Applied Mathematical Modelling 36 (2012) 1347–1358

Contents lists available at SciVerse ScienceDirect

Applied Mathematical Modelling

journal homepage: www.elsevier.com/locate/apm

Recursive state estimation for hybrid systems

Evgenia Suzdaleva [a,*], Ivan Nagy [b]

[a] Department of Adaptive Systems, Institute of Information Theory and Automation of the ASCR, Pod vodárenskou věží 4, 18208 Prague, Czech Republic
[b] Faculty of Transportation Sciences, Czech Technical University, Na Florenci 25, 11000 Prague, Czech Republic

ARTICLE INFO

Article history:
Received 12 January 2011
Received in revised form 15 August 2011
Accepted 16 August 2011
Available online 26 August 2011

Keywords:
Recursive state estimation
Hybrid systems
State-space model
Filtering
Mixed data

ABSTRACT

The paper deals with recursive state estimation for hybrid systems. An unobservable state of such systems is changed both in a continuous and a discrete way. Fast and efficient online estimation of hybrid system state is desired in many application areas. The presented paper proposes to look at this problem via Bayesian filtering in the factorized (decomposed) form. General recursive solution is proposed as the probability density function, updated entry-wise. The paper summarizes general factorized filter specialized for (i) normal state-space models; (ii) multinomial state-space models with discrete observations; and (iii) hybrid systems. Illustrative experiments and comparison with one of the counterparts are provided.

1. Introduction

Systems whose state is changing dynamically continuously in time and also switching among several discrete values are understood as hybrid systems. A state of a hybrid system is modeled by continuous variables within several discrete modes, among them a system is switching. Usually system parameters are changing according to a particular mode. Hybrid systems are widely used in many fields of signal processing (target tracking, medicine, speech recognition, traffic control etc.). Fast and efficient online estimation of their state is desired in some of these areas.

A lot of works are devoted to state estimation of hybrid systems. One of the well-known approaches dealing with switching systems with Gaussian linear and discrete states is the interactive multiple model (IMM) algorithm [1]. It performs classical Kalman filter [2] for each mode under assumption that this particular mode is a right one at current time step. Then the IMM algorithm computes a weighted combination of updated state estimates produced by all the filters yielding a final Gaussian mean and covariance. This mixed state estimate is taken as the initial one for the next time step. The weights are chosen according to the probabilities of the models, which are computed in filtering step of the algorithm.

The paper [3] proposes the exact filter for a specialized hybrid system state. The reference probability method for hidden Markov models (HMM) is employed. The solution is presented as Gaussian sum with explicitly computed specific weights, means and variances. However, a number of statistics grows geometrically in time, and provided results are restricted only by 15 time steps. The approach [4] considers another special case of a dynamic linear state-space model with measurement matrices switching according to time-varying independent random process. The updating of probabilities is derived as an application of Bayes rule to the weighted observation model. The estimation of the normal state is shown as extension of the classical Kalman filter with involved weighted combinations of the gain-adjusted innovations.

Iterative techniques for jump Markov linear systems are nicely presented in [5]. The algorithms are derived to obtain the marginal maximum a posteriori sequence estimate of the finite state Markov chain. The paper [6] is concerned with optimal

* Corresponding author.
E-mail addresses: suzdalev@utia.cas.cz (E. Suzdaleva), nagy@utia.cas.cz (I. Nagy).

0307-904X/$ - see front matter
doi:10.1016/j.apm.2011.08.042

filtering for hybrid systems with non-Gaussian noises. The derived filter is optimal in the sense of the most probable trajectory (MPT) estimate. The state and the observation are considered as a pair of deterministic processes with switching coefficient as a random process. Despite the claimed generality of solution, this can restrict application in practice. The paper [7] proposes mixture Kalman filter based on a special sequential Monte Carlo method using a random mixture of Gaussian distributions for approximation of target posterior distribution. The approach deals with conditional dynamic linear models (CDLM) with mixed Gaussian noises defined via known indicator process. The weighted sample of the indicators is used within the proposed effective filter. A series of other research in the field of nonlinear hybrid systems [8] and online real-time state estimation [9] can be also found.

The presented paper is focused on modeling of system states as conditionally dependent entries of the state vector. Their entry-wise *recursive* estimation is subsequently reached via factorization of the state-space model and prior distributions for Bayesian filtering [10]. A part of the work dealing with estimation of discrete state is also closely related to algorithms based on hidden Markov models (HMM) theory [11]. However, these algorithms run in offline mode supported by Monte Carlo computations. Important features of the proposed theory are that:

- the algorithms used run in online mode,
- numerical procedures are applied only in that parts, which cannot be computed analytically. In this way the amount of computations as well as the risk of collapsing is minimized,
- general probabilistic approach is universal for the distributions used,
- it opens a way to recursive estimation of discrete system modes dependent on evolution of continuous states. This is planned for future research.

The paper is structured as follows. Necessary preliminaries are provided in Section 2. Section 3 presents general probabilistic solution of the factorized form of Bayesian filtering. The paper summarizes general factorized filter specialized for (i) normal state-space models in Section 4; (ii) multinomial state-space models with discrete-valued observations in Section 5 and (iii) hybrid systems in Section 6. Section 7 demonstrates examples with real data and comparison with the IMM filter. Remarks in Section 8 close the paper. Derivations of the proposed formulas are provided in Appendix A.

2. Preliminaries

2.1. State-space model

The system is described by the state-space model in the form of the following conditional probability (density) functions (p(d)fs) for simplicity denoted as pdfs within this paper

$$\textit{observation model} \quad f(y_t|x_t, u_t), \tag{1}$$

$$\textit{state evolution model} \quad f(x_{t+1}|x_t, u_t), \tag{2}$$

where the system output y_t and the control input u_t are measured at discrete time moments $t = \{1, \ldots, T\} \equiv t^*$. In general, the variables are column vectors such that $y_t = [y_{1;t}, \ldots, y_{Y;t}]'$, $u_t = [u_{1;t}, \ldots, u_{U;\,t}]'$. The system state $x_t = [x_{1;t}, \ldots, x_{X;t}]'$ is not directly observed and has to be estimated in an online (*recursive*) mode.

2.2. Bayesian filtering

Bayesian filtering, estimating the system state, includes the following coupled formulas.

Data updating

$$f(x_t|D(t)) = \frac{f(y_t|x_t, u_t) f(x_t|D(t-1))}{\int f(y_t|x_t, u_t) f(x_t|D(t-1)) dx_t} \propto f(y_t|x_t, u_t) f(x_t|D(t-1)), \tag{3}$$

incorporates information contained in observations $D(t) = (d_1, \ldots, d_t)$, where $d_t \equiv (y_t, u_t)$. Relation (3) also comprises the natural conditions of control [12], according to those

$$f(x_t|u_t, D(t-1)) = f(x_t|D(t-1)).$$

Time updating

$$f(x_{t+1}|D(t)) = \int f(x_{t+1}|x_t, u_t) f(x_t|D(t)) dx_t, \tag{4}$$

fulfills state prediction. The prior pdf $f(x_1|D(0))$ which expresses the subjective prior knowledge on the system initial state starts the recursions. Application of (3,4) to linear Gaussian state-space model provides Kalman filter [12].

2.3. Chain rule

An operation intensively used throughout the paper is:

Chain rule

$$f(a,b|c) = f(a|b,c)f(b|c), \tag{5}$$

which decomposes the joint pdf $f(a,b|c)$ into a product of conditional pdfs for any random variables a, b and c.

3. General solution in a factorized form

Bayesian filtering (3,4) is proposed to be done in one integration step, i.e.,

$$f(x_{t+1}|D(t)) \propto \int f(x_{t+1}|x_t,u_t)\left\{\underbrace{f(y_t|x_t,u_t)f(x_t|D(t-1))}_{\propto f(x_t|D(t))}\right\}dx_t, \tag{6}$$

which is obtained by a trivial substitution of the state estimate updated by measurements (3) in the time updating (4).

A basic idea of the approach is to apply the chain rule (5) to models (1,2) and to (6). Afterwards, models (1,2) are factorized as

$$f(y_t|x_t,u_t) = \prod_{j=1}^{Y} f(y_{j;t}|y_{(j+1):Y;t}, x_{1:X;t}, u_{1:U;t}), \tag{7}$$

$$f(x_{t+1}|x_t,u_t) = \prod_{i=1}^{X} f(x_{i;t+1}|x_{(i+1):X;t+1}, x_{1:X;t}, u_{1:U;t}), \tag{8}$$

that is a product of factors that are conditional pdfs of corresponding distributions. A notation of the form $x_{(i+1):X;t}$ denotes a sequence $\{x_{i+1;t}, x_{i+2;t}, \ldots, x_{X;t}\}$ for current time instant t, which is empty, when $(i+1) \geqslant X$.

Substitution of (7,8) in (6) and application of the chain rule to the prior pdf $f(x_t|D(t-1))$ provide the following factorized form of (6), i.e.,

$$\prod_{i=1}^{X} f(x_{i;t+1}|x_{(i+1):X;t+1}, D(t)) \propto \int \prod_{i=1}^{X} f(x_{i;t+1}|x_{(i+1):X;t+1}, x_{1:X;t}, u_{1:U;t}) \prod_{j=1}^{Y} f(y_{j;t}|y_{(j+1):Y;t}, x_{1:X;t}, u_{1:U;t}) \prod_{i=1}^{X} f(x_{i;t}|x_{(i+1):X;t}, D(t-1))dx_t, \tag{9}$$

where integration is assumed to be done over $x_t = [x_{1;t}, \ldots, x_{X;t}]'$. Formal factorization into the factors helps in designing the resulting algorithms as all the factors are scalar pdfs of respective distributions.

4. Factorized filter for linear normal models

Let us apply the proposed factorized solution (9) to linear normal state-space model. In this field, the sequential Kalman filter [13] can be found closely related to the proposed one. In contrast to the sequential filter, the factorized solution is not restricted by a diagonal measurement covariance matrix (as well as the process one). This is a significant benefit of the approach, since full covariances contribute to a better quality of estimation of normally distributed state. Furthermore, factorization of covariance matrices for Kalman filtering is often aimed at more computational stability via a lesser rank of the matrix, e.g. the Square-Root and U-D Kalman filters [13]. The presented algorithm exploits matrix factorization for reaching the entry-wise updating of state estimate.

The normal observation model (1) has the form

$$f(y_t|u_t,x_t) \equiv \mathcal{N}_{y_t}\left(\underbrace{Cx_t + Hu_t}_{mean}, \overbrace{R_v}^{covariance}\right) = (2\pi)^{-\frac{Y}{2}}|R_v|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}\underbrace{[y_t - Cx_t - Hu_t]'R_v^{-1}[y_t - Cx_t - Hu_t]}_{Q_y}\right\}, \tag{10}$$

where $\mathcal{N}(\cdot)$ denotes normal distribution; C and H are parameters supposed to be known or estimated offline; R_v is a known covariance matrix of the measurement Gaussian noise with zero mean; Q_y denotes a quadratic form inside the exponent. Similarly, the state evolution model (2) is

$$f(x_{t+1}|u_t,x_t) \equiv \mathcal{N}_{x_{t+1}}(Ax_t + Bu_t, R_w) = (2\pi)^{-\frac{X}{2}}|R_w|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}\underbrace{[x_{t+1} - Ax_t - Bu_t]'R_w^{-1}[x_{t+1} - Ax_t - Bu_t]}_{Q_x}\right\}, \tag{11}$$

where A and B are known parameters of appropriate dimensions; R_w is a known covariance matrix of the process Gaussian noise with zero mean; and Q_x is a quadratic form inside the exponent.

Application of recursion (9) to normal models (10,11) leads to a factorized version of Kalman filter. For normal distribution, the posterior state estimate preserves its form

$$\prod_{i=1}^{X} f(x_{i;t+1}|x_{(i+1):X;t+1}, D(t)), \tag{12}$$

for factors via LDL' decomposition [10] of the precision (i.e., inverse covariance) matrices. Such a decomposition supposes L to be a lower triangular matrix with unit diagonal, D to be a diagonal one and $'$ denoting transposition. This type of matrix decomposition is an analogue of factorization (7,8) via the chain rule for normal models (10,11).

The factorization of (10,11) can be clearly demonstrated via exploitation of the quadratic forms Q_y and Q_x. Let us firstly factorize the observation model (10). Matrix R_v is inverted into a precision matrix and decomposed so that

$$R_v^{-1} = L_v D_v L_v'. \tag{13}$$

The resulted quadratic form corresponding to normal distribution (10) is

$$Q_y = \left[L_v' y_t - \underbrace{L_v' H u_t}_{\rho_t} - \underbrace{L_v' C}_{\mathcal{A}} x_t \right]' D_v \left[L_v' y_t - \rho_t - \mathcal{A} x_t \right], \tag{14}$$

which helps to express the jth output factor as scalar pdf

$$\mathcal{N}_{y_{j;t}} \left(\underbrace{\rho_{j;t} - \sum\nolimits_{k=j+1}^{Y} L_{v;kj} y_{k;t} + \sum\nolimits_{l=1}^{X} \mathcal{A}_{jl} x_{l;t}}_{mean\ \ value}, \overbrace{\frac{1}{D_{v;jj}}}^{variance} \right), \tag{15}$$

where $L_{v;kj}$, $\mathcal{A}_{jl}$ and $D_{v;jj}$ are elements of matrices L_v, $\mathcal{A}$ and D_v respectively.

Normal model (11) is factorized quite similarly via the following operations with matrix R_w and quadratic form Q_x, i.e.,

$$R_w^{-1} = L_w D_w L_w', \tag{16}$$

$$Q_x = \left[L_w' x_{t+1} - \underbrace{L_w' B u_t}_{z_t} - \underbrace{L_w' A}_{\Xi} x_t \right]' D_w \left[L_w' x_{t+1} - z_t - \Xi x_t \right], \tag{17}$$

resulting into normal factor of the ith state

$$\mathcal{N}_{x_{i;t+1}} \left(z_{i;t} - \sum_{k=i+1}^{X} L_{w;ki} x_{k;t+1} + \sum_{l=1}^{X} \Xi_{il} x_{l;t}, \frac{1}{D_{w;ii}} \right), \tag{18}$$

where $L_{w;ki}$, Ξ_{il} and $D_{w;ii}$ are elements of matrices L_w, Ξ and D_w respectively.

The prior state distribution is chosen as the normal one with mean μ_t and covariance matrix P_t for $t = 1$. It is transformed to a similar form as follows.

$$P_t^{-1} = L_{p|t} D_{p|t} L_{p|t}', \tag{19}$$

$$Q_{p|t} = \left[L_{p|t}' x_t - \mu_t^f \right]' D_{p|t} \left[L_{p|t}' x_t - \mu_t^f \right], \quad \text{where } \mu_t^f = L_{p|t}' \mu_t, \tag{20}$$

where $Q_{p|t}$ is a resulted quadratic form for the normal distribution of the initial state x_t, which enables expressing the prior factorized state estimate as

$$\mathcal{N}_{x_{i;t}} \left(\mu_{i;t}^f - \sum_{k=i+1}^{X} L_{p|t;ki} x_{k;t}, \frac{1}{D_{p|t;ii}} \right), \tag{21}$$

where $L_{p|t;ki}$ and $D_{p|t;ii}$ are elements of matrices $L_{p|t}$ and $D_{p|t}$ respectively.

Usage of the quadratic forms (14), (17) and (20) allows to represent an elegant form of solution (9) for normal models. Substituting the factorized distributions in (9) and after all rearrangements, one obtains the posterior state estimate in the preserved form (12), or precisely (21), for the ith factor, i.e.,

$$\mathcal{N}_{x_{i;t+1}} \left(\mu_{i;t+1}^f - \sum_{k=i+1}^{X} L_{p|t+1;ki} x_{k;t+1}, \frac{1}{D_{p|t+1;ii}} \right) \tag{22}$$

for that it holds

$$Q_{p|t+1} = \Bigg[\underbrace{L'_{u|t}L'_{w}}_{L'_{p|t+1}}x_{t+1} - \mu^f_{t+1}\Bigg]' \underbrace{D_{u|t}}_{D_{p|t+1}} \Big[L'_{p|t+1}x_{t+1} - \mu^f_{t+1}\Big], \tag{23}$$

$$\mu^f_{t+1} = L'_{u|t}\Big(z_t + \widetilde{D}_t^{-1}\Big(D_w \Xi \Gamma_t^{-1}\Big(\mathcal{A}' D_v(L'_v y_t - \rho_t) + L_{p|t}D_{p|t}\mu^f_t\Big)\Big)\Big) \tag{24}$$

and

$$D_w - D_w \Xi \Gamma_t^{-1} \Xi' D_w = \widetilde{D}_t = L_{u|t}D_{u|t}L'_{u|t}, \tag{25}$$

$$\Gamma_t = \Big[\Xi; \mathcal{A}; L'_{p|t}\Big]' \Omega_t \Big[\Xi; \mathcal{A}; L'_{p|t}\Big], \tag{26}$$

$$\Omega_t = \operatorname{diag}[D_w, D_v, D_{p|t}]. \tag{27}$$

Detailed derivations can be found in Appendix A.

4.1. Algorithm 1

The obtained results can now be summarized in the form of an algorithm.

Initial part of the algorithm

1. Load data y_t, u_t and parameters A, B, C, H, R_w and R_v.
2. Set prior values μ_t and P_t.
3. Factorize the observation model (10) according to (13) and obtain

$$\rho_t = L'_v H u_t,$$
$$\mathcal{A} = L'_v C.$$

4. Factorize the state evolution model (11) according to (16) and compute

$$z_t = L'_w B u_t,$$
$$\Xi = L'_w A.$$

5. Factorize the prior distribution according to (19,20) to obtain $L_{p|t}$, $D_{p|t}$ and μ^f_t.

Online part of the algorithm

For time t from 1 to T

1. Make diagonal matrix $\Omega_t = \operatorname{diag}[D_w, D_v, D_{p|t}]$.
2. Compute $\Gamma_t = \Big[\Xi; \mathcal{A}; L'_{p|t}\Big]' \Omega_t \Big[\Xi; \mathcal{A}; L'_{p|t}\Big]$.
3. Compute matrix $\widetilde{D}_t = D_w - D_w \Xi \Gamma_t^{-1} \Xi' D_w$.
4. Factorize matrix $\widetilde{D}_t = L_{u|t}D_{u|t}L'_{u|t}$.
5. Compute the factorized state estimate (22) according to (23) and (24), i.e.,

$$\mu^f_t = L'_{u|t}\Big(z_t + \widetilde{D}_t^{-1}\Big(D_w \Xi \Gamma_t^{-1}\Big(\mathcal{A}' D_v(L'_v y_t - \rho_t) + L_{p|t}D_{p|t}\mu^f_t\Big)\Big)\Big),$$
$$L'_{p|t} = L'_{u|t}L'_w,$$
$$D_{p|t} = D_{u|t}.$$

End of the cycle for t.

5. Factorized filter for discrete models

Let us apply the proposed factorized solution (9) to discrete models with multinomial distribution. In this area the HMM approaches are widely used. However, the presented online filter is based on explicit solution and avoids Monte Carlo computations. Here factors are obtained naturally since multivariate discrete variables are reduced to scalars with finite number of possible values.

The multinomial observation model (1), i.e.,

$$f(y_t|x_t, u_t) \tag{28}$$

is provided by the output transition table and a known (or estimated offline) probability $\alpha_{q|l,n}$ with multi-index $q|l,n$. This multi-index denotes realizations $q \in \{1,\ldots,Q\}$ of random discrete variable y_t at time instant t according to a set of its possible values $\{1,\ldots,Q\}$, where Q is a finite number. Realization q in the multi-index $q|l,n$ is conditioned by realizations $l \in \{1,\ldots,L\}$ of discrete state x_t and $n \in \{1,\ldots,N\}$ of discrete input u_t from their sets of possible values with finite numbers L and N. Notation $\alpha_{q|l,n}$ reflects probability of transition of output y_t to the discrete value q, i.e., $y_t = q$ conditioned by $x_t = l$ and $u_t = n$. It holds

$$\sum_{q=1}^{Q} \alpha_{q|l,n} = 1, \quad \text{and} \quad \alpha_{q|l,n} \geqslant 0 \quad \forall q, l, n.$$

Similarly, the state evolution model (2)

$$f(x_{t+1}|x_t, u_t), \tag{29}$$

is the multinomial distribution presented by the state transition table containing known probability $\beta_{l|m,n}$ with a multi-index $l|m,n$. Here the multi-index is evolved in a similar way as for the observation model but the condition $m \in \{1,\ldots,L\}$, which relates to value of the discrete state x_t at time instant t, while l here belongs to x_{t+1}. It holds

$$\sum_{l=1}^{L} \beta_{l|m,n} = 1, \quad \text{and} \quad \beta_{l|m,n} \geqslant 0 \quad \forall l, m, n.$$

The prior distribution of the discrete state is chosen as the multinomial one

$$f(x_t|D(t-1)) = p_{x_t}, \tag{30}$$

that has the form of a vector containing the initial probabilities $p_l\ \forall l \in \{1,\ldots,L\}$ at time instant t, and it has to be recursively estimated for time $t+1$. It holds

$$\sum_{l=1}^{L} p_l = 1, \quad \text{and} \quad p_l \geqslant 0 \quad \forall l.$$

Substituting models (28,29) in (9) (here precisely (6)) with incorporation of the prior distribution (30), one obtains the following expression, which simultaneously updates the estimate by actual measurements and predicts the state, i.e.,

$$f(x_{t+1}|D(t)) \propto \sum_{x^*} f(x_{t+1}|x_t, u_t) f(y_t|x_t, u_t) f(x_t|D(t-1)), \tag{31}$$

where integration is replaced by regular summation. For each value $l \in \{1,\ldots,L\}$ of x_{t+1} and with discrete observations $y_t = q \in \{1,\ldots,Q\}$ and $u_t = n \in \{1,\ldots,N\}$ available at time instant t the predicted probability p_l for time instant $t+1$ is explicitly computed as

$$p_l = \beta_{l|1n}\alpha_{q|1n}p_1 + \beta_{l|2n}\alpha_{q|2n}p_2 + \cdots + \beta_{l|Ln}\alpha_{q|Ln}p_L, \tag{32}$$

and then normalized, i.e.,

$$p_l = \frac{p_l}{\sum_{l=1}^{L} p_l},$$

resulting in the multinomial distribution

$$f(x_{t+1}|D(t)) = p_{x_{t+1}}, \tag{33}$$

which preserves the original form (30) and can be used for the next step of recursive estimation.

6. Factorized filter for hybrid systems

Let us consider a hybrid system with the observed output $y_t = [y_t^c, y_t^d]'$ with $y_t^c = \left[y_{1;t}^c, \ldots, y_{Y-1;t}^c\right]$ and $y_t^d = y_{Y;t}$, where superscript c denotes a continuous type of a variable, while superscript d belongs to a discrete variable. Here the case both with normally distributed and multinomial variables is considered. The control input is similarly $u_t = [u_t^c, u_t^d]' = \left[u_{1;t}^c, \ldots, u_{U-1;t}^c, u_t^d\right]'$, and the unobserved state to be estimated is

$$x_t = \left[x_t^c, x_t^d\right]' = \left[x_{1;t}^c, \ldots, x_{X-1;t}^c, x_{X;t}\right]', \quad \text{where } x_{X;t} = x_t^d.$$

Factorization of pdfs shown in (9) allows to represent it in the following way

$$\begin{aligned}
&\prod_{i=1}^{X-1} f\left(x_{i;t+1}^c | x_{(i+1):X;t+1}, D(t)\right) f\left(x_{t+1}^d | D(t)\right) \\
&\quad \propto \int_{x^{c*}} \sum_{x^{d*}} \underbrace{\prod_{i=1}^{X-1} f\left(x_{i;t+1}^c | x_{(i+1):X;t+1}, x_{1:X-1;t}^c, u_t^c\right) f\left(x_{t+1}^d | x_t^d, u_t^d\right)}_{f(x_{t+1}|x_t, u_t)} \\
&\quad \times \underbrace{\prod_{j=1}^{Y-1} f\left(y_{j;t}^c | y_{(j+1):Y;t}, x_{1:X-1;t}^c, u_t^c\right) f\left(y_t^d | x_t^d, u_t^d\right)}_{f(y_t|x_t, u_t)} \\
&\quad \times \underbrace{\prod_{i=1}^{X-1} f\left(x_{i;t}^c | x_{(i+1):X;t}, D(t-1)\right) f\left(x_t^d | D(t-1)\right)}_{prior\ \ pdf}, dx_t^c
\end{aligned} \tag{34}$$

with assumptions that continuous entries can be omitted from the condition for discrete state and output, and the past discrete state and a discrete input – from the condition for y_t^c, as well as u_t^d for x_{t+1}^c. Using relations (3,4), (6) and (31), one can see that the prescribed original form of the posterior pdf is destroyed in (34): it is a sum of distributions

$$\sum_{x^{d*}} f\left(x_{t+1}^d|D(t)\right)\prod_{i=1}^{X-1} f\left(x_{i;t+1}^c|x_{(i+1):X;t+1},D(t)\right). \tag{35}$$

It is necessary to restore the original form to use it for the next step of estimation. An approximation based on Kerridge inaccuracy [14] is an explicit solution, which restores the original form of the pdf via computation of a specific weighted combination of the pdfs involved in (35). Kerridge inaccuracy is a part of Kullback–Leibler divergence [15] adopted as a theoretically justified proximity measure. This divergence is known to be an optimal tool within the Bayesian approach [10]. For any random variable a, Kerridge inaccuracy is used to measure the proximity of pdfs $f(a)$ and $\hat{f}(a)$

$$K_a(f(a)\|\hat{f}(a)) = \int_{a*} f(a)\ln\frac{1}{\hat{f}(a)}da, \tag{36}$$

and its minimization allows to find the approximated pdf $\hat{f}(a)$. According to this approximation [10], the original form of pdf is restored and the product

$$\prod_{i=1}^{X-1}\hat{f}\left(x_{i;t+1}^c|x_{(i+1):X;t+1},D(t)\right)f\left(x_{t+1}^d|D(t)\right), \tag{37}$$

is used as the prior pdf for the next step of recursive estimation (34).

Let us apply the presented solution for the system with normal factors provided by (15), (18) and (21) and discrete factors from (28)–(30).

Solution (34) related to normal factors coincides with that proposed in Section 4 running for each value l of discrete state. A part of (34) outside the integral corresponds to discrete factors and is explained in Section 5.

Relation (35) in this case is the mixture distribution

$$\sum_{l=1}^{L} p_l \prod_{i=1}^{X-1}\mathcal{N}_{x_{i;t+1}}\left(\mu_{i;t+1}^f - \sum_{k=i+1}^{X-1} L_{p|t+1;ki}x_{k;t+1}^c, \frac{1}{D_{p|t+1;ii}}\right).$$

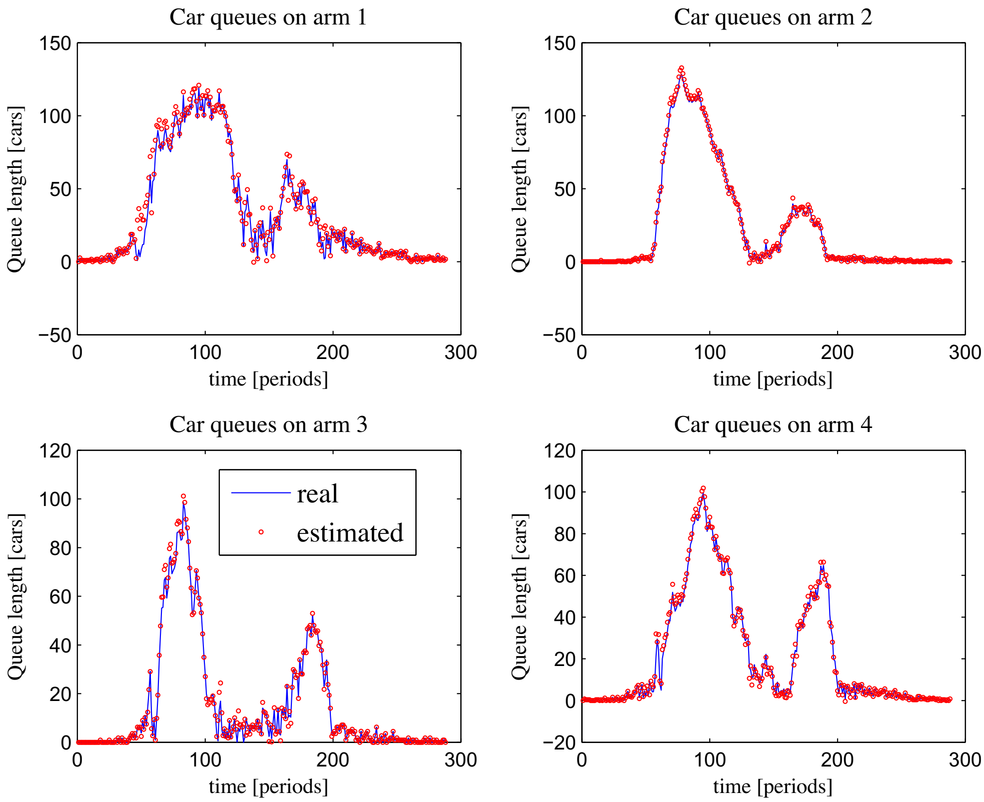

Fig. 1. Queue length estimation with the proposed filter.

Restoring the original normal form needs to use the approximation based on Kerrigde inaccuracy [14]. According to [10], for the case of normal pdfs the Kerridge inaccuracy (36) is minimized with the following mean and covariance matrix of the approximated distribution

$$\hat{\mu}_{t+1} = \sum_{l=1}^{L} p_l \mu_{l,t+1}, \quad \text{where } \mu_{l,t+1} = \left(L'_{l(p|t+1)}\right)^{-1} \mu^f_{l(t+1)}, \tag{38}$$

$$\widehat{P}_{t+1} = \sum_{l=1}^{L} p_l P_{l,t+1} + \sum_{l=1}^{L} p_l (\hat{\mu}_{t+1} - \mu_{l,t+1})^2, \tag{39}$$

where

$$P_{l,t+1} = \left(L_{l(p|t+1)} D_{l(p|t+1)} L'_{l(p|t+1)}\right)^{-1},$$

where subscript l denotes results obtained for each value l of discrete state. The approximation (38) is then factorized according to (19,20) and used as the prior normal distribution for the next step of the recursion.

To summarize the obtained solution, one can structure it as follows.

1. Compute the state estimate for discrete factors, see Section 5.
2. Compute the normal state estimates, see Section 4.1, running the algorithm for each discrete state value.
3. Restore the original form via (38) and factorize it.

Note that the above specialization is shown for vector

$$x_t = \left[x^c_{1;t}, \ldots, x^c_{X-1;t}, x_{X;t}\right]',$$

where $x_{X;t} = x^d_t$. For another case, for instance, $\left[x^c_{1;t}, \ldots, x_{X-1;t}, x^c_{X;t}\right]'$ with $x_{X-1;t} = x^d_t$ and $x^c_{X;t}$ one should use distributions modeling discrete variables dependent on continuous ones. The proposed factorization enables to consider this task that will be presented soon.

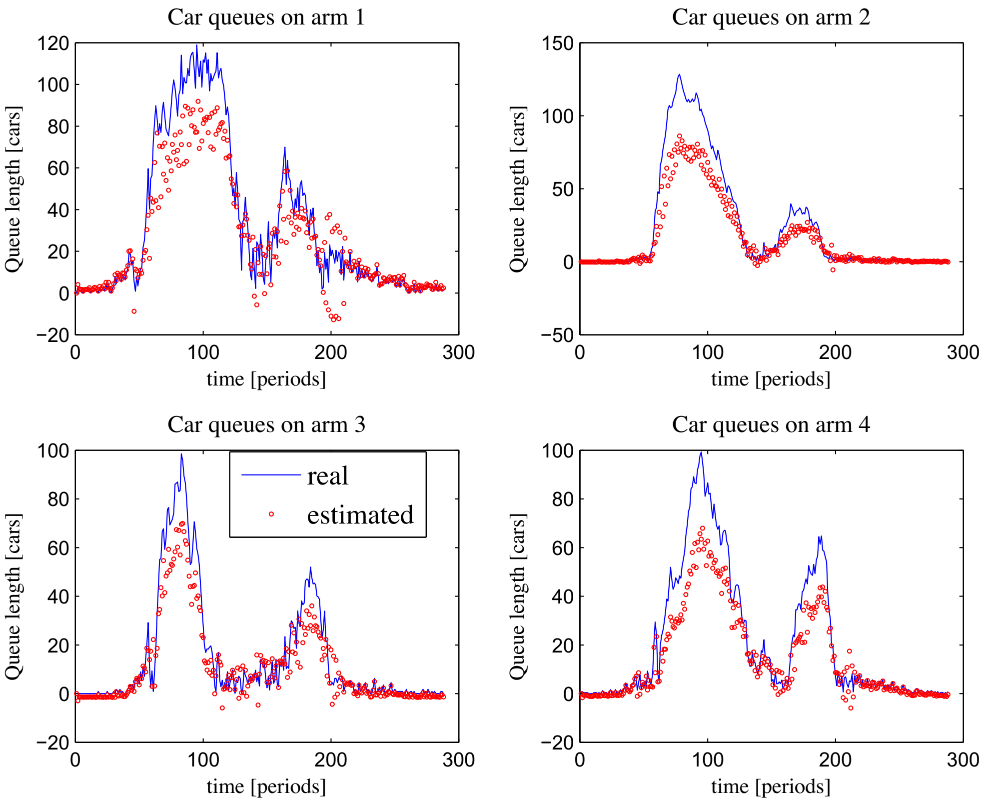

Fig. 2. Queue length estimation with the IMM filter.

7. Experiments

To test the proposed approach, a real data sample containing intensity (number of cars per time unit) of the traffic flow in a chosen point of traffic communications in Prague has been taken. In practice in the field of traffic-flow control, fast online

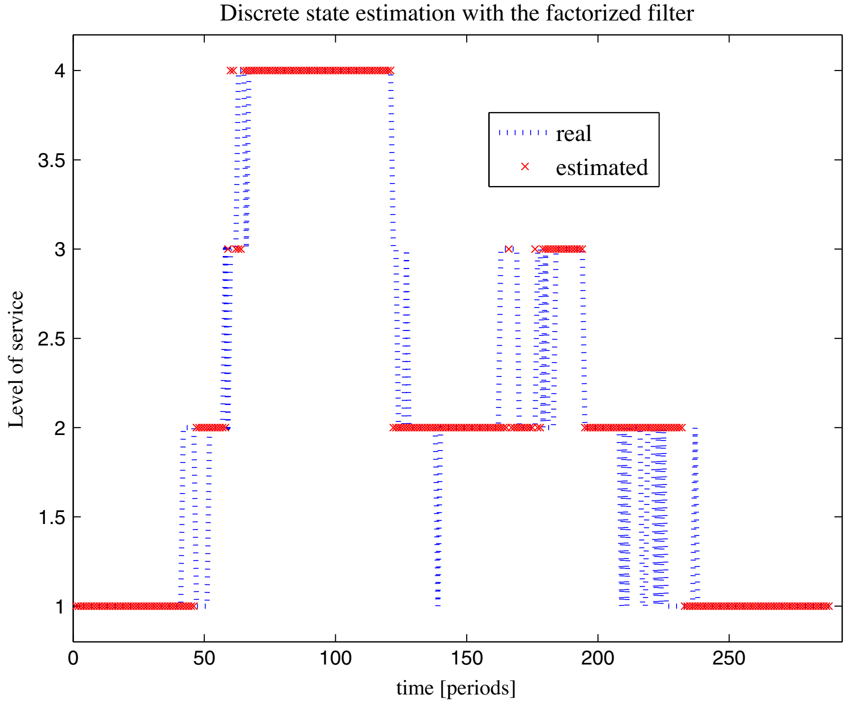

Fig. 3. LoS estimation with the proposed filter.

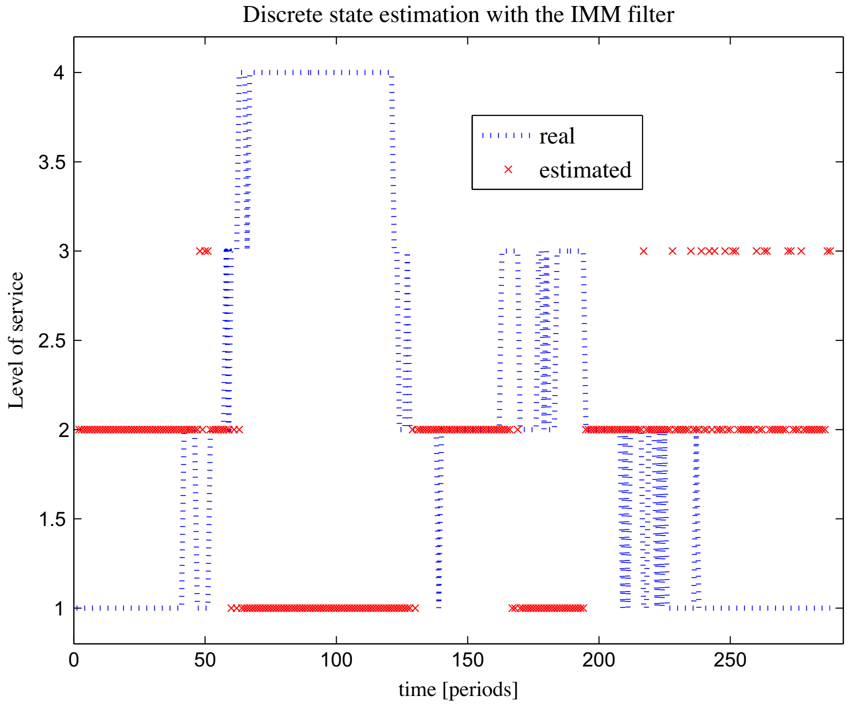

Fig. 4. LoS estimation with the IMM filter.

Table 1
Estimation error and correct point estimates.

	EE	CPE
Factorized filter	0.0296	240
IMM filter	5.5762	70

state estimation is important: it can influence the adaptive control of the microregion via adequate green light time. A normally distributed state x_t^c of the considered hybrid system is a four-dimensional queue length of cars waiting for passing through a traffic microregion. A full dimension of the taken normal state is eight, since occupancy of a measured detector is added to the vector to ensure observability of the model. A discrete state x_t^d is a level of service (LoS) of the microregion. It expresses a degree of traffic saturation in that sense how easy cars can pass through the microregion with 4 possible values: from 1 (the best) to 4 (the worst).

The measured data used were: y_t^c – car outgoing intensity along with occupancy of a measured detector; y_t^d – a time mode of a workday (morning peak-hour time, lunch, late afternoon peak-hour time, evening); u_t^c – a relative time of the green light; u_t^d – a discrete variable, reflecting whether the saturated strategy of the adaptive control is used or not. A duration of the online filtering was 1 workday, which corresponds to 288 time periods. The filtering started at midnight that simplifies a choice of prior distributions (i.e., zero queue length and LoS = 1).

The state estimation was performed via the presented approach and, to compare, with the help of the IMM filter implemented in toolbox [16]. Comparison of these filters provided the following results.

Significant difference between these methods is that the proposed one considers the probabilistic state-space model in a general form both for the normal and the discrete states and takes into account hybrid observations and control inputs. The discrete state estimation in the IMM filter is based mostly on the state transition table, i.e., other discrete variables bringing some information are not taken into account. This caused a worsened stability of the IMM filter during the testing.

Results of the online queue length estimation for four arms of the considered traffic microregion (here an intersection) are shown in Figs. 1, 2, where the first figure corresponds to the proposed hybrid filter, and the second figure – to the IMM filter. The LoS estimation for both the filters is demonstrated in Figs. 3, 4. The estimation error (EE) computed as

$$EE = \frac{1}{T}\sum_t \left(x_t^c - \mu_{t+1}\right)'\left(x_t^c - \mu_{t+1}\right),$$

where $T = 288$ is the duration of the estimation and x_t^c is the state identified with the real one, is provided in Table 1 for both the methods. A number of correctly point-estimated states (CPE) from the total 288-data sample was evaluated for both the filters and shown in Table 1. It is assumed that for better quality of estimation, EE should have a minimal value, and CPE on the contrary – a maximal (from 288) value. An advantage of the factorized hybrid filter is rather significant.

It should be also noted that both the output and the state transition tables used for the factorized filter were given as rather uncertain models. However, usage of more deterministic transition table for the IMM filter does not improve its results.

8. Conclusion

The paper is devoted to recursive state estimation that can be applied to hybrid systems. The proposed solution is based on the factorized form of Bayesian filtering. The paper summarizes application of general factorized filter to normal, discrete and hybrid models. The presented algorithms run in online mode avoiding numerical procedures as far possible. A number of statistics does not grow in time and the risk of collapsing is minimized. An important contribution is that the presented approach can be evolved for recursive estimation of discrete modes dependent on evolution of continuous states. The proposed method demonstrates better stability and quality of estimation in comparison with the IMM filter.

Acknowledgements

The research was supported by projects MŠMT 1M0572 and TAČR TA01030123.

Appendix A. Derivations for normal models

Substituting the factorized normal distributions with quadratic forms (14), (17) and (20) into (9), one obtains the following function to be integrated

$$\int \exp\left\{-\frac{1}{2}\underbrace{\left[Q_x + Q_y + Q_{p|t}\right]}_{Q_e}\right\} dx_t, \tag{A.1}$$

where to facilitate algebraic rearrangements, the following additional notations in the quadratic form Q_e appear, i.e.,

$$Q_e = \left[\underbrace{L'_w x_{t+1} - z_t}_{\beta_1} - \Xi x_t\right]' D_w \left[L'_w x_{t+1} - z_t - \Xi x_t\right] + \left[\underbrace{L'_v y_t - \rho_t}_{\beta_2} - \mathcal{A} x_t\right]' D_v \left[L'_v y_t - \rho_t - \mathcal{A} x_t\right] + \left[\underbrace{\mu^f_t}_{\beta_3} - L'_{p|t} x_t\right]' D_{p|t} \left[\mu^f_t - L'_{p|t} x_t\right], \quad \text{(A.2)}$$

where β_1, β_2, β_3 and $\beta = [\beta_1; \beta_2; \ \beta_3]$ are column vectors. To have the variable x_t in (A.1) integrated out, one has to fulfill the completion of squares [12] in (A.2) for x_t. After that and subsequent integration of non-normalized Gaussian pdf [17], the variable x_t is being integrated out. The computational result of the filtering (A.1) is proportional to $\exp\left\{-\frac{1}{2}\lambda\right\}$ with the following remainder λ obtained after integration

$$\lambda = \beta'\left(\Omega_t - \Omega_t \left[\Xi; \mathcal{A}; L'_{p|t}\right] \Gamma_t^{-1} \left[\Xi; \mathcal{A}; L'_{p|t}\right]' \Omega_t\right)\beta, \quad \text{(A.3)}$$

$$\text{where} \quad \Omega_t = \operatorname{diag}[D_w, D_v, D_{p|t}], \quad \text{(A.4)}$$

$$\Gamma_t = \left[\Xi; \mathcal{A}; L'_{p|t}\right]' \Omega_t \left[\Xi; \mathcal{A}; L'_{p|t}\right]. \quad \text{(A.5)}$$

With the help of algebraic rearrangement of the remainder (A.3) using completion of squares for x_{t+1}, one obtains the following quadratic form

$$\left[L'_w x_{t+1} - z_t - \widetilde{D}_t^{-1}\left(D_w \Xi \Gamma_t^{-1}\left(\mathcal{A}' D_v (L'_v y_t - \rho_t) + L_{p|t} D_{p|t} \mu^f_t\right)\right)\right]' \widetilde{D}_t \times \left[L'_w x_{t+1} - z_t - \widetilde{D}_t^{-1}\left(D_w \Xi \Gamma_t^{-1}\left(\mathcal{A}' D_v (L'_v y_t - \rho_t) + L_{p|t} D_{p|t} \mu^f_t\right)\right)\right], \quad \text{(A.6)}$$

where

$$\widetilde{D}_t = D_w - D_w \Xi \Gamma_t^{-1} \Xi' D_w. \quad \text{(A.7)}$$

The matrix $\widetilde{D}_t$, obtained in (A.7) is decomposed so that

$$\widetilde{D}_t = L_{u|t} D_{u|t} L'_{u|t}. \quad \text{(A.8)}$$

The decomposition (A.8) and factorization of the quadratic form (A.6) (i.e., its multiplication by triangular matrix $L_{u|t}$) enable to preserve the prior form (20) and obtain the following result

$$Q_{p|t+1} = \left[\underbrace{L'_{u|t} L'_w}_{L'_{p|t+1}} x_{t+1} - \mu^f_{t+1}\right]' \underbrace{D_{u|t}}_{D_{p|t+1}} \left[L'_{p|t+1} x_{t+1} - \mu^f_{t+1}\right], \quad \text{(A.9)}$$

where

$$\mu^f_{t+1} = L'_{u|t}\left(z_t + \widetilde{D}_t^{-1}\left(D_w \Xi \Gamma_t^{-1}\left(\mathcal{A}' D_v (L'_v y_t - \rho_t) + L_{p|t} D_{p|t} \mu^f_t\right)\right)\right). \quad \text{(A.10)}$$

References

[1] Y. Bar-Shalom, T. Kirubarajan, X.-R. Li, Estimation with Applications to Tracking and Navigation, Wiley, New York, NY, USA, 2002.
[2] M. Grewal, A. Andrews, Kalman Filtering: Theory and Practice Using MATLAB, Second ed., Wiley, 2001.
[3] F.D.R.J. Elliot, D. Sworder, Exact hybrid filters in discrete time, IEEE Transactions on Automatic Control 41 (12) (1996) 1807–1810.
[4] R. Shumway, D. Stoffer, Dynamic linear models with switching, Journal of the American Statistical Association 86 (415) (1991) 763–769.
[5] A. Doucet, C. Andrieu, Iterative algorithms for state estimation of jump markov linear systems, IEEE Transactions on Signal Processing 49 (6) (2001) 1216–1227.
[6] Q. Zhang, Optimal filtering of discrete-time hybrid systems, Journal of Optimization Theory and Applications 100 (1).
[7] R. Chen, J.S. Liu, Mixture kalman filters, Journal of the Royal Statistical Society Series B 62 (2000) 493–508.
[8] J.S.-H. Tsai, C.-T. Wang, C.-C. Kuang, S.-M. Guo, L.-S. Shieh, C.-W. Chen, A narmax model-based state-space self-tuning control for nonlinear stochastic hybrid systems, Applied Mathematical Modelling 34 (10) (2010) 3030–3054.
[9] J.-B. Sheu, Y.-H. Chou, A. Chen, Stochastic modeling and real-time prediction of incident effects on surface street traffic congestion, Applied Mathematical Modelling 28 (5) (2004) 445–468.
[10] M. Kárný, J. Böhm, T.V. Guy, L. Jirsa, I. Nagy, P. Nedoma, L. Tesař, Optimized Bayesian Dynamic Advising: Theory and Algorithms, Springer, London, 2005.
[11] M.J. Beal, Z. Ghahramani, C.E. Rasmussen, The infinite hidden markov model, in: Advances in Neural Information Processing Systems, Vol. 14, (2002).
[12] V. Peterka, Bayesian system identification, in: P. Eykhoff (Ed.), Trends and Progress in System Identification, Pergamon Press, Oxford, 1981, pp. 239–304.
[13] D. Simon, Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches, Wiley, 2006.
[14] D. Kerridge, Inaccuracy and inference, Journal of Royal Statistical Society B 23 (1961) 284–294.

[15] S. Kullback, R. Leibler, On information and sufficiency, Annals of Mathematical Statistics 22 (1951) 79–87.
[16] J. Hartikainen, S. Särkkä, Optimal filtering with kalman filters and smoothers – a manual for matlab toolbox ekf/ukf, <http://www.lce.hut.fi/research/mm/ekfukf/> (February 2008).
[17] E. Suzdaleva, M. Kárný, Factorized Filtering, Tech. Rep. 2161, ÚTIA AV ČR, Praha, <http://www.library.utia.cas.cz/prace/20060109.pdf> (2006).