6) and stochastic ones (11) and

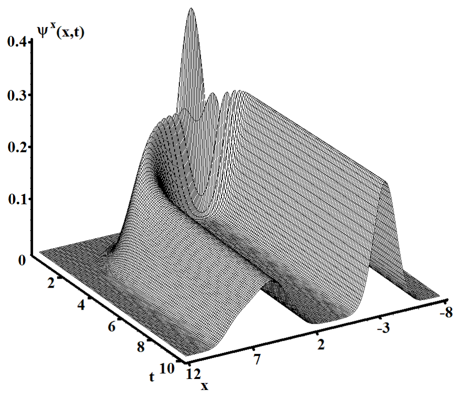

Fig. 1. Evolution of marginal pdf $\psi^x(x,t)$

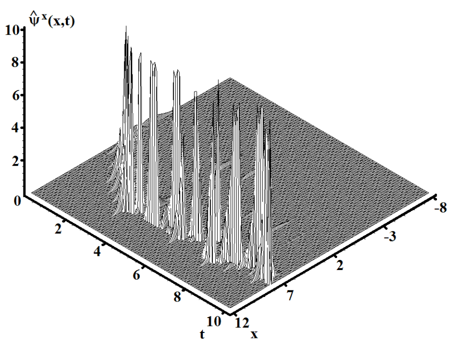

Fig. 2. Evolution of conditional marginal pdf $\widehat{\psi}^x(x,t)$

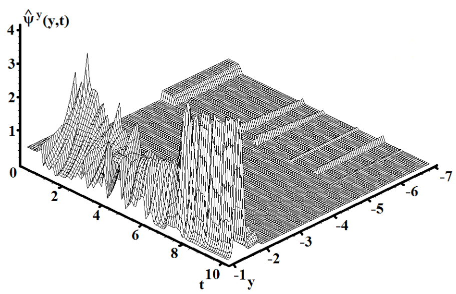

Fig. 3. Evolution of conditional marginal pdf $\widehat{\psi}^y(y,t)$

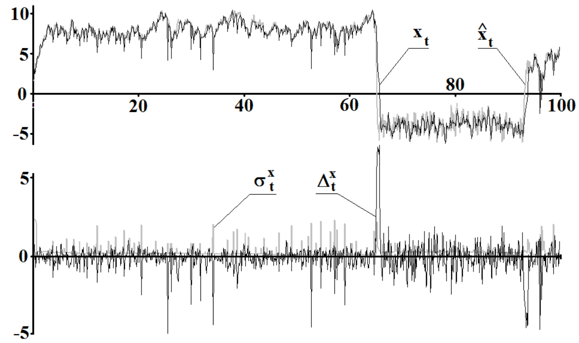

Fig. 4. Filtering diffusion state estimate $\widehat{x}_t$ vs its exact value x_t, filtering error Δ_t^x and conditional MS deviation σ_t^x

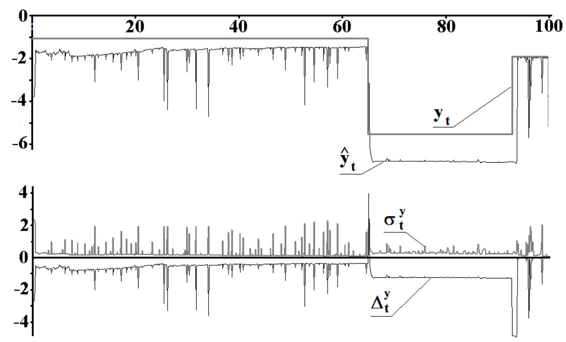

Fig. 5. Filtering jump state estimate $\widehat{y}_t$ vs its exact value y_t, filtering error Δ_t^y and conditional MS deviation σ_t^y

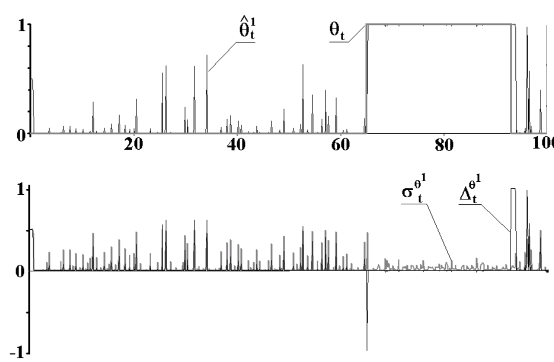

Fig. 6. Filtering generating process estimate $\widehat{\theta}_t^1$ vs its exact value $\theta_t^* e_1$, filtering error $\Delta_t^{\theta^1}$ and conditional MS deviation $\sigma_t^{\theta^1}$

(13) is also presented. Nevertheless, there are several interesting problems that deserve further investigation. First, one can see the assumptions 1)–6) needed to make equations (6), (11) and (13) meaningful are substantionally stronger than those that ensure the existence both the state process z_t and its conditional expectation. Obviously, relaxation of these conditions, e.g. in the same way as in [15] would be highly appreciable. Second, as was mentioned in the previous section, mathematical framework for preliminary identifiability analysis in system (2) state might be very useful. Third, utilization of rather simple fractional step method for solution of (6), (11) and (13) is intended to demonstration of the ability to solve them in principle. At the same time, development of corresponding numerical schemes based on particle method or differential geometry approach looks more prospective and effective. All these problems can be objects of subsequent research.

REFERENCES

[1] Y. Bar-Shalom, L. Campo, X. R. Li, *Control of Discrete-Time Hybrid Stochastic Systems*, International Series on Advances in Control and Dynamic Systems, vol. 76, (C.T. Leondes, ed.), Academic Press, 1996.

[2] Y. Bar-Shalom, X. R. Li, Multiple-model estimation with variable structure, *IEEE Trans. Autom. Contr.*, vol. 41(4), pp. 478–493, 1996.

[3] F. Bernard, F. Dufour, P. Bertrand, Systems with Markovian Jump Parameters: Approximations for the Nonlinear Filtering Problem, in Proc. ECC 97, Brussels, Belgium, 1997.

[4] T. Björk, Finite Optimal Filters for a Class of Nonlinear Diffusions with Jumping Parameters, *Stochastics*, vol. 2, pp. 121–138, 1982.

[5] A. Borisov, Analysis and Estimation of the States of Special jump Markov Processes I: Martingale Representation, *Autom. Remote Contr.* vol. 65, no. 1, 2004, pp. 44–57.

[6] A. Borisov, Fokker-Plank Like Equation for Hidden Markov Models Governed by Special Jump Processes, in 43 Proc. Conf. Dec. Contr., pp. 4151–4156, 2004.

[7] J. Cvitanič, R. Sh. Liptser, B. Rozovskii, Tracking volatility, in 39th Proc. Conf. Dec. Contr., pp. 1189–1193, 2000.

[8] R. J. Elliott, *Stochastic Calculus and Applications*, Springer-Verlag, NY, 1982.

[9] R. J. Elliott, L. Aggoun, J.B. Moore, *Hidden Markov Models: Estimation and Control*, Springer-Verlag, Berlin, 1995.

[10] R. J. Elliott, W. P. Malcolm, A. Tsoi, HMM Volatility Estimation, in Proc. 41st IEEE Conf. Dec. Contr., Las Vegas, pp. 398–404, 2002.

[11] P. Kloeden, E. Platen, *The numerical Solution of Stochastic Differential Equations*, Springer-Verlag, Berlin, 1992.

[12] V. Krishnamurthy, J. B. Moore, On-Line Estimation of Hidden Markov Model Parameters Based on the Kullback-Leibler Information Measure, *IEEE Trans. Sign. Proc.*, vol. 40(8), pp. 2557–2573, 1993.

[13] R. S. Liptser, A. N. Shiryayev, *Statistics of Random Processes*, Springer-Verlag, NY, 1977.

[14] M. Mariton, *Jump Linear Systems in Automatic Control*, Marcel Decker, NY, 1990.

[15] R. Mikulevicius, B. Rozovskii, Linear Parabolic Stochastic PDEs and Wiener Chaos, *SIAM J. Math. Anal.*, vol. 29, no. 2, pp. 452–480, 1998.

[16] B. Miller, K. Avrachenkov, K. Stepanyan, G. Miller, Flow control as stochastic optimal control problem with incomplete information, *INRIA Research Report No. 5239*, 2004, `http://www.inria.fr/rrrt/rr-5239.html`.

[17] B. M. Miller, W. J. Runggaldier, Kalman filtering for linear systems with coefficients driven by a hidden Markov jump process *Syst. & Control Lett.*, vol. 31, pp. 93–102, 1997.

[18] W. J. Runggaldier, Jump Diffusion Models, in *Handbook of Heavy Tailed Distributions in Finance* (Ed. S.T. Rachev). Handbooks in Finance. B. 1 (Series Ed. W.Ziemba) North-Holland: Elesevier, pp. 169–209, 2003.

[19] A. N. Shiryaev, *Essential of Stochastic Finance: Facts, Models, Theory*, vol.3, World Scientific, 1999.

[20] E. Wong and B. Hajek, *Stochastic Processes in Engineering Systems*, Springer Verlag, NY, 1985.

[21] J. Ying, T. Kirubarajan and K. R. Pattipati, A Hidden Markov Model-based Algorithm for Online Fault Diagnosis with Partial and Imperfect Tests, *IEEE Trans. Systems, Man and Cybernetics*, SMC(C)-30(4), pp. 463–473, 2000.

[22] X. Y. Zhou, G. Yin, Markowitz's mean-variance portfolio selection with regime switching: a continuous-time model, *SIAM J. Control Optim.*, vol. 42, no. 4, pp. 1466–1482, 2003.