Towards Filtering with Mixed-type States

Evgenia Suzdaleva

Abstract— **The paper deals with state estimation, organized in an entry-wise manner. The entry-wise updating of the posterior state estimates is reached via application of the chain rule and factorization of covariance matrices. Such the filtering provides distributions of entries of the state vectors individually in the factorized form. Application to Gaussian linear state-space model with Gaussian observations and Gaussian prior distribution provides Kalman filter. A series of works has been devoted to preparation of solution of the entry-wise filtering. The recently proposed modification of the entry-wise algorithm with a simultaneous fulfillment of data and time updating steps is used at the present paper. The paper considers its application to the mixed-type (continuous and discrete-valued) data. A special case with an involved last discrete-valued state entry, unlike the previous continuous ones, is described.**

I. INTRODUCTION

The paper deals with the entry-wise organized fitlering and proposes its application to mixed-type (continuous and discrete-valued) data. Despite a number of existing approaches in the field of mixed data modelling [1], [2], [3], [4], the estimation with mixed-type states still calls for a reliable solution. The entry-wise filtering is a potential solution of this problem. Via factorization of a state-space model such the filter provides the state estimates of individual state vector entries in the factorized form and allows to update them entry-wise. The entry-wise updating indicates a chance to estimate the mixed-type states. The paper presents the algorithmic solution of a special case of the estimation with the mixed-type state, where the last state entry is of a discrete-valued nature, unlike the previous continuous ones.

A general motivation for the research is a traffic situation in cities with the overloaded traffic systems and long queues at intersections. The long queues of waiting cars strongly affect ecology of the cities. Extension of traffic network (new roads, flyovers, tunnels, etc) is very expensive and often impossible, especially in historical cities. Such a situation motivated to consider one of the steps to its improving – an adequate modelling of a traffic area and exploitation of modern adaptive traffic control systems. The state-space models with a queue length as the main state variable have made a good showing for traffic control systems. But in a traffic control domain some of the state variables are of discrete-valued nature (signal lights, level of service, visibility, road surface, etc). Their involvement calls for the joint modelling of mixed-type (continuous and discrete-valued) variables.

Mixed-type data modelling is known to be hard problem addressed within completely different context of logistic

regression. A number of flexible models for such data is rather limited. One of the most known works in this field was [5], which proposed the general "location model", based on multinomial model for the discrete data, and Gaussian multivariate model for the continuous data, conditional on the discrete data. This work inspired many studies in the discussed area. The paper [1] extends the "location model" and proposes the likelihood-based approach for analyzing mixed continuous and discrete data. A series of papers [2], [3] is devoted to the latent variable models with mixed data and different modifications of the expectation-maximization algorithm, which is exploited for finding maximum likelihood estimates of model parameters. The research work [6] describes the mixed-type data modelling based on discriminant analysis. The paper [7] deals with Bayesian latent variable models for clustered mixed data and uses a Markov chain Monte Carlo sampling algorithm for estimating the posterior distributions of the parameters and latent variables. More details about existing approaches can be found in [8].

As regards the state of the art of the entry-wise (factorized) filtering, it is as follows. Most works [9], [10] found in the field are devoted to the factorized implementations of Kalman filtering. However, the global aim of the mentioned works is, primarily, reduction of the computational complexity via a lesser rank of the covariance matrix. Exploitation of matrix factorization to approach the entry-wise filtering under Bayesian methodology [11] was proposed in [12] with a reduced form of the state-space model. The paper [13] removed this restriction and proposed the solution of factorized Bayesian prediction and filtering, based on applying the chain rule to the single output state-space model. The work [14] offered the version of factorized Kalman filtering with Gaussian models, which was based on the $L'DL$ decomposition of the covariance matrices. The paper [15] expanded the line with $L'DL$-factorized covariance matrices and demonstrated the application of the solution to the traffic-control area. However, the mentioned works had problems with preserving of the distribution form of state entries and consequently with the entry-wise updating. The recent paper [16] proposes the novel algorithm with LDL'-factorized covariance matrices and simultaneous data and time updating of the posterior state entry estimates. This modified form of filtering is used at the present paper. The solution proposed at the paper is expected to be applied in the traffic control area, but not restricted to.

The paper has the following layout. Section II is devoted to basic facts about used system model, factorization by the chain rule and Bayesian and Kalman filtering. Section III provides a general derivation of Kalman filter with simul-

taneous data and time updating and presents the algorithm of its factorized version. Subsection III-C considers the involvement of the discrete-valued state entry at the end of the state vector. The illustrative experiments with a simple simulated system are shown in Section IV. Summary and plans of future work in Section V close the paper.

II. PRELIMINARIES

A. Models

The system is described by the joint probability density function (pdf)

$$f\left(x_t, y_t \mid x_{t-1}, u_t\right), \tag{1}$$

where x_t is the unobservable system state at discrete time moments $t \in t^* \equiv \{0, \ldots, \mathring{t}\}$, where $\mathring{t}$ is the cardinality of the set t^*, $\equiv$ means equivalence, and y_t and u_t are the system output and input respectively. When the pdf (1) is Gaussian, it includes the *state evolution model* (2) and *observation model* (3)

$$
\begin{aligned}
x_t &= A x_{t-1} + B u_t + \omega_t, \tag{2} \\
y_t &= C x_{t-1} + H u_t + v_t, \tag{3}
\end{aligned}
$$

where $[A, C]$ and $[B, H]$ are known joint matrices of appropriate dimensions; ω_t is a process (Gaussian) noise with zero mean and known covariance matrix R_w; v_t is a measurement (Gaussian) noise with zero mean and known covariance matrix R_v.

The joint pdf (1) can be decomposed, according to the chain rule [17], into the following factorized form.

$$\prod_{i=1}^{\mathring{x}} f(x_{i;t} \mid x_{i+1:\mathring{x};t}, u_t, x_{t-1}, y_t) \prod_{j=1}^{\mathring{y}} f(y_{j;t} \mid y_{j+1:\mathring{y};t}, u_t, x_{t-1}),$$
$$\tag{4}$$

where $\mathring{x}$ and $\mathring{y}$ are numbers of entries of column vectors x_t and y_t respectively, $i = \{1, \ldots, \mathring{x}\}$, $j = \{1, \ldots, \mathring{y}\}$, and such a notation as $x_{i+1:\mathring{x};t}$ denotes a sequence $\{x_{i+1;t}, x_{i+2;t}, \ldots, x_{\mathring{x};t}\}$.

B. Bayesian filtering

Bayesian filtering [17] includes two coupled formulae.

$$
\begin{aligned}
f(x_t \mid u_t, d^{t-1}) &= \int f(x_t \mid u_t, x_{t-1}), \tag{5} \\
&\times \quad f(x_{t-1} \mid d^{t-1}) dx_{t-1}, \\
f(x_t \mid d^t) &\propto f(y_t \mid u_t, x_t) f(x_t \mid u_t, d^{t-1}), \tag{6}
\end{aligned}
$$

where $\propto$ means proportionality, data are defined as $d^t = (d_0, \ldots, d_{\mathring{t}})$, $d_t \equiv (y_t, u_t)$. Relation (5) represents the time updating of the state estimate, while (6) – the data updating. The application of (5) and (6) to Gaussian state-space model with Gaussian prior distribution and Gaussian observations provides Kalman filter [18].

III. ENTRY-WISE FILTERING WITH INVOLVED DISCRETE-VALUED STATE

A. Simultaneous Data and Time Updating

Manupilation with the state-space model in the form of joint pdf (1) allows to rewrite Bayesian filtering so that the data and time updating are fulfilled simultaneously. Generally, with the help of Bayes rule [17], one obtains the following relation

$$
\begin{aligned}
f(y_t, x_t \mid u_t, d^{t-1}) &= f(y_t \mid u_t, d^{t-1}, x_t) f(x_t \mid u_t, d^{t-1}), \\
&= f(x_t \mid d^t) f(y_t \mid u_t, d^{t-1}), \tag{7}
\end{aligned}
$$

which provides the data updating (6) in the form

$$f(x_t \mid d^t) = \frac{f(y_t, x_t \mid u_t, d^{t-1})}{f(y_t \mid u_t, d^{t-1})}, \tag{8}$$

with Bayesian predictor [17] as a denominator. With the help of operation of marginalization [17] and according to the used model, one can obtain the expression of filtering with relations (5) and (6) to be fulfilled simultaneously (i.e. simultaneous data and time updating).

$$
\begin{aligned}
f(x_t \mid d^t) &= \frac{f(y_t, x_t \mid u_t, d^{t-1})}{f(y_t \mid u_t, d^{t-1})}, \tag{9} \\
&= \frac{\int f(x_t, y_t, x_{t-1} \mid u_t, d^{t-1}) dx_{t-1}}{f(y_t \mid u_t, d^{t-1})}, \\
&= \frac{\int f(x_t, y_t \mid u_t, x_{t-1}) f(x_{t-1} \mid d^{t-1}) dx_{t-1}}{f(y_t \mid u_t, d^{t-1})}, \\
&\propto \int f(x_t, y_t \mid u_t, x_{t-1}) f(x_{t-1} \mid d^{t-1}) dx_{t-1}.
\end{aligned}
$$

Substitution of (4) in (9) and decomposition of the initial state according to the chain rule provide the following factorized form of (9).

$$
\begin{aligned}
f(x_t \mid d^t) &\propto \int \prod_{i=1}^{\mathring{x}} f(x_{i;t} \mid x_{i+1:\mathring{x};t}, u_t, x_{t-1}, y_t), \tag{10} \\
&\times \prod_{j=1}^{\mathring{y}} f(y_{j;t} \mid y_{j+1:\mathring{y};t}, u_t, x_{t-1}), \\
&\times \prod_{i=1}^{\mathring{x}} f(x_{i;t-1} \mid x_{i+1:\mathring{x};t-1}, d^{t-1}) dx_{t-1}.
\end{aligned}
$$

The left-hand side of (10) is assumed to be $f(x_t \mid d^t) = \prod_{i=1}^{\mathring{x}} f(x_{i;t} \mid x_{i+1:\mathring{x};t}, d^t)$, and a presence of vector x_{t-1} in all pdfs in (10) means involvement of all entries of the respective vector in integration. Such the notations are used only for shortening of the equation.

B. Algorithm of Factorized Kalman Filtering

The entry-wise updating assumed in (10) and preserving of the factorized form (4) of the posterior state estimate $f(x_t \mid d^t) = \prod_{i=1}^{\mathring{x}} f(x_{i;t} \mid x_{i+1:\mathring{x};t}, d^t)$ can be reached via LDL' decomposition [11] of the precision (i.e. inverse covariance) matrices. Such the decomposition supposes L to be a lower triangular matrix with unit diagonal, D to

be a diagonal one, $'$ – transposition. This kind of matrix decomposition is used throughout the paper.

The key moments of the entry-wise organized Kalman filter (10), applied to Gaussian models (2-3) are as follows.

1) Factorization of State Evolution Model: Gaussian state evolution model (2) is factorized in the following way. The process noise covariance matrix R_w is inverted into a precision matrix and decomposed so that

$$R_w^{-1} \quad = \quad L_w D_w L_w'. \tag{11}$$

The factorized Gaussian quadratic form, corresponding to the model (2), becomes now

$$[L_w' x_t - z_t - \Xi x_{t-1}]' D_w [L_w' x_t - z_t - \Xi x_{t-1}], \tag{12}$$

where

$$z_t \quad = \quad L_w' B u_t, \tag{13}$$
$$\Xi \quad = \quad L_w' A. \tag{14}$$

Gaussian distribution of an individual state entry has the following form.

$$\mathcal{N}_{x_{i;t}}(z_i - \sum_{k=i+1}^{\mathring{x}} L_{w;ki} x_{k;t} + \sum_{l=1}^{\mathring{x}} \Xi_{il} x_{l;t-1}, \frac{1}{D_{w;ii}}). \tag{15}$$

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

2) Factorization of Observation Model: Factorization of Gaussian observation model (3) is fulfilled similarly. The measurement noise covariance matrix R_v is inverted into the precision matrix and decomposed so that

$$R_v^{-1} \quad = \quad L_v D_v L_v', \tag{16}$$

The factorized quadratic form, corresponding to the model (3), is as follows.

$$[L_v' y_t - \rho_t - \mathcal{A} x_{t-1}]' D_v [L_v' y_t - \rho_t - \mathcal{A} x_{t-1}], \tag{17}$$

where

$$\rho_t \quad = \quad L_v' H u_t, \tag{18}$$
$$\mathcal{A} \quad = \quad L_v' C. \tag{19}$$

Gaussian distribution of an individual output entry takes the form

$$\mathcal{N}_{y_{j;t}}(\rho_j - \sum_{k=j+1}^{\mathring{y}} L_{v;kj} y_{k;t} + \sum_{l=1}^{\mathring{x}} \mathcal{A}_{jl} x_{l;t-1}, \frac{1}{D_{v;jj}}). \tag{20}$$

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

3) Initial State Factorization: The initial state Gaussian distribution $f(x_{t-1}|d^{\,t-1}) \sim \mathcal{N}(\mu_0, P_0)$, where μ_0 is a known vector of mean values and P_0 is a known covariance matrix, is also factorized in the similar way. Decomposition of the precision matrix is done so that

$$P_0^{-1} \quad = \quad L_{p|0} D_{p|0} L_{p|0}'. \tag{21}$$

The factorized quadratic form, corresponding to Gaussian distribution of the initial state, looks like

$$[L_{p|0}' x_0 - \mu_0^f]' D_{p|0} [L_{p|0}' x_0 - \mu_0^f], \tag{22}$$

with

$$\mu_0^f \quad = \quad L_{p|0}' \mu_0. \tag{23}$$

Via (21) the initial state entries have the following factorized form of Gaussian distribution.

$$\mathcal{N}_{x_{i;0}}(\mu_{i;0}^f - \sum_{k=i+1}^{\mathring{x}} L_{p|0;ki} x_{k;0}, \frac{1}{D_{p|0;ii}}). \tag{24}$$

where $L_{p|0;ki}$ and $D_{p|0;ii}$ are the elements of matrices $L_{p|0}$ and $D_{p|0}$ respectively.

4) Factorized Simultaneous Data & Time Updating: The simultaneous data and time updating in the factorized form is proposed as follows. After integration in (10) and completion of squares the following quadratic form for the factorized state is obtained.

$$[L_w' x_t - \mu_t^*]' \tilde{D}_t [L_w' x_t - \mu_t^*], \tag{25}$$

where

$$\mu_t^* \quad = \quad z_t + \tilde{D}^{-1}, \tag{26}$$
$$\times \quad (D_w \Xi \tilde{A}_t^{-1}(\mathcal{A}' D_v (L_v' y_t - \rho_t),$$
$$+ \quad L_{p|t-1} D_{p|t-1} \mu_{t-1}^f)),$$
$$\Omega_t \quad = \quad \mathrm{diag}\left[D_w, \; D_v, \; D_{p|t-1}\right], \tag{27}$$
$$\tilde{A}_t \quad = \quad [\Xi; \; \mathcal{A}; \; L_{p|t-1}']' \Omega_t [\Xi; \; \mathcal{A}; \; L_{p|t-1}'], \tag{28}$$
$$\tilde{D}_t \quad = \quad D_w - D_w \Xi \tilde{A}_t^{-1} \Xi' D_w. \tag{29}$$

The matrix $\tilde{D}_t$, obtained in (29) is decomposed so that

$$\tilde{D}_t = L_{u|t} D_{u|t} L_{u|t}'. \tag{30}$$

After decomposition and factorization of the quadratic form (25), the updated factorized Gaussian quadratic form is obtained.

$$\left[\underbrace{L_{u|t}' L_w'}_{L_{p|t}'} x_t - \underbrace{L_{u|t}' \mu_t^*}_{\mu_t^f}\right]' \underbrace{D_{u|t}}_{D_{p|t}} \left[L_{p|t}' x_t - \mu_t^f\right]. \tag{31}$$

It means, that the updating of the decomposed matrices and the factorized mean value is as follows.

$$D_{p|t} \quad = \quad D_{u|t}, \tag{32}$$
$$L_{p|t} \quad = \quad L_{u|t} L_w, \tag{33}$$
$$\mu_t^f \quad = \quad L_{u|t}' \mu_t^*, \tag{34}$$

which allows to preserve the form (22) of the prior state

$$\left[L_{p|t}' x_t - \mu_t^f\right]' D_{p|t} \left[L_{p|t}' x_t - \mu_t^f\right]. \tag{35}$$

Finally, Gaussian distribution of the i-th state entry keeps its factorized form (24), i.e.

$$\mathcal{N}_{x_{i;t}}(\mu_{i;t}^f - \sum_{k=i+1}^{\mathring{x}} L_{p|t;ki} x_{k;t}, \frac{1}{D_{p|t;ii}}). \tag{36}$$

The obtained results are proved by direct calculation of the integral (10). Verification of the entry-wise updating by means of comparison with solution of (9) is available in [16].

Remark 1. *Exploitation of the joint pdf (1) as a system model and, therefore, presence of the state x_{t-1} in both state evolution and observation models (2-3) enables a full factorization of the observation model (3). It means, that in practice the proposed algorithm is not restricted by a single-output model, as its previous versions [13].*

Remark 2. *The proposed algorithm can be sensitive to preserving of positive-definiteness of matrix $\tilde{D}_t$, used in calculation of final variances of the state entries. For more stability it would be useful to try a QR factorization, where Q is an orthogonal matrix and R is an upper triangular one. However, the present paper is focused on the proposed LDL' factorization due to a lower computational complexity.*

C. Involvement of Discrete State Entry

The proposed factorization gives a chance to consider the state estimation entry-wise, i.e. by rows. The execution of the algorithm begins at the end of the state vector, i.e. $i = \mathring{x}$ in (36). Let's replace a continuous state entry at the end of the state entry by the discrete-valued one so that it is a scalar one with a set of possible discrete values $\{\mathbf{0}, \mathbf{1}\}$. To facilitate calculations, the last entry at the end of the output vector is also replaced by the discrete one, i.e. $y_{\mathring{y};t} \in \{\mathbf{0}, \mathbf{1}\}$. Respectively it is transformed into the discrete-valued variable via chosen thresholds. Let's assume that, according to the chain rule, the joint pdf (1) takes a form $f(x_{\mathring{x};t}, y_{\mathring{y};t}|u_t, x_{\mathring{x};t-1}) = f(x_{\mathring{x};t}|u_t, x_{\mathring{x};t-1}, y_{\mathring{y};t})f(y_{\mathring{y};t}|u_t, x_{\mathring{x};t-1})$ for the case of the last discrete entries. The involved discrete state entry is described it by an individual model with the alternative distribution shown in Table I, where p with respective indices denotes a probability of taking the possible values. The system input $u_t \in u^*$ is a known constant. The alternative

TABLE I
ALTERNATIVE DISTRIBUTION $f(x_{\mathring{x};t}|u_t, x_{\mathring{x};t-1}, y_{\mathring{y};t})$

	$x_{\mathring{x};t} = \mathbf{0}$	$x_{\mathring{x};t} = \mathbf{1}$
$x_{\mathring{x};t-1} = \mathbf{0}, y_{\mathring{y};t} = \mathbf{0}$	$p_{\mathbf{0}|\mathbf{00}}$	$p_{\mathbf{1}|\mathbf{00}}$
$x_{\mathring{x};t-1} = \mathbf{1}, y_{\mathring{y};t} = \mathbf{0}$	$p_{\mathbf{0}|\mathbf{10}}$	$p_{\mathbf{1}|\mathbf{10}}$
$x_{\mathring{x};t-1} = \mathbf{0}, y_{\mathring{y};t} = \mathbf{1}$	$p_{\mathbf{0}|\mathbf{01}}$	$p_{\mathbf{1}|\mathbf{01}}$
$x_{\mathring{x};t-1} = \mathbf{1}, y_{\mathring{y};t} = \mathbf{1}$	$p_{\mathbf{0}|\mathbf{11}}$	$p_{\mathbf{1}|\mathbf{11}}$

distribution from Table I can be written in the product form

$$f(x_{\mathring{x};t}|u_t, x_{\mathring{x};t-1}, y_{\mathring{y};t}), \tag{37}$$
$$= \prod_{x_{\mathring{x};t}|u_t, x_{\mathring{x};t-1}, y_{\mathring{y};t}} p_{x_{\mathring{x};t}|u_t, x_{\mathring{x};t-1}, y_{\mathring{y};t}}^{\delta(x_{\mathring{x};t}|u_t, x_{\mathring{x};t-1}, y_{\mathring{y};t}, \hat{x}_{\mathring{x};t}|u_t, \hat{x}_{\mathring{x};t-1}, \hat{y}_{\mathring{y};t})},$$

where δ is Dirac delta and $\hat{x}_{\mathring{x};t}$, $x_{\mathring{x};t-1}$ and $\hat{y}_{\mathring{y};t}$ denote possible values from Table I. Similarly, the discrete-valued output entry is described by the alternative distribution shown in Table II. The product form of the alternative distribution

TABLE II
ALTERNATIVE DISTRIBUTION $f(y_{\mathring{y};t}|u_t, x_{\mathring{x};t-1})$

	$y_{\mathring{y};t} = \mathbf{0}$	$y_{\mathring{y};t} = \mathbf{1}$
$x_{\mathring{x};t-1} = \mathbf{0}$	$p_{\mathbf{0}|\mathbf{0}}$	$p_{\mathbf{1}|\mathbf{0}}$
$x_{\mathring{x};t-1} = \mathbf{1}$	$p_{\mathbf{0}|\mathbf{1}}$	$p_{\mathbf{1}|\mathbf{1}}$

TABLE III
PRIOR ALTERNATIVE DISTRIBUTION

$x_{\mathring{x};t-1} = \mathbf{0}$	$x_{\mathring{x};t-1} = \mathbf{1}$
$p_{\mathbf{0}(t-1)}$	$p_{\mathbf{1}(t-1)}$

from Table II is as follows.

$$f(y_{\mathring{y};t}|u_t, x_{\mathring{x};t-1}), \tag{38}$$
$$= \prod_{y_{\mathring{y};t}|u_t, x_{\mathring{x};t-1}} p_{y_{\mathring{y};t}|u_t, x_{\mathring{x};t-1}}^{\delta(y_{\mathring{y};t}|u_t, x_{\mathring{x};t-1}, \hat{y}_{\mathring{y};t}|u_t, \hat{x}_{\mathring{x};t-1})},$$

The prior alternative distribution of the pdf $f(x_{\mathring{x};t-1}|d^{t-1})$ with the chosen probabilities of taking the possible values is shown in Table III, where $p_{\mathbf{1}(t-1)} = (1 - p_{\mathbf{0}(t-1)})$. It takes the following product form.

$$f(x_{\mathring{x};t-1}|d^{t-1}) = \prod_{k=0}^{1} p_{k(t-1)}^{\delta(x_{\mathring{x};t-1}, k)}, \tag{39}$$

where $\sum_{k=0}^{1} p_{k(t-1)} = 1$, $p_{k(t-1)} > 0 \; \forall \; k$.

The discrete state entry estimation is performed according to the technique, proposed in [19]. The integral (10) is replaced a regular summation for the discrete state.

$$f(x_{\mathring{x};t}|d^t), \tag{40}$$
$$= \sum_{x_{\mathring{x};t-1} \in \{\mathbf{0},\mathbf{1}\}},$$
$$\prod_{x_{\mathring{x};t}|u_t, x_{\mathring{x};t-1}, y_{\mathring{y};t}} p_{x_{\mathring{x};t}|u_t, x_{\mathring{x};t-1}, y_{\mathring{y};t}}^{\delta(x_{\mathring{x};t}|u_t, x_{\mathring{x};t-1}, y_{\mathring{y};t}, \hat{x}_{\mathring{x};t}|u_t, \hat{x}_{\mathring{x};t-1}, \hat{y}_{\mathring{y};t})},$$
$$\times \prod_{y_{\mathring{y};t}|u_t, x_{\mathring{x};t-1}} p_{y_{\mathring{y};t}|u_t, x_{\mathring{x};t-1}}^{\delta(y_{\mathring{y};t}|u_t, x_{\mathring{x};t-1}, \hat{y}_{\mathring{y};t}|u_t, \hat{x}_{\mathring{x};t-1})},$$
$$\times \prod_{k=0}^{1} p_{k(t-1)}^{\delta(x_{\mathring{x};t-1}, k)}.$$

The probabilities for the values $f(x_{\mathring{x};t} = \mathbf{0}|d^t) = p_{\mathbf{0}(t)}$ and $f(x_{\mathring{x};t} = \mathbf{1}|d^t) = (1 - p_{\mathbf{0}(t)})$ can be calculated with the help of (40).

$$p_{\mathbf{0}(t)}, \tag{41}$$
$$= \sum_{x_{\mathring{x};t-1} \in \{\mathbf{0},\mathbf{1}\}},$$
$$\prod_{x_{\mathring{x};t}|u_t, x_{\mathring{x};t-1}, y_{\mathring{y};t}} p_{x_{\mathring{x};t}|u_t, x_{\mathring{x};t-1}, y_{\mathring{y};t}}^{\delta(x_{\mathring{x};t}|u_t, x_{\mathring{x};t-1}, y_{\mathring{y};t}, \mathbf{0}|u_t, \hat{x}_{\mathring{x};t-1}, \hat{y}_{\mathring{y};t})},$$
$$\times \prod_{y_{\mathring{y};t}|u_t, x_{\mathring{x};t-1}} p_{y_{\mathring{y};t}|u_t, x_{\mathring{x};t-1}}^{\delta(y_{\mathring{y};t}|u_t, x_{\mathring{x};t-1}, \hat{y}_{\mathring{y};t}|u_t, \hat{x}_{\mathring{x};t-1})},$$
$$\times \prod_{k=0}^{1} p_{k(t-1)}^{\delta(x_{\mathring{x};t-1}, k)},$$

$$
\begin{aligned}
=\ & \underbrace{p_{0|00}p_{0|01}p_{0|0}p_{1|0}p_{0(t-1)}}_{sum\ for\ x_{\mathring{x};t-1}=0}, \\
+\ & \underbrace{p_{0|10}p_{0|11}p_{0|1}p_{1|1}(1-p_{0(t-1)})}_{sum\ for\ x_{\mathring{x};t-1}=1}.
\end{aligned}
$$

The result is obtained by substitution of probabilities from Table I, corresponding to the state $x_{\mathring{x};t}=0$ and all probabilities from Tables II-III and subsequent summation over respective values of the state $x_{\mathring{x};t-1}$. The probability $p_{1(t)}$ is calculated quite similarly with substitution of probabilities from Table I, corresponding to the state $x_{\mathring{x};t}=1$ and should be a complement of (41). It means, that the updated pdf $f(x_{\mathring{x};t}|d^{\,t})$ preserves the form (39)

$$
f(x_{\mathring{x};t}|d^{\,t}) = \prod_{k=0}^{1} p_{k(t)}^{\delta(x_{\mathring{x};t},k)}, \tag{42}
$$

with $\sum_{k=0}^{1} p_{k(t)} = 1,\ p_{k(t)} > 0\ \forall\ k$.

The filtering begins from the described discrete state entry estimation. The mean value to be involved in the consequent processing is calculated as a sum of possible values of the entry multiplied by the updated probabilities, i.e.

$$
\mu_{\mathring{x};t}^{f} = \sum_{k=0}^{1} x_{\mathring{x};t,k}\, p_{k(t)}. \tag{43}
$$

Then the algorithm, proposed in Subsection III-B, is executed, starting at the last state entry, i.e. for $i = \{\mathring{x}, \mathring{x}-1, \ldots, 1\}$.

IV. ILLUSTRATIVE EXPERIMENTS

A physical interpretation of the state vector of the mixed type composed so that it contains several continuous entries and the last discrete one can be explained as follows. The expected application of the research is a traffic control area, where the continuous Gaussian state entries are interpreted as the car queue lengths at the intersection lanes [20]. The discrete state entry is a two-valued variable, indicating, for example, signal lights ($0 = green$, $1 = red$), visibility ($0 = good$, $1 = worse$), etc. Before to implementation with a traffic state-space model [20], the proposed filtering with mixed states should be tested on simple simulated data. This section provides the illustrative experiments with a simple simulated system with two state entries and two output entries in order to verify a correct performance of the proposed algorithm.

The last discrete state entry estimation is tested on the following data. According to (39), the prior probabilities are chosen as $p_{0(t-1)} = 0.5$ and $p_{1(t-1)} = 0.5$. The simulated probabilities of the model from Tables I-II are demonstrated in Tables IV-V. The system input is a constant $u_t = 0.5$. The discrete output entry has been simulated by a discrete system generator for 100 time moments. The estimation of the discrete state is shown at Fig. 1 (top). For more illustrative plotting, the results of the filtering are shown only for 50 time moments. The rest of the estimates are similar. The mean value of the last state entry has been calculated according to

TABLE IV
SIMULATED PROBABILITIES FOR $f(x_{2;t}|u_t, x_{2;t-1}, y_{2;t})$

	$x_{2;t} = 0$	$x_{2;t} = 1$
$x_{2;t-1} = 0, y_{2;t} = 0$	0.7711	0.2289
$x_{2;t-1} = 1, y_{2;t} = 0$	0.6702	0.3298
$x_{2;t-1} = 0, y_{2;t} = 1$	0.3583	0.6417
$x_{2;t-1} = 1, y_{2;t} = 1$	0.9960	0.0040

TABLE V
SIMULATED PROBABILITIES FOR $f(y_{2;t}|u_t, x_{2;t-1})$

	$y_{2;t} = 0$	$y_{2;t} = 1$
$x_{2;t-1} = 0$	0.1071	0.8929
$x_{2;t-1} = 1$	0.8939	0.1061

(43) and used in the subsequent factorized Kalman filtering with the continuous state entries.

The following simulated data are used in the models (2-3).

$$
A = \begin{bmatrix} 0.1 & -0.9 \\ 0.9 & 0.01 \end{bmatrix}, B = \begin{bmatrix} -0.4 \\ -0.4 \end{bmatrix}, \tag{44}
$$

$$
C = \begin{bmatrix} -0.1 & 1 \\ 0.1 & -0.5 \end{bmatrix}, H = \begin{bmatrix} 1 \\ -1 \end{bmatrix}. \tag{45}
$$

The noise covariances R_w and R_v are computed as a mean of squares of differences between the state (or output respectively) value and its conditional mean. The mean is substituted by the samples of a periodic course of the state (or output), which is constructed as a spline approximation of several last periodic courses. The resulted covariance matrices are as follows.

$$
R_w = \begin{bmatrix} 0.3974 & -0.1060 \\ -0.1060 & 0.4011 \end{bmatrix}, \tag{46}
$$

$$
R_v = \begin{bmatrix} 0.4844 & 0.3598 \\ 0.3598 & 0.9887 \end{bmatrix}. \tag{47}
$$

The chosen initial state mean vector and covariance matrix are

$$
\mu_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, P_0 = \begin{bmatrix} 2.029 & -2.784 \\ -2.784 & 5.432 \end{bmatrix}. \tag{48}
$$

The estimation of the continuous state entry is demonstrated at Fig. 1 (bottom). Good correspondence between simulated and estimated values at Fig. 1 verifies the adequate performance of the proposed version of the filtering.

V. CONCLUSIONS AND FUTURE WORKS

The paper deals with the entry-wise organized filtering. The proposed filtering enables the entry-wise updating of the states and allows to describe them individually. Such a version of the state estimation is directed towards processing of the mixed-type (continuous and discrete-valued) states. The application to continuous data, described by Gaussian linear state-space model with Gaussian observations and Gaussian prior distribution provides Kalman filter. The entry-wise updating of the posterior state estimates is reached via application of the chain rule and factorization of covariance matrices. The resulted Gaussian distributions are obtained
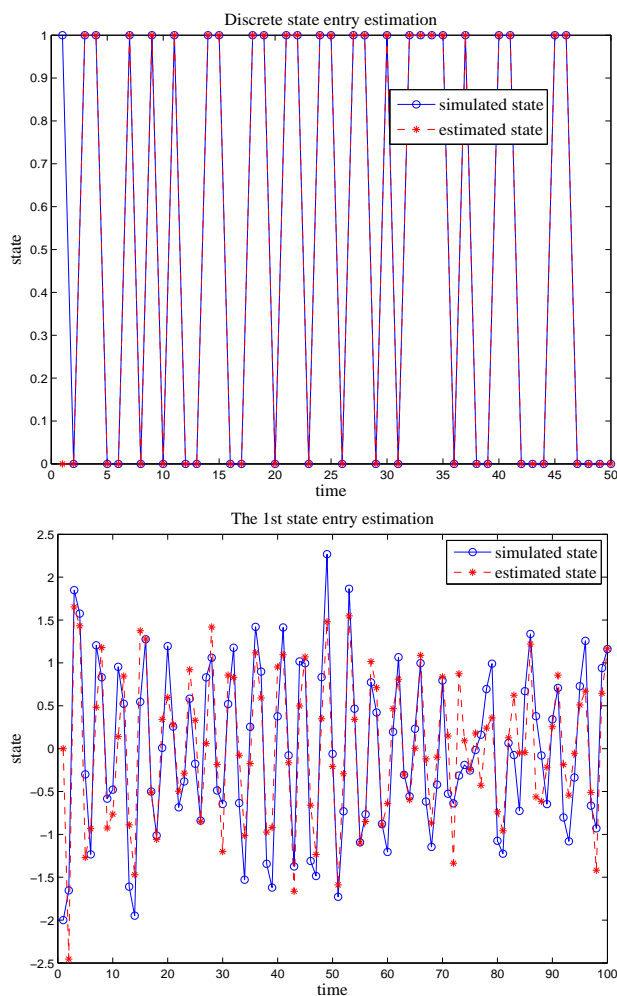
Fig. 1.   Filtering with discrete state (top) and with continuous state (bottom)

in the factorized form. The paper considers a special case of the joint estimation of the mixed-type state with the discrete-valued state entry, involved at the end of the state vector. The discrete state entry is described by the alternative distribution. Its estimation is realized via filtering for the discrete-valued variables, where integration is replaced by a regular summation of probabilities. The practical application of the research is expected at the traffic control area. However, the proposed solution is not restricted by this application and is, in general, an universal one. Future work in the discussed field includes a solution of the joint filtering of mixed states, not restricted by the presented special case with the last discrete entry. The involvement of the discrete entry among the continuous ones is planned to be the next problem to be considered. The experimental part of the work will be devoted to implementation and tests with realistic traffic data provided by the traffic microsimulator AIMSUN [21].

VI. ACKNOWLEDGMENTS

REFERENCES

[1] G.M. Fitzmaurice and N.M. Laird, "Regression models for mixed discrete and continuous responses with potentially missing values", *Biometrics*, vol. 53, no. 1, pp. 110–122, March 1997.

[2] Jian-Qing Shi and Sik-Yum Lee, "Latent variable models with mixed continuous and polytomous data", *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, vol. 62, no. 1, pp. 77–87, 2000.

[3] M.D. Sammel, L.M. Ryan, and J.M. Legler, "Latent variable models for mixed discrete and continuous outcomes", *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 59, no. 3, pp. 667–678, 1997.

[4] Kevin M. Quinn, "Bayesian factor analysis for mixed ordinal and continuous responses", *Political Analysis*, vol. 12, pp. 338–353, 2004.

[5] I. Olkin and R.F. Tate, "Multivariate correlation models with mixed continuous and discrete variables", *Annals of Mathematical Statistics*, vol. 32, pp. 448–465, 1961.

[6] Klaus-D. Wernecke, "A coupling procedure for the discrimination of mixed data", *Biometrics*, vol. 48, no. 2, pp. 497–506, June 1992.

[7] David B. Dunson, "Bayesian latent variable models for clustered mixed outcomes", *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, vol. 62, no. 2, pp. 355–366, 2000.

[8] E. Suzdaleva, "Overview of models with mixed-type data", Tech. Rep. 2208, ÚTIA AV ČR, Praha, December 2007.

[9] G. Dimitriu, "Implementation issues related to data assimilation using Kalman filtering", in *Proceedings for 3th World Conference on Computational Statistics & Data Analysis*, Limassol, Cyprus, October 28-31 2005, International Association for Statistical Computing.

[10] I. Patras and M. Pantic, "Particle filtering with factorized likelihoods for tracking facial features", in *Proceedings of the 6th IEEE International Conference on Automatic Face and Gesture Recognition (FGR04)*, Seoul, Korea, May 17-19 2004.

[11] M. Kárný, J. Böhm, T. V. Guy, L. Jirsa, I. Nagy, P. Nedoma, and L. Tesař, *Optimized Bayesian Dynamic Advising: Theory and Algorithms*, Springer, London, 2005.

[12] E. Suzdaleva, "On entry-wise organized filtering", in *Proceedings of 15th International Conference on Process Control'05*, High Tatras, Slovakia, June 7-10 2005, pp. 1–6, Department of Information Engineering and Process Control, FCFT STU, ISBN 80-227-2235-9.

[13] E. Suzdaleva, "Towards state estimation in the factorized form", in *Proceedings of the 2nd International Conference on From Scientific Computing to Computational Engineering*, D. Tsahalis, Ed., Athens, July 2006, pp. 1–6, University of Patras.

[14] E. Suzdaleva, "Factorized Kalman filtering", in *Proceedings of the 7th International PhD Workshop on Interplay of societal and technical decision-making*, J. Přikryl, J. Andrýsek, and V. Šmidl, Eds., Hruba Skala, Czech Republic, September 25-30 2006, pp. 226–233, ÚTIA AV ČR.

[15] E. Suzdaleva, "Early experiments with traffic system model and factorized Kalman filtering", in *Proceedings of the IEEE Fourth International Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications IDAACS'2007*, Dortmund, Germany, September 6-8 2007, IEEE, ISBN 1-4244-1348-6.

[16] E. Suzdaleva, "On new design of Kalman filter with entry-wise updating", in *Proceedings of the 8th WSEAS International Conference on Simulation, Modelling and Optimization (SMO'08)*, Santander, Spain, September 23-25 2008, accepted.

[17] V. Peterka, "Bayesian system identification", in *Trends and Progress in System Identification*, P. Eykhoff, Ed., pp. 239–304. Pergamon Press, Oxford, 1981.

[18] Greg Welch and Gary Bishop, "An Introduction to the Kalman Filter", Tech. Rep. 95-041, UNC-CH Computer Science, 1995.

[19] E. Suzdaleva, "Filtering with discrete-valued states", Tech. Rep. 2229, ÚTIA AV ČR, Praha, August 2008.

[20] J. Homolová and I. Nagy, "Traffic model of a microregion", in *Preprints of the 16th World Congress of the International Federation of Automatic Control*, P. Horáček, M. Šimandl, and P. Zítek, Eds., Prague, July 2005, pp. 1–6, IFAC.

[21] TSS Transport Simulation Systems, "AIMSUN microscopic traffic simulator: a tool for the analysis and assessment of ITS systems", Tech. Rep., *www.tss-bcn.com*.