A Hybrid Factor Graph Model for Biomedical Activity Detection

Mareike Stender *, Jan Graßhoff †, Tanya Braun ‡, Ralf Möller *, Philipp Rostalski †
*German Research Center for Artificial Intelligence, Lübeck {mareike.stender, ralf.moeller}@dfki.de
†Institute for Electrical Engineering in Medicine, University of Lübeck, Lübeck {j.grasshoff, philipp.rostalski}@uni-luebeck.de
‡Institute of Information Systems, University of Lübeck, Lübeck braun@ifis.uni-luebeck.de

***Abstract*—For activity detection on biomedical time-series data, biomedical signals are modeled as a switching linear dynamical system with random variables, including discrete and continuous dynamics. We present a formalism for representing a system's joint probability density function as a hybrid factor graph. Solving inference problems is based on belief propagation using message passing. Inference results yield the activity estimations in terms of probability distributions instead of binary decisions. This work builds on previous efforts to consolidate factor graphs as unifying representations for signal processing algorithms. We show that the formalism can be successfully applied to detect activities in surface electromyography data acquired during walking. The modularity of factor graphs enables the straightforward adoption and extension of the formalism expanding its scope of application.**

***Index Terms*—factor graphs, probabilistic graphical models, inference algorithm, activity estimation, activity detection, belief propagation, biomedical engineering**

I. Introduction

An important goal of biomedical engineering is the interpretation of biomedical signals, for which activity changes and abnormalities must be detected. Robust activity detection on biomedical time-series data may then permit the interpretation for diagnostic purposes, the derivation of clinical decisions, or even a closed-loop control of medical devices.

In practice, many biomedical signals exhibit switching behavior, i.e., the underlying physiological system changes rapidly or the data acquisition is corrupted due to a sudden failure. We consider a concrete example using surface electromyography (sEMG) data, reflecting the activity of separate muscles or groups of muscles. Activities can be modeled as discrete random variables with the categories *active* and *inactive*, and sEMG data can be modeled as a continuous random variable. To combine discrete and continuous random variables, we model a switching linear dynamical system (SLDS) in hybrid factor graphs. As a subset of probabilistic graphical models, factor graphs provide a consistent framework, modular setup, and high clarity due to the graphical representation. Furthermore, inference algorithms can be directly derived from belief propagation for calculating marginal distributions.

SLDSs have been used in a variety of domains: In control theory, they were investigated as *Markov Jump Linear Systems* [1]. In the context of state estimation and target tracking, the SLDS concept was introduced as *Interacting Multiple Model* estimator [2] or *Switching Kalman Filter* [3].

In the machine learning community, several approximate inference algorithms have been introduced, see for instance [4]. In the context of factor graphs, SLDSs have so far been considered by Zoeter and Heskes [5] to solve inference problems on respective factor graphs by applying Expectation Propagation. The study presents the Markov structure of the model graphically. However, it does not include the derivation of nodes and message types required [5]. In comparison, we use the sum-product algorithm [6] to derive the inference algorithm on the modeled SLDS.

Within the scope of this work, we implement the nodes and message types needed to facilitate modular construction of hybrid factor graphs. We describe the derivation of the inference algorithm on the resulting hybrid factor graph based on forward and backward message-passing and show that the resulting inference algorithm enables activity detection on sEMG data where activities are determined as probability distributions instead of binary decisions.

We extend previous efforts aiming to introduce factor graphs as a unifying approach to signal processing, see [7]. In the same spirit, we provide straightforward factor graph building blocks, which can be easily combined and connected with the previous factor graph nodes in [7]. Messages can be passed over the graph using simple-to-use tabulated propagation rules.

The rest of the paper is structured as follows: In Section II the fundamentals of factor graphs including operation rules are depicted. Following, the mathematical description of the SLDS and its modeling in hybrid factor graphs is outlined in Section III. Section IV presents the application of the developed inference algorithm to sEMG data sets recorded during walking. Lastly, the conclusion summarizes the findings of the work and points out limitations.

II. Fundamentals

A main goal in probabilistic models is to determine the marginal probability density function (MPDF) of a subset of variables using the joint probability density function (JPDF) over all variables in the model. Generally, for computing MPDFs, we need to solve the integral over all variables. Message passing on factor graphs allows to simplify computationally expensive calculations by exploiting the conditional independence structure of the probability density function:

978-1-6654-0358-0/21/$31.00 ©2021 IEEE

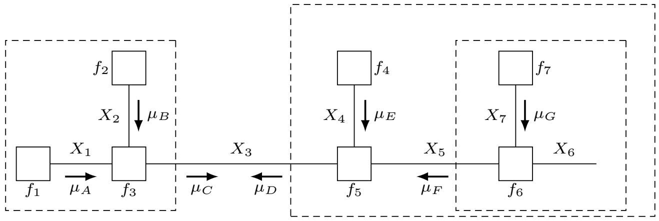


Figure 1: Graphical representation of Equation (1) with $\overrightarrow{\mu}$ as forward and $\overleftarrow{\mu}$ backward messages on the edges towards X_3. Dashed boxes highlight the combination of nodes during message passing by eliminating variables.

Global functions (dependent on all variables) are factorized by a product of local functions, each depending only on a subset of variables [6]. This in turn allows to eliminate variables before multiplying factors during inference.
The following declaration of the functionalities of factor graphs is based on the work of Loeliger et al. [7].
Figure 1 depicts a factor graph for the factorized JPDF

$$f(x_1,\dots,x_7) = f_1(x_1)f_2(x_2)f_3(x_1,x_2,x_3)f_4(x_4) \cdot f_5(x_3,x_4,x_5)f_6(x_5,x_6,x_7)f_7(x_7) \quad (1)$$

of continuous random variables $X_1,\dots,X_7$. The messages μ on the edges of the factor graph are determined by the sum-product algorithm: By applying the sum-product rule, messages are computed from leaves inward, when the factor graph is a tree and all required incoming messages are given, see [7] for more details on this convention. The direction of the messages is given by the direction of the mathematical operation indicated by the used nodes. In Figure 1, the MPDF for the variable/edge X_3

$$\overline{f}_3(x_3) \propto \overrightarrow{\mu}_C(x_3)\overleftarrow{\mu}_D(x_3) \quad (2)$$

is determined as the product of the forward message $\overrightarrow{\mu}_C(x_3)$ and the backward message $\overleftarrow{\mu}_D(x_3)$ on the edge X_3.
Passing messages over a factor graph combines nodes to subgraphs (dashed boxes in Figure 1) by eliminating variables. Despite high model complexity, this leads to efficient inference algorithms. The classical message-passing algorithm applied to cyclic factor graphs usually does not converge to the exact solution. Section III includes an approach for dealing with specific cyclic factor graphs, however.

Message Types In factor graphs, various types of distributions can be used as messages. In Figure 1, the factors f represent abstract nodes without specific functions for illustrating the basic functionalities of factor graphs. For the use of Gaussian distributions, Loeliger et al. [7] define basic nodes for e.g. multiplication and summation according to tabulated rules. In this work, we additionally use discrete distributions, Gaussian mixture distributions, and their combination as messages on a hybrid factor graph.
A discrete message $\mu_X(x)$ – representing the distribution of a discrete random variable X – is defined by the parameters α_{X_i} specifying the probability of the respective discrete category i with $i = 1,\dots,k$. The semantics of a discrete message is:

$$p_X(x) = \sum_{x_i \in R_X} \alpha_{X_i}\, \delta(x - x_i) \text{ with } \alpha_{X_i} \geq 0, \quad (3)$$

where δ defines the Dirac delta function. The range of the discrete variable X is $R_X = \{x_1, x_2, \dots, x_k\}$. A Gaussian mixture message $\mu_X(x)$ – representing the distribution of a continuous normal distributed random variable X – is defined as a tuple of the parameters weights w_{X_i}, means m_{X_i}, and variances V_{X_i} depicting a sum of n Gaussian distributions, where i represents the respective Gaussian distribution in the sum. The semantics of a Gaussian mixture message is:

$$p_X(x) = \sum_{i=1}^{n} w_{X_i} \cdot \mathcal{N}(x; m_{X_i}, V_{X_i}), \text{ with } w_{X_i} \geq 0. \quad (4)$$

III. SLDS as Hybrid Factor Graph

The combination of discrete and continuous variables calls for a hybrid representation. For this, we model an SLDS that combines various linear dynamical systems, switching dependent on the discrete state. Hereafter, we give the mathematical description of an SLDS following the definition in [4].

SLDS The hybrid dynamics of the system is given by the discrete and the continuous latent variables, H_t and X_t in each time step $t = 1,\dots,T$. The observation is modeled as a continuous variable Y_t. The transition distribution

$$p_{H_t|H_{t-1}}\left(h^{(l)} \middle| h^{(k)}\right) = P\left(H_t = h^{(l)} \middle| H_{t-1} = h^{(k)}\right) \quad (5)$$

represents a table, which contains the probabilities with which the system switches between discrete categories of H_t.
In the probabilistic sense, the distribution $p_{H_t|H_{t-1}}\left(h^{(l)} \middle| h^{(k)}\right)$ describes a conditional joint probability density function of H_t given the distribution of H_{t-1}. Superscripts indicate the category of a discrete variable. The range of values of H_t is defined as $R_H = \left\{h^{(1)}, h^{(2)}, \dots, h^{(q)}\right\}$. The continuous state transition equation for an SLDS is

$$X_t = A(H_t)X_{t-1} + W \text{ with } W \sim \mathcal{N}(0, Q(H_t)) \quad (6)$$

and the observation equation is

$$Y_t = C(H_t)X_t + V \text{ with } V \sim \mathcal{N}(0, R(H_t)). \quad (7)$$

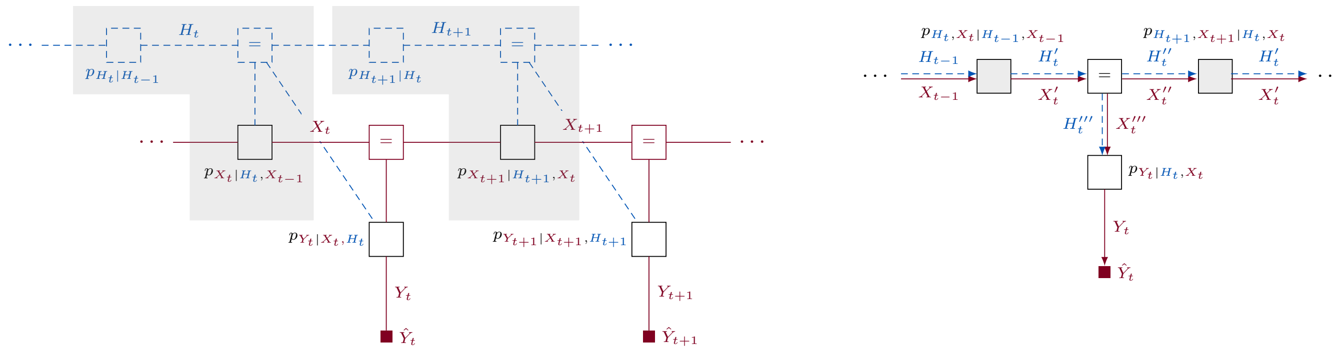
The discrete variable H_t defines which linear dynamical system is to be used in time step t. In case $H_t = h^{(l)}$, the linear dynamical system is defined by matrices A_l, Q_l, C_l and R_l. The observed variable Y_t depends on latent variables H_t and X_t. The initial discrete distribution is

$$p_{H_0}\left(h^{(l)}\right) = P\left(H_0 = h^{(l)}\right) \quad (8)$$

and the initial continuous distribution is

$$p_{X_0|H_0}\left(x \middle| h^{(l)}\right) = \mathcal{N}\left(x; m_{X_0}^{(l)}, V_{X_0}^{(l)}\right). \quad (9)$$

The JPDF of both discrete and continuous variables over T



(a) Factor graph of an common SLDS with discrete-continuous dynamics. The gray boxes depict the nodes that are combined to one single node for eliminating the loops from the graph.

(b) An SLDS transformed as a loop-free hybrid factor graph. The gray shaded nodes represent the combined nodes of the factor graph in (a).

Figure 2: SLDS in hybrid factor graphs (a) and SLDS in hybrid factor graphs with tree structure (b). Blue dashed edges/nodes are discrete, red edges/nodes are continuous. Arrows highlight the direction of the arithmetic operation of the used nodes, nevertheless, factor graphs are still undirected graphs.

time steps is:

$$p_{Y_0,\ldots,Y_T,X_0,\ldots,X_T,H_0,\ldots,H_T} = p_{H_0} p_{X_0|H_0} p_{Y_0|X_0,H_0} \cdot \prod_{t=1}^{T} p_{H_t|H_{t-1}} p_{X_t|X_{t-1},H_t} p_{Y_t|X_t,H_t}. \tag{10}$$

SLDS as Hybrid Factor Graph The graphical representation of the SLDS – described mathematically above – is shown in Figure 2a as a hybrid factor graph containing discrete and continuous nodes. Since the hybrid factor graph contains loops, inference cannot be solved directly by applying the sum-product rule. To deal with loops, we exploit the property of factor graphs that nodes can be summarized by a compound node. We eliminate the loops by summarizing the nodes highlighted by gray boxes. Combining the nodes requires the propagation of mixed discrete/continuous messages over the gray nodes. To this end, we introduce the cluster message type, which contains a discrete and a continuous part and is based on the cluster distributions used in [4].

Cluster Messages We define a cluster message – representing the combined distribution of the discrete random variable H_t and the continuous normal distributed random variable X_t – as a tuple of the parameters weights $w_{X_i}^{(l)}$, means $m_{X_i}^{(l)}$, and variances $V_{X_i}^{(l)}$ depicting l sums of n Gaussian distributions, where i represents the respective Gaussian distribution in the sum, and l is the category of the discrete variable H_t. The semantics of a cluster message is:

$$p_{X,H}(x,h) = \sum_{i,l} w_{X_i}^{(l)} \delta\Big(h - h^{(l)}\Big) \mathcal{N}\Big(x; m_{X_i}^{(l)}, V_{X_i}^{(l)}\Big), \quad \text{with } w_{X_i}^{(l)} \geq 0. \tag{11}$$

Figure 3 visualizes a cluster message including the marginalized discrete and continuous distributions whose discrete state contains three categories. The index l depicts the category of the discrete variable H_t. For each discrete category l the cluster message contains a Gaussian mixture distribution. To

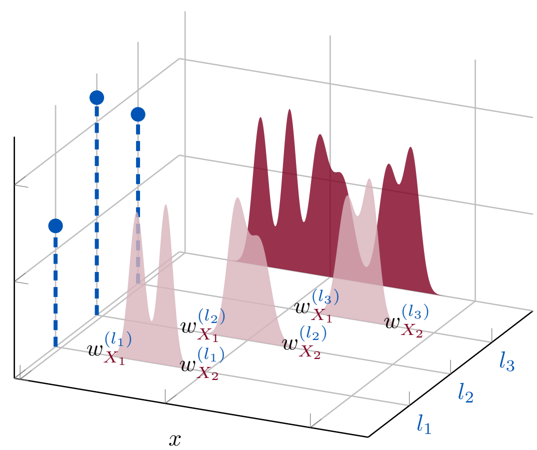


Figure 3: Graphical representation of a cluster message with a discrete state containing three categories l_1, l_2 and l_3. For each discrete category l the cluster message contains a Gaussian mixture distribution. The components of the discrete distribution of the cluster message – marginalized over the continuous variable – are shown as blue dashed lines. In dark red, the Gaussian mixture distribution is shown marginalized over all discrete categories.

determine the discrete distribution of the cluster message the continuous variable is marginalized.

By using cluster messages and combining the highlighted nodes in the gray boxes, the hybrid factor graph in Figure 2a is transferred to the hybrid factor graph in Figure 2b.

Inference Now that the resulting factor graph is a tree, the inference problem is solvable by message passing: Inference is performed by passing cluster messages forward and backward over the whole factor graph. The MPDFs are calculated on the edges H''_t, X''_t by combining the forward and backward message. The combination is performed by multiplying each part of one message with each part of the other message.

Pruning When combining Gaussian mixture distributions, the number of Gaussian distributions increases in a quadratic way in each time step t and thus exponentially over time. Large numbers of Gaussian distributions obtained with a large T lead to high computation times, which call for a approximation step to preserve computational performance. We account for this by applying pruning methods, which trim the number of Gaussian distributions dependent on fixed parameters. As in any kind of approximation, empirical errors occur with pruning and have to be traded off against accuracy. Other methods for pruning Gaussian mixtures are described in [8].

IV. Application on sEMG data

This section presents the application of the inference algorithm on the hybrid factor graph of Figure 2b, on sEMG data to showcase its applicability in biomedical signal processing. A key problem in activity detection on sEMG signals is noise. Since the amplitude of sEMG signals is low, they are characterized by a small signal to noise ratio (SNR) due to the high noise components, making activity detection more difficult. The processes occurring during the generation of the sEMG signals can be approximated by means of an autoregressive model (AR model) [9]. In an AR model, each sample is represented as a linear combination of previous samples and a noise term. The model presented in Section III can be understood as an AR model and will be used in the following to perform activity detection on real sEMG data.

We use sEMG data sets of ten subjects from the study in [10], which contains data of the subjects right musculus tibialis anterior, recorded during walking. The goal was to evaluate the estimates of the discrete state – corresponding to the muscle activity/inactivity – and the continuous state. We use a $30\,\mathrm{s}$ segment of the raw sEMG signal, which is sampled with a rate of $1500\,\mathrm{Hz}$, down-sampled to $500\,\mathrm{Hz}$.

Table I illustrates quantitative results, by means of precision and recall using the proposed inference algorithm on the hybrid factor graph of Figure 2b. We empirically set the model parameters and artificially added noise to the raw sEMG data to generate different SNRs. For determining precision and recall, the average is reported over the ten subjects. Precision and recall provide accurate estimation results for high SNRs compared to small SNRs. Decreasing SNRs < 15 lead to decreasing estimation accuracy. Each activity detection on a sEMG data set takes a computation time of $\approx 5\,\mathrm{min}$.

TABLE I: Evaluation of the estimation results of the inference algorithm on the hybrid factor graph of Figure 2b on sEMG data. Different artificial SNRs are modeled by adding noise to the raw sEMG data.

SNR	**Precision [%]**		**Recall [%]**	
[dB]	**inactive**	**active**	**inactive**	**active**
15	19.52	91.86	11.24	99.78
25	54.16	98.19	38.57	99.82
35	95.62	99.69	91.23	99.88
45	97.45	99.89	98.30	99.93
55	98.45	99.96	98.90	99.97

V. Conclusion

We successfully modeled a switching linear dynamical system in hybrid factor graphs. We solved inference problems by applying message-passing based on the sum-product algorithm, combined with a pruning method to account for the exponential growing of the Gaussian mixture distributions. The inference algorithm over the factor graph delivers suitable estimation results on sEMG data recorded during walking, where small SNRs cause higher estimation errors than high SNRs. All required nodes and message types have been implemented. Hence, a modular construction of various types of hybrid factor graphs becomes possible. Due to the modularity, factor graphs can be easily extended by factors integrating reference signals or various measurements.

There are four major limitations in this study that could be addressed in future research: First, the empirical tuning of the model parameters lead to inaccuracy. Second, the used pruning method discards Gaussians, which causes loss of information. Third, estimates with SNRs < 15 become mostly useless. Finally, the high computational time $\approx 5\,\mathrm{min}$ is impractical.

Despite mentioned limitations, our work extends previous factor graph frameworks [7] and can be easily combined with other factor graphs in a mix-and-match style. It is thus a further step towards unifying signal processing algorithms in a single framework. We have motivated the usefulness in a first biomedical application and plan to extend these first results to larger data sets and other biomedical signals, such as electrocardiography and electroencephalography in the future.

References

[1] S. L. Nogueira, A. A. G. Siqueira, R. S. Inoue, and M. H. Terra, “Markov jump linear systems-based position estimation for lower limbs exoskeletons,” in American Control Conference, 2014, pp. 4095-4100.

[2] E. Mazor, A. Averbuch, Y. Bar-Shalom, and J. Dayan, “Interacting multiple model methods in target tracking: A survey,” IEEE Transactions on Aerospace and Electronic Systems, vol. 34, no. 1, pp. 103-123, 1998.

[3] K. P. Murphy, “Switching Kalman filters,” University of California, Berkeley, Tech. Rep., 1998.

[4] D. Barber, “Expectation correction for smoothed inference in switching linear dynamical systems,” Journal of Machine Learning Research, vol. 7, pp. 2515-2540, Nov. 2006.

[5] O. Zoeter and T. Heskes, “Change point problems in linear dynamical systems,” Journal of Machine Learning Research, vol. 6, pp. 1999-2026, 2005.

[6] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, “Factor graphs and the sum-product algorithm,” IEEE Transactions on Information Theory, vol. 47, no. 2, pp. 498-519, 2001.

[7] H.-A. Loeliger, J. Dauwels, J. Hu, S. Korl, L. Ping, and F. R. Kschischang, “The factor graph approach to model-based signal processing,” Proceedings of the IEEE, vol. 95, no. 6, pp. 1295-1322, 2007.

[8] I. Arasaratnam, S. Haykin, and R. J. Elliott, “Discrete-time nonlinear filtering algorithms using Gauss-Hermite quadrature,” Proceedings of the IEEE, vol. 95, no. 5, pp. 953-977, 2007.

[9] O. Paiss and G. F. Inbar, “Autoregressive modeling of surface EMG and its spectrum with application to fatigue,” IEEE Transactions on Biomedical Engineering, vol. 34, no. 10, pp. 761-770, 1987.

[10] B. M. P. Wayne and, V. Novak, M. Costa, J. M. Hausdorff, A. L. Goldberger, A. C. Ahn, G. Y. Yeh, C.-K. Peng, M. Lough, R. B. Davis, M. T. Quilty, L. A. Lipsitz , “A systems biology approach to studying Tai Chi, physiological complexity and healthy aging: Design and rationale of a pragmatic randomized controlled trial,” Contemporary clinical trials, vol. 34, pp. 21-34, Sep. 2012.