The 8th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications
24-26 September 2015, Warsaw, Poland

Recursive Estimation of Mixtures of Exponential and Normal Distributions

Evgenia Suzdaleva[1] , Ivan Nagy[2, 1] , T. Mlynářová[1]
[1] Department of Signal Processing,
The Institute of Information Theory and Automation of the Czech Academy of Sciences,
Pod vodárenskou věží 4, 18208 Prague, Czech Republic,
suzdalev@utia.cas.cz, www.utia.cas.cz/people/suzdaleva
[2] Faculty of Transportation Sciences, Czech Technical University,
Na Florenci 25, 11000 Prague, Czech Republic, nagy@utia.cas.cz, www.utia.cas.cz/people/nagy

***Abstract* – The paper deals with estimation of a mixture of normal and exponential distributions with the dynamic model of their switching. A separate estimation of normal or exponential mixtures is solved by various approaches in many papers over the world. However, in some application areas, data are of such a nature that they should be described by a combination of exponential and normal models. The paper proposes a recursive Bayesian algorithm of estimation of such a mixture based on continuously measured data. Specific tasks the paper solves are: (i) parameter estimation of both the types of components; (ii) parameter estimation of the dynamic switching model and (iii) detection of the currently active component. Results of experiments with real data are demonstrated.**

***Keywords* – recursive mixture estimation; mixture of different distributions; dynamic switching model; exponential distribution**

I. INTRODUCTION

The presented paper deals with a problem of the recursive estimation of a mixture of normal and exponential distributions with dynamic model of their switching. In some application areas it is necessary to consider a combination of exponential and normal models because of the nature of measured data. It is extremely important, for example, in traffic control, for a task of on-line modeling a speed of cars passing the intersections, where its description by normal distributions is rather limited. In observing some longer stop of driving the speed values remain near zero, and after a start of driving they gradually grow. Thus, a combination of the exponential distribution in the beginning of modeled measurements and normal distributions for the rest of data is suitable. A similar task is considered in real-time analyzing data measured on a driven vehicle (for instance, speed, fuel consumption, pressing the gas pedal, etc.). This issue was a motivation of the presented research.

A separate estimation of normal mixtures is solved by different approaches in many papers over the world. The approaches are primarily based on (i) the EM algorithm [1], see, e.g., [2]–[6]; (ii) the Variational Bayes (VB) approach [7]–[9]; (iii) Markov Chain Monte Carlo (MCMC) techniques, e.g., [10]–[13]. Exponential mixtures are considered in, e.g., [14] based on the EM algorithm. A problem a bit close to the presented one is discussed in [15]. However, to avoid the off-line numerical computations, a solution presented in this paper is based on the entirely different philosophy and inspired by papers [16]–[19], which propose the recursive Bayesian estimation algorithms oriented at using the predominantly explicit solutions with the easily computable approximations.

Mixture models consist of components, describing different modes of the considered system behavior, and a model of switching the active modes. The random process describing the switching is called the pointer, which indicates the active component. This paper is also based on the experience from [20], which is devoted to the recursive mixture estimation with the dynamic switching model.

The paper proposes a recursive Bayesian estimation algorithm of the mixture of exponential and normal distributions based on continuously measuring the new data. This is the main contribution of the presented paper. Specific tasks the paper solves are: (i) parameter estimation of both the types of components; (ii) parameter estimation of the dynamic switching model and (iii) detection of the currently active component.

The layout of the paper is as follows. Section II introduces notations used in the text and the considered models and recalls existing algorithms of their estimation. Section III is the main emphasis of the paper and presents the estimation algorithm for a mixture of the exponential and normal components. Section IV provides results of experiments with real data. Conclusions can be found in Section V.

II. PRELIMINARIES

A. Notations

The paper uses the following notations:

The research was supported by projects GAČR GA15-03564S, MŠMT 7H14004 and MŠMT 7H14005.

- Both the probability density function and the probability function are replaced by the abbreviation *pdf*.
- For continuous random variables A, B and their realizations a, b the conditional pdf is denoted by

$$f_{A|B}(a|b) \equiv f(a|b).$$

- For categorical random variables C, D and their realizations c, d the conditional pdf is denoted by

$$f_{C|D}(c|d) \equiv f(C = c|D = d).$$

- Mixed distributions of random variables A, B, C, D with realizations a, b, c, d are denoted by

$$f_{A,C|B,D}(a, c|b, d) \equiv f(a, C = c|b, D = d).$$

- t used as a subscript of a random variable denotes discrete time instants, where the variable is measured, $t = \{1, 2, \dots\}$.
- x_t denotes the random variable x measured at the discrete time instant t.
- x^* is a set of all possible values of the random variable x.
- $x(t)$ is a collection of all available measurements of the random variable x up to the time instant t, i.e., $x(t) = \{x_0, x_1, \dots, x_t\}$, including prior data x_0.
- $\propto$ denotes equality up to the normalization constant.
- $\equiv$ means equality by definition.
- $y_t \in \mathbb{R}^{k_y}$ is the output vector of the dimension k_y measured on the observed system.
- In general, all variables are column vectors.

B. Components

In this paper the mixture model consists of m_c static components in the form of the following pdf

$$f(y_t|\Theta, c_t = i), \ \forall i \in \{1, 2, \dots, m_c\} \equiv c^*, \quad (1)$$

where Θ are parameters of components, and $c_t \in c^*$ is the random unmeasured categorical variable called the pointer, which indicates the component, active at time t.

One of the components (1) is the exponential pdf

$$\left(\prod_{l=1}^{k_y} a_l\right) \exp\{-a y_t\}, \quad (2)$$

where a is its parameter. It is a vector of the dimension k_y, and a_l are its entries, $l = 1, 2 \dots, k_y$.

The rest $m_c - 1$ components (1) are the normal pdfs specified as the static regression models with the normally distributed noise with the zero mean vector and the covariance matrix r_i, $\forall i \in \{1, 2, \dots, m_c - 1\}$, i.e.,

$$(2\pi)^{-k_y/2}|r_i|^{-1/2} \exp\left\{-\frac{1}{2}[y_t - \theta_i]' r_i^{-1}[y_t - \theta_i]\right\}, \quad (3)$$

where θ_i and r_i are parameters of the i-th normal component.

$$\Theta \equiv \{\{\theta_i, r_i\}_{i=1}^{m_c-1}\}, a\} \quad (4)$$

is a collection of parameters of all components.

C. Dynamic switching model

Switching the components is described by the dynamic model

$$f(c_t = i|c_{t-1} = j, \alpha), \ i, j \in c^*, \quad (5)$$

which is represented by the transition table

	$c_t = 1$	$c_t = 2$	$\cdots$	$c_t = m_c$
$c_{t-1} = 1$	$\alpha_{1\|1}$	$\alpha_{2\|1}$	$\cdots$	$\alpha_{m_c\|1}$
$c_{t-1} = 2$	$\alpha_{1\|2}$		$\cdots$	
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$c_{t-1} = m_c$	$\alpha_{1\|m_c}$		$\cdots$	$\alpha_{m_c\|m_c}$

where the parameter α is the $(m_c \times m_c)$-dimensional matrix, and its entries $\alpha_{i|j}$ are the probabilities of the pointer $c_t = i$ (expressing that the i-th component is active at time t) under condition that the previous pointer $c_{t-1} = j$, for $i, j \in c^*$ and it holds

$$\alpha_{i|j} \geq 0, \sum_{k=1}^{m_c} \alpha_{k|j} = 1, \ \forall i, j, k \in c^*. \quad (6)$$

D. Estimation of individual models

The individual Bayesian estimation of (2), (3) and (5) is briefly recalled here. It is based on Bayes rule [19]

$$f(\Theta|y(t)) \propto f(y_t|\Theta) f(\Theta|y(t-1)), \quad (7)$$

where $f(\Theta|y(t-1))$ is the prior pdf to be reproduced.

For the individual exponential model (2), substitution of (2) and the prior exponential pdf with the statistics S_{t-1} and N_{t-1} (initially chosen) in (7) enables to recursively update them for the time t as follows:

$$S_t = S_{t-1} + y_t, \quad (8)$$

$$N_t = N_{t-1} + 1. \quad (9)$$

The point estimate of the parameter a is obtained from

$$\hat{a}_{l;t} = \frac{N_t}{S_{l;t}}, \ l = 1, 2 \dots, k_y, \quad (10)$$

where $\hat{a}_{l;t}$ are entries of the vector $\hat{a}_t$, and $S_{l;t}$ are entries of the vector S_t from (8), see, e.g., [21].

For estimation of the individual normal model (3) via (7), the conjugate prior Gauss-inverse-Wishart pdf with the recomputable (initially chosen) statistics V_{t-1} and k_{t-1} of the appropriate dimensions is used according to [17], [19]. The statistics are updated as follows:

$$V_t = V_{t-1} + \begin{bmatrix} y_t \\ 1 \end{bmatrix} [y_t, 1], \quad (11)$$

$$\kappa_t = \kappa_{t-1} + 1, \quad (12)$$

and the point estimates of parameters are:

$$\hat{\theta}_t = V_1^{-1} V_y, \quad \hat{r}_t = \frac{V_{yy} - V_y' V_1^{-1} V_y}{\kappa_t} \quad (13)$$

with the help of partition

$$V_t = \begin{bmatrix} V_{yy} & V_y' \\ V_y & V_1 \end{bmatrix}, \quad (14)$$

where V_{yy} is the square matrix of the dimension k_y of the vector y_t, V_y' is k_y-dimensional column vector and V_1 is scalar [19].

The individual categorical model (5) in the case of the measured values of c_t and c_{t-1} is estimated via (7) using the conjugate prior Dirichlet pdf according to [18] with the recomputable statistics ϑ_{t-1}, which is here the square m_c-dimensional matrix, whose entries for $c_t = i$ and $c_{t-1} = j$ are updated in the following way:

$$\vartheta_{i|j;t} = \vartheta_{i|j;t-1} + 1. \quad (15)$$

The point estimate of α is then obtained by

$$\hat{\alpha}_{i;t} = \frac{\vartheta_{i|j;t}}{\sum_{k=1}^{m_c} \vartheta_{k|j;t}}, \;\; i, j \in c^*. \quad (16)$$

E. Problem formulation

The task is: based on the available data collection, estimate recursively a mixture of one exponent (2) and $m_c - 1$ components (3) switching according to (5) with the unmeasured pointer, i.e., parameters α, θ_i, r_i, a and the pointer c_t.

III. Recursive estimation of mixtures of exponential and normal components

The algorithm proposed below is inspired by [17], where the Bayesian recursive estimation of normal mixtures with the static pointer model was proposed, and by [20] solved the problem for the dynamic pointer model. The derivations are based on construction of the joint pdf of all variables to be estimated and application of the Bayes rule and of the chain rule [19]. Here, to save space, they are explained briefly.

Under assumption of the mutual independence of Θ and α, and y_t and α, and c_t and Θ, the joint pdf of all variables to be estimated takes the form

$$\begin{aligned}
&\underbrace{f(\Theta, c_t = i, c_{t-1} = j, \alpha | y(t))}_{joint\ posterior\ pdf} \\
&\propto \underbrace{f(y_t, \Theta, c_t = i, c_{t-1} = j, \alpha | y(t-1))}_{via\ chain\ rule\ and\ Bayes\ rule} \\
&= \underbrace{f\left(y_t | \Theta, c_t = i\right)}_{(1)} \underbrace{f(\Theta | y(t-1))}_{prior\ pdf\ of\ \Theta} \\
&\times \underbrace{f\left(c_t = i | c_{t-1} = j, \alpha\right)}_{(5)} \underbrace{f(\alpha | y(t-1))}_{prior\ pdf\ of\ \alpha} \\
&\times \underbrace{f(c_{t-1} = j | y(t-1))}_{prior\ pointer\ pdf}, \forall i, j \in c^*. \quad (17)
\end{aligned}$$

To obtain recursive formulas for estimation of c_t, Θ and α with the help of (17), it is necessary to marginalize it firstly over the parameters Θ and α. It will give the posterior pdf $f(c_t = i, c_{t-1} = j | y(t))$, which is joint for both c_t and c_{t-1}. In order to obtain the posterior pdf $f(c_t = i | y(t))$, this joint pdf should be again marginalized over the values of c_{t-1}.

For marginalization of (17) over parameters Θ, the integral of (17) over Θ is evaluated by substituting the point estimate (10) and the i-th point estimates (13) available from the previous time instant $t - 1$ and the currently measured y_t into the exponential component (2) and into each i-th normal component (3) respectively. This substitution provides the proximity of each component to the current output y_t.

Similarly, the integral of (17) over α provides the computation of its point estimate (16) using the previous-time statistics ϑ_{t-1}.

After the above marginalization of (17) over Θ and α, the posterior pdf $f(c_t = i, c_{t-1} = j | y(t))$ is obtained by entry-wise multiplying the proximity obtained from each component, the previous-time point estimate of α and the prior pointer pdf $(c_{t-1} = j | y(t-1))$. The last is denoted by $w_{j;t-1}$ and represents the (initially chosen) probability of the activity of the j-th component at time $t - 1$. For all $i, j \in c^*$, posterior pdf $f(c_t = i, c_{t-1} = j | y(t))$ is the square m_c-dimensional matrix denoted by $W_{i,j;t}$, which is normalized and summed up over rows to obtain the posterior pdf $f(c_t = i | y(t))$. The last provides the updated probabilities $w_{i;t}$ of activity of each i-th component at time t for $i \in c^*$. The maximal probability $w_{i;t}$ defines the currently active component, i.e., the point estimate of the pointer c_t at time t.

The probability $w_{i;t}$ for $i \in \{1, 2, \ldots, m_c - 1\}$ is used in the updates (11), (12) of the statistics of the i-th normal component according to [17], i.e.,

$$\begin{aligned}
V_{i;t} &= V_{i;t-1} + w_{i;t} \begin{bmatrix} y_t \\ 1 \end{bmatrix} [y_t, \, 1], \quad (18) \\
\kappa_{i;t} &= \kappa_{i;t-1} + w_{i;t}, \quad (19)
\end{aligned}$$

and now in the update (8) – (9) of the exponential component as follows:

$$\begin{aligned}
S_t &= S_{t-1} + w_{i;t} y_t, \quad (20) \\
N_t &= N_{t-1} + w_{i;t}, \quad (21)
\end{aligned}$$

where $i \notin \{1, 2, \ldots, m_c - 1\}$ used in (18)–(19), but $i \in c^*$.

The update (15) is performed for the matrix-form statistics with entries $\vartheta_{i|j;t}$ in the following way. In [20] the solution was introduced with the approximation based on the Kerridge inaccuracy [22]. Here, for simplicity, it is updated similarly to [17], but modified for the dynamic case:

$$\vartheta_{i|j;t} = \vartheta_{i|j;t-1} + W_{i,j;t}, \; i, j \in c^*. \quad (22)$$

The initial statistics $V_{i;0}$, $\kappa_{i;0}$, S_0, N_0 and ϑ_0 can be chosen as the small-valued matrices (respectively vectors) of the appropriate dimensions. For the real data application it is suitable to construct them from some part of the prior data.

The algorithm can be summarized as follows.

Algorithm

Initial part (for t=1)

- Specify the exponential component (2), $m_c - 1$ normal components (3) and the switching model (5).
- Set initial statistics of all components $V_{i;0}$, $\kappa_{i;0}$, S_0, N_0 and ϑ_0.
- Using these initial statistics, compute the initial point estimates of all parameters and for all components according to (10), (13) and (16).
- Set the initial m_c-dimensional vector w_0.

On-line part (for t=2,...)

1) Measure the new data y_t.
2) For all components, substitute y_t and the point estimates $\hat{\theta}_{i;t-1}$ and $\hat{r}_{i;t-1}$ into each normal component, and $\hat{a}_{t-1}$ into the exponential component. Construct the m_c-dimensional vector of proximities from results from all components.
3) Multiply entry-wise the resulted vector from the previous step, the prior weighting vector w_{t-1} and the point estimate matrix $\hat{\alpha}_{t-1}$.
4) The result of this entry-wise multiplication is the matrix with entries $W_{i,j;t}$. Normalize this matrix.
5) Perform the summation of the normalized matrix over rows and obtain the updated vector w_t with entries $w_{i;t}$.
6) Detect the currently active component according to the maximal probability $w_{i;t}$, if necessary.
7) Update all statistics, using $w_{i;t}$ and $W_{i,j;t}$ according to (18), (19), (20), (21) and (22).
8) Recompute the point estimates of all parameters according to (10), (13) and (16) and use them for Step 1 of the on-line part of the algorithm.

IV. Results

Here the proposed algorithm is validated by testing on real data measured on a vehicle during driving each 0.2 seconds. Two measured variables create the vector y_t. One of them expresses pressing the gas pedal [%], and the second one is the engine torque [Nm]. The whole number of the used data is 17000 data items.

The chosen number of normal components is 5, which means that the whole number of components together with the exponential one is 6. One of the most illustrative visualization of the approach is the evolution of the parameter estimates during the on-line part of the algorithm. However, to save space it is not shown here. The finally obtained stabilized point estimates are demonstrated below. The point estimate of the parameter a of the exponential component takes the form

$$\hat{a}_t = \left[\begin{array}{c} 1.80 \\ 0.11 \end{array} \right]. \quad (23)$$

The point estimates of the regression coefficients of 5 normal components are

$$\hat{\theta}_{1;t} = \left[\begin{array}{c} 18.08 \\ 44.76 \end{array} \right], \hat{\theta}_{2;t} = \left[\begin{array}{c} 39.07 \\ 84.36 \end{array} \right],$$

$$\hat{\theta}_{3;t} = \left[\begin{array}{c} 52.77 \\ 124.32 \end{array} \right], \hat{\theta}_{4;t} = \left[\begin{array}{c} 72.21 \\ 180.1 \end{array} \right],$$

$$\hat{\theta}_{5;t} = \left[\begin{array}{c} 93.58 \\ 224.67 \end{array} \right]. \quad (24)$$

The point estimates of the covariance matrices $\hat{r}_{i;t}$ of each normal component of the dimension (2×2) are not shown here to save space. The obtained point estimate $\hat{\alpha}_t$ of the switching model is the matrix of the dimension (6×6).

A selected fragment of results of the estimation of switching the components is shown in Figure 1. 500 data items are chosen for the clear visible presentation. It can be seen that the activity of the components switch among all 6 of them. This confirms that the model is well established.

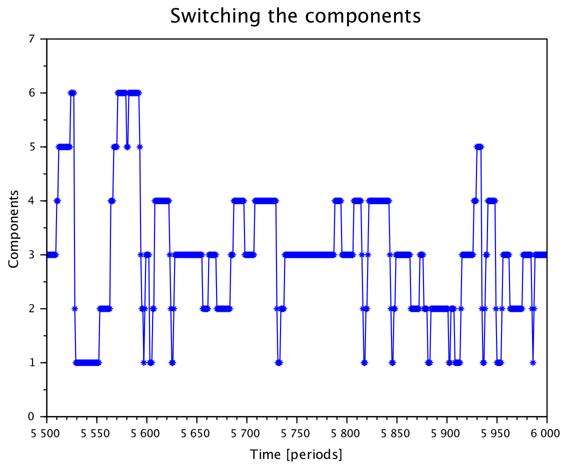

Figure 1. Switching the active components

Value 1 corresponds to the exponential component, while the rest of values denote normal components. Notice that all the components are switching.

With the aim of validation of the approach, the obtained point estimates of parameters are used for predicting the data by substituting them into corresponding models. Figure 2 shows a selected fragment in the beginning of the on-line part of the algorithm, where pressing the gas pedal is plotted against the engine torque. 6 clusters (here components) are visible. Figure 3 demonstrates the same results approximately in the middle of the on-line part of the algorithm with the already accumulated data and the sufficiently updated statistics. It leads to refining the point estimates and better predicting.

The illustrative presentation of the obtained results is shown in Figure 4 in the form of the histogram plot. Frequencies corresponding to values of pressing the gas pedal near zero represent the exponential component. The rest of values are covered by 5 normal distributions.

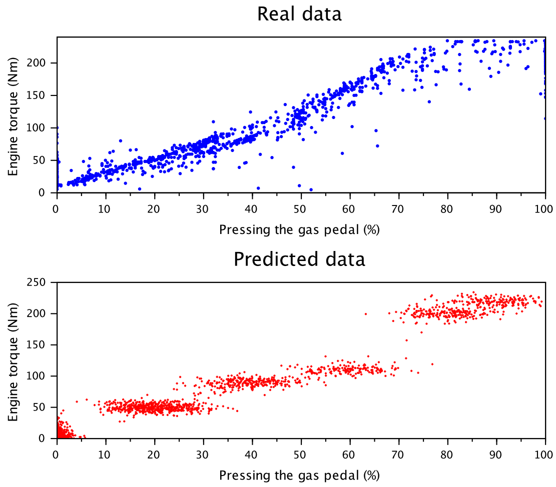

Figure 2. Clusters in the beginning of the estimation
The top figure plots the real measurements of pressing the gas pedal against the engine torque. The bottom figure shows the same for the predicted data. Notice 6 visible clusters in the bottom figure.

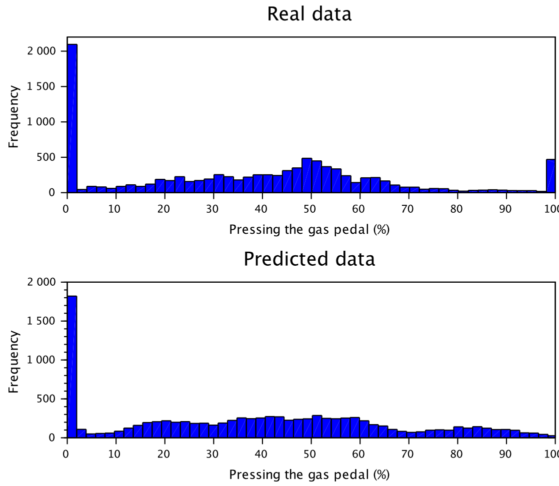

Figure 4. Histogram of real and predicted data
Notice frequencies of the clearly visible one exponential component near zero values. The rest of them correspond to five normal components. Notice correspondence of the predicted data in the bottom figure to the original real data in the top figure.

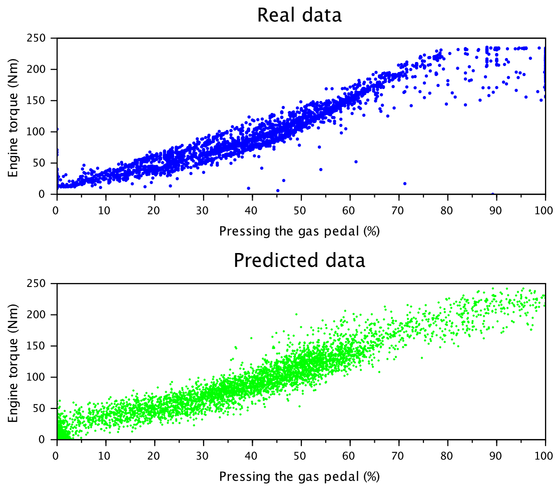

Figure 3. Comparison of real and predicted data
Here the same results as in Figure 2 are shown with the already accumulated data and the sufficiently updated statistics. Notice that the predicted data in the bottom figure correspond to the real data in the top figure.

V. Conclusions

The paper considers a problem of the recursive estimation of mixtures of exponential and normal distributions. The task can be significant for the on-line analysis of specific data, where their description by one type of distributions (here normal) is not sufficient and leads to limitations. The paper explores a possibility to describe data by the combination of exponential and normal distributions and presents the recursive Bayesian estimation algorithm. The results obtained during the experimental part of the work with real measurements validate the approach and show it as competitive.

Plans of the future work in the considered context are directed at further systematic developing the recursive mixture estimation theory from the viewpoint of clustering and classification of data. They include exploring the following issues: (i) extension of the estimation algorithms for different combinations of components (categorical, uniform and other (for instance, general triangular) distributions); (ii) the mixture estimation with the dynamic data dependent switching model; (iii) multi-step prediction of activity of components, etc.

References

[1] M. R. Gupta and Y. Chen, *Theory and Use of the EM Method. (Foundations and Trends(r) in Signal Processing)*. Now Publishers Inc, 2011.

[2] O. Boldea and J. Magnus, "Maximum likelihood estimation of the multivariate normal mixture model," *Journal Of The American Statistical Association*, vol. 104, no. 488, pp. 1539–1549, 2009.

[3] J. Cuesta-Albertos, C. Matran, and A. Mayo-Iscar, "Robust estimation in the normal mixture model based on robust clustering," *Journal Of The Royal Statistical Society Series B-Statistical Methodology*, vol. 70, no. 4, pp. 779–802, 2008.

[4] H. Wang, B. Luo, Q. Zhang, and S.Wei, "Estimation for the number of components in a mixture model using stepwise split-and-merge em algorithm," *Pattern Recognition Letters*, vol. 25, no. 16, 2004.

[5] H. Zeng and Y. Cheung, "Learning a mixture model for clustering with the completed likelihood minimum message length criterion," *Pattern Recognition*, vol. 47, no. 5, pp. 2011–2030, 2014.

[6] S. Ng and G. McLachlan, "Mixture models for clustering multi-level growth trajectories," *Computational Statistics & Data Analysis*, vol. 71, pp. 43–51, 2014.

[7] C. A. McGrory and D. M. Titterington, "Variational Bayesian analysis for hidden Markov models," *Australian & New Zealand Journal of Statistics*, vol. 51, p. 227–244, 2009.

[8] V. Šmídl and A. Quinn, *The Variational Bayes Method in Signal Processing*. Springer, 2005.

[9] Z. Ghahramani and G. E. Hinton, "Variational learning for switching state-space models," *Neural Computation*, vol. 12, no. 4, pp. 831–864, 2000.

[10] S. Frühwirth-Schnatter, *Finite Mixture and Markov Switching Models. 2nd Edition*. Springer New York, 2006.

[11] S. Früwirth-Schnatter, "Fully bayesian analysis of switching gaussian state space models," *Annals of the Institute of Statistical Mathematics*, vol. 53, no. 1, pp. 31–49, 2001.

[12] A. Doucet and C. Andrieu, "Iterative algorithms for state estimation of jump markov linear systems," *IEEE Transactions on Signal Processing*, vol. 49, no. 6, pp. 1216–1227, 2001.

[13] R. Chen and J. S. Liu, "Mixture kalman filters," *J. R. Statist. Soc.B*, vol. 62, pp. 493–508, 2000.

[14] K.Barger, "Mixtures of exponential distributions to describe the distribution of poisson means in estimating the number of unobserved classes," Ph.D. dissertation, Cornell University, 2006.

[15] J. Silver, M. E. Ritchie, and G. Smyth, "Microarray background correction: maximum likelihood estimation for the normal–exponential convolution," *Biostat*, vol. 10, no. 2, pp. 352–363, 2009.

[16] D. M. Titterington, A. Smith, and U. Makov, *Statistical analysis of finite mixture distributions. Wiley series in probability and mathematical statistics: Applied probability and statistics*. Wiley, 1985.

[17] M. Kárný, J. Kadlec, and E. Sutanto, "Quasi-bayes estimation applied to normal mixture," in *Preprints of the 3rd European IEEE Workshop on Computer-Intensive Methods in Control and Data Processing*, J. Rojíček, M. Valečková, M. Kárný, and K. Warwick, Eds. CMP '98 /3./, Prague, CZ, September 1998, pp. 77–82.

[18] M. Kárný, J. Böhm, T. V. Guy, L. Jirsa, I. Nagy, P. Nedoma, and L. Tesař, *Optimized Bayesian Dynamic Advising: Theory and Algorithms*. London, Springer, 2006.

[19] V. Peterka, *Trends and Progress in System Identification*. Oxford: Pergamon Press, 1981, ch. Bayesian system identification, pp. 239–304.

[20] I. Nagy, E. Suzdaleva, M. Kárný, and T. Mlynářová, "Bayesian estimation of dynamic finite mixtures," *Int. Journal of Adaptive Control and Signal Processing*, vol. 25, no. 9, pp. 765–787, 2011.

[21] L.Yang, H. Zhou, and S.Yuan, "Bayes estimation of parameter of exponential distribution under a bounded loss function," *Research Journal of Mathematics and Statistics*, vol. 5, no. 4, pp. 28–31, 2013.

[22] D. Kerridge, "Inaccuracy and inference," *Journal of Royal Statistical Society B*, vol. 23, pp. 284–294, 1961.