Applications of graphical models in signal processing

Brain-Style Information Systems Research Group, RIKEN BSI, Saitama, Japan

Justin Dauwels

dauwels@isi.ee.ethz.ch

Signal and Information Processing Laboratory

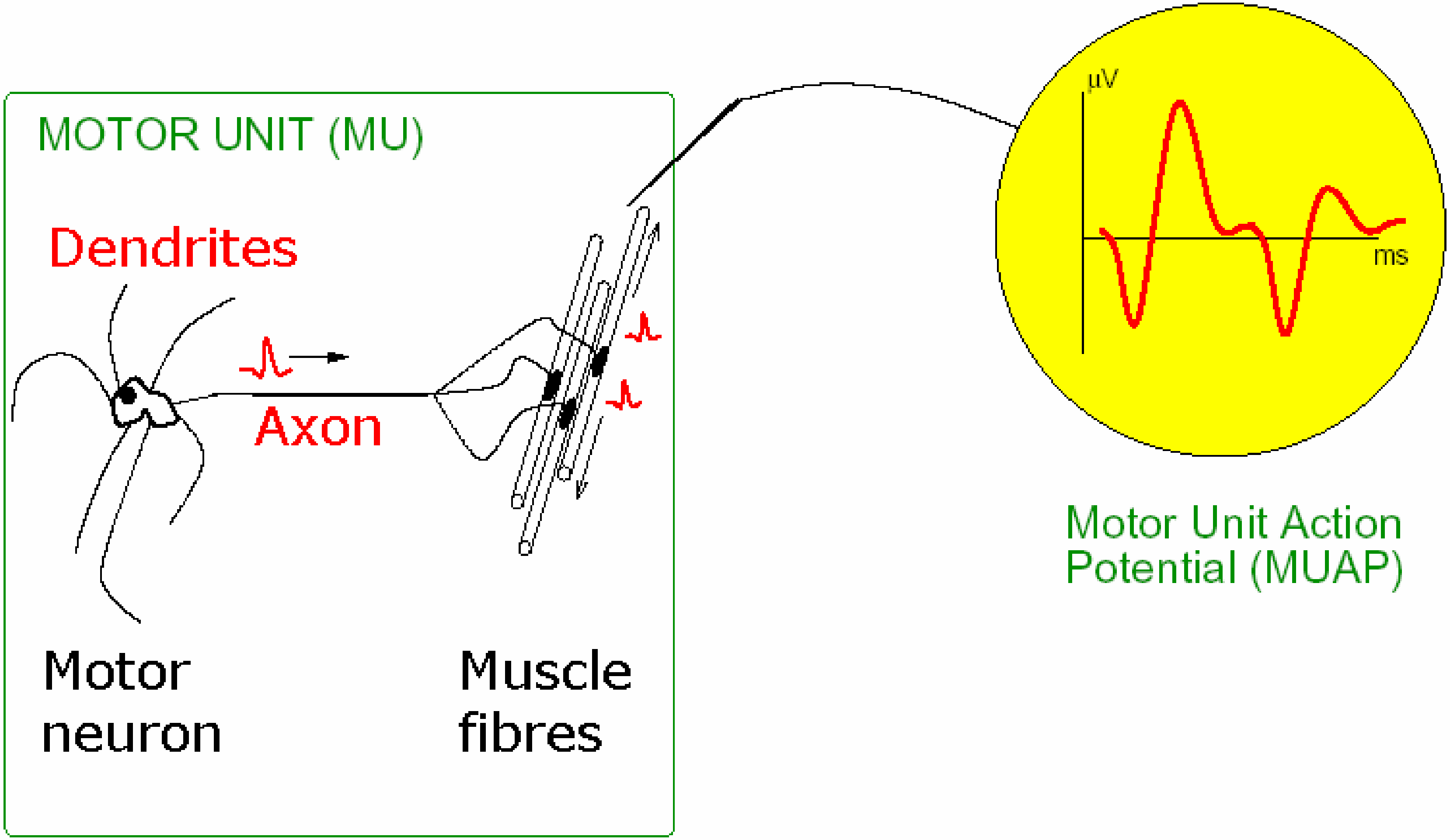
Department of Information Technology and Electrical Engineering

ETH Zurich, Switzerland

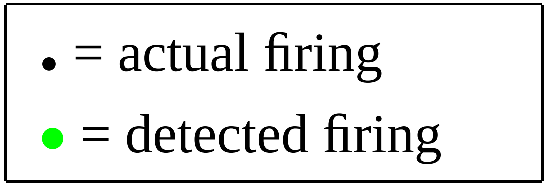
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Signal and Information Processing Laboratory
Institut für Signal- und Informationsverarbeitung

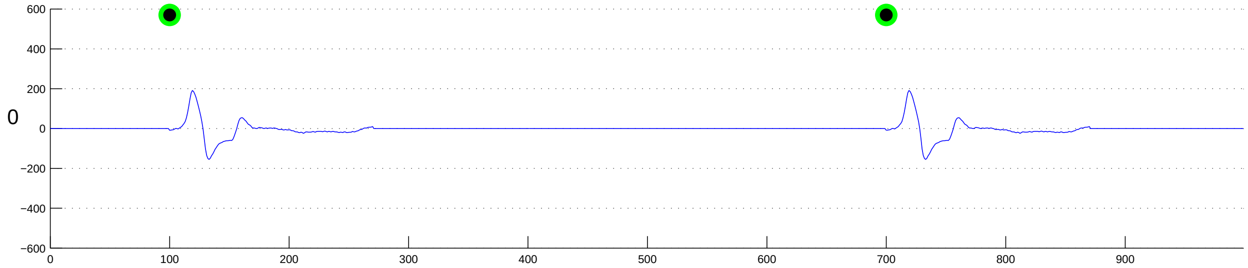
Electromyography



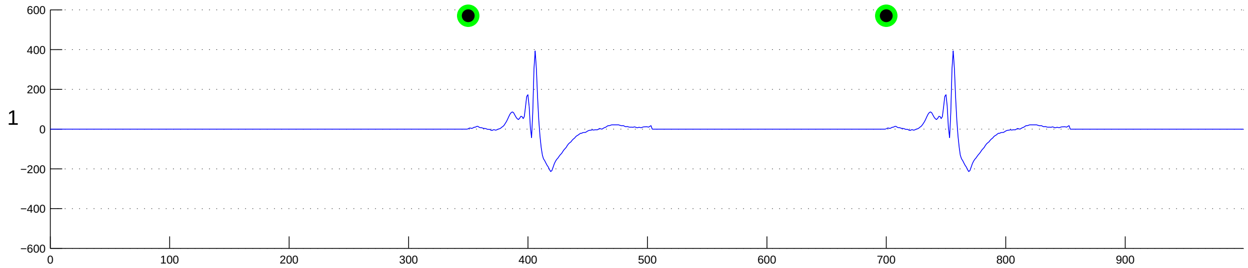
EMG Signal Decomposition



Signal from MU 0:

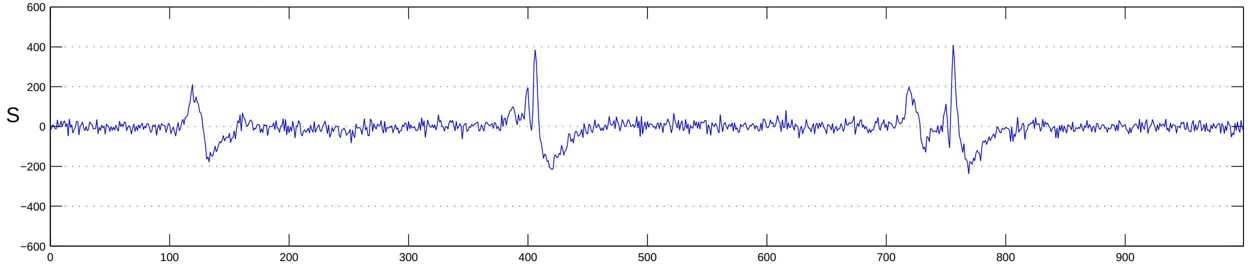


Signal from MU 1:



EMG signal:

$(\sigma_{\text{noise}} = 20)$



Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Signal and Information Processing Laboratory
Institut für Signal- und Informationsverarbeitung

Estimation Theory

Optimal estimators

- Block MAP (Maximum a posteriori)

$$\hat{\mathbf{x}} = \text{argmax}_{\mathbf{x}}\, p(\mathbf{x}|\mathbf{y}) = \text{argmax}_{\mathbf{x}}\, p(\mathbf{x}, \mathbf{y})$$

- Symbol MAP

$$\hat{x}_k = \text{argmax}_{x_k}\, p(x_k|\mathbf{y}) = \text{argmax}_{x_k}\, p(x_k, \mathbf{y}) = \text{argmax}_{x_k} \sum_{\sim x_k} p(\mathbf{x}, \mathbf{y})$$

Key observations

- Summation and maximization are the involved operations.
- The probability function $p(\mathbf{x}, \mathbf{y})$ factorizes in most practical systems . . .
- ... which leads to more efficient computations.
- Despite this fact, optimal estimators are often infeasible . . .
- ... and hence one typically resorts to approximations.

Our methodology

Our approach rests on two pillars

1. A graphical representation of $p(\mathbf{x}, \mathbf{y})$, a so-called factor graph.
2. An algorithm (sum-product algorithm) that computes marginals $p(x_k)$ and operates on a factor graph that represents the system at hand.

Within this framework

- the factorization of $p(\mathbf{x}, \mathbf{y})$ is intensively exploited and, as a consequence, marginals $p(x_k)$ are computed efficiently.
- approximations can be designed in a systematic fashion.

Motivation

- Many algorithms in coding, signal processing, and machine learning may be viewed as instances of the sum-product algorithm.

Motivation

- Many algorithms in coding, signal processing, and machine learning may be viewed as instances of the sum-product algorithm.
- Specific instances of such algorithms include Kalman filtering and smoothing, the forward-backward algorithm for hidden Markov models, Viterbi algorithm, probability propagation in Bayesian networks, decoding algorithms for error correcting codes (in particular iterative decoding of turbo codes and low-density parity check codes), . . .

Motivation

- Many algorithms in coding, signal processing, and machine learning may be viewed as instances of the sum-product algorithm.
- Specific instances of such algorithms include Kalman filtering and smoothing, the forward-backward algorithm for hidden Markov models, Viterbi algorithm, probability propagation in Bayesian networks, decoding algorithms for error correcting codes (in particular iterative decoding of turbo codes and low-density parity check codes), . . .
- Applications at ISI: Decomposition of EMG signals, signal processing for hearing aids, synchronization in digital communications, decoding, estimation/detection with analog circuits, computation of capacities of communication channels, smart fire detection, body-state estimation, seismosomnography, . . .

Motivation

- Many algorithms in coding, signal processing, and machine learning may be viewed as instances of the sum-product algorithm.
- Specific instances of such algorithms include Kalman filtering and smoothing, the forward-backward algorithm for hidden Markov models, Viterbi algorithm, probability propagation in Bayesian networks, decoding algorithms for error correcting codes (in particular iterative decoding of turbo codes and low-density parity check codes), . . .
- Applications at ISI: Decomposition of EMG signals, signal processing for hearing aids, synchronization in digital communications, decoding, estimation/detection with analog circuits, computation of capacities of communication channels, smart fire detection, body-state estimation, seismosomnography, . . .
- Factor graphs and message-passing algorithm have been proven to be very powerful in decoding and signal processing.

Overview

- Factor graphs
- Sum-product algorithm
- Application: decomposition of EMG-signals
- Conclusions

Factor Graphs (FGs)

- Factor graphs represent the factorization of a function.
- Example

$$f(x_1, x_2, x_3, x_4, x_5) = f_A(x_1, x_2, x_3) f_B(x_3, x_4, x_5) f_C(x_4).$$

Factor Graphs (FGs)

- Factor graphs represent the factorization of a function.
- Example

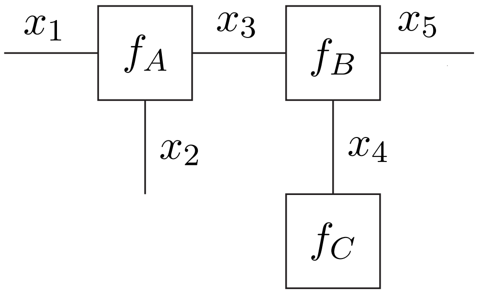
$$f(x_1, x_2, x_3, x_4, x_5) = f_A(x_1, x_2, x_3) f_B(x_3, x_4, x_5) f_C(x_4).$$

- x_1, x_2, x_3, x_4, and x_5 are the variables.
- f is called the global function.
- f_A, f_B, and f_C are called local functions.

Factor Graphs (FGs)

- Factor graphs represent the factorization of a function.
- Example

$$f(x_1, x_2, x_3, x_4, x_5) = f_A(x_1, x_2, x_3) f_B(x_3, x_4, x_5) f_C(x_4).$$

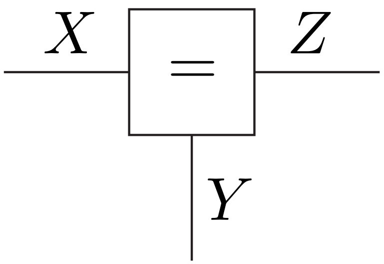


- Rules for drawing a factor graph
  - A node for every factor
  - An edge for every variable
  - Node g is connected to edge x iff variable x appears in factor g

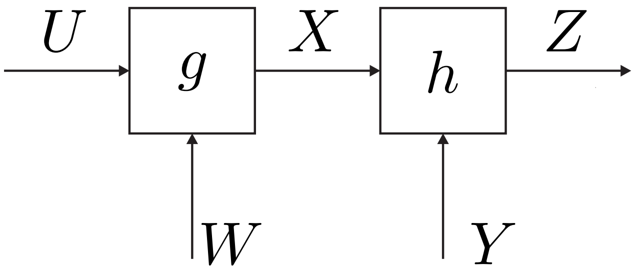
Equality constraint node

$$f_{=}(x, y, z) \triangleq \begin{cases} 1, & \text{if } x = y = z \\ 0, & \text{else} \end{cases}$$

$$= \delta[x - y] \cdot \delta[x - z]$$

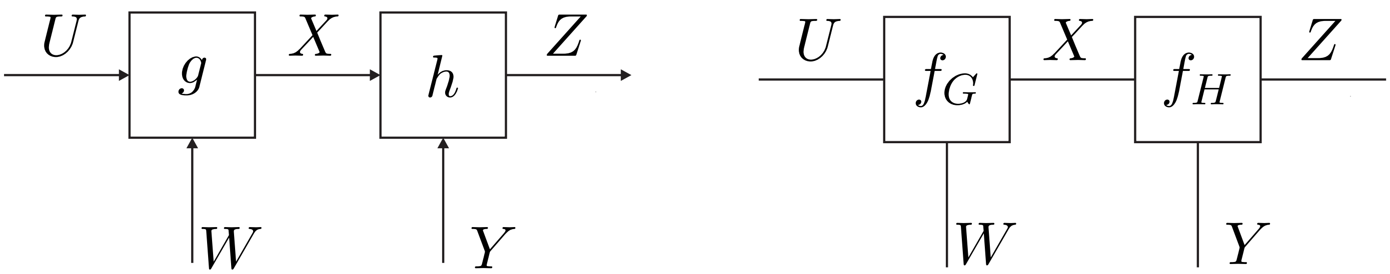


Block Diagrams and Factor Graphs



- Block diagram: $X = g(U, W)$, $Z = h(X, Y)$

Block Diagrams and Factor Graphs



- Block diagram: $X = g(U, W)$, $Z = h(X, Y)$

- Factor graph: $f_G = \delta(x - g(u, w))$, $f_H = \delta(z - h(x, y))$
- Global function $f(u, w, x, y, z) = \delta(x - g(u, w)) \cdot \delta(z - h(x, y))$

- $f \neq 0$ (configuration valid) iff the configuration is consistent with the functions g and h of the block diagram.

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

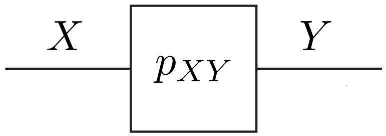
Signal and Information Processing Laboratory
Institut für Signal- und Informationsverarbeitung

Joint Probability Distribution

- Factor graphs are used mostly to represent probabilistic models.
- Given is a joint probability distribution

$$p_{XY}(x, y)$$

of the two discrete random variables X and Y.

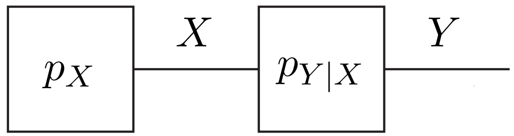


Factorization of Joint Probability Distribution

- Chain rule

$$p_{XY}(x,y) = p_X(x)p_{Y|X}(y|x)$$

- Factor graph

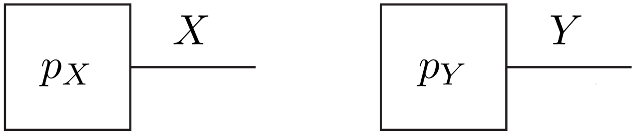


Independent Random Variables

- If the two random variables X and Y are independent, then $p_{Y|X}(y|x) = p_Y(y)$ and we get the factorization

$$p_{XY}(x, y) = p_X(x)p_Y(y).$$

- Factor graph



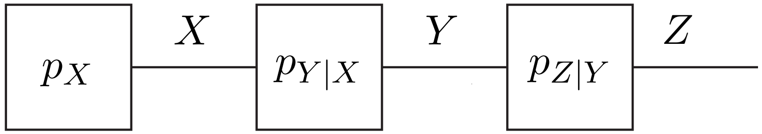
- If two components are unconnected, then every random variable of one component is independent of every random variable in the other component.

Markov Chain

- Markov chain of the form

$$p_{XYZ}(x, y, z) = p_X(x) \cdot p_{Y|X}(y|x) \cdot p_{Z|Y}(z|y)$$

- Factor graph



Summary: Factor Graphs

- Factor graphs
- Equal node
- Block diagrams and factor graphs
- Probabilistic models represented by factor graph

Overview

- Factor graphs
- Sum-product algorithm
- Application: decomposition of EMG-signals
- Conclusions

Introduction (4)

Our approach rests on two pillars

1. A graphical representation of $p(\mathbf{x}, \mathbf{y})$, a so-called factor graph.
2. An algorithm (sum-product algorithm) that computes marginals $p(x_k)$ and operates on a factor graph that represents the system at hand.

Within this framework

- the factorization of $p(\mathbf{x}, \mathbf{y})$ is intensively exploited and, as a consequence, marginals $p(x_k)$ are computed efficiently.
- approximations can be designed in a systematic fashion.

Computing marginals

- Given: Discrete probability mass function

$$f(x_1, \ldots, x_8) = \Big(f_1(x_1) f_2(x_2) f_3(x_1, x_2, x_3, x_4)\Big) \cdot \Big(f_4(x_4, x_5, x_6) f_5(x_5) \big(f_6(x_6, x_7, x_8) f_7(x_7)\big)\Big)$$

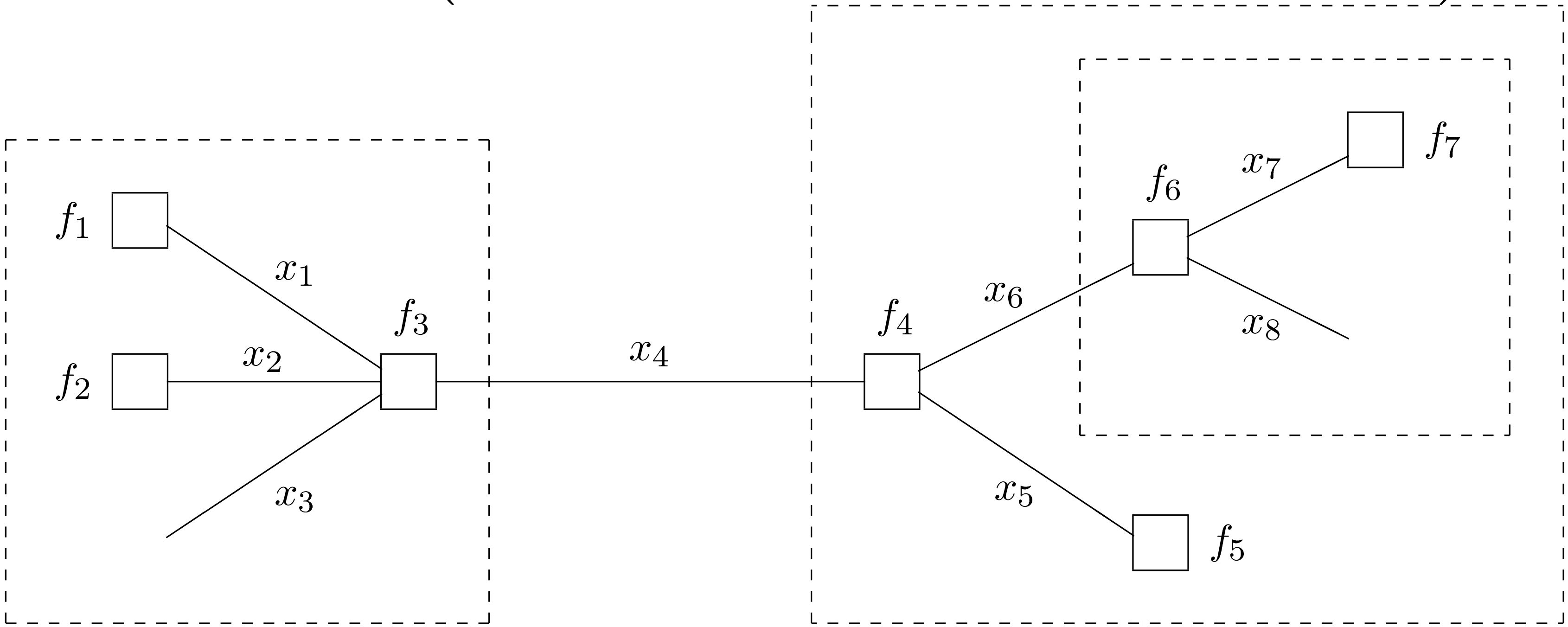
- Wanted: Marginal probability

$$p(x_4) = \sum_{x_1, x_2, x_3, x_5, x_6, x_7, x_8} f(x_1, \ldots, x_8)$$

- This factorization can be represented by a factor graph.

Elimination of Variables

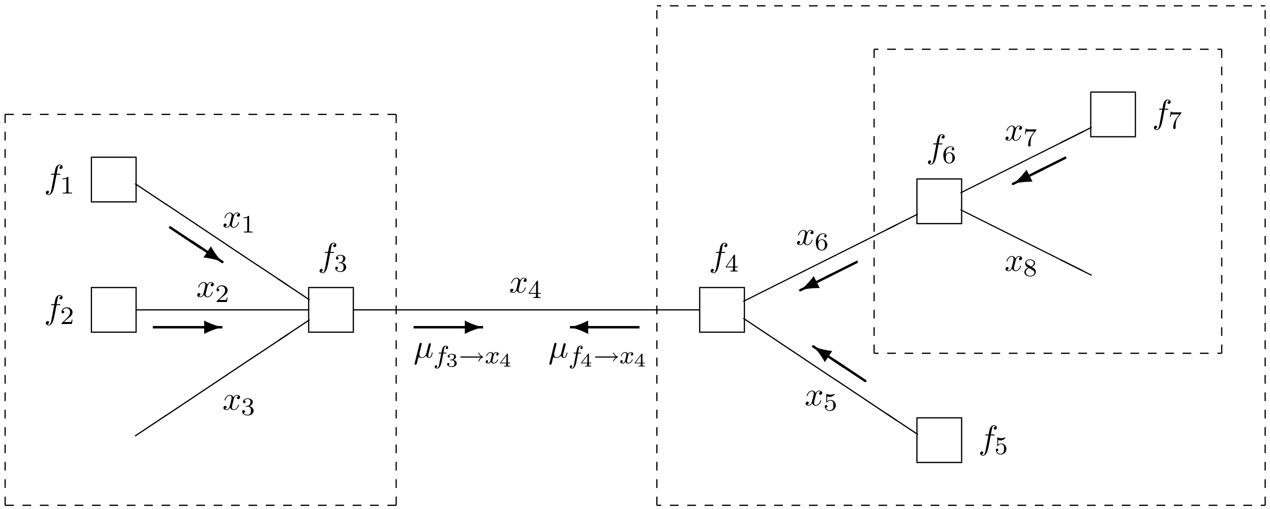
$$f(x_1, ..., x_8) \quad = \quad \big(f_1(x_1) f_2(x_2) f_3(x_1, x_2, x_3, x_4)\big) \cdot$$
$$\Big(f_4(x_4, x_5, x_6) f_5(x_5) \big(f_6(x_6, x_7, x_8) f_7(x_7)\big)\Big)$$



Elimination of Variables

$$p(x_4) \quad = \quad \underbrace{\left(\sum_{x_1} \sum_{x_2} \sum_{x_3} f_3(x_1, x_2, x_3, x_4) f_1(x_1) f_2(x_2) \right)}_{\mu_{f_3 \to x_4}} \cdot$$

$$\underbrace{\left(\sum_{x_5} \sum_{x_6} f_4(x_4, x_5, x_6) f_5(x_5) \underbrace{\left(\sum_{x_7} \sum_{x_8} f_6(x_6, x_7, x_8) f_7(x_7) \right)}_{\mu_{f_6 \to x_6}} \right)}_{\mu_{f_4 \to x_4}}$$

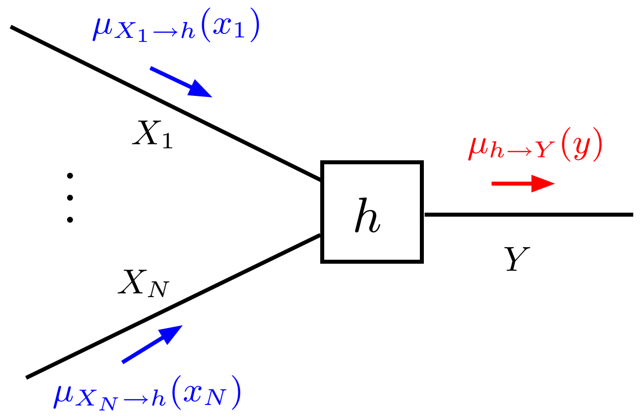


ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Signal and Information Processing Laboratory
Institut für Signal- und Informationsverarbeitung

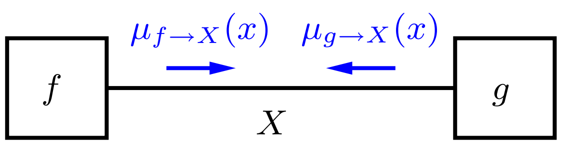
Sum-product Algorithm

Sum-product update rule



$$\mu_{h\to Y}(y) = \sum_{x_1,\ldots,x_N} h(x_1,\ldots,x_N,y)\mu_{X_1\to h}(x_1)\ldots\mu_{X_N\to h}(x_N).$$

Computing marginals



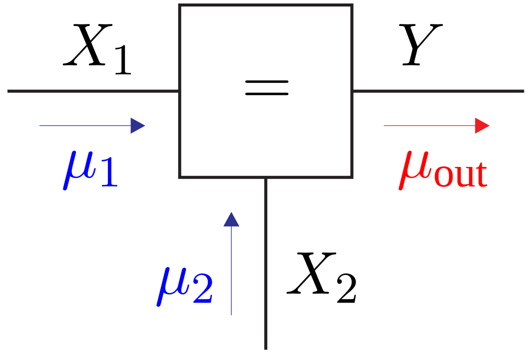
$$p(X) = \gamma\mu_{f\to X}(x)\mu_{g\to X}(x)$$

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Signal and Information Processing Laboratory
Institut für Signal- und Informationsverarbeitung

Sum-product Algorithm: Equal Node

$$f_=(x_1, x_2, y) \triangleq \delta[x_1 - x_2] \cdot \delta[x_1 - y]$$



For binary variables X_1, X_2, and Y we get:

$$\begin{aligned}
\mu_{\text{out}}(y) &= \sum_{x_1} \sum_{x_2} \delta[x_1 - x_2]\delta[x_1 - y]\mu_1(x_1)\mu_2(x_2) \\
&= \mu_1(y)\mu_2(y) \\
\mu_{\text{out}}(0) &= \mu_1(0)\mu_2(0) \\
\mu_{\text{out}}(1) &= \mu_1(1)\mu_2(1)
\end{aligned}$$

Sum-product Algorithm

- Factor graph without loops
  - The sum-product algorithm computes all marginals simultaneously and exactly.
- Factor graph with loops
  - The sum-product rule can be applied repeatedly: iterative algorithm.
  - The iterative algorithm might not converge.
    - Then another factor graph (factorization) can be tried.
    - There are several factor graph design choices.
  - The algorithm computes only an approximation of the marginals $p(x_k)$.
  - This approximation is often (but not always!) good enough.
  - In some very demanding applications the performance is nearly optimal.
  - The algorithm is suboptimal - but it can handle very large systems.

Intermediate Summary

- A factor graph is a graphical representation of a mathematical model.

Intermediate Summary

- A factor graph is a graphical representation of a mathematical model.
- The sum-product algorithm
  - runs on a factor graph and calculates messages that are sent along the edges of the graph and
  - is suboptimal on graphs with cycles.

Intermediate Summary

- A factor graph is a graphical representation of a mathematical model.
- The sum-product algorithm
  - runs on a factor graph and calculates messages that are sent along the edges of the graph and
  - is suboptimal on graphs with cycles.
- However, iterative algorithms on graphs with cycles have been proven to be very powerful.

Intermediate Summary

- A factor graph is a graphical representation of a mathematical model.
- The sum-product algorithm
  - runs on a factor graph and calculates messages that are sent along the edges of the graph and
  - is suboptimal on graphs with cycles.
- However, iterative algorithms on graphs with cycles have been proven to be very powerful.
- Applications of iterative algorithms on graphs are, e.g.,
  - Decoding of error correcting codes,
  - Signal processing.

Overview

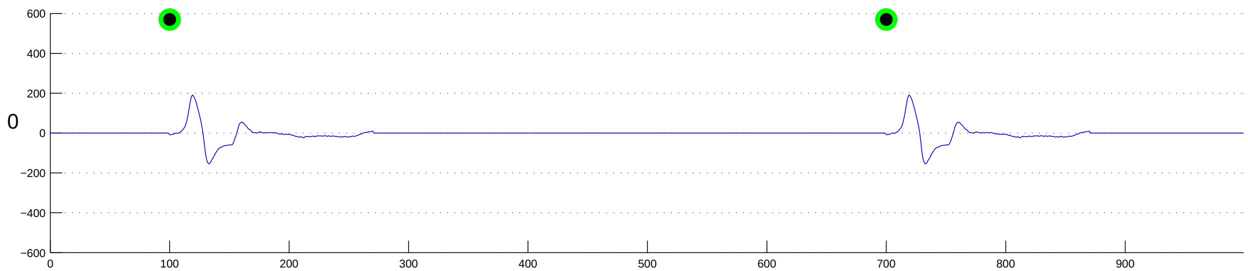
- Factor graphs
- Sum-product algorithm
- Application: decomposition of EMG-signals
- Conclusions

EMG Signal Decomposition

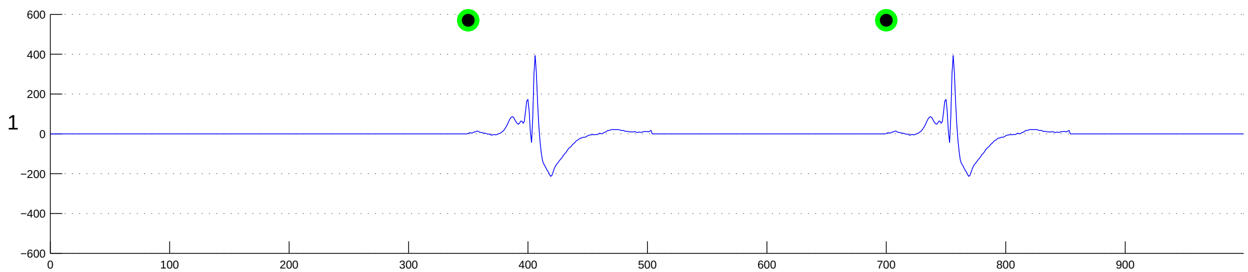
• = actual firing

● = detected firing

Signal from MU 0:

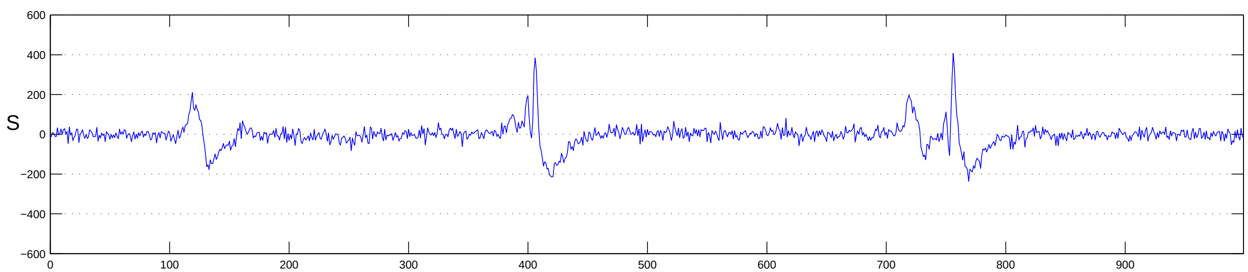


Signal from MU 1:



EMG signal:

$(\sigma_{\text{noise}} = 20)$

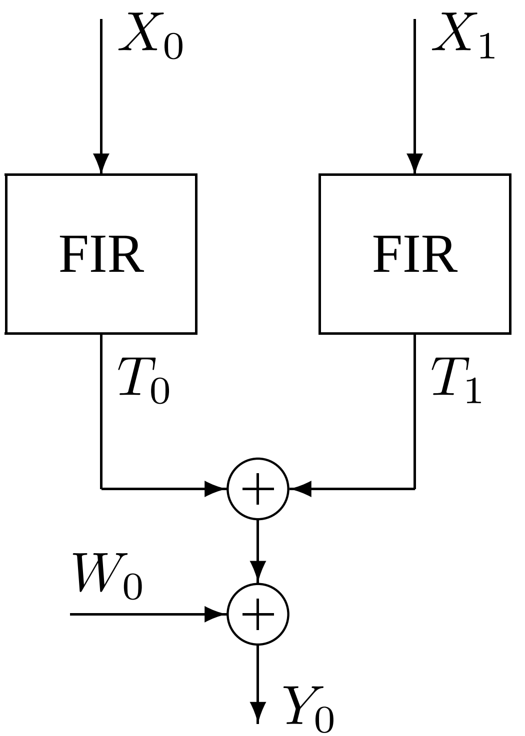


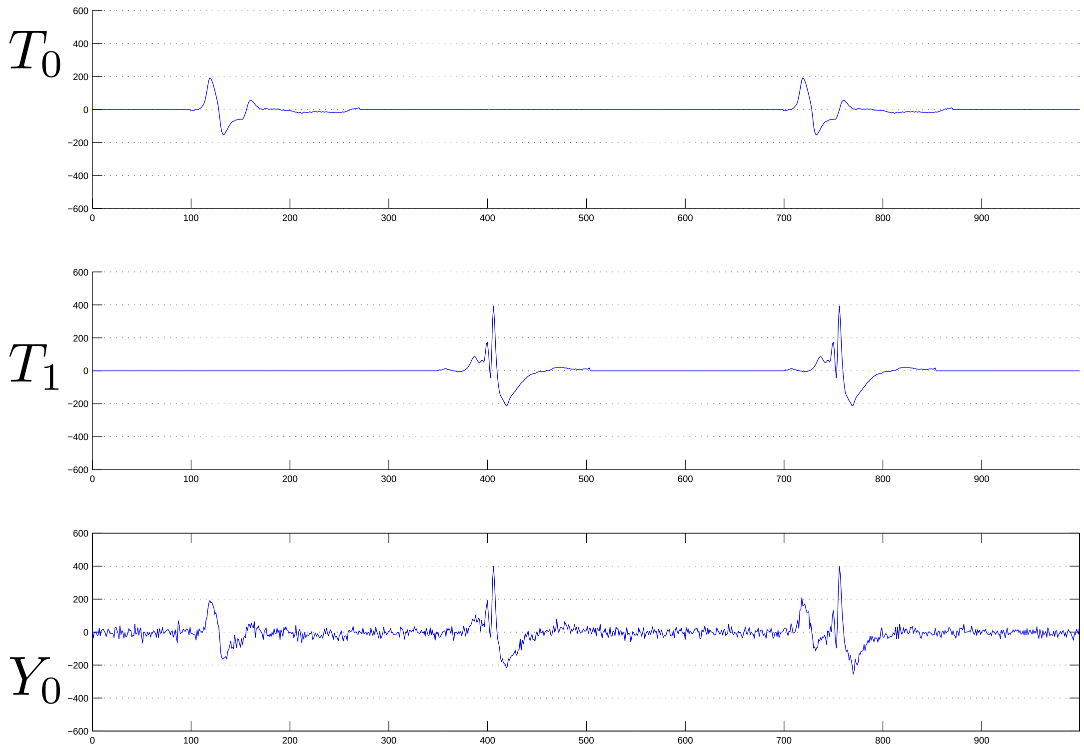
Lab History and Objective

- Our lab: 8 PhD students in the field of EMG since 1975
- *1992: Richard Gut (Prof. Moschytz)*
- *2000: Peter Wellig (Prof. Moschytz)*
- *200?: Volker Koch (Prof. Loeliger)*
  - Decomposition algorithm based on a
    - message-passing algorithm (belief propagation) in a
    - graphical model (factor graph)
  - *Preliminary result*: decomposition of more than 8 overlapping MUAPs
  - Information on the web
    - Volker's Webpage: `http://www.volker-koch.de/`
    - EMG forum: `http://www.emg.ethz.ch`

Block Diagram → Factor Graph

Block Diagram



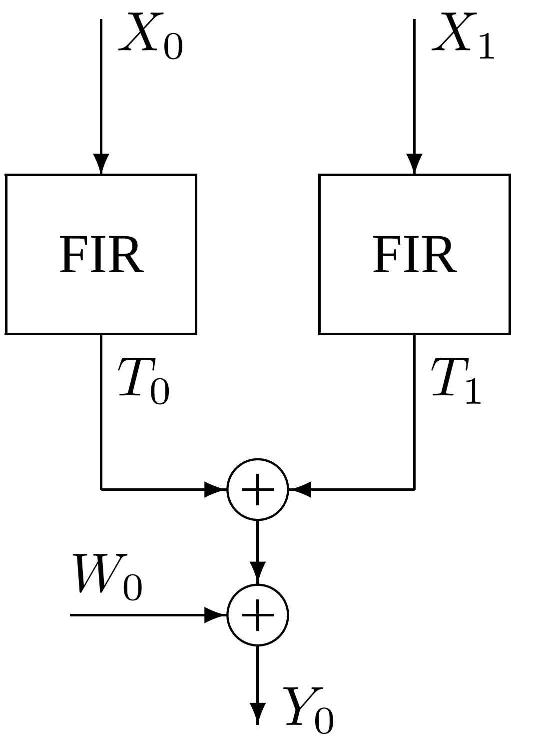


$$Y_{0,k} = \sum_{i=0}^{1} \sum_{\ell=0}^{M} X_{i,k-\ell} \cdot h_{i,0,\ell} + W_{0,k}$$

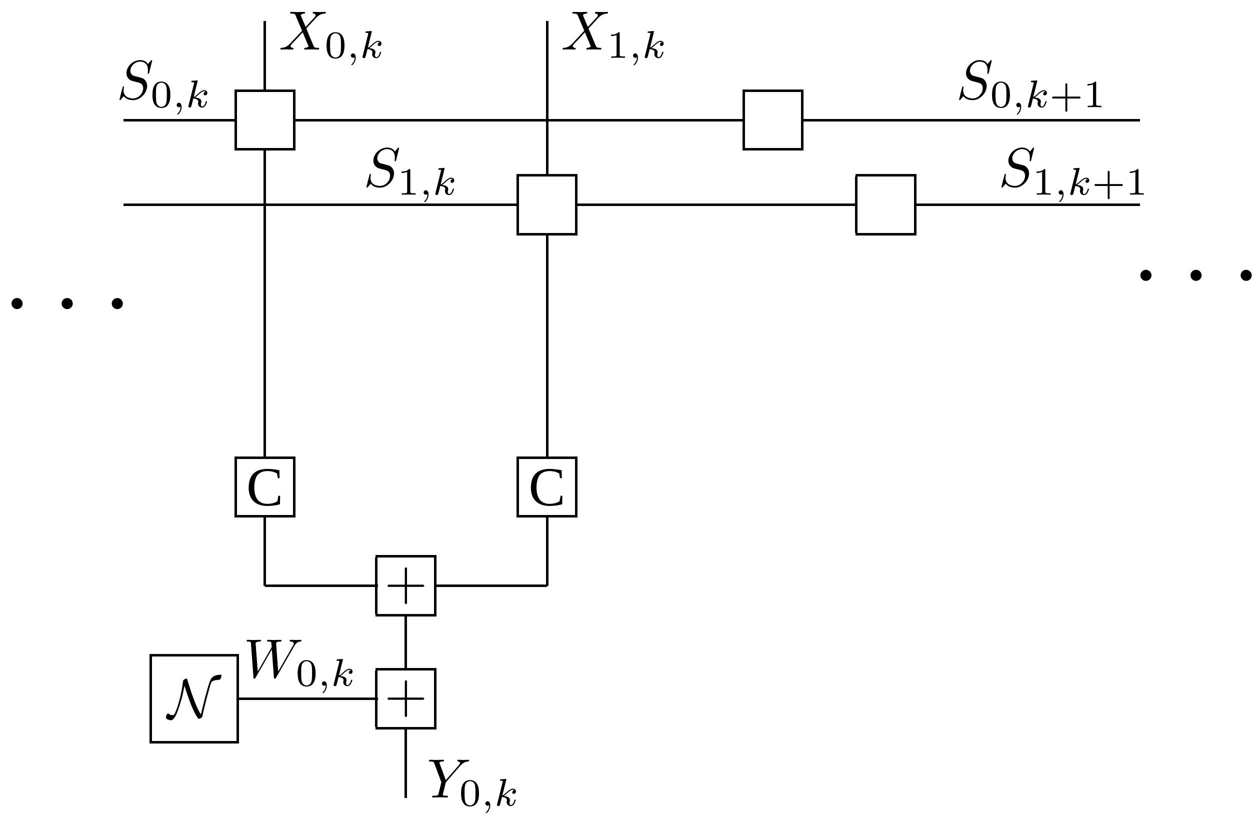
$$Y_0 \longrightarrow X_0, X_1 \quad ?$$

Block Diagram → Factor Graph

Block Diagram



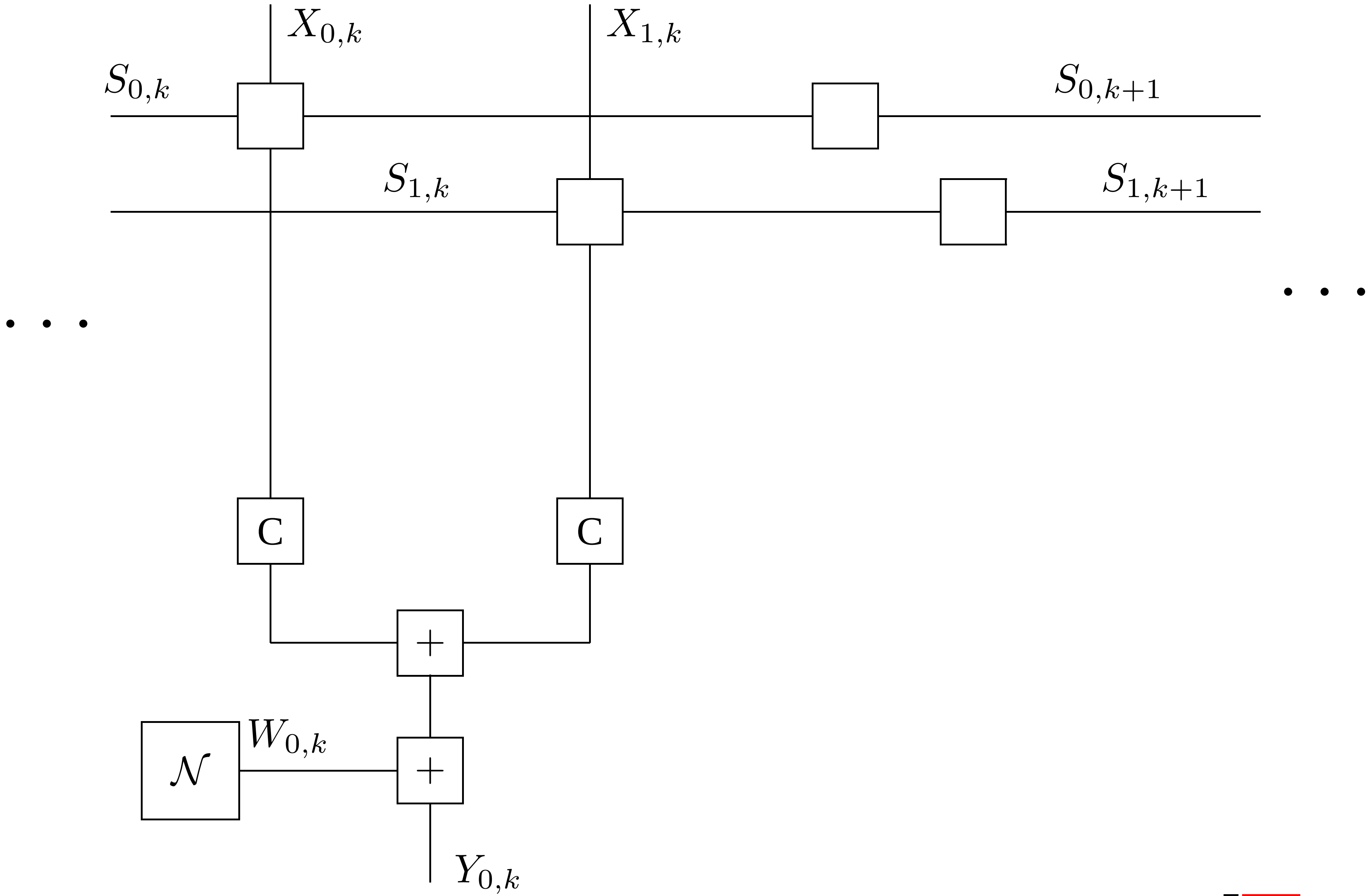
Factor Graph



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

$$Y_0 \rightarrow X_0, X_1 \quad \textbf{?}$$

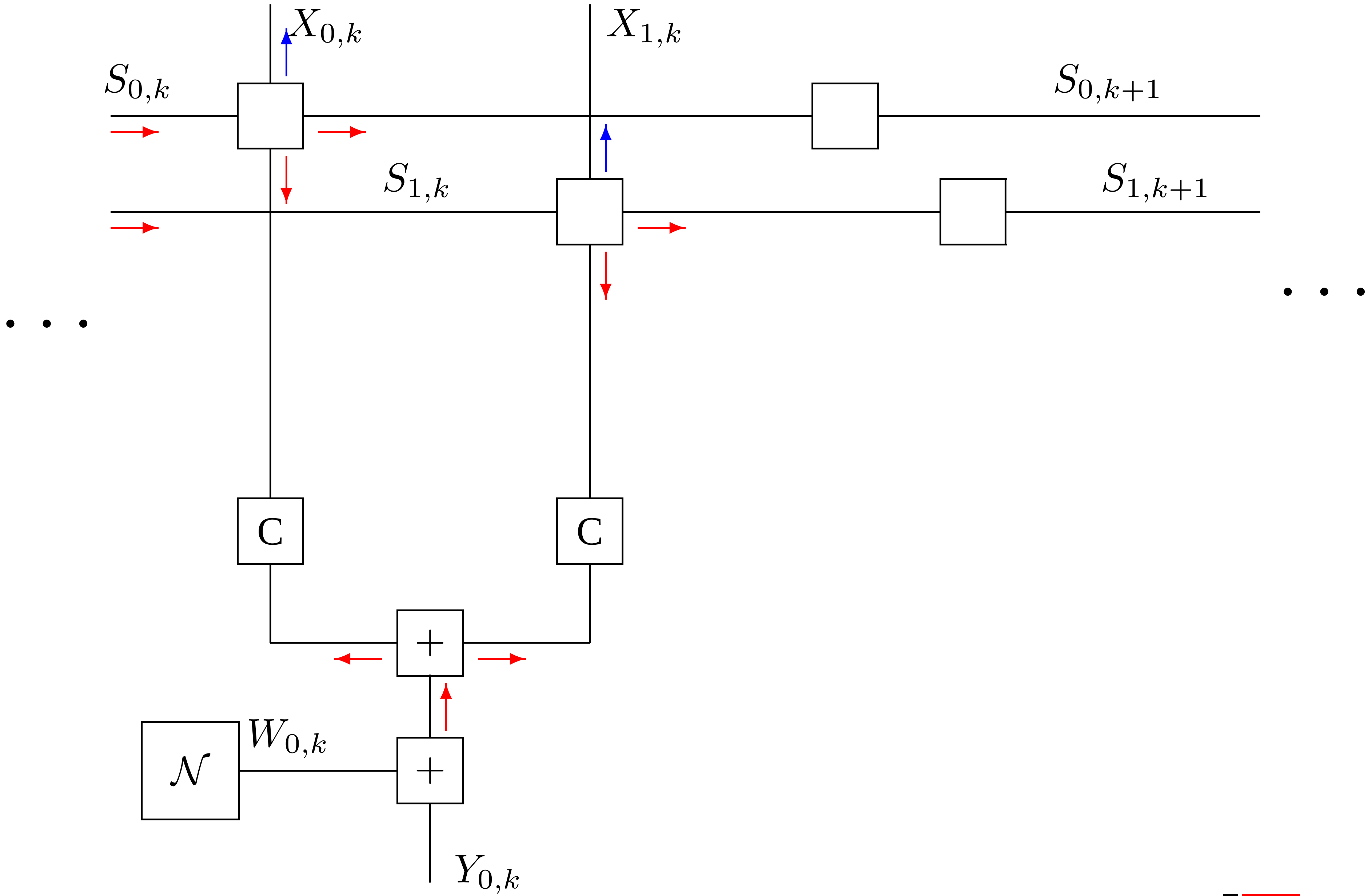
Estimating Motor Unit Firing Times



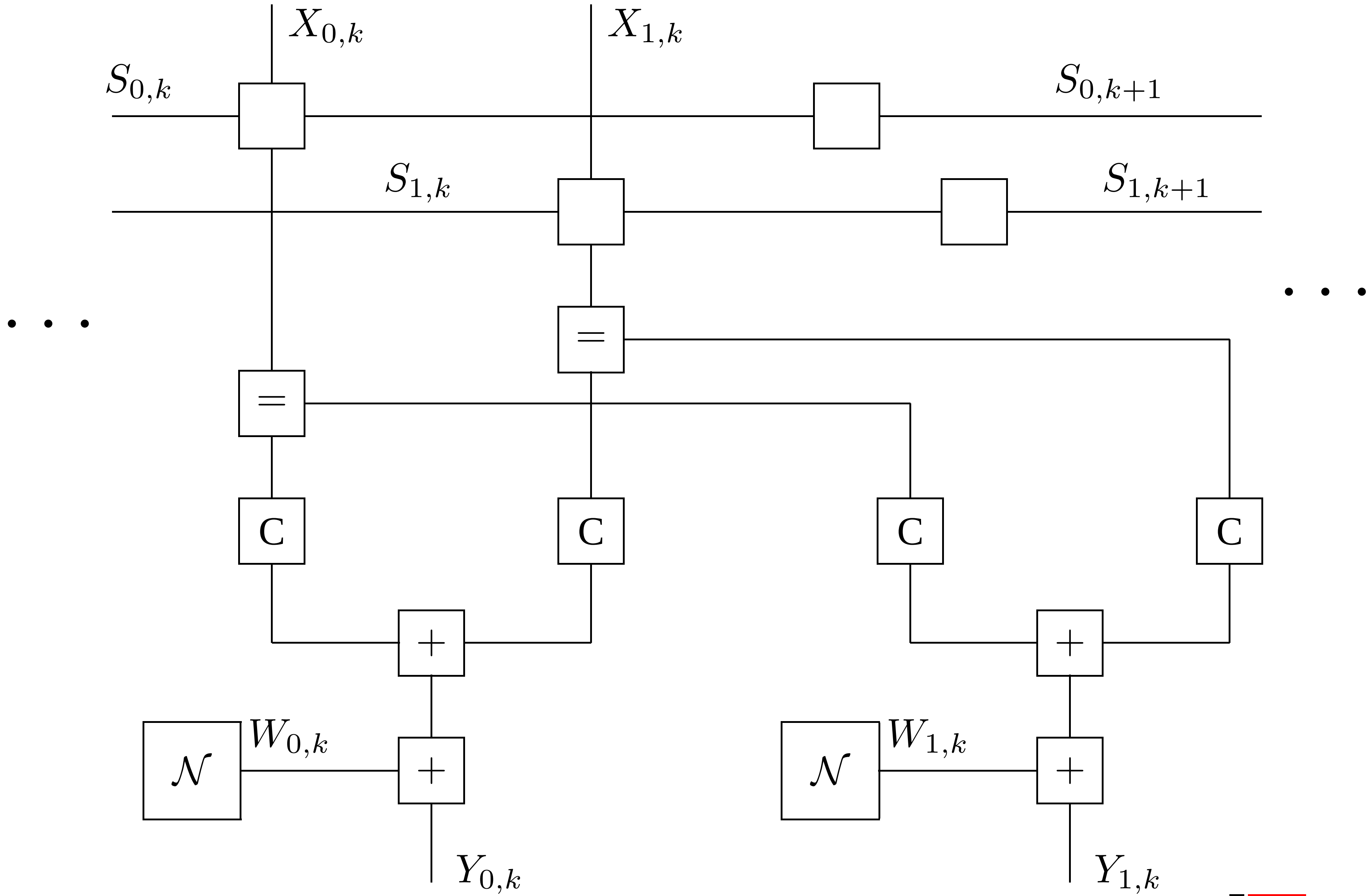
ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Signal and Information Processing Laboratory
Institut für Signal- und Informationsverarbeitung

Estimating Motor Unit Firing Times



Factor Graph for Two Channels



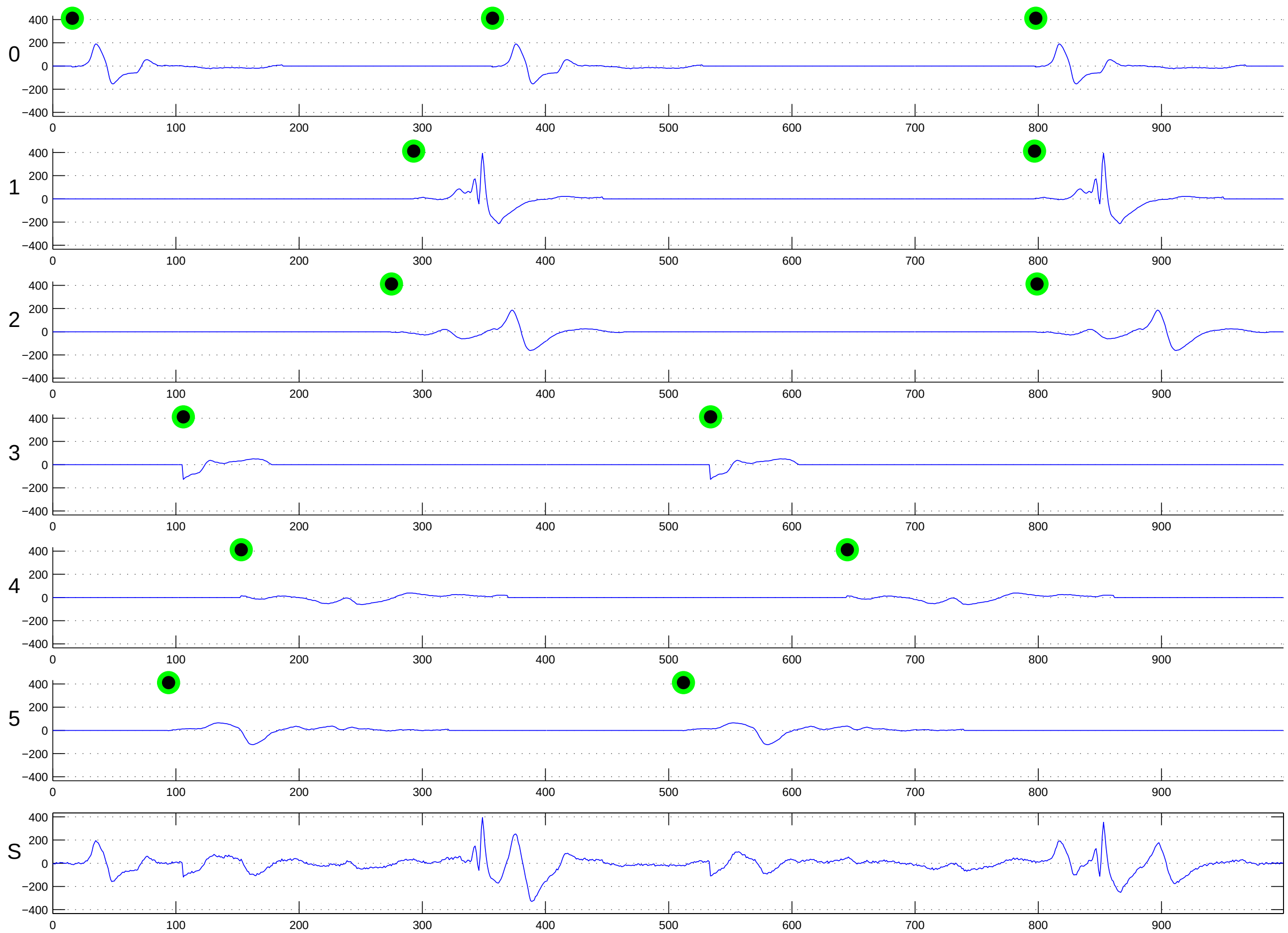
Preliminary Results

- **Simulated signals**
  - Decomposition of superpositions of more than 8 overlapping MUAPs
- **Measured signals**
  - First tests done
- **Changing MUAP shapes in simulated long-term recordings**
  - First tests done
- **MUAP shapes**
  - Need to be given to our algorithm
- **Noise model**
  - AWGN, tried others

Results: 1 Simulated Signal, 6 MU

• = actual firing

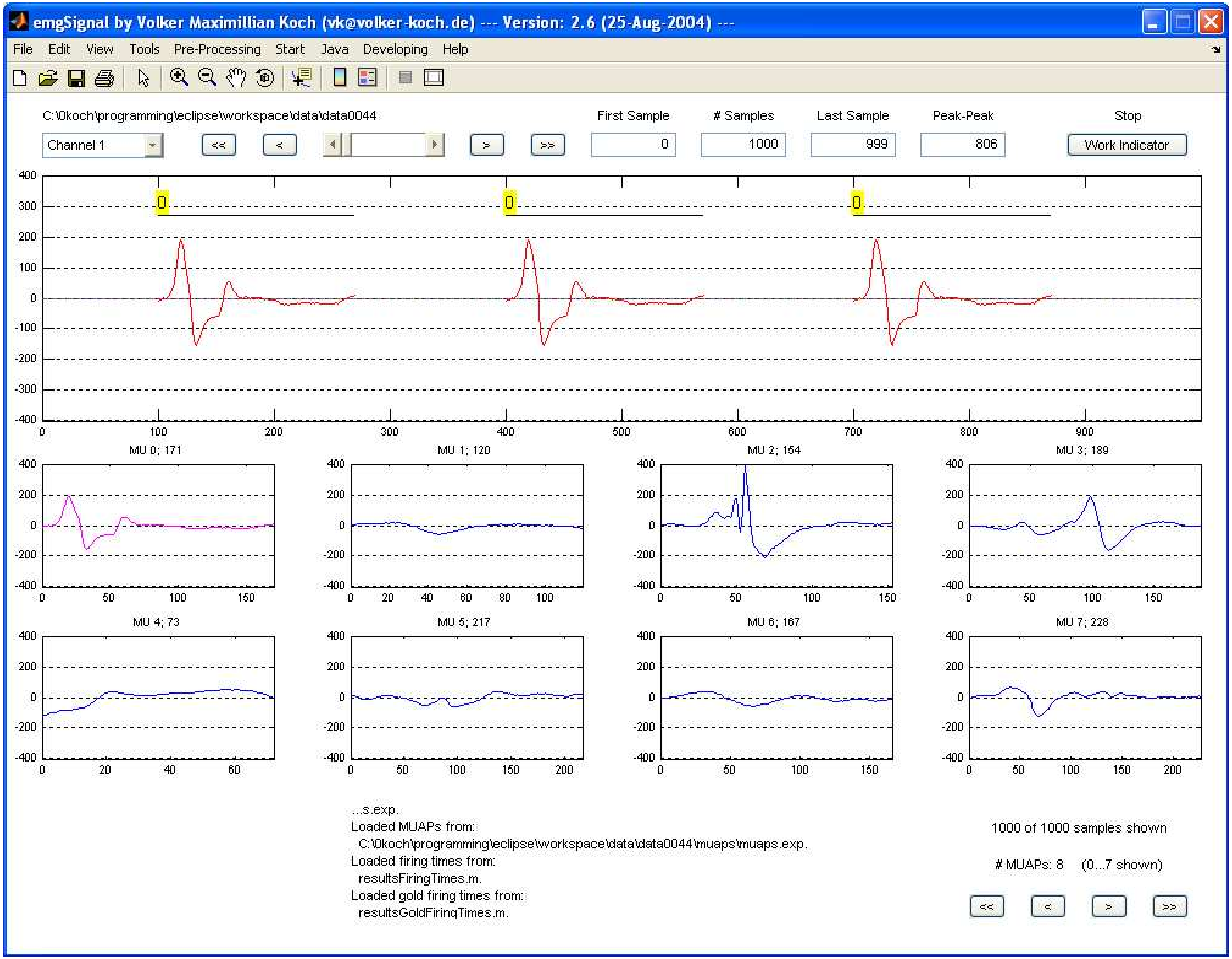
● = detected firing



Results: 3000 Simulated Signals, 6 MU

σ_{noise}	Number of completely correct decomposed signals / total number of signals
0	$899/1000$
5	$836/1000$
10	$720/1000$

Results: 1 Simulated Signal, 8 MU



ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

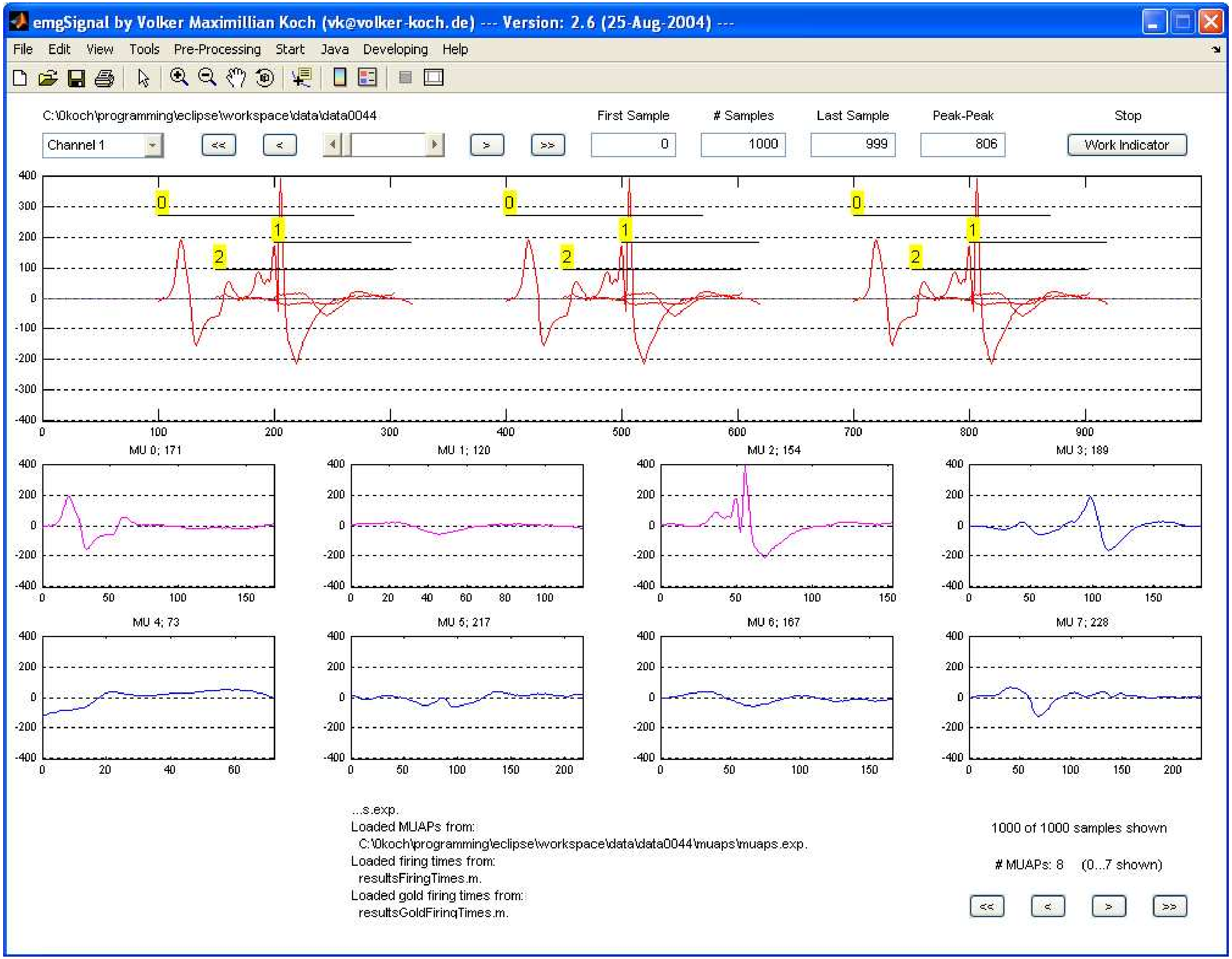
Signal and Information Processing Laboratory
Institut für Signal- und Informationsverarbeitung

Results: 1 Simulated Signal, 8 MU

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Signal and Information Processing Laboratory
Institut für Signal- und Informationsverarbeitung

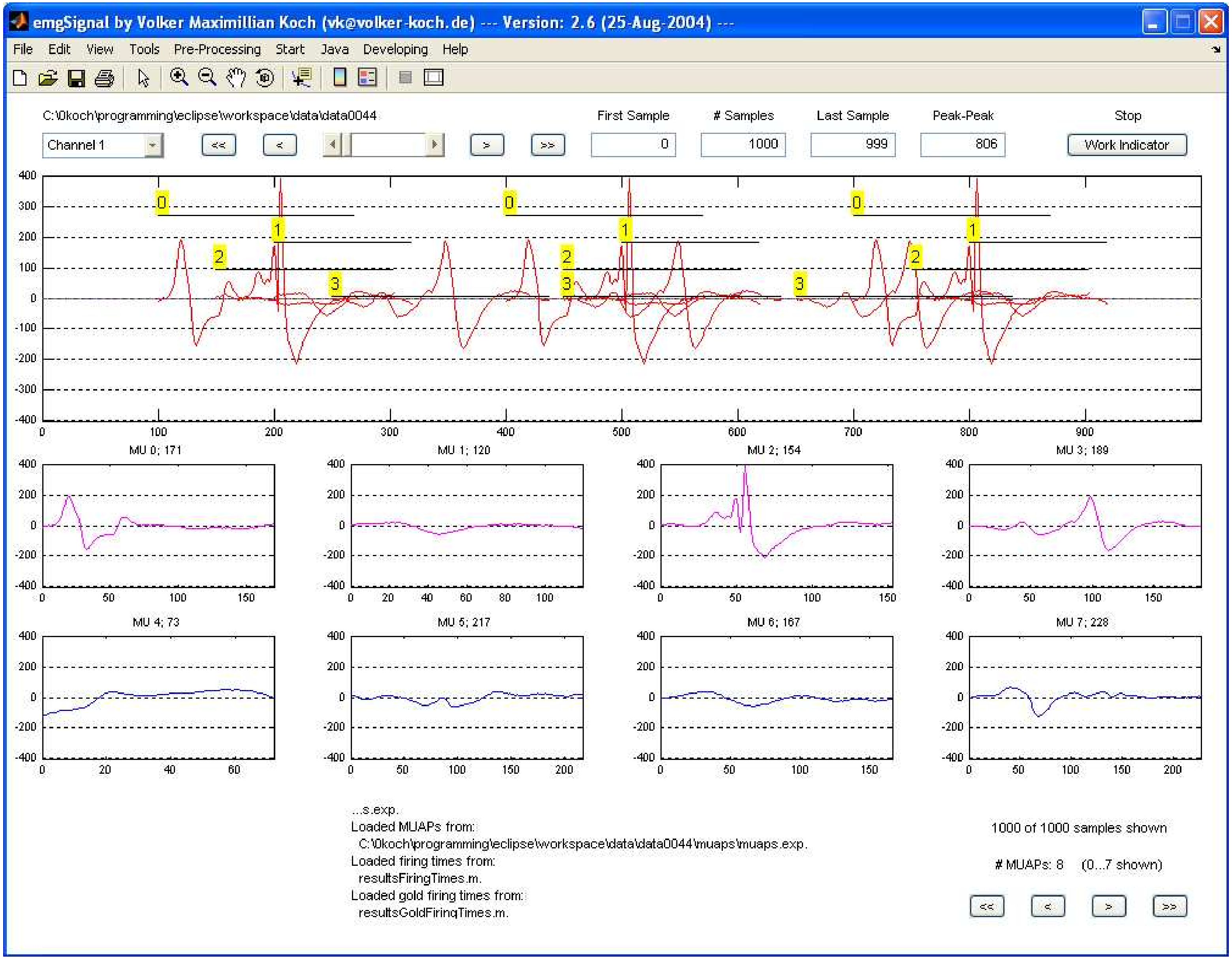
Results: 1 Simulated Signal, 8 MU



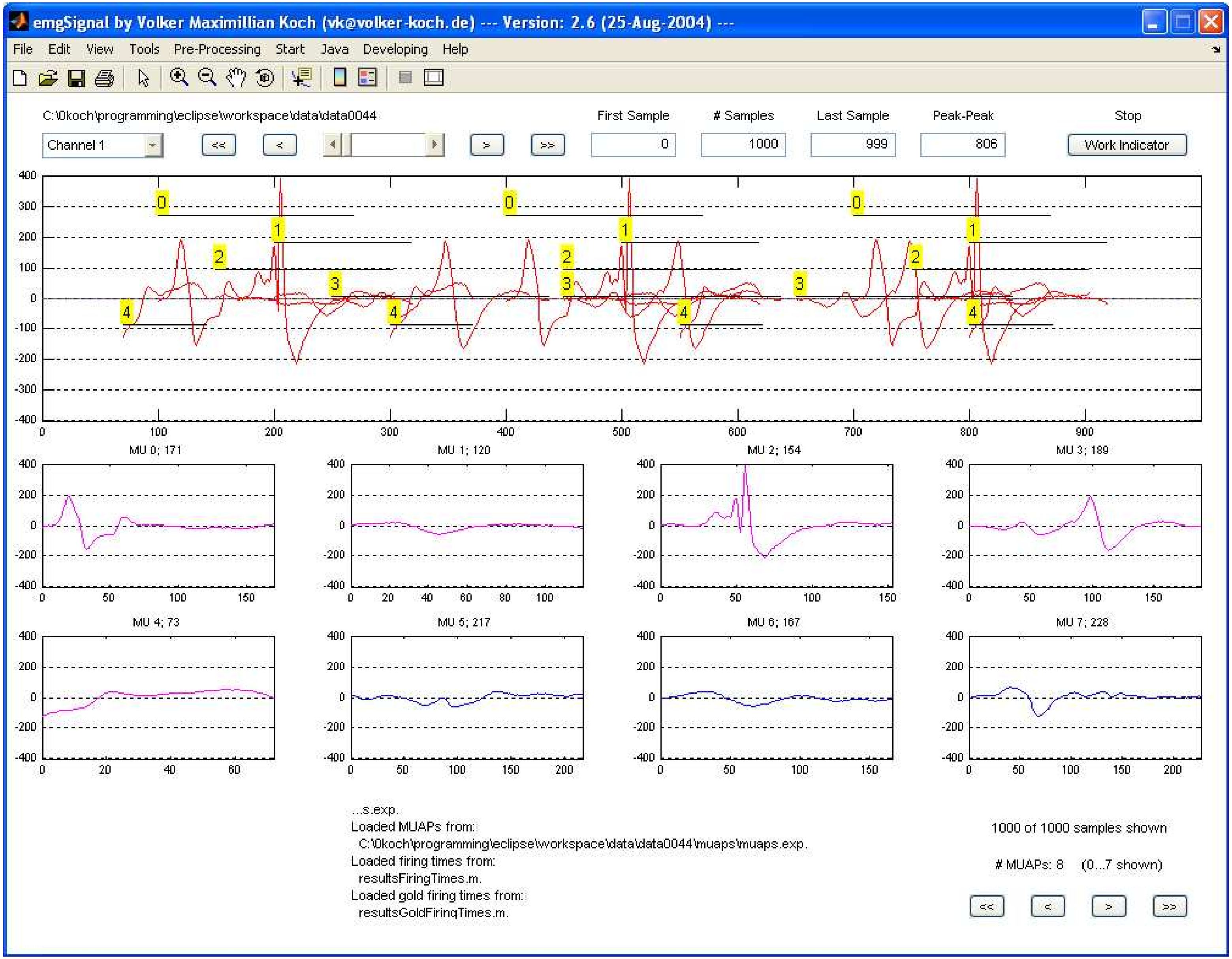
ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Signal and Information Processing Laboratory
Institut für Signal- und Informationsverarbeitung
ISI

Results: 1 Simulated Signal, 8 MU



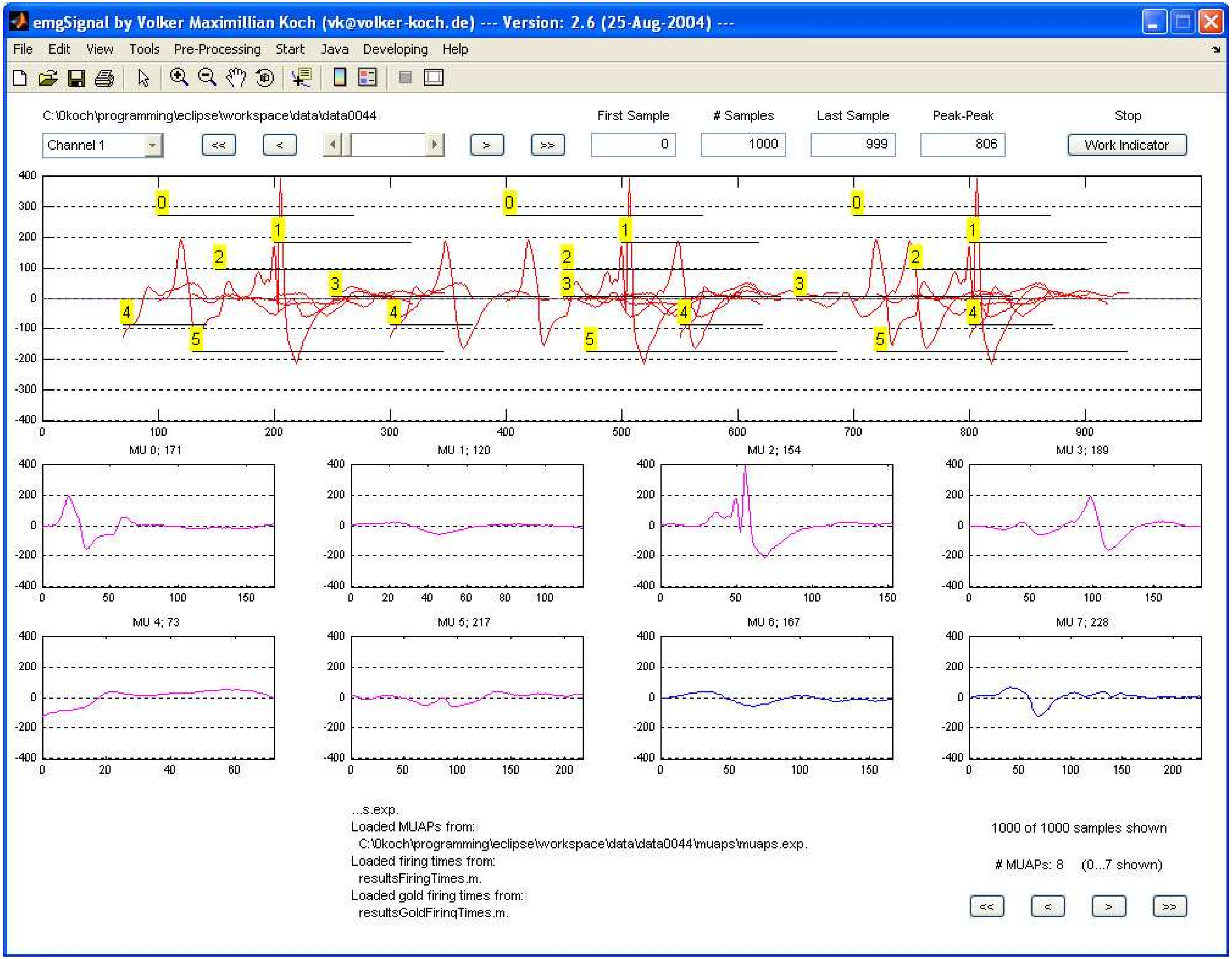
Results: 1 Simulated Signal, 8 MU



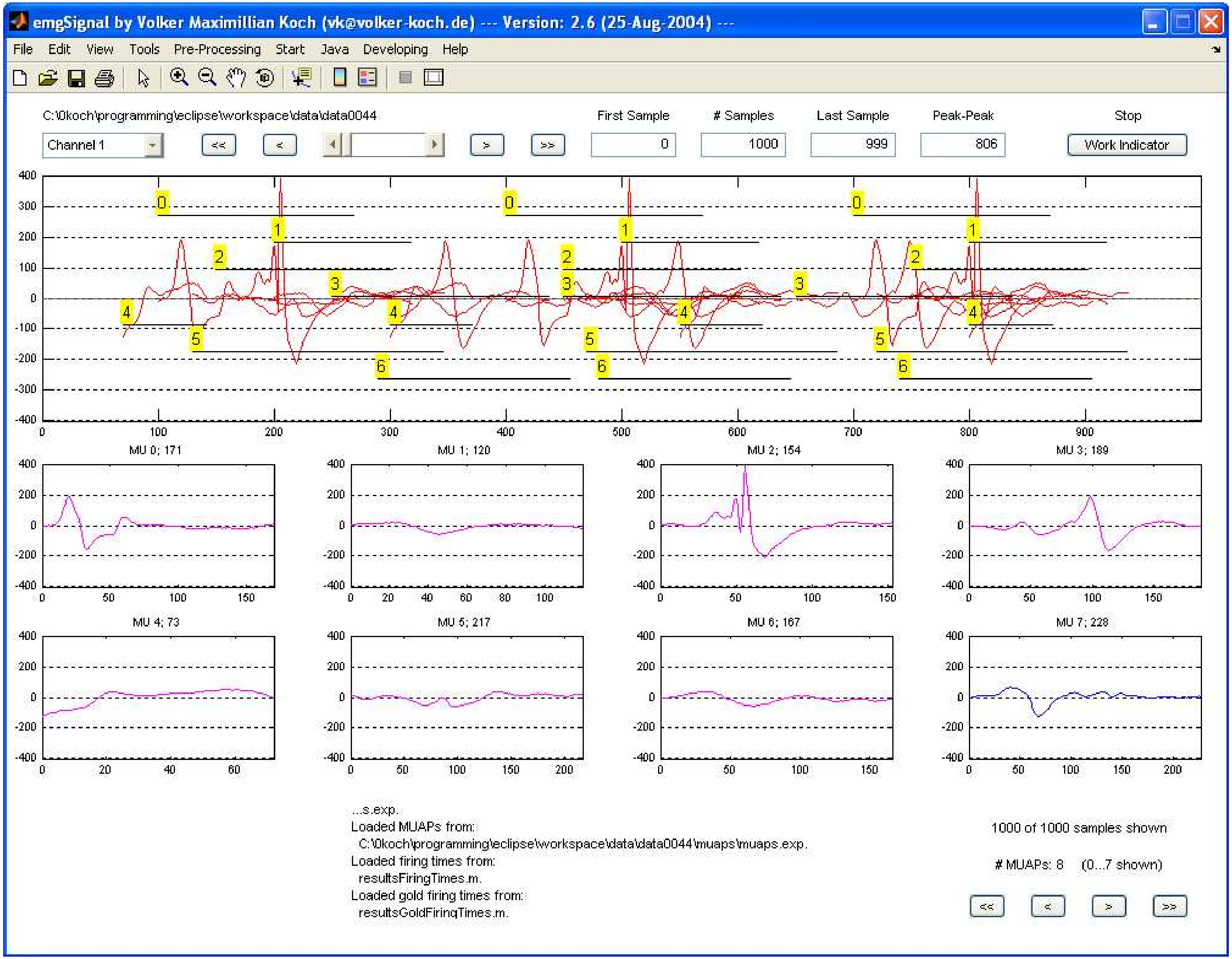
ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Signal and Information Processing Laboratory
Institut für Signal- und Informationsverarbeitung

Results: 1 Simulated Signal, 8 MU



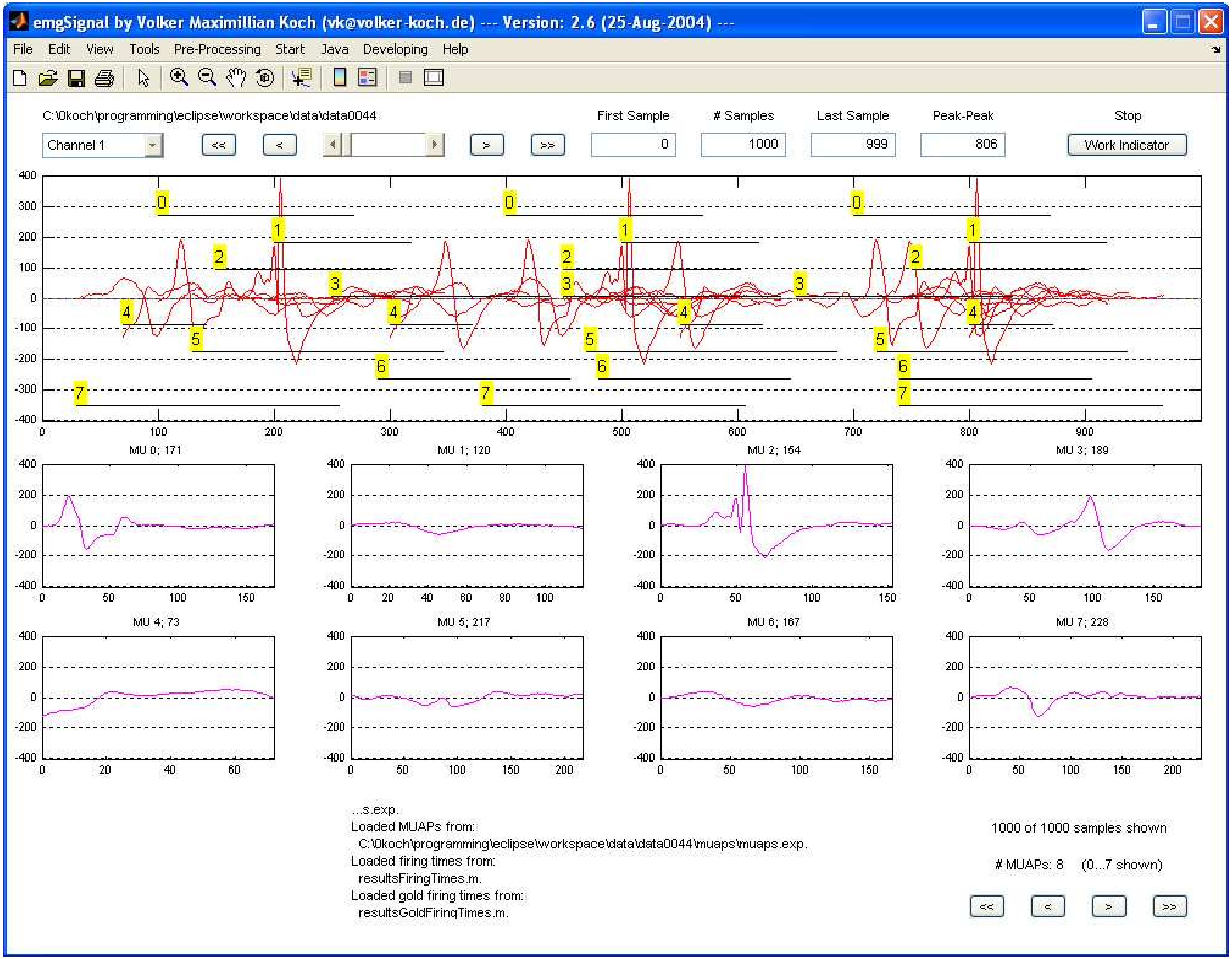
Results: 1 Simulated Signal, 8 MU



Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Signal and Information Processing Laboratory
Institut für Signal- und Informationsverarbeitung

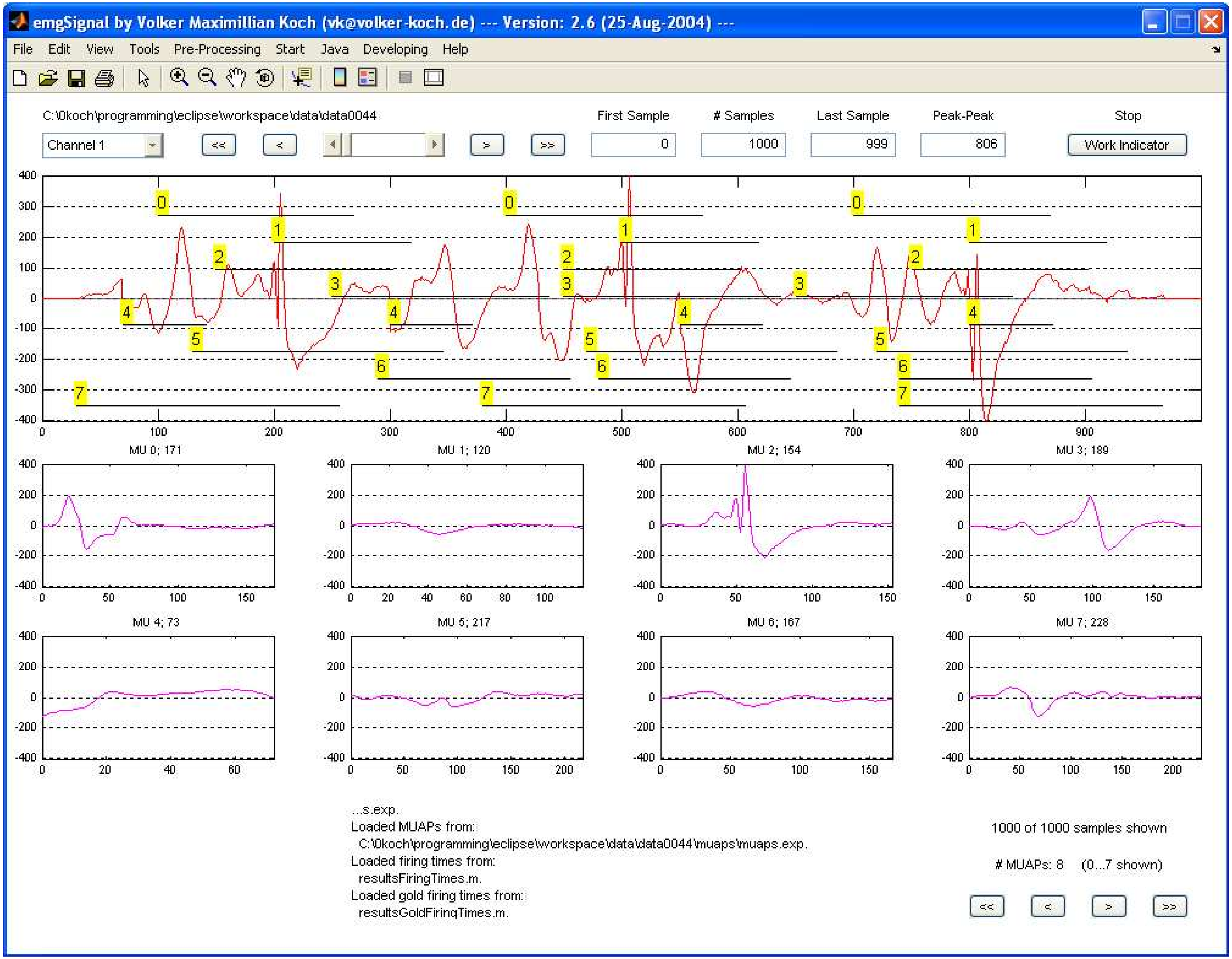
Results: 1 Simulated Signal, 8 MU



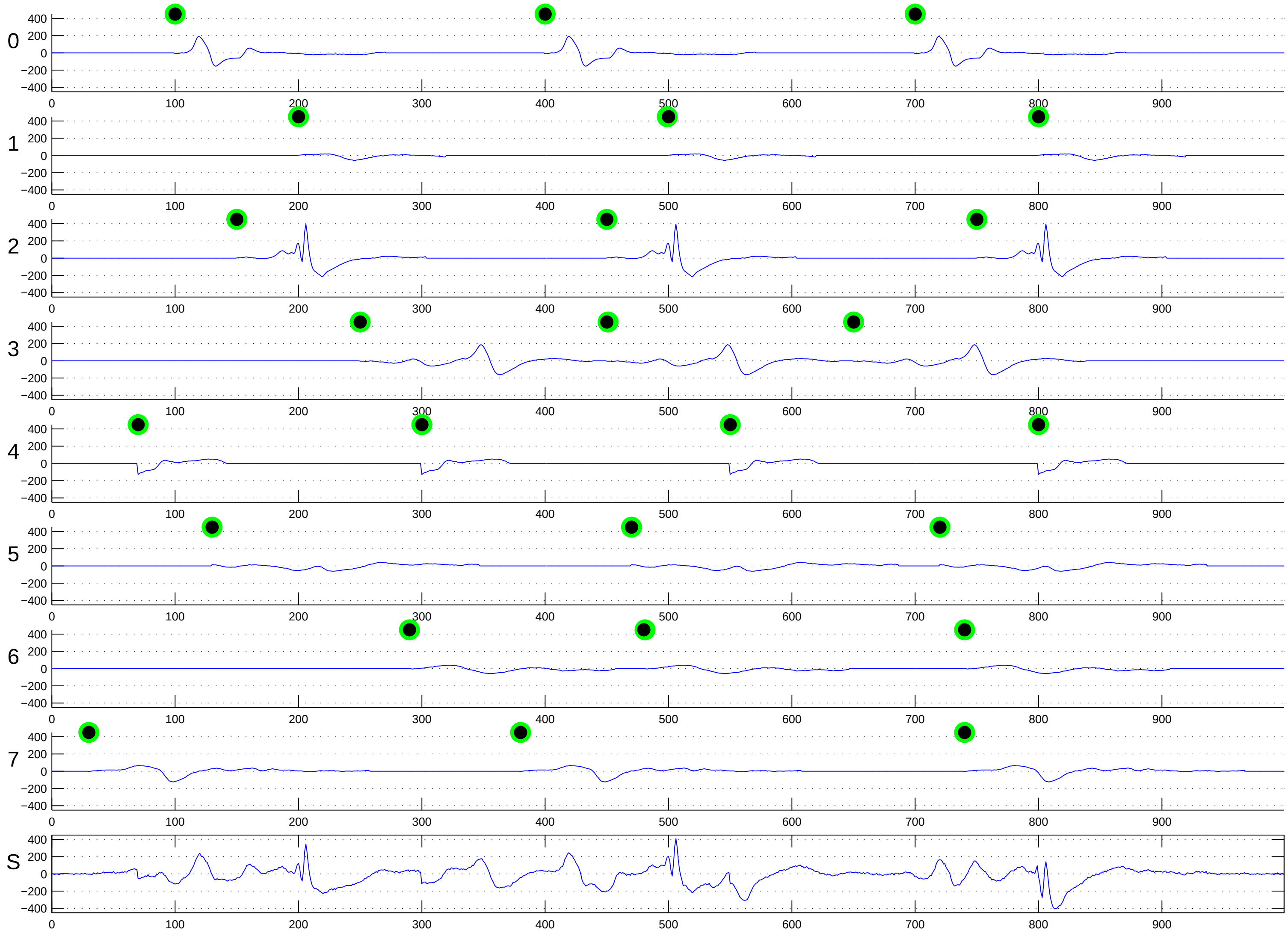
ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Signal and Information Processing Laboratory
Institut für Signal- und Informationsverarbeitung

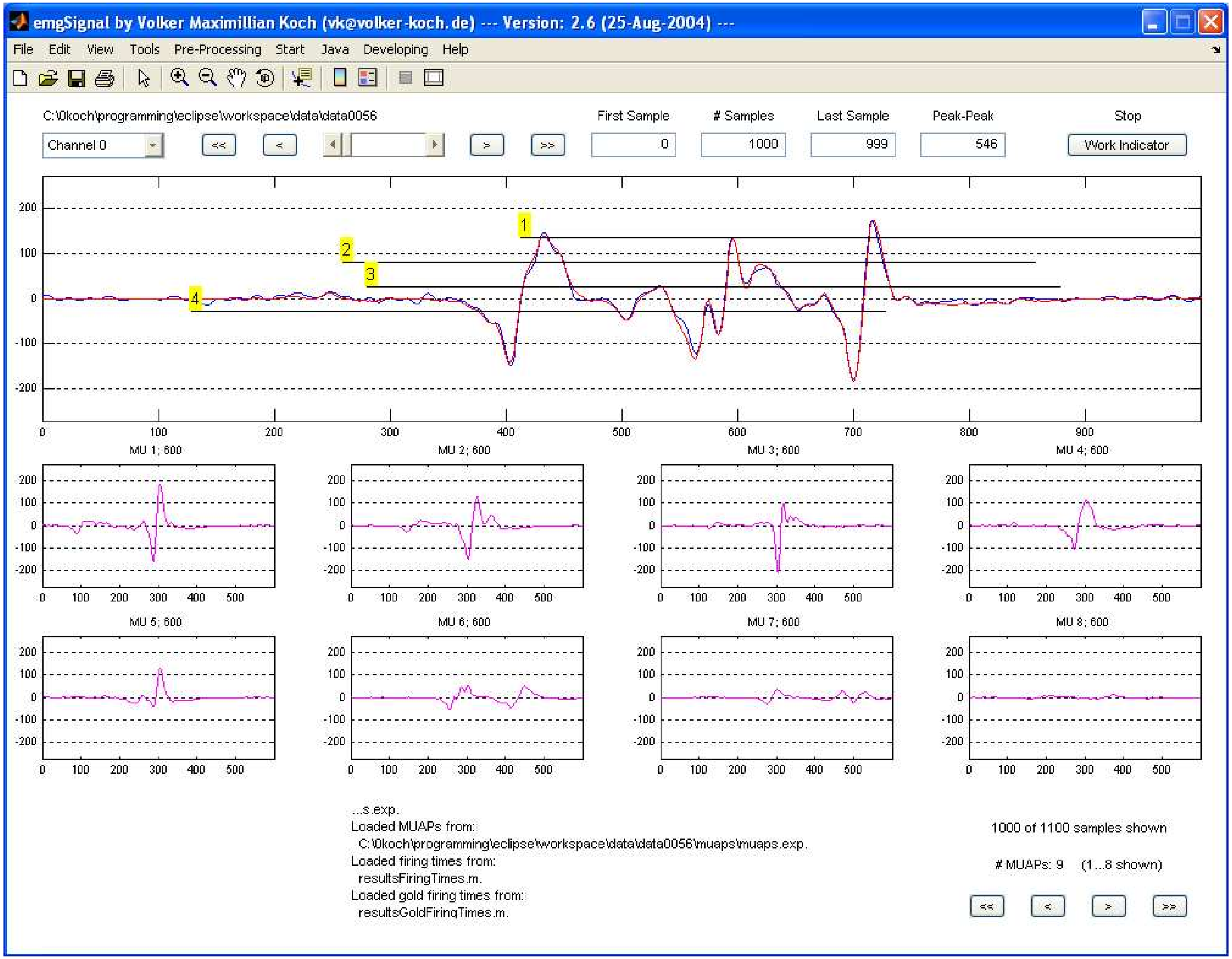
Results: 1 Simulated Signal, 8 MU



Results: 1 Simulated Signal, 8 MU



Results: 1 Measured Signal, 9 MU



Thanks to Kevin McGill for providing the EMG signal!

Overview

- Factor graphs
- Sum-product algorithm
- Application: decomposition of EMG-signals
- Conclusions

Conclusions

- **EMG signal decomposition algorithm**
- **Based on belief propagation in a graph**
- **Aim: Resolving difficult superpositions without much a-priori information**
  - e.g., no firing statistics
  - e.g., dynamic contractions
  - e.g., doublets

Applications of graphical models in signal processing

Brain-Style Information Systems Research Group,
RIKEN BSI, Saitama, Japan

Justin Dauwels

dauwels@isi.ee.ethz.ch

Signal and Information Processing Laboratory

Department of Information Technology and Electrical Engineering

ETH Zurich, Switzerland

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Signal and Information Processing Laboratory
Institut für Signal- und Informationsverarbeitung