THE MINIMAX POSTERIOR WONHAM FILTERING/IDENTIFICATION

Andrey V. Borisov [*,1]

[*] *Institute of Informatics Problems of the Russian Academy of Sciences, 44/2 Vavilova st., 119333, Moscow, Russia* ABorisov@ipiran.ru

Abstract: The paper investigates a filtering/identification problem of finite-state Markov processes given continuous and/or counting observations. All the transition intensity matrix, observation plan and counting intensity are parameterized by a random vector with uncertain distribution belonging to a known class. An assertion concerning existence of saddle point in the considered minimax optimization problem as well as a form of corresponding estimate is presented. Monitoring of TCP link status under uncertainty is proposed as an illustrative numerical example of application of obtained theoretical results.

Keywords: Markov models, uncertainty, filtering theory, identification, minimax techniques

1. INTRODUCTION

For stochastic differential observation systems the Kalman-Bucy filter (Kalman and Bucy, 1961) and the Wonham one (Wonham, 1964) provide the most famous and applicable optimal estimates having a finite-dimensional form. A priori uncertainty in observation systems, occurring in many practical situations, generates numerous variations of these algorithms as well as the whole realms of estimation theory in stochastic systems, and the minimax approach, particularly (Anan'ev, 1993), (Borisov and Pankov, 1994), (Kats and Kurzhanskii, 1978), (Martin and Mintz, 1983), (Pankov and Miller, 2005) and (Semenikhin *et al.*, 2005). The most attention in this area was paid to the Kalman-Bucy filter, and basic results are tightly connected with paradigm of the combination *"linear observation system – Gaussianity – linear estimate"*. In fact, admissibility of Gaussian noises in a linear observation system causes for an optimal estimate to have a linear form. In the case Gaussianity is unacceptable, the set of feasible estimates, as a rule, is bounded forcibly by the class of linear ones. Without pursuing to write an exhaustive review of minimax filtering, only the papers closely related to the proposed approach are mentioned here.

The minimax filtering problem in linear stochastic systems with a parametric uncertainty in both the noises intensity and dynamics/observation was investigated in (Martin and Mintz, 1983). The authors considered optimality criterion in the form of the (unconditional) mathematical expectation of quadratic and/or generalized quadratic loss functions. Using game-theoretic framework, the filtering problem was transformed to a zero-sum game, which solution existed, generally, in terms of mixed strategies. This meant, artificial random nature was a burden to the uncertain parameters, and the corresponding saddle point made sense as the pair "the worst distribution of uncertain pa-

[1] This work was supported by Russian Foundation for Basic Research, grant No. 05-01-00508-a and the Program OITVS RAN "Fundamental Algorithms of Information Technology" (project 1.5).

rameters — the best estimator in the worst case". Notably, the obtained best estimate represented a weighted average of the classic Kalman filter estimates. Further, these results were extended in (Anan'ev, 1993) to the case of nonparametric uncertainty in the investigated observation systems.

A case of minimax filtering problem for linear Gaussian systems with conditional mean square (MS) criterion was considered in (Kats and Kurzhanskii, 1978). The uncertainty was only in the noise parameters. The criterion form, system linearity and Gaussian noises led to existence of a saddle point in pure strategies. The initial problem was reduced to searching for the least favorable noise parameters (i.e. the mean and intensity), and the best estimate was delivered by the classic Kalman filter designed for this choice of parameters.

Aforementioned papers were directed to the discrete - time observation systems. The authors of (Pankov and Miller, 2005) and (Semenikhin *et al.*, 2005) investigated linear stochastic differential systems (SDS) with uncertainty in the noise intensity. As the objectives were considered traditional MS criterion and its integral analogue, respectively. In turn, linearity of obtained minimax filters was predetermined by admissability of Gaussian noises in the considered systems.

The Wonham filtering problem, i.e. the one of MS-optimal filtering of finite-state Markov jump process (FSMJP) given its indirect noised continuous and counting observations, differs from the filtering problem in linear Gaussian systems (i.e. the Kalman-Bucy case). First, although the observation system with Markov jump state can be rewritten formally as a linear SDS (Elliott *et al.*, 1995), it is non-Gaussian. Second, the Wonham filtering estimate is essentially nonlinear and represents conditional distribution of FSMJP given available observations. Naturally, the observation (linear stochastic differential!) system under investigation can be studied to design the optimal linear filter, but the obtained estimate would not possess the property of non-negativity, by contrast with any of "genuine" probability distributions. This means, refusal of filter nonlinearity depreciates potential estimation results.

The aim of this paper is correct formulation and solution of estimation problem for FSMJPs under mutual statistical parameter uncertainty in both the state and observation equations.

The paper is organized as follows. Section 2 contains description of investigated observation system. The crucial feature of the system is the triple "transition intensity matrix – continuous observation plan – counting observation intensity" is parameterized by a random vector with known distribution. The section also contains solution of the Bayes estimation problem for both the parameter value and system state. Note, this result is a ground for subsequent minimax inferences.

Section 3 is devoted to detailed formulation of minimax estimation problem. To the best of author's knowledge, analogous problem statement is not presented in literature earlier. First, the optimality criterion contains a conditional expectation of generalized quadratic loss function given obtained observations. Second, the class of admissible estimates includes nonlinear ones. Third, the uncertainty class contains all probability distributions of random parameter with a known fixed support, and the considered observation system is necessarily non-Gaussian. The section also includes the arguments of practical significance of presented "exotic" setting.

Section 4 contains solution of stated minimax estimation problem. The main results concern the saddle point existence in the corresponding minimax optimization problem, characterization of minimax estimate and properties of the least favorable distributions.

Section 5 demonstrates applicability of the presented filter by an illustrative numerical example of TCP link status monitoring under uncertainty.

2. BAYESIAN ESTIMATION IN MARKOV OBSERVATION SYSTEMS

Before consideration of the filtering problem in minimax setting, a problem of mutual Bayesian estimation of FSMJP state and identification of observation system parameters is investigated.

On the finite time interval $[0, T]$ let us consider the observation system

$$\begin{cases} \theta_t = \theta_0 + \int_0^t \Lambda^* \theta_{s-} ds + M_t^{\theta}, \\ N_t = \int_0^t \mu \theta_{s-} ds + M_t^N, \\ U_t = \int_0^t A \theta_{s-} ds + \varepsilon W_t. \end{cases} \tag{1}$$

Here

• $\theta_t \in S_n$ is a unobservable homogenous FSMJP with the state space $S_n = \{e_1, \ldots, e_n\}$ (e_k denotes the k-th unit vector in Euclidian space $\mathbb{R}^n$), with the known initial distribution p_0 and transition intensity matrix $\Lambda \in \mathbb{R}^{n \times n}$,

• $N_t \in \mathbb{R}$ is an observable counting process with intensity $\mu\theta$ dependent on the state θ ($\mu \in \mathbb{R}^{1 \times n}$ is a row vector of rates),

• $U_t \in \mathbb{R}^{m \times 1}$ is an observable continuous process ($A \in \mathbb{R}^{m \times n}$ is an observation plan),

• $W_t \in \mathbb{R}^{m\times 1}$ is a Wiener process representing errors in continuous observations,

• $\varepsilon\varepsilon^* > 0$ is a known nondegenerated observation noise intensity.

The two first equations in (1) define martingale representation of θ and N (Elliott *et al.*, 1995): processes M_t^θ and M_t^N are $\mathcal{F}_t^{\theta,N}$-adapted square integrable martingales with predictable characteristics

$$\langle M^\theta, M^\theta\rangle_t = \int_0^t (\operatorname{diag}(\Lambda^*\theta_{s-}) - \\ -\Lambda^* \operatorname{diag}(\theta_{s-}) - \operatorname{diag}(\theta_{s-})\Lambda)ds,$$
$$\langle M^N, M^N\rangle_t = \int_0^t \mu\theta_{s-}ds.$$

Further in the paper it is supposed that all the transition intensity matrix $\Lambda = \Lambda(\gamma(\omega))$, observation plan $A = A(\gamma(\omega))$ and rate vector $\mu = \mu(\gamma(\omega))$ are known bounded functions of random parameter $\gamma(\omega) \in \mathbb{R}^k$. The functions $\Lambda(v)$ and $\mu(v)$ satisfy usual conditions for intensities: $\lambda_{ij}(v) \geqslant 0$ if $i \neq j$, $\sum_{j=1}^n \lambda_{ij}(v) \equiv 0$ and $\mu_i(v) \geqslant 0$ for $i = 1, 2, \ldots, n$.

To define a probability triplet with filtration for the observation system (1) the following notation is reservated:

• $\mathcal{F}_t \overset{\text{def}}{=} \sigma\{\gamma(\omega), \theta_s(\omega), N_s(\omega), W_s(\omega),\ 0 \leqslant s \leqslant t\}$ is the natural filtration commonly induced by θ, N, W and γ; $\mathcal{F} \overset{\text{def}}{=} \bigvee_{t\in[0,T]} \mathcal{F}_t$,

• $\mathcal{U}_t \overset{\text{def}}{=} \sigma\{N_s(\omega), U_s(\omega)\ 0 \leqslant s \leqslant t\}$ is the natural filtration commonly induced by the observations N and U.

The probability measure $\mathbf{P}_F$, defined on the underlying space $(\Omega, \mathcal{F})$, is known and satisfies the following conditions:

1) the initial condition $\theta_0(\omega)$, random parameter $\gamma(\omega)$ and observation noise $W_t(\omega)$ are mutually independent;

2) $\gamma(\omega) \in \mathcal{C} \subseteq \mathbb{R}^k$ for any $\omega \in \Omega$;

3) $M^N \perp\!\!\!\perp M^\theta$;

4)

$$\mathbf{E}_F\left\{\exp\left[\int_0^t A^*(\gamma)(\varepsilon\varepsilon^*)^{-1}dU_s + \right.\right. \\ +\frac{1}{2}\int_0^t A^*(\gamma)(\varepsilon\varepsilon^*)^{-1}A(\gamma)ds + \\ \left.\left. +\int_0^t \ln(\mu(\gamma)\theta_{s-})dQ_s - \int_0^t (\mu(\gamma)\theta_{s-} - 1)ds\right]\right\} = 1,$$

5) $\mathbf{P}_F\{\omega : \gamma(\omega) \in \mathcal{L}\} = F(\mathcal{L})$, where $F(\cdot)$ is known a priori distribution of γ.

The index $_F$ in the notation $\mathbf{P}_F$ and $\mathbf{E}_F$ indicates for the probability measure and mathematical expectation to be dependent on a priori distribution F of $\gamma(\omega)$.

The Bayesian estimation problem for the vector $z_t \overset{\text{def}}{=} \operatorname{col}(\theta_t, \gamma)$ is to find $\widehat{z}_t^F \overset{\text{def}}{=} \mathbf{E}_F\{z_t|\mathcal{U}_t\}$.

As is known, the required estimate $\widehat{z}_t^F$ is optimal both in the sense of *unconditional MS-criterion:*

$$\widehat{z}_t^F \in \operatorname{Arg}\min_{\overline{z}_t\in\mathcal{M}_t} \mathbf{E}_F\left\{\|z_t - \overline{z}_t\|^2\right\}, \qquad (2)$$

and *its conditional version:*

$$\widehat{z}_t^F \in \operatorname{Arg}\min_{\overline{z}_t\in\mathcal{M}_t} \mathbf{E}_F\left\{\|z_t - \overline{z}_t\|^2|\mathcal{U}_t\right\}. \qquad (3)$$

Above, the set $\mathcal{M}_t$ of admissible estimates $\overline{z}_t$ consists of all $\mathcal{U}_t$-measurable functions. The optimality of $\widehat{z}_t^F$ in the sense of conditional criterion (3) means, the inequality

$$\mathbf{E}_F\left\{\|z_t - \widehat{z}_t^F\|^2|\mathcal{U}_t\right\} \leqslant \mathbf{E}_F\left\{\|z_t - \overline{z}_t\|^2|\mathcal{U}_t\right\}$$

holds $\mathbf{P}_F$-a.s. for any estimate $\overline{z}_t \in \mathcal{M}_t$.

Evidently, the stated problem can be transformed into one of optimal filtering. Note, the martingale representations of θ and N are still valid after replacing of the nonrandom matrix Λ and vector μ by the random ones $\Lambda(\omega)$ and $\mu(\omega)$. The formulae for characteristics $\langle M^\theta, M^\theta\rangle_t$ and $\langle M^N, M^N\rangle_t$ also keep the form.

Consider the observation system with the extended state $z_t \overset{\text{def}}{=} \operatorname{col}(\theta_t, \gamma_t)$:

$$\begin{cases} \theta_t = \theta_0 + \displaystyle\int_0^t \Lambda^*(\gamma_{s-})\theta_{s-}ds + M_t^\theta, \\ \gamma_t = \gamma, \\ N_t = \displaystyle\int_0^t \mu(\gamma_{s-})\theta_{s-}ds + M_t^N, \\ U_t = \displaystyle\int_0^t A(\gamma_{s-})\theta_{s-}ds + \varepsilon W_t, \end{cases} \qquad (4)$$

and define the conditional distributions

$$\widehat{P}_i^F(\mathcal{L}, t) \overset{\text{def}}{=} \mathbf{P}_F\{\gamma_t \in \mathcal{L}, \theta_t = e_i \mid \mathcal{U}_t\}, \quad i = \overline{1,n},$$
$$\widehat{P}^F(\mathcal{L}, t) \overset{\text{def}}{=} \operatorname{col}(\widehat{P}_1^F(\mathcal{L}, t), \ldots, \widehat{P}_n^F(\mathcal{L}, t)).$$

Theorem 1. If conditions 1) — 5) hold for the system (4), then

i) conditional distribution $\widehat{P}^F(\mathcal{L}, t)$ is defined as

$$\widehat{P}^F(\mathcal{L}, t) = K\int_{\mathcal{L}} \widetilde{\theta}_t(q)F(dq), \\ K = \left(\int_{\mathcal{C}} \mathbf{1}\widetilde{\theta}_t(q)F(dq)\right)^{-1}, \qquad (5)$$

where $\widetilde{\theta}_t(q)$ is the unnormalized conditional distribution of θ_t given $\mathcal{U}_t$, calculated by the Wonham filtering algorithm under the assumption $\gamma(\omega) = q$:

$$\widetilde{\theta}_t(q) = p_0 + \int_0^t \Lambda^*(q)\widetilde{\theta}_{s-}(q)ds + \\ + \int_0^t \operatorname{diag}(\widetilde{\theta}_{s-}(q))A^*(q)(\varepsilon\varepsilon^*)^{-1}dU_s + \\ + \int_0^t [\operatorname{diag}\mu(q) - I_{n\times n}]\widetilde{\theta}_{s-}(q)(dN_s - ds), \qquad (6)$$

where $I_{n\times n}$ is the $n \times n$ unit matrix, and $\mathbf{1}$ is a row vector formed by units;

ii) the Bayesian estimate $\widehat{z}_t^F$ of z_t is defined by the formulae

$$\widehat{\theta}_t^F = K \int_{\mathcal{C}} \widetilde{\theta}_t(q) F(dq), \qquad \widehat{\gamma}_t^F = K \int_{\mathcal{C}} q \mathbf{1} \widetilde{\theta}_t(q) F(dq). \tag{7}$$

Note, that (6), being a linear SDS, has a unique strong solution for any fixed parameter $q \in \mathcal{C}$.

3. STATEMENT OF MINIMAX ESTIMATION PROBLEM

Let us consider the observation system (4), for which conditions 1)—4) of previous section are valid, and condition 5) is replaced by

5') the distribution $F(\mathcal{L}) = \mathbf{P}_F\{\omega \in \Omega : \gamma(\omega) \in \mathcal{L}\}$ is a priori unknown. The uncertainty set $\mathbb{F}$ consists of all distributions F concentrated on the fixed known convex compact $\mathcal{C} \subseteq \mathbb{R}^k$.

In view of this uncertainty there is a family of canonical spaces $\mathcal{P}_{\mathbb{F}} \stackrel{\text{def}}{=} \{(\Omega, \mathcal{F}, \mathbf{P}_F, \{\mathcal{F}_t\}_{t \in [0,T]})\}_{F \in \mathbb{F}}$, parameterized by the distribution $F \in \mathbb{F}$.

The set $\mathcal{Z}_t$ of admissible estimators consists of $\mathcal{B}_t$-measurable functions $\varphi : \mathbf{C}^m[0,t] \times \mathbf{B}[0,t] \to \mathbb{R}^{(n+k)\times 1}$ such that $\sup_{F \in \mathbb{F}} \mathbf{E}_F \left\{\|\varphi(O^t)\|^2\right\} < \infty$ (here $\mathbf{B}[0,t]$ denotes a Blackwell space, and $O^t = \{U_s,\ N_s :\ 0 \leqslant s \leqslant t\}$ denotes an observation trajectory occurring on the time interval $[0,t]$).

Let $g = g(x) :\ \mathcal{C} \to \mathbb{R}^l$ be a function such that $\sup_{F \in \mathbb{F}} \mathbf{E}_F \left\{\|g(\gamma(\omega))\|^2\right\} < \infty$. We introduce an auxiliary estimate $\mathbf{g}_t$ for the function $g(\gamma(\omega))$ of the random parameter $\gamma(\omega)$. Namely, $\mathbf{g}_t :\ C^m[0,t] \times \mathbf{B}[0,t] \to \mathbb{R}^{l \times 1}$ is a fixed known $\mathcal{B}_t$-measurable function of observations, such that $\sup_{F \in \mathbb{F}} \mathbf{E}_F \left\{\|\mathbf{g}_t(O^t)\|^2\right\} < \infty$. For example, additional a priori information like $\mathbf{E}_F\{\gamma(\omega)\} = \mathbf{g}$ can be considered as a sort of the auxiliary estimate.

The minimax posterior estimation problem for the state z_t is to find an estimate $\widehat{\mathbf{z}}_t$ such that

$$\widehat{\mathbf{z}}_t \in \operatorname{Arg} \min_{\overline{z}_t \in \mathcal{Z}_t} \sup_{F \in \mathbb{F}} \mathbf{E}_F\{\|z_t - \overline{z}_t(O^t)\|^2_{\Sigma_1} - \|g(\gamma_t) - \mathbf{g}_t(O^t)\|^2_{\Sigma_2} | \mathcal{U}_t\}. \tag{8}$$

Here $\|x\|^2_\Sigma \stackrel{\text{def}}{=} x^* \Sigma x$, and Σ_1 and Σ_2 are known nonnegative weight matrices.

The conditional expectation in the criterion above has transparent interpretation. The point is, any practical estimation problem should be investigated ultimately with respect to the realized observation (trajectory, sample, etc.) Utilization of unconditional minimax criterion implies searching for the worst system parameters irrelative of the available observations. By contrast with this case, conditional minimax criterion (8) is more pessimistic, because it forces to find the worst parameters regarding to the observation system and realized observation trajectory as well.

From the practical point of view, allotment of the parameter γ by the random nature is hardly reasonable. Actually, in most of applied problems γ is a priori unknown but *nonrandom*. Ascription of randomness is only an artificial trick, because under nonrandom settings the corresponding criterion has no saddle point at the feasible set of arguments (see, e.g. (Martin and Mintz, 1983) for the minimax estimation in linear discrete-time systems). On the other hand, in many real problems additional a priori or statistical information $\mathbf{g}_t$ concerning functions $g(\gamma)$ is often available: prior information, estimates, guess values, etc. Note, the auxiliary value $\mathbf{g}_t$ could be either nonrandom or $\mathcal{U}_t$-measurable. We need only to precise it at one time with the Markov state estimation. It is also remarkable, the generalized quadratic loss function in (8) penalizes the distributions F for the deviations of some fixed function $g(\gamma(\omega))$ from its auxiliary value $\mathbf{g}_t$.

4. SOLUTION FOR MINIMAX ESTIMATION PROBLEM

Let us fix some distribution $F \in \mathbb{F}$ and consider the estimate $\widehat{z}_t^F$ calculated by (6) and (7), i.e. conditional expectation corresponding to F. Further, $\widehat{\mathcal{Z}}_t = \{\widehat{z}_t^F :\ F \in \mathbb{F}\}$ denotes a set of all these estimates calculated for each distribution $F \in \mathbb{F}$ given a fixed observation trajectory O^t. Below we also reserve the general notation $\widehat{f}_t^F = \widehat{f}_t^F(O^t) = \mathbf{E}_F\{f_t | \mathcal{U}_t\}$ for conditional expectation of any random value $f_t = f(\theta_t, \gamma_t)$ calculated by the given trajectory O^t:

$$\widehat{f}_t^F = \int_{\mathcal{C}} \sum_{j=1}^n f(e_j, q) \widehat{P}_j^F(dq, t). \tag{9}$$

Theorem 2. i) If the dual optimization problem

$$\widehat{\mathrm{F}} \in \operatorname{Arg} \max_{F \in \mathbb{F}} \left\{ \left(\widehat{\|z_t\|^2_{\Sigma_1}}^F - \|\widehat{z}_t^F\|^2_{\Sigma_1} \right) - \left(\widehat{\|g_t\|^2_{\Sigma_2}}^F - \|\widehat{g}_t^F\|^2_{\Sigma_2} \right) - \|\mathbf{g}_t - \widehat{g}_t^F\|^2_{\Sigma_2} \right\} \tag{10}$$

has a solution (dependence on O^t and γ in the above criterion is omitted), then the function

$$J(\overline{z}_t, F) \stackrel{\text{def}}{=} \mathbf{E}_F\left\{\|z_t - \overline{z}_t\|^2_{\Sigma_1} - \|g_t - \mathbf{g}_t\|^2_{\Sigma_2} | \mathcal{U}_t\right\}$$

has the saddle point $(\widehat{\mathbf{z}}_t, \widehat{\mathrm{F}})$ on the set $\mathcal{Z}_t \times \mathbb{F}$: the least favorable distribution $\widehat{\mathrm{F}}$ is a solution of (10) and $\widehat{\mathbf{z}}_t = \widehat{z}_t^{\widehat{\mathrm{F}}}$ is the estimate calculated by (6), (7) and (9) under the least favorable distribution $\widehat{\mathrm{F}}$; in this case the estimate $\widehat{\mathbf{z}}_t$ is a solution of the minimax posterior estimation problem (8) for the state z_t;

ii) if $\widetilde{\theta}_t(q)$ (6) depends on $q \in \mathcal{C}$ continuously, then the solution of (10) does exist, moreover there exists a version of the least favorable distribution concentrated at most at $n+k+l+2$ points of $\mathcal{C}$.

The uncertainty set $\mathbb{F}$ of admissible distributions in Theorem 2 can be replaced by any subset of $\mathbb{F}$ closed in the weak topology. In this case item i) is still valid, meanwhile fulfillment of item ii) can not be guaranteed.

Item ii) provides existence of some discrete version of the least favorable distribution. In general, this distribution is neither unique nor necessarily discrete. Characterization and further investigation of the whole set of the worst distributions is very complicated problem out of the scope of this paper.

5. NUMERICAL EXAMPLE: MONITORING OF TCP LINK STATE UNDER UNCERTAINTY

Let us consider the Gilbert model (Gilbert, 1960) of TCP link functioning. It is supposed, the unobservable link status is described by a FSMJP θ_t with two possible states: the "good" ($\theta_t = e_1$) and the "bad" ($\theta_t = e_2$) ones. Corresponding observation system, being a specific case of (1), has the following form:

$$\begin{cases} \theta_t = \theta_0 + \displaystyle\int_0^t \begin{bmatrix} -\lambda_1 & \lambda_1 \\ \lambda_2 & -\lambda_2 \end{bmatrix}^* \theta_{s-} ds + M_t^{\theta}, \\ N_t = \displaystyle\int_0^t \begin{bmatrix} \mu_1 & \mu_2 \end{bmatrix} \theta_{s-} ds + M_t^N, \\ U_t = \displaystyle\int_0^t \begin{bmatrix} A_1 & A_2 \end{bmatrix} \theta_{s-} ds + \varepsilon W_t, \end{cases} \quad (11)$$

The transition intensity matrix $\Lambda = \|\lambda_i\|_{i=1,2}$ of θ_t is a priori unknown, but the bounds of elements of Λ are usually available: $\lambda_i \in [\underline{\lambda}_i, \overline{\lambda}_i]$, $i = 1, 2$.

TCP protocol allows to observe a process of packet losses N_t, assumed to be a counting process with intensity $\mu\theta$ dependent on the current link state θ_t. Evidently, $\mu_2 > \mu_1$, i.e. the rate of packet losses in the "bad" state is higher than one in the "good" state. Exact values of the rates are also unknown, but their bounds are available: $\mu_i \in [\underline{\mu}_i, \overline{\mu}_i]$, $i = 1, 2$. Continuous observation U_t represents integral statistical data concerning the Round Trip Time (RTT) history: A_1 and A_2 are unknown expected RTT values in the "good" and "bad" states, respectively ($A_2 > A_1$), and the Wiener process εW_t describes RTT disturbances. A priori uncertainty of vector A is in the same form as ones of Λ and μ: $A_i \in [\underline{A}_i, \overline{A}_i]$, $i = 1, 2$.

The problem is to find a filtering estimate of the TCP link state θ_t under a priori uncertainty of observation system parameters, i.e. to solve a special case of minimax estimation problem (8), when $\Sigma_1 = \mathrm{diag}[I_{2\times 2},\ 0]$ and $\Sigma_2 = 0$.

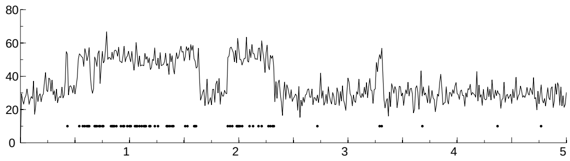

Fig. 1. Available observations: scaled increments of continuous observations $\Delta U_t/\Delta t$ (*solid black line*), and moments of packet losses N_t (*black dots*).

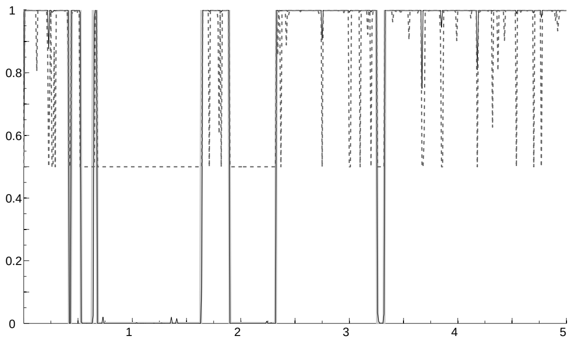

Fig. 2. The indicator of "good" state (*solid gray line*), Wonham estimate given true value of (Λ, μ, A) (*solid black line*) and minimax estimate (*dotted line*).

For simulation the following true parameter values are assigned: $\lambda_1 = 1.9$, $\lambda_2 = 5.1$, $A_1 = 29$, $A_2 = 51$, $\mu_1 = 0.5$, $\mu_2 = 60$, $\varepsilon = 0.5$. The uncertainty is given by the bounds: $\underline{\lambda}_1 = 0.5$, $\overline{\lambda}_1 = 2$, $\underline{\lambda}_2 = 5$, $\overline{\lambda}_2 = 25$, $\underline{A}_1 = 20$, $\overline{A}_1 = 30$, $\underline{A}_2 = 50$, $\overline{A}_2 = 120$, $\underline{\mu}_1 = 0.1$, $\overline{\mu}_1 = 1.1$, $\underline{\mu}_2 = 50$, $\overline{\mu}_2 = 100$.

Figure 1 contains available observations: scaled increments of continuous observations $\frac{\Delta U_t}{\Delta t}$ ($\Delta t = 0.01$), and moments of packet losses N_t. Figure 2 presents the indicator function of the "good" state $\theta_t^* e_1$ in comparison with both the Wonham filtering estimate, calculated given true value of the triple (Λ, μ, A), and presented minimax estimate. Figure 3 presents the indicator $\theta_t^* e_1$ in comparison with both the minimax estimate and the Wonham filtering estimate, calculated with uncertain parameters replaced by the guess values – centers of corresponding uncertainty sets: $\lambda_i^{gv} = \frac{(\underline{\lambda}_i + \overline{\lambda}_i)}{2}$, $\mu_i^{gv} = \frac{(\underline{\mu}_i + \overline{\mu}_i)}{2}$, $A_i^{gv} = \frac{(\underline{A}_i + \overline{A}_i)}{2}$, $i = 1, 2$. Obviously, performance of the minimax filter and the Wonham one, calculated with guess values, can not be correctly compared with one of the Wonham filter calculated under the perfect knowledge of (Λ, μ, A), because of different a priori information used in estimation. For an estimate ν of the state process θ let us consider the averaged $\mathcal{L}_2$ norm of error $\frac{1}{T}\int_0^T \|\theta_t - \nu_t\|^2 dt$. Being calculated for

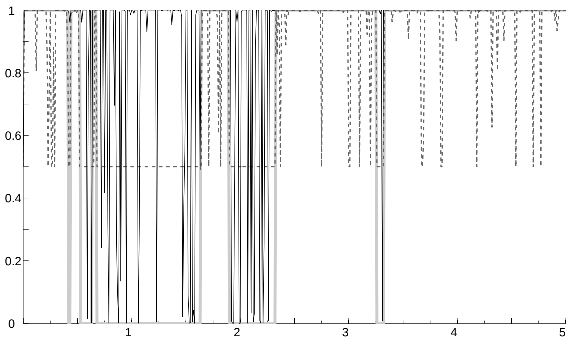

Fig. 3. The indicator of "good" state (*solid gray line*), Wonham estimate calculated under guess values of (Λ, μ, A) (*solid black line*) and minimax estimate (*dotted line*).

$T = 10$ in the example, this performance index is equal to 0.00438 for the "perfect" Wonham filter, 0.1753 for the minimax one, and 0.46246 for the "guess" Wonham filter. The point is, the uncertainty of parameters, corresponding to the "bad" state is wider, and the chosen guess values occur to be far from exact ones. This is a reason the "guess" Wonham filter identifies the "bad" state poorly, demonstrating unstable oscillating behavior (see, e.g., intervals (0.68, 1.63) and (1.9, 2.32) in Figure 3). This means, the only observations are not enough for the "guess" Wonham filter in the "bad" state to estimate this current link status surely under missing a priori information. In this situation the proposed minimax filter turns out to be more effective, suggesting a "case-neutral" uniform estimate $\widehat{\theta}_t = \mathrm{col}(0.5, 0.5)$.

6. CONCLUSIONS

The contributions of this paper are as follows.

1. The problem of estimation in Markov jump observation system under a priori uncertainty is properly stated in terms of game-theoretic framework.

2. The work contains an assertion, specifying a solution of stated minimax optimization problem: saddle point existence, dual optimization problem, defining the least favorable distribution, and the form of desired minimax estimate.

3. Applicability of the minimax estimate is demonstrated by the example of TCP link monitoring given observations of RTT and packet losses under a priori uncertainty of the link parameters.

At the same time we can point following open problems. First, practical implementation of the proposed results requires development of effective numerical schemes, realizing the minimax estimation procedure. Second, it is worth to specify conditions guaranteeing existence of the dual optimization problem solution. Third, investigation of minimax versions for optimal control problems (see, e.g. (Miller *et al.*, 2005)) looks very promising.

REFERENCES

Anan'ev, B.I. (1993). Minimax linear filtering in discrete-time systems with uncertain disturbances. *Automat. Remot. Contr.* **54**(10), 131–139.

Borisov, A.V. and A.R. Pankov (1994). Optimal filtering in stochastic discrete-time systems with unknown inputs. *IEEE Trans. Autom. Contr.* **39**(12), 2461–2464.

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

Gilbert, E.M. (1960). Capacity of a burst-noise channel. *Bell Syst. Tech. J.* **5**, 1253–1265.

Kalman, R.E. and R.C. Bucy (1961). New results in linear filtering and prediction problems. *Trans. ASME.* **D**, 95–111.

Kats, I.Ya. and A.B. Kurzhanskii (1978). Minimax discrete-time filtering under statistical uncertainty. *Automat. Remot. Contr.* **39**(11), 79–87.

Martin, C.J. and M. Mintz (1983). Robust filtering and prediction for linear systems with uncertain dynamics: A game-theoretic approach. *IEEE Trans. Autom. Contr.* **28**(9), 888–896.

Miller, B.M., K.E. Avrachenkov, K.V. Stepanyan and G.B. Miller (2005). Flow control as stochastic optimal control problem with incomplete information. In: *Proceedings of INFOCOM'2005.* Miami. pp. 1328–1337.

Pankov, A.R. and G.B. Miller (2005). Random process in a statistically uncertain linear stochastic differential system. *Automat. Remot. Contr.* **66**(1), 53–64.

Semenikhin, K.V., M.V. Lebedev and E.N. Platonov (2005). Kalman filtering by minimax criterion with uncertain noise intensity functions. In: *Proceedings of IEEE CDC-ECC'2005.* Seville. pp. 1929–1934.

Wonham, W.N. (1964). Some applications of stochastic differential equations to optimal nonlinear filtering. *SIAM J. Contr.* **2**(3), 347–369.