Model selection for the extraction of EMG synergies

Poster · April 2014

4 authors, including:

Dominik Endres
Philipps University of Marburg

Andrea D'Avella
Università degli Studi di Messina

Martin A. Giese
University of Tuebingen

Hertie-Institut für klinische Hirnforschung
UKT Eberhard-Karls-Universität Universitätsklinikum Tübingen
FONDAZIONE SANTA LUCIA

Model selection for the extraction of EMG synergies

Enrico Chiovetto[1], Dominik Endres[1], Andrea d'Avella[2] and Martin Giese[1]

1- Section for Computational Sensomotorics, Department of Cognitive Neurology, Hertie Institute for Clinical Brain Research, Centre for Integrative Neuroscience, University Clinic Tübingen, Tübingen, Germany. 2- Laboratory of Neuromotor Physiology, Santa Lucia Foundation, Rome, Italy.

CIN Werner Reichardt Centre for Integrative Neuroscience
EBERHARD KARLS UNIVERSITÄT TÜBINGEN

Introduction

It has been shown that the electromyographic (EMG) patterns underlying complex movements can be approximated by the combinations of a small number of muscle synergies [1, 2]. Different definitions of muscle synergies have been given, which translate into either synchronous or non-instantaneous generative models [3]. However, how to choose a specific model and how to identify its parameters, such as the number of synergies, still remain open questions [4]. For this study we extracted different kinds of synergies from simulated EMG data sets with known statistical properties. We then applied a series of model selection criteria to discriminate the actual synergistic models underlying the ground-truth data sets and to identify the actual number of synergies. We could in this way compare model selection performance across the different criteria.

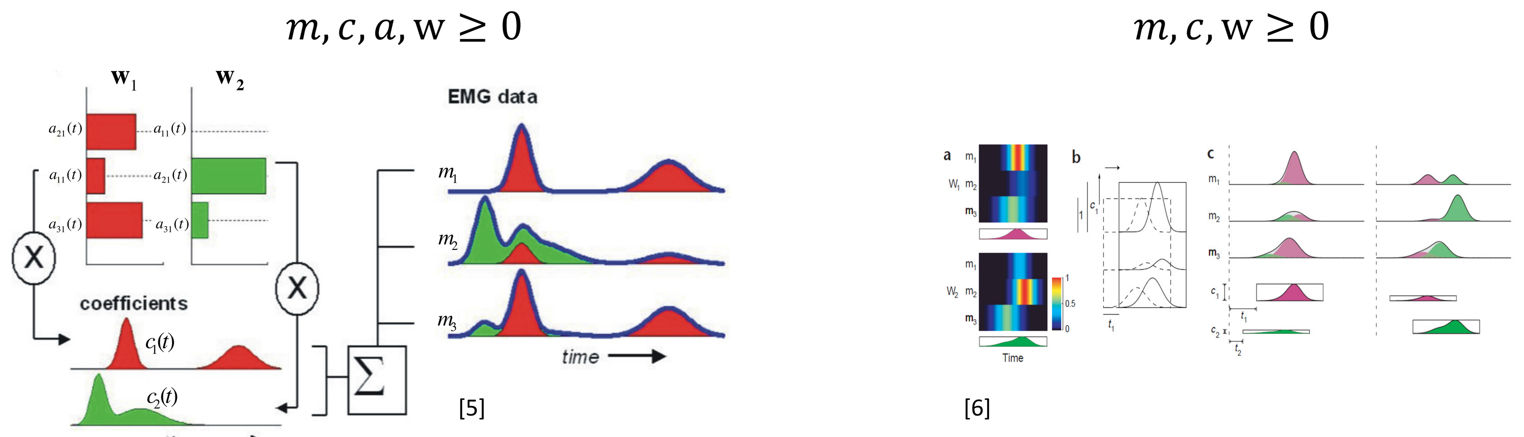
Methods

Models of muscle synergies and synthetic EMG data sets

- **Synchronous synergies**
- **Temporal synergies**
- **Time-varying synergies (TV)**

$$\mathbf{m}(t)=\sum_{j=1}^{N}\mathbf{w}_j\cdot c_j(t) \qquad m_i(t)=\sum_{j=1}^{N}a_{ij}\cdot c_j(t) \qquad \mathbf{m}(t)=\sum_{j=1}^{N}c_j\cdot\mathbf{w}_j(t-\tau_j)$$

$$m, c, a, \mathrm{w}\geq 0 \qquad m, c, \mathrm{w}\geq 0$$



For each synergistic organization we generated 20 artificial EMG data sets. Temporal waveforms generated starting from renewal stochastic processes. Weighting coefficients and temporal delays withdrawn from random exponential distributions. Data were generated by combining N = **4** artificial synergies according to the models above. Each data set consisted in the simulated activity of **10** muscles collected during **25** trials. T = **100** time samples. Data corrupted with signals dependent noise (R^2 = 0.85, between noisy and noiseless data).

FADA algorithm

All definitions of primitives can be derived from a unique generative model, usually referred to as **anechoic mixture model**

$$x_i(t)=\sum_{j=1}^{N}\alpha_{ij}\cdot s_j(t-\tau_{ij})$$

They differ from each other only by additional constraints imposed on the parameters of the model (e.g. non-negativity or equality constraints, presence of delays). We started from previous work developed in our lab [7] to design a new and more efficient algorithm for the identification of motor synergies based on Fourier series decomposition of the EMG signals[8-10].
Approximating the signal and the delayed sources by truncated Fourier series we obtain

$$x_i(t)=\sum_{k=-M}^{M}c_{ik}\,e^{ikt} \qquad s_j(t-\tau_{ij})=\sum_{k=-M}^{M}v_{jk}e^{-ik\tau_{ij}}e^{ikt}$$

where M is an integer and c_{ik} and v_{jk} belong to the complex space. The last equations allows to derive the following iterative EM algorithm

A
$$|c_{ik}|^2=\sum_{J=1}^{N}|\alpha_{ij}|^2|v_{jk}|^2$$

Identification of the **absolute values** of the Fourier coefficients of the sources

B
$$c_{ik}=\sum_{J=1}^{N}\alpha_{ij}|v_{jk}|e^{i\varphi_{v_{jk}}}e^{-ik\tau_{ij}}$$

Identification of phases of the Fourier coefficients of the sources (non-linear quadratic optimization, non-negativity constraints are imposed here)

C
$$[\overline{\tau}_j, A_j]=\mathrm{argmin}_{[\overline{\tau}_j, A_j]}\left\|x_j-A_j\cdot S(\overline{\tau}_j)\right\|_2$$

Identification of **weights and delays** assuming that the sources are known (standard cross-correlation procedure, non-negativity of the weights are imposed here)

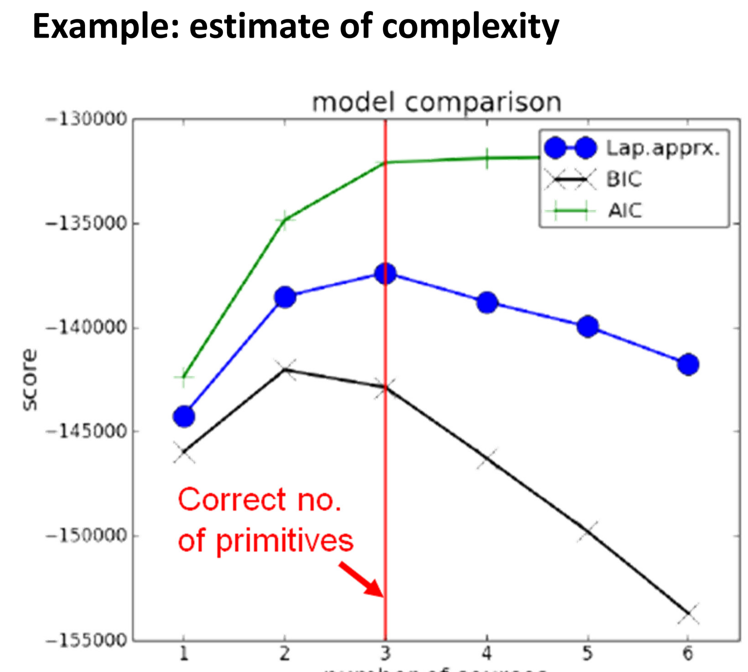
Model selection criteria

Development of a new Bayesian criterion [4] based on Laplace approximation (**LAP**) for:
- Model selection
- Estimation of model complexity (# of primitives)
- Most likely type of smoothness prior
- Non-negativity is imposed with a rectifying function $y=\frac{1}{\alpha}\log(1+e^{\alpha y})$

Log likelihood of data (Laplace approximation):

$$\log(p(D\mid S,M))\approx\log(p(D\mid\Theta^*,S,M))+\log(p(\Theta^*\mid S,M))+\frac{F}{2}\log(2\pi)-\frac{1}{2}\log(|H|)$$

Example: estimate of complexity

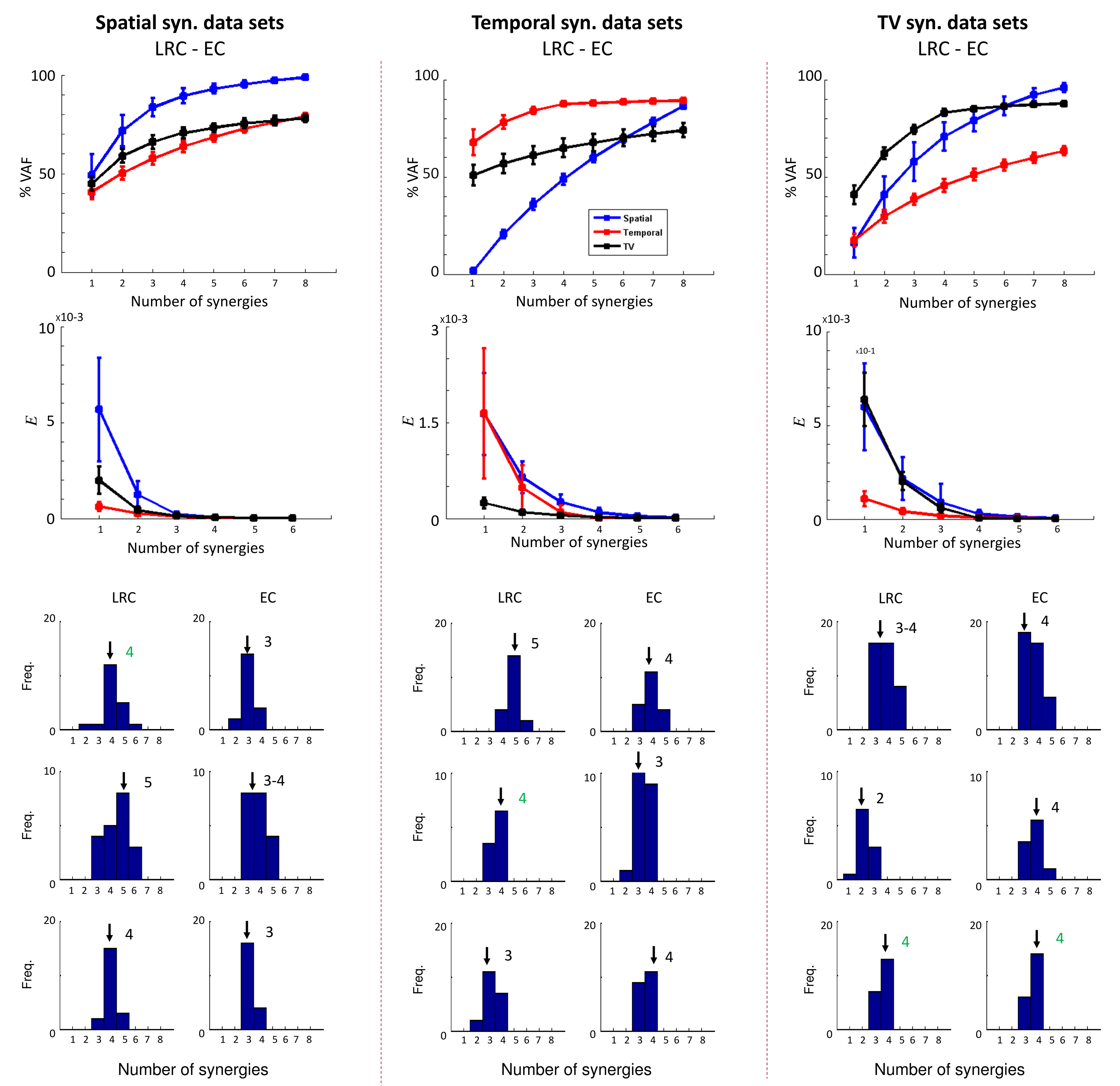


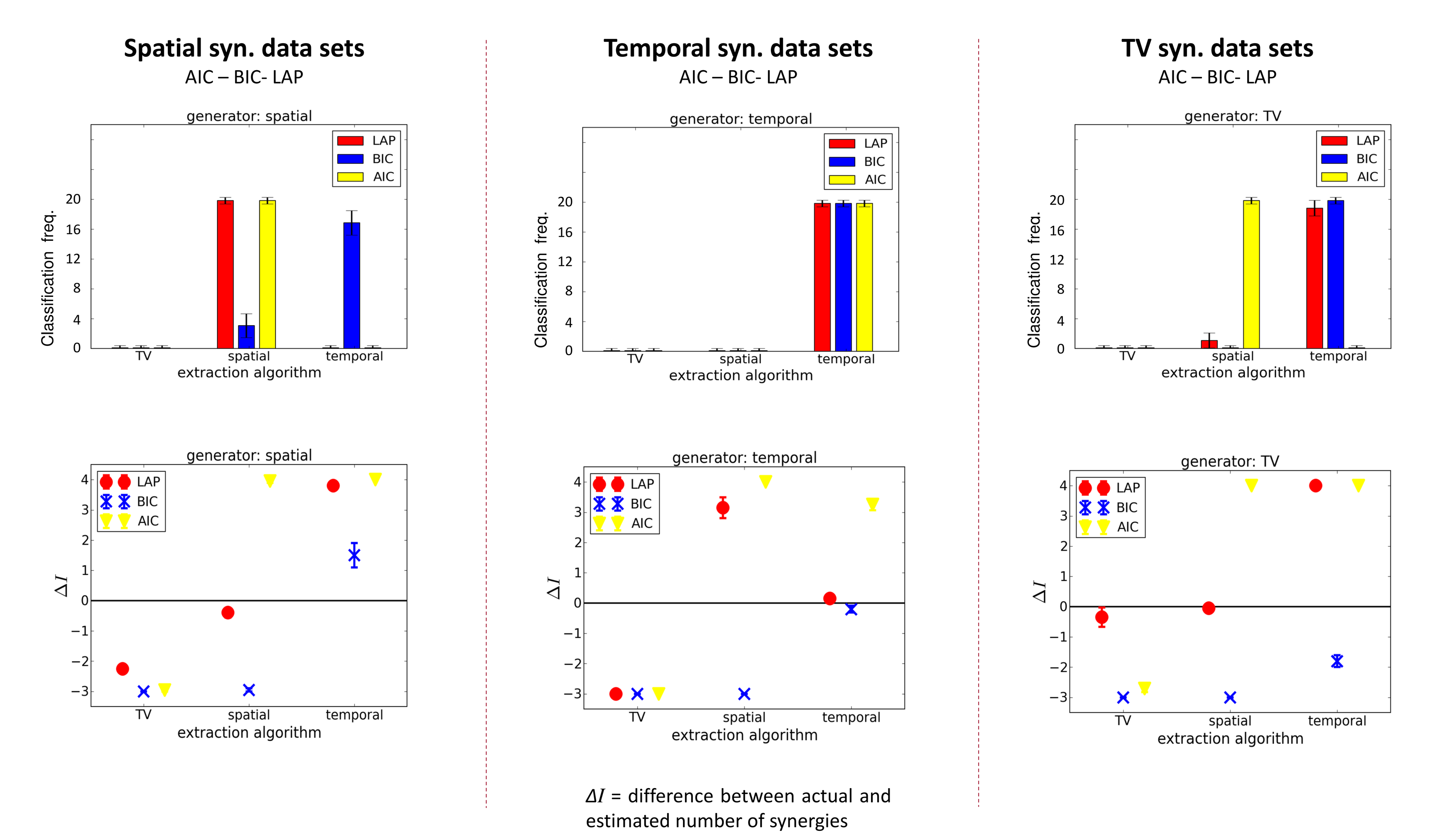
D: data
S: number of primitives
M: model type (PCA, ICA, Anechoic, …)
Θ: parameter vector
F: number of estimated parameters
H: Hessian (from Laplace approximation)

Additional model selection criteria

- **Linear Regression Criterion (LRC):** Linear regression procedure to identify the value of N above which the VAF curve is essentially straight. A series of linear regressions is performed, starting from a regression on the entire VAF curve and progressively removing the smallest N value from the regression interval. The mean square residual errors of the different regressions are then compared and the optimal number of synergies N^* is selected as the first N^* value corresponding to a regression line from N to $Nmax$ with a mean square error $< Emin$, $Emin$ = 1x10-4 .
- **Elbow Criterion (EC):** Approximation of the curve indicating the amount of variance accounted for (VAF) with 2 line segments. Position of the 'elbow' determined by the minimum least square error computed over all N values.
- **Akaike Information Criterion:** $AIC=-2(\log(p(D\mid\Theta^*,S,M))-\frac{1}{2}\dim(\Theta)$
- **Bayesian Information Criterion:** $BIC=-2(\log(p(D\mid\Theta^*,S,M))-\frac{1}{2}\dim(\Theta\log(N))$

Results





Conclusions

MODEL SELECTION

- The graphs of the VAF provide information about the generative model underlying a given EMG data set
- AIC and LAP provide reliable model selection performance when the modular organizations underlying the EMG data sets rely on either spatial or temporal synergies.

DETERMINATION OF THE NUMBER OF SYNERGIES

- Once the underlying model is chosen LRC can always identify, on average, the exact number of synergies. EC revealed less reliable than LRC
- LAP always outperformed AIC and BIC and provides the right number of synergies

References

[1] Bizzi E, Cheung VCK, d'Avella A, Saltiel P and Tresch M. Combining modules for movement. *Brain Research Reviews*, vol. 57, no. 1, pp. 125-133, 2008.
[2] Flash T and Hochner B. Motor primitives in vertebrates and invertebrates. *Curr Opin Neurobiol*, vol. 15, no. 6, pp. 660-666, 2005.
[3] Chiovetto E, Berret B, Delis I, Panzeri S and Pozzo T. Investigating reduction of dimensionality during single-joint elbow movements: a case study on muscle synergies. *Front Comput Neurosci*, vol. 7, p. 11, 2013.
[4] Endres DM, Chiovetto E, Giese MA. Model selection for the extraction of movement primitives. *Front Comput Neurosci,* vol. 7, pag 185, 2013.
[5] Tresch MC, Jarc A. The case for and against muscle synergies. *Curr Opin Neurobiol,* vol. 19, no. 6, pp. 601-7, 2009.
[6] d'Avella A, Saltiel P and Bizzi E. Combinations of muscle synergies in the construction of a natural motor behavior. *Nat Neurosci,* vol. 6, no. 3, pp. 300-308, 2003.
[7] Omlor L and Giese MA. Anechoic Blind Source Separation using Wigner Marginals. *Journal of Machine Learning Research*, vol. 12, pp. 1111-1148, 2011.
[8] Chiovetto E, d'Avella A, Giese MA. A comprehensive framework for the identification of motor primitives. Submitted.
[9] Chiovetto E, d'Avella A, Giese MA. A unifying framework for the identification of kinematic and electromyographic motor primitives. *NMC meeting, Puerto Rico, 2013.*
[10] Chiovetto E and Giese MA. Kinematics of the coordination of pointing during locomotion. *Plos One*, vol. 8, no. 11:e79555. doi: 10.1371/journal.pone.0079555, 2013.

Acknowledgements: This work was supported with funds from EU project AMARSi (FP7-ICT-248311), Koroibot (FP7-ICT-611909) and by the German Federal Ministry of Education and Research (BMBF; FKZ: 01GQ1002)