SIGNAL PROCESSING WITH FACTOR GRAPHS: EXAMPLES

Hans-Andrea Loeliger, Justin Dauwels, Volker M. Koch, and Sascha Korl

Dept. of Information Technology and Electrical Engineering (ISI-DITET),
ETH, CH-8092 Zürich, Switzerland.

1. INTRODUCTION

Graphical models such as factor graphs allow to model complex systems and help to derive practical detection/estimation algorithms as message passing in the graph. In this paper, we outline three examples of on-going work of this type. For an introduction to factor graphs, we refer to [1] and [2]. We will use the notation of [2].

2. EMG SIGNAL DECOMPOSITION

All muscular activity in human bodies is accompanied by electrical signals inside the muscle fibers. Such signals can be measured by electrodes (either on the skin or inside the muscle). The art of measuring and analyzing such signals is called electromyography (EMG).

A discrete-time model of such signals may be described as follows (see Fig. 1). There are N_{src} independent "sources" (typically, $2 \leq N_{\text{src}} < 50$), each emitting some binary signal. Specifically, source i emits the signal $X_i \triangleq (X_{i,1}, X_{i,2}, X_{i,3}, \ldots)$ with $X_{i,k} \in \{0, 1\}$. These signals are sparse: the fraction of ones in each signal is usually well below 10^{-2} (assuming a sampling rate of 25 kHz).

Each of N_{chn} electrodes (with $1 \leq N_{\text{chn}} \leq 128$) picks up a noisy and heavily filtered superposition of all these sources. Specifically, electrode j picks up the signal $Y_j \triangleq (Y_{j,1}, Y_{j,2}, Y_{j,3}, \ldots)$ with

$$Y_{j,k} = \sum_{i=1}^{N_{\text{src}}} \sum_{\ell=0}^{M} X_{i,k-\ell} \cdot h_{i,j,\ell} + W_{j,k} \tag{1}$$

where $h_{i,j,\ell} \in \mathbb{R}$ are the filter coefficients, where $M \approx 100$ is the maximal memory of all the FIR filters and where $W_j = (W_{j,1}, W_{j,2}, W_{j,3}, \ldots)$ is white Gaussian noise. A (simulated) example of such a signal with $N_{\text{src}} = 8$ and $N_{\text{chn}} = 1$ is shown in Fig. 3.

A basic task in electromyography is to estimate the source signals X_i from the measured signals Y_j. We will assume here that the filter coefficients $h_{i,j,\ell}$ are known; in reality, they have to be estimated as well (and we will address this problem in future work).

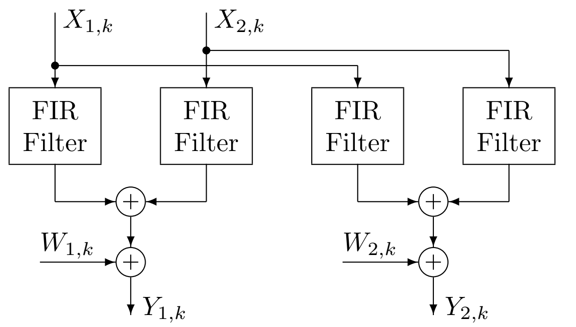

Fig. 1. Model of EMG signals for $N_{\text{src}} = N_{\text{chn}} = 2$.

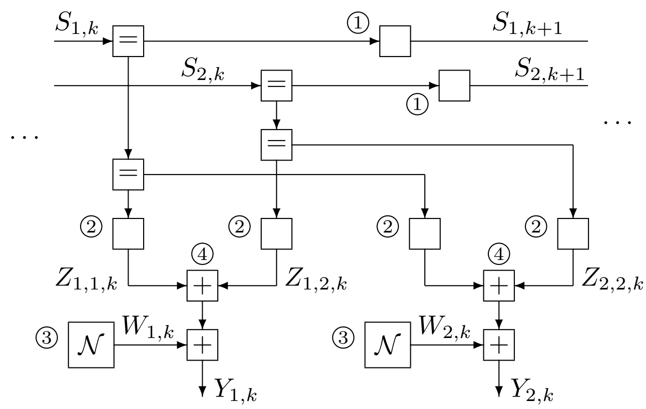

Fig. 2. The time-k section of a factor graph (in the style of [2]) corresponding to Fig. 1.

The maximum-likelihood (or MAP) estimate of $X_{i,k}$ appears to be computationally untractable. Linear estimators will not work well: the individual FIR filters are highly non-orthogonal, and the finite-alphabet constraint of the sources cannot be exploited by linear methods. In fact, the algorithms described in the survey article [3] fail wherever more than two (or, with extreme effort, more than three) source signals overlap.

A factor graph of the system model is shown in Fig. 2. The variables $S_{i,k}$ in this figure represent the vector $(X_{i,k}, X_{i,k-1}, \ldots, X_{i,k-M})$—the state of the FIR

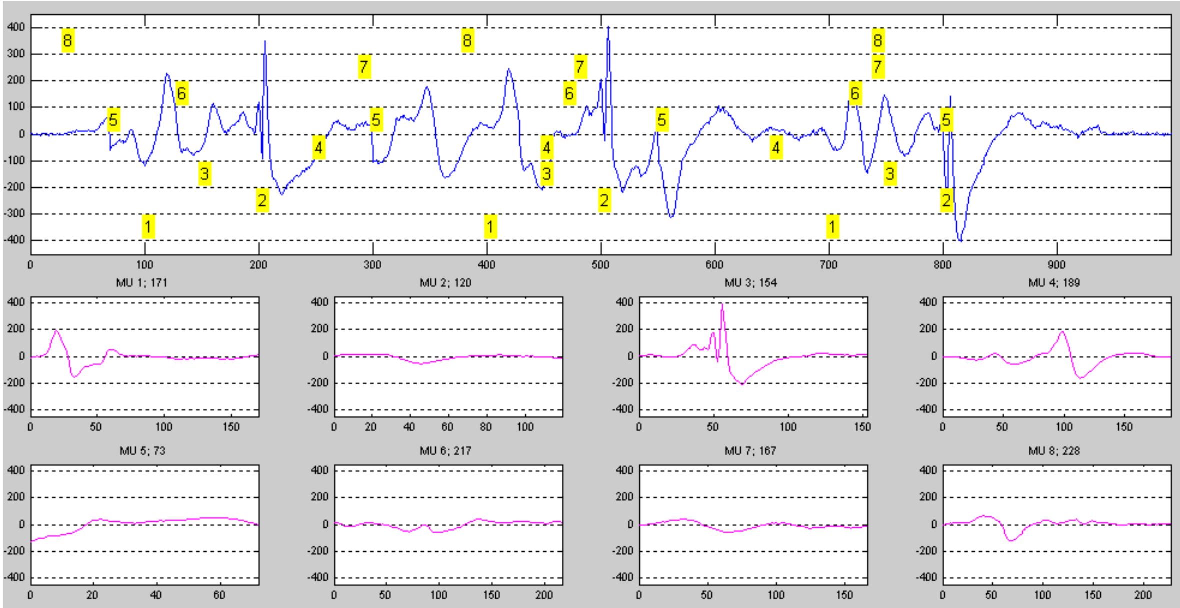

Fig. 3. Single-channel EMG signal annotated with the "firing times" of the $N_{\mathrm{src}} = 8$ sources (top) and the impulse responses of the sources (bottom).

filters fed by X_i—which we assume to contain at most one "1". (EMG experts[1] tell us that this assumption is innocent.) Due to this assumption, we can define the range of $S_{i,k}$ as the set $\{0, 1, 2, \ldots, M{+}1\}$; $S_{i,k} = n$ means $X_{i,k-n} = 1$ if $0 \le n \le M$ and $X_{i,k} = \ldots = X_{i,k-M} = 0$ if $n = M + 1$.

In Fig. 2, the boxes labeled ① represent the function $p(s_{i,k+1}|s_{i,k})$ defined in Table 1, where the parameter $\varepsilon \approx 10^{-2}$ models the sparseness of the sources. The variables

$$Z_{i,j,k} \triangleq \sum_{\ell=0}^{M} X_{i,k-\ell} \cdot h_{i,j,\ell} \qquad (2)$$

(the output of the FIR filters) are deterministic functions of $S_{i,j,k}$, which is represented by the boxes labeled ②. The nodes labeled ③ represent Gaussian distributions.

By iterative sum-product message passing (cf. [1], [2]), we obtain a practical algorithm to estimate $X_{i,k}$ (simultaneously for all i and k) with a computational complexity that is roughly linear in the number of sources N_{src}, in the number of electrodes N_{chn}, and in the maximal filter memory M. Message passing through the sum-constraint node ④ (involving $N_{\mathrm{src}} + 1$ variables) is handled as outlined in [4, Section 3.5].

[1] Specifically, Thomas Läubli and Daniel Zennaro from IHA, ETH Zurich, with whom we have been collaborating.

$s_{i,k}$	$s_{i,k+1}$	$p(s_{i,k+1}\|s_{i,k})$
$n \in \{0, 1, \ldots, M\}$	$n+1$	1
$M+1$	$M+1$	$1-\varepsilon$
$M+1$	0	ε
everything else		0

Table 1. State transition probabilities $p(s_{i,k+1}|s_{i,k})$ (nodes ① in Fig. 2).

Although straightforward in principle, the development of the specific algorithm involves many design choices: details of the system model and of the factor graph, quantization issues, the scheduling of the message updates, etc. Much experimentation was necessary to make the algorithm work well (and fast), and we expect further improvements as our experience grows. The message passing algorithm now allows to decompose heavily superimposed EMG signals (such as the example in Fig. 3) that appear to be far beyond the reach of other published methods.

3. PHASE ESTIMATION IN A COMMUNICATION RECEIVER

Consider a communication channel of the form

$$Y_k = X_k \, e^{i\Theta_k} + N_k \qquad (3)$$

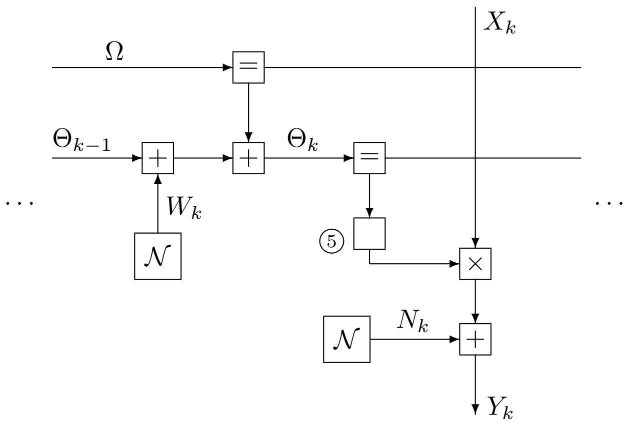

Fig. 4. Phase estimation: time-k section of factor graph of (3) and (5).

where X_k is the complex channel input symbol at time k, Y_k is the corresponding received symbol, Θ_k is the unknown phase, and N_k is white Gaussian noise with known variance σ_N^2. In [5] and [6], three different models for the phase Θ_k are considered:

Constant Phase: $\Theta_k = \Theta$, an unknown constant.

Random Walk:

$$\Theta_k = \Theta_{k-1} + W_k, \tag{4}$$

where W_k is white Gaussian noise with known variance σ_W^2.

Random Walk with Unknown Drift:

$$\Theta_k = \Theta_{k-1} + W_k + \Omega \tag{5}$$

with an unknown drift parameter Ω and with W_k as above.

A factor graph of the last case is shown in Fig. 4. The node labeled ⑤ represents the deterministic function $\Theta_k \mapsto e^{i\Theta_k}$. As described in [6], we obtain various message passing algorithms suitable for joint iterative channel estimation and decoding. (For the random walk phase models, no such estimator seems to have been proposed before.) The proposed message passing algorithms consist of various combinations of quantized messages, gradient methods, Kalman filters (cf. [7]), generalizations of Kalman filters that work with Gaussian mixtures, and particle-filter methods.

4. AR MODEL PARAMETER ESTIMATION

Let $X_1, X_2, \ldots$ be real-valued random variables defined by

$$X_k = \sum_{\ell=1}^{M} X_{k-\ell} \cdot a_\ell + U_k \tag{6}$$

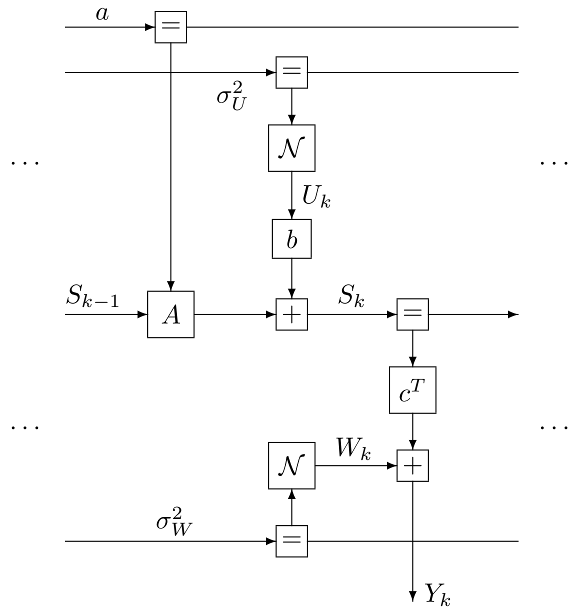

Fig. 5. AR parameter estimation: time-k section of the factor graph.

with $a_\ell \in \mathbb{R}$ and where $U_1, U_2, \ldots$ is white Gaussian noise with variance σ_U^2. We observe the process $Y_1, Y_2, \ldots$ with

$$Y_k = X_k + W_k, \tag{7}$$

where $W_1, W_2, \ldots$ is white Gaussian noise with variance σ_W^2, and we wish to estimate the unknown coefficients

$$a \stackrel{\triangle}{=} (a_1, \ldots, a_M)^T. \tag{8}$$

In a first version of the problem, the parameters σ_U^2 and σ_W^2 are known; in a second version, they are unknown and need to be estimated as well.

It is convenient to write (6) and (7) in state-space form as

$$S_k = AS_{k-1} + bU_k \tag{9}$$

$$Y_k = c^T S_k + W_k \tag{10}$$

with

$$S_k \stackrel{\triangle}{=} (X_k, \ldots, X_{k-M+1})^T \tag{11}$$

$$A \stackrel{\triangle}{=} \begin{pmatrix} & a^T & \\ I & & 0 \end{pmatrix} \tag{12}$$

$$b \stackrel{\triangle}{=} c \stackrel{\triangle}{=} (1, 0, \ldots, 0)^T. \tag{13}$$

The factor graph corresponding to (9)–(10) is shown in Fig. 5. Using the recipes from [2], we obtain message passing algorithms for the simultaneous estimation of S_k, a, σ_U^2, and σ_W^2. The estimation of S_k amounts to Kalman filtering (and smoothing), which uses a hard-decision estimate $\hat{a}$ of the coefficient vector a. The estimation of a itself may be carried out either by another Kalman filter of by an LMS-type gradient method. Note that, if σ_U^2 is known and $\sigma_W^2 = 0$ (i.e., without observation noise), the problem reduces to the classical LPC parameter estimation problem; in this case, the message passing algorithms reduce to standard gradient methods or RLS algorithms. The estimation of the variances σ_U^2 and σ_W^2 is accomplished by particle filters. The message passing algorithms may thus be roughly described as two coupled Kalman filters (or a Kalman filter coupled with an LMS-type algorithm) coupled with two particle filters. A more detailed description of the algorithms and some simulation results are given in [8] (and a full report is in preparation). The extension of the message passing algorithms to time-varying model parameters should be straightforward.

5. DISCUSSION

We have outlined three examples of ongoing work in signal processing with factor graphs. Using the general recipes described in [2] and [4], we have obtained practical algorithms for complex detection/estimation problems; these algorithms either outperform previously published algorithms or are actually the first working estimators for the respective problem. It should be noted, however, that the design of such algorithms involves a large number of design choices; much experimentation is usually necessary to obtain the best performance or to make the algorithm work at all.

6. REFERENCES

[1] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Information Theory,* vol. 47, pp. 498–519, Feb. 2001.

[2] H.-A. Loeliger, "An introduction to factor graphs," *IEEE Signal Proc. Mag.,* to appear in Jan. 2004.

[3] D. Stashuk, "EMG signal decomposition: how can it be accomplished and used?" *Journal of Electromyography & Kinesiology,* vol. 11, pp. 151–173, 2001.

[4] H.-A. Loeliger, "Some remarks on factor graphs," *Proc. 3rd Int. Symp. on Turbo Codes and Related Topics,* Sept. 1–5, 2003, Brest, France, pp. 111–115.

[5] J. Dauwels and H.-A. Loeliger, "Joint decoding and phase estimation: an exercise in factor graphs," *Proc. 2003 IEEE Int. Symp. Information Theory,* Yokohama, Japan, June 29 – July 4, 2003, p. 231.

[6] J. Dauwels and H.-A. Loeliger, "Phase estimation by message passing," *Proc. 2004 IEEE Int. Conf. on Communications,* Paris, France, June 20–24, 2004, to appear.

[7] H.-A. Loeliger, "Least squares and Kalman filtering on Forney graphs," in *Codes, Graphs, and Systems,* R. E. Blahut and R. Koetter, eds., Kluwer, 2002, pp. 113–135.

[8] S. Korl, H.-A. Loeliger, and A. G. Lindgren, "AR model parameter estimation: from factor graphs to algorithms," *Proc. Int. Conf. on Acoustics, Speech, and Signal Processing,* May 17–21, 2004, Montreal, Canada, to appear.