Linear and non-linear regression techniques for simultaneous and proportional myoelectric control

J. M. Hahne*[1,3], F. Bießmann[1,5], N. Jiang[2,3], H. Rehbaum[2,3], D. Farina*[2], F. C. Meinecke[1], K.-R. Müller*[1,4,5], L. C. Parra*[6]

Abstract—In recent years the number of active controllable joints in electrically powered hand-prostheses has increased significantly. However, the control strategies for these devices in current clinical use are inadequate as they require separate and sequential control of each degree of freedom (DoF). In this study we systematically compare linear and non-linear regression techniques for an independent, simultaneous and proportional myoelectric control of wrist movements with two DoF.

These techniques include linear regression (LR), mixture of linear experts (ME), multilayer-perceptron (MLP) and kernel ridge regression (KRR). They are investigated offline with electromyographic (EMG) signals acquired from ten able-bodied subjects and one person with congenital upper limb deficiency. The control accuracy is reported as a function of the number of electrodes and the amount and diversity of training data providing guidance for the requirements in clinical practice

The results showed that KRR, a non-parametric statistical learning method, outperformed the other methods. However, simple transformations in the feature space could linearize the problem, so that linear models could achieve similar performance as KRR at much lower computational costs.

Especially ME, a physiologically inspired extension of linear regression represents a promising candidate for the next generation of prosthetic devices.

Index Terms—Electromyography, regression, prosthetic hand, robust control.

I. INTRODUCTION

In recent years there have been substantial advances in constructing electrically powered hand prostheses that could perform complex movements involving many simultaneously controlled degrees of freedom (DoF), including independent finger movements [1], [2]. However, so far there exists no electro-myographic (EMG)-based controller that can extract the required control information needed to make full use of these prostheses. Clinically available controllers are based on very simple techniques that control only one DoF at a time.

Multiple dimensions have to be controlled sequentially, requiring slow and cumbersome mode-switching initiated by co-contractions. Significant research has been devoted to directly control many DoFs with classification based approaches (see e.g. [3] for a recent review). The reported accuracy of recent approaches is very high and also robustness issues under real world conditions have been addressed [4], [5]. Yet, most classification based approaches control only one function at a time, precluding intuitive control of smooth movements. Recent efforts have also extended the classification into more than one class (movement) at a time [6], [7]. However these approaches still limit the type of movements because the speed of the related DoFs cannot be controlled independently if two functions are activated at the same time. Conversely, natural movements can only be achieved with independent proportional control of the related DoFs.

To achieve an independent proportional and simultaneous control, regression techniques can be applied. The major difference to classification is that a regressor does not decide for a certain class but instead a continuous output value is estimated for each DoF. This allows for an *independent* simultaneous and proportional estimation and can facilitate a fluent and natural control, given a good regression performance. Lacking of this natural control is indeed one of the main limitations of the current myoelectric control approach based on classification [8].

Relative little work has been done on this in the context of myoelectric control, mostly focusing on multilayer perceptrons (MLPs) for regression ([9], [10], [11]). This study aims at a comprehensive and systematic comparison of state-of-the-art regression methods for independent proportional and simultaneous myoelectric control of multiple DoF. We compare simple linear models with state-of-the-art non-linear and non-parametric machine learning methods. For a clinical application, a method should require little user training, be computationally efficient and also perform well with few electrodes. Those aspects are addressed as well in the present study by reducing the amount of training data, reducing the number of EMG channels and by evaluating the processing times of the algorithms.

A major challenge for regression methods in myocontrol is to obtain accurate movement and force data for training in the absence of the missing limb. Jiang et al. [12] approached this problem by applying a semi-supervised algorithm, where only information about the active DoF and desired direction are needed to learn the relationship between muscle forces and EMG features. This approach can only exploit training data

with individual DoFs active. Nielsen et al. [9] investigated a bilateral training strategy that can be applied for unilateral amputees who represent the majority of hand prostheses users. The subjects performed bilateral mirrored contractions and the forces were estimated from EMG signals using artificial neuronal networks trained with force labels from the contra-lateral hand. Muceli [10] and Jiang [11] showed that it is also possible to estimate wrist angles instead of forces performing free dynamic movements from EMG with neuronal networks using this contra-lateral training strategy.

Most studies on simultaneous myoelectric control used the variance of the EMG (also denoted as mean square value or band power) [12], [11] or, similarly, the lowpass-filtered, down-sampled squared raw EMG-signal [10]. Nielsen et al. ([9],[13]) discovered that other features, like the time domain-feature set (mean absolute value, zero-crossings, slope sing changes, waveform length) perform significantly better than the variance.

In this study we demonstrate that the relationship between the variance and the wrist angle is highly non-linear and that simple transformations in feature space can simplify the problem. This allows to use linear methods, which are computationally efficient. We compare four linear and non-linear regression techniques, namely, linear regression (LR), mixture of linear experts (ME), multilayer perceptrons (MLPs) and kernel ridge regression (KRR). To our knowledge, KRR and ME have not previously been applied to myoelectric control. This comparison provides an evaluation of the potential use of EMG for simultaneous and proportional control and indications on the main factors of influence for regression performance.

II. METHODS

A. Experimental Setup

This study involved ten able bodied subjects (3 females, 7 males, age 19-30) and one person with congenital upper limb deficiency (male, age 39) performing a series of wrist movements. Accurate data labels were gained by tracking the wrist angles with a motion tracking system (Xsens with MTx sensors, Fig. 1b). EMG was recorded with a high density 192-channel electrode grid (ELSCH064NM 3-3, OT Bioelettronica, 8 x 24 channels, 10 mm inter-electrode-distance) in a monopolar configuration. The electrode array was placed on the proximal portion of the left forearm, covering a range of 8 cm. The biosignal amplifier was a 12 bit "OT Bioelettronica EMGUSB-2", configured to a sampling rate of 2048Hz. The reference electrode was a disposable Ag/AgCl electrode placed on the elbow. Ground was formed by an electrode band placed at the distal end of the forearm. Synchronization between kinematic and EMG signals was performed offline via a square-wave synchronization signal provided by the motion tracking system that was recorded as an additional (auxiliary) channel. Previous studies involved all three DoFs of wrist contractions. In this study we focus only on two DoFs, namely, flexion/extension and radial/ulnar deviation (Fig. 1 a). This restriction helped to prevent long recording times and difficulties with recording stability (pronation/supination can

(a) Target Trajectories
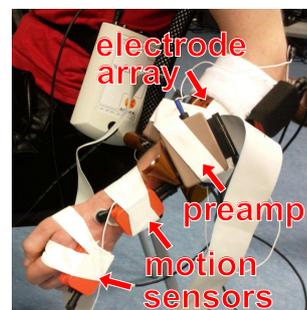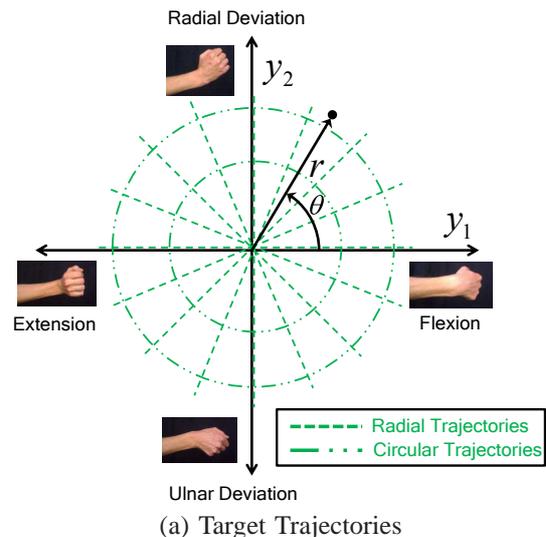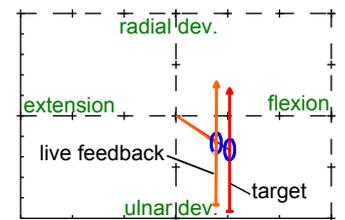
(b) Setup   (c) Feedback

Fig. 1. Experimental setup: (a) Subjects were instructed to follow radial and circular trajectories (dashed and dash-doted lines). Coordinates spanned by the two wrist angles **y** and in polar coordinates (r and θ). (b) placement of electrodes and motion sensors. (c) feedback during recording.

lead to shifting of muscles relative to skin and electrodes in able-bodied subjects – it is not known if this complication occurs in persons with limb deficiency).

The target movement trajectories (Fig. 1a) included moving the wrist in 16 (radial) directions, and drawing circles of two different diameters (clockwise and counter-clockwise). The subjects were instructed to keep the fingers in a relaxed position and not to rotate the wrist (keeping the thumb pointing upwards). At the beginning of each session, the individual range of motion in both DoFs of the subject was measured. The experimental paradigm was calibrated in such a way that the radial trajectories would start at the center (rest position) and reach the maximal range of motion for each direction. The circular trajectories were located at 90% and 60% of the maximal range of motion. The time from the center position to the maximal position was 3 s, followed by 2 s at the maximal position and 3 s for returning to the center position. The time for a full circular trajectory was 10 s. The completion of one trajectory will be referred in the following as a trial. The experiment was divided into several runs, where each run contained each type of trajectory (16 radial and 4 circular trials) exactly once. During the recordings, the target wrist angles were displayed on a computer screen together with the actual angles obtained by the motion tracking system (Fig. 1c).

This online feedback assisted subjects in better matching the target trajectories. Six able bodied subjects and the subject with congenital deficiency performed 15 runs and 4 subjects stopped after 10 runs because of fatigue. The time to record one session with 15 runs was about one hour plus another hour for placing the electrodes and motion sensors and familiarizing with the system.

To investigate the transferability of the results to the contra-lateral training strategy, for five of the ten subjects motion data was recorded from both sides while the subjects performed bilateral mirrored movements [9]. This allows for comparing the performance of ipsi-lateral training (motion data from the EMG side were used as training labels) with contra-lateral training (motion data from the other side were used as training labels). The contra-lateral training is relevant, particularly for future applications in uni-lateral amputees, where motion data can only be obtained from the intact side. The feedback for all able bodied subjects was given for the EMG side. An example of the recorded motion data is shown in figure 2.

To prove that the applied methods are also suitable for users of upper limb prostheses, we included one subject with congenital deficiency. The subject's forearm terminates at the wrist level. This subject performed also bilateral mirrored contractions. The EMG signals were recorded from the side with deficiency (right side) and the motion data were obtained from the contra-lateral side with intact limb. All experiments were in accordance with the declaration of Helsinki and were approved by the local ethics commission. (Ethikkommission d. Med. Fak. Göttingen, approval number 8/2/11)

(a) radial trajectories (b) circular trajectories
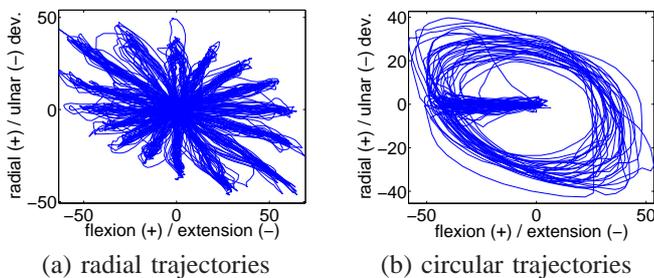
Fig. 2. Motion traces obtained by the motion tracking system (in degree) for both types of trajectories. The motion signals form the data labels $y(t)$ used to train and test the regressors.

B. Preprocessing

The data were filtered using a 4th order Butterworth high-pass filter ($f_c = 20Hz$) to remove movement artifacts, a lowpass ($fc = 500Hz$) to remove high frequency noise and a 50 Hz comb filter to remove power-line interference, including harmonics. Sample-wise common mean subtraction was performed to remove correlated noise and distortion that might be caused by activity at the reference electrode.

C. Feature Extraction

The features were extracted from non-overlapping intervals of 200 ms. This window duration is within the acceptable time delay between user command and prosthesis action [14],[15].

To obtain good estimation results when using linear methods the relationship between the features and the target labels (i.e., the motion data) should be as linear as possible. As the first feature we choose the *variance*. As we will show in section III-A, the variance is increasing monotonically with the deflection of the wrist in any direction, but the relationship between deflection and *variance* is not linear (see Fig. 3a). Therefore, we investigated two non-linear transformations, $f(x) = \sqrt{x}$ and $f(x) = log(x)$, to linearize the relationship between EMG and wrist angle. The transformed features are denoted by rms and log-var, respectively. All dimensions in feature space were normalized to have on average unit variance. This is useful for methods with parameters that depend on the numerical range of the features. The scaling factors were calculated based on the training data sets only.

D. Regression Models

The set of C dimensional feature vectors for T time instances is given as $\mathbf{X} \in \mathbb{R}^{C \times T}$, and $\mathbf{Y} \in \mathbb{R}^{D \times T}$ contains the corresponding wrist angles for D DoFs as data labels. The goal of all regression techniques is to find a mapping $\widehat{\mathbf{Y}} = f(\mathbf{X})$, where $\widehat{\mathbf{Y}}$ is an approximation of $\mathbf{Y}$.

1) Linear Regression (LR): In LR [16], [17] this mapping function is linear:

$$\widehat{\mathbf{Y}} = \mathbf{W}^{\top}\mathbf{X} + \mathbf{w}_0 \qquad (1)$$

where $\mathbf{W} \in \mathbb{R}^{C \times D}$ contains the weight vectors and $\mathbf{w}_0$ the bias that can compensate for possible offsets. By convention $\mathbf{w}_0$ is included in $\mathbf{W}$, thus extending $\mathbf{X}$ by an additional dimension including T ones.

The least mean squares solution for equation 1 including regularization is obtained by minimizing the following error function:

$$err(\mathbf{w}^{(d)}) = \frac{1}{2}\sum_{T}[\mathbf{y}^{(d)}(t) - \mathbf{w}^{(d)\top}\mathbf{x}(t)]^2 + \frac{\lambda}{2}\mathbf{w}^{(d)\top}\mathbf{w}^{(d)} \quad (2)$$

The closed form solution is given by:

$$\mathbf{W} = (\mathbf{X}\mathbf{X}^{\top} + \lambda\mathbf{I})^{-1}\mathbf{X}\mathbf{Y}^{\top} \qquad (3)$$

where $\mathbf{I}$ is the identity matrix and the regularization constant λ is optimized by grid-search in a nested cross-validation (sec. II-E).

2) Mixture of linear experts (ME): In LR each column vector $\mathbf{w}^{(d)}$ of $\mathbf{W}$ is responsible for the mapping from $\mathbf{X}$ to one DoF in $\mathbf{Y}$. This means that in LR the same coefficients are used for both antagonistic wrist movements which is physiologically not reasonable, since the antagonistic movements involve different muscles.

Therefore an extension of LR was applied which uses two different weight vectors $\mathbf{w}^{(d+)}$ and $\mathbf{w}^{(d-)}$ for each DoF d that are individually trained using only time intervals with positive or negative labels, respectively. The outputs of both filters are combined smoothly according to the probability to which direction the current feature sample belongs to, estimated by penalized logistic regression (PLR) [18], [17]:

$$\widehat{y}_d(t) = \left[\begin{array}{c} \mathbf{w}^{(d+)} \\ \mathbf{w}^{(d-)} \end{array} \right]^{\top} \left[\begin{array}{c} \mathbf{x}(t)\,p_d(t) \\ \mathbf{x}(t)\,(1 - p_d(t)) \end{array} \right] \qquad (4)$$

With

$$p_d(t) = \sigma\left(\boldsymbol{\beta_d}^\top \mathbf{x}(t)\right) \quad (5)$$

$\sigma(s) = \frac{e^s}{1+e^s}$ is the sigmoid function and the coefficients $\boldsymbol{\beta_d}$ are obtained by iterative reweighted least squares. The penalty term of PLR and the regularization parameter λ of the LRs are optimized in a nested cross-validation (sec. II-E).

For a steep sigmoid function the model can be seen as piece-wise linear with some smoothing around the origin. In this article we will refer to it as a linear method, even though this is not correct in a strict sense.

3) Multilayer perceptrons (MLPs): MLPs have been often used in the present context [9], [10], [17] and will be analyzed here for comparison. Each DoF was estimated by an individual network. Each MLP had one hidden layer with sigmoidal transfer functions and a single output neuron with linear transfer function. The number of inputs were defined by the dimensionality of the feature space (i.e 192 for the full channel-set).

The number of hidden neurons in each MLP was optimized with cross-validation. A grid search on a range between one and 20 hidden neurons per DoF have shown that the performance did not increase with more than three neurons and decreases when using more than eight neurons. Similar results were also reported by other studies ([9], [10]). Thus we fixed the number of hidden neurons to three per DoF.

The MLPs were trained with the Levenberg-Marquardt back-propagation algorithm. All MLP training was implemented with the Matlab neural network toolbox. In previous studies where MLP were applied with a high number of features, the dimensionality of the feature-space (and thus the number of network inputs) was reduced using principal component analysis (PCA) [10], [19]. The number PCA components was defined by a threshold on the fraction of variance captured by those components. This can speed up the training of the MLPs but leads to a reduced performance. For a fair comparison with the other methods no dimensionality reduction was applied in this study.

4) Kernel Ridge Regression (KRR): Another simple but powerful non-linear regression method is *kernel ridge regression*. In KRR the same error function as in LR is minimized. The decisive difference to LR is that the error function is not minimized in the input space of the data. Instead the data in X is mapped through a (potentially non-linear) mapping ϕ into a *kernel feature space*. Applying the *kernel trick* [20], [21], [22], [23], [24] this mapping does not have to be computed explicitly. The kernel trick is based on a *kernel function* $k(.,.)$ that takes two data points as arguments and computes the inner product $\langle .,.\rangle_\phi$ in the kernel feature space

$$k(x(i), x(j)) = \langle \phi(x(i)), \phi(x(j))\rangle_\phi. \quad (6)$$

In this study we used a Gaussian kernel function

$$k(x(i), x(j)) = e^{-(x(i)-x(j))^2/\sigma}, \quad (7)$$

where σ is the width of the Gaussian kernel function. Given a fixed data set $X \in \mathbb{R}^{C\times T} = [x(t_1), x(t_2), \ldots, x(t_T)]$ the kernel function is evaluated for each pair $x(t=i)$, $x(t=j)$ of data points; the output of the kernel function $k(x(t=i), x(t=j))$ is then stored in the K_{ij}th entry of the *kernel matrix K*. The essence of the kernel trick is that one can express the prediction of the target labels $\hat{y}(T+1)$ as a linear combination α of the similarity in kernel feature space between the new data point $x(T+1)$ and all training data points $x(1), x(2), \ldots, x(T)$

$$\hat{y}(T+1) = \sum_{t=1}^{T} k(x(T+1), x(t))\alpha_t. \quad (8)$$

The so called *dual coefficients* α can be computed by inverting the kernel matrix and multiplying each column with the respective label

$$\alpha = (K + I\lambda)^{-1}y, \quad (9)$$

where I denotes a $T \times T$ identity matrix and λ is a regularization constant. For a detailed review of kernel ridge regression see e.g. [17], [25].

The hyper-parameters σ and λ have to be optimized using appropriate model selection techniques. We used a grid search in the inner fold of a nested cross-validation to find optimal parameters (sec. II-E).

E. Cross-validation

To evaluate performance, five-fold cross-validation was applied. The folds were formed by entire runs. This was done in order to keep training and test set not only disjoint but as independent as possible [26] and to guarantee a balanced appearance of movements within both sets.

The performance was in all cases evaluated on test sets including all trajectory types. Training was usually also done with all trajectory types; only the results shown in figure 8 were based on training with subsets of trajectories.

As a performance metric we used the r-square value [27]:

$$r^2 = 1 - \frac{\sum_d \mathrm{Var}(y^d - \hat{y}^d)}{\sum_d \mathrm{Var}(y^d)} \quad (10)$$

where y^d is the wrist deflection angle of the dth DoF, measured by the motion tracking system, and $\hat{y}^d$ its estimate predicted by the models. The numerator is the mean squared error, which is normalized by the variance of the correct labels in the denominator. Thus, the r-square value is not influenced by the numerical range of the labels. The maximal r-square value at perfect estimation is one. Note that also negative r-square values are possible for estimation errors larger than the variance of the targets.

For methods with parameters that have to be optimized, a nested cross-validation was applied. I.e. with the training set of each fold, a second (inner) cross-validation was done to determine the performance for a certain parameter configuration. This inner cross-validation was repeated for a number of parameter configurations and the best configuration was used to train the algorithm for the outer cross-validation [21], [26]. The reported performance was measured on the test sets of the outer cross-validation, which was not used to determine the parameters. Simply repeating a normal cross-validation with different parameter settings would lead to a

wrong performance estimation, since the parameters would over-fit to the test data sets.

A typical session with 15 runs contained 14700 feature samples whereof 11760 were used in each outer fold for training and parameter optimization and 2940 for testing. For the investigations in section III-D the training sets where reduced while the testing-sets were kept unchanged.

F. Amount of training data

All presented methods need data to learn the relationship between EMG features $x(t)$ and labels $y(t)$. For a clinical application the amount of training data required for calibrating the controller is an important factor because it determines the time to fit the prosthesis. To the best of our knowledge it was never explored in a systematical way how much data is needed for a proper model fitting. The influence of the amount of training data was investigated in two ways. First, by decreasing the training data set of each fold within the cross-validation by entire runs. Second by removing training-trials corresponding to certain trajectory-types within each run by defining the following subsets:

a) all trajectories (20 trials per run)
b) all radial trajectories (22.5° steps, 16 trials per run)
c) half of radial trajectories (45° steps, 8 trials per run)
d) quarter of radial trajectories (90° steps, 4 trials per run)
e) all circular trajectories (4 trials per run)

Both ways were combined and for a fair comparison the total number of training samples was logged. The aim of this investigation was to assess if it is better to reduce the density of combining the DoFs or to reduce the number of repetitions if the time for collecting training data is limited.

If the feature space is also linear with respect to the DoFs (i.e., if the features sum when activating more than one DoF at a time) we would expect that it is not necessary to train with all trajectories. Conversely, if this linearity does not occur, eliminating trajectories would negatively impact the performance.

III. RESULTS

A. Effect of feature transformation

Figure 3 illustrates the linearization of the feature space and the impact on the estimation by LR. Since it is impossible to visualize the relationship between the labels and the feature space in full dimension, the features were averaged over all channels: $x_{intensity}(t) = \frac{1}{C} \sum_{c=1}^{C} x_c(t)$. Although this "feature intensity" does not contain enough information for the regression task, it can give insights to the complexity of the underlying relationship. The top row (a-c) illustrates the relationship between wrist inclination r and EMG feature intensity. Several trials of the radial trajectories are plotted. The x-axis shows the distance from center position, the y-axis shows feature intensity, and different target directions are distinguished by different colors. The curves are obtained by polynomial fitting with a model complexity limited to third order.

Prediction with variance features: Plot (a) in figure 3 illustrates the nonlinear relationship between EMG variance and wrist inclination. When estimating the labels with LR, the predicted wrist angles cannot be modeled well, as depicted in figure 3d. For wrist angles close to the origin, the predicted angle is *underestimated* while at wrist angles far from the origin, the predicted angles tend to be *overestimated*.

Prediction with rms features: The panels in the middle column of figure 3 show data and results for the square root of the variance features. Panel b illustrates that the non-linearity between wrist inclination and EMG features is not as pronounced as in the case of the variance features in panel a. This leads to a better prediction, as visualized in figure 3e.

Prediction with log-var features: The results obtained when taking the log of the EMG variance are depicted in the panels in the right column of figure. 3. In contrast to the other two features, the relationship between wrist angles and EMG log-var is almost linear, as illustrated in panel c. This leads to a significantly better prediction with less under- or overestimation at small or large targets as shown qualitatively in figure 3f.

B. Cross-validation results

The effect of linearization is also seen in the cross-validation performance measured by the r-square value (Fig. 4).To check for statistical significance 3-way ANOVA ($p = 0.05$) was performed. The three factors were regressor, feature and subject. Subjects 8 and 9 had large negative r-square values (at LR with var, < -10) and were excluded from the test as outliers. The full model ANOVA (with all two-way interactions and the three-way interaction) revealed no significant three-way interaction ($p = 0.98$), and two-way interactions including subject ($p = 0.09$ with regressor and p=0.11 with feature, respectively). These interaction terms were pooled and a three-way ANOVA with only the two-way interaction between regressor and feature was performed, from which significant interaction was detected ($p < 10^{-3}$).

Subsequently, compartmentalized two-way ANOVA tests were performed by fixing the level of one of the two interacting factors. When the level of regressor was fixed at LR, ME, MLP, and KRR, the 2-way ANOVA tests found that feature was significant ($p < 10^{-3}$, $p < 10^{-3}$, $p = 0.026$, and $p = 0.031$ respectively), regardless of the regressor. Post-hoc Tukey-Kramer tests showed that var was always significantly worse than log-var in all cases, while rms was never significantly different from log-Var. Further, for the two non-linear methods, rms was not significantly better from var, but log-var was (Fig. 4a).

When the level of features was fixed at var, rms, and log-var, the 2-way ANOVA tests found that regressor was not significant for log-var (p=0.14), while it was significant for var and rms($p < 10^{-3}$ for both cases). Post-hoc Tukey-Kramer tests showed that, for the var feature, LR was significantly worse than the other three regressors, ME was significant worse than MLP and KRR, while there was no significant difference between MLP and KRR. For the rms feature, LR was significantly different from MLP and KRR, and no other
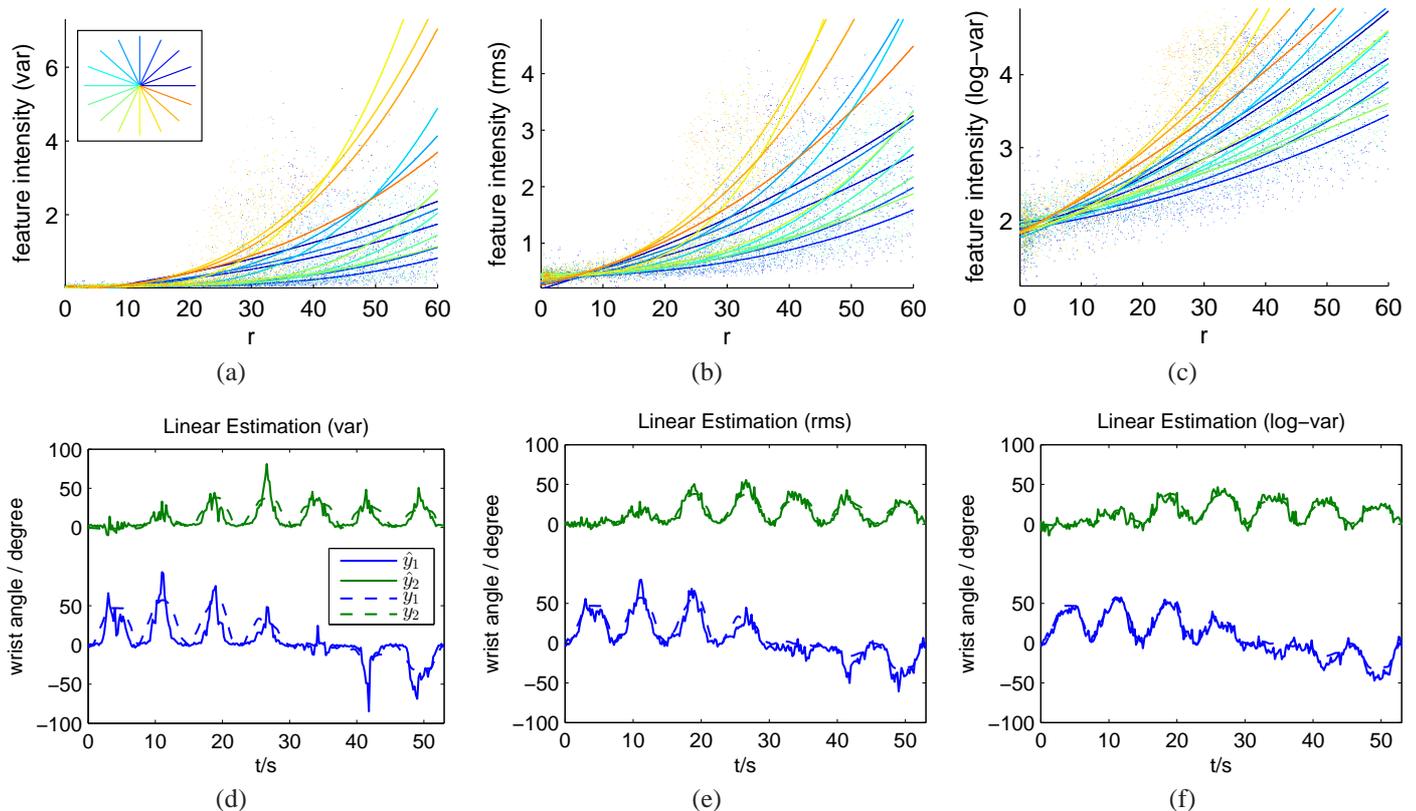
Fig. 3. Visualization of feature intensity (features averaged over all channels) vs. wrist inclination r for radial trajectories in polar coordinates (a-c). Each line was obtained by polynomial fitting of the intensities for one direction of wrist inclination θ. For this illustration, only radial trajectories were used and the color of each curve indicates the direction of the trajectory as illustrated in the legend in panel a. The lower panels (d-f) show an example of the estimations $\hat{y}$ by linear regression (solid lines) and the true labels y (dashed lines) for all features. For the log-var feature the relationship between wrist inclination and feature intensity is almost linear (a) which results in the best estimation (f).

significant different pairs were found. For log-var features, no significant differences were found among all the regressors.

All in all the linear methods performed significantly worse than nonlinear methods with variance features. It is very clear that here the feature transformations had the largest effect. But even for the non-linear methods MLP and KRR the log-transformation led to a small but significant improvement. Because for log-var features all regressors perform equally well, throughout the rest of this study all results are based on the log-var feature.

For the subject with congenital deficiency, the effect of feature transformation was similar to able-bodied subjects (Fig. 5). With the log-var feature, the r-square value was 0.7 to 0.8, which is almost as good as the average able-bodied subjects.

C. Contra-lateral training

In order to assess the ability of all methods to be applied to uni-lateral amputees, we trained each model with the contra-lateral labels and tested with the ipsi-lateral labels (available for five subjects, Fig. 6). The performance decreased from approximately 0.8 - 0.9 (ipsi-lateral training, upper panel) to 0.6 - 0.7 (contra-lateral training, lower panel) for four subjects and to 0.3 - 0.4 for one subject. This is to be compared to the reproducibility of the left and right hand mirror movements

(black lines in Fig. 6, lower panel). Evidently the performance drop is largely a result of the inability of the subject to perform exact mirror movements.

D. Impact of reduced training data

For a clinical application a method should be calibrated with as few training data as possible and generalize from a small amount of training data to as many possible motor actions as possible. We quantified the generalization performance of all methods by successively reducing the amount of training data and the regions in data space from which training data was obtained. These results are based on the six subjects for whom 15 runs are available.

1) Reduction by runs: As expected, performance decreases when the amount of available training data is reduced (Fig. 7). KRR and ME and LR are similarly robust to a reduction in data set size, whereas the MLP does require a large set of examples.

2) Reduction by runs and trials per run: The cross-validation performance of a combined reduction of the number of training-runs and the types of motor actions performed within each run are shown for the ME in figure 8 (similar results were obtained with the other regressors). The performance depends mainly on the amount of training data. When enough sample are used (e.g. more than 1500) the type of

(a) factor regressor fixed

(b) factor feature fixed
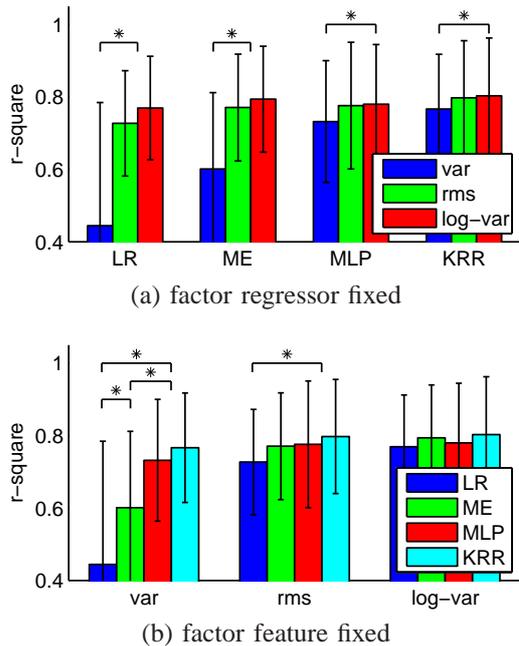
Fig. 4. Mean cross-validation performance of ipsi-lateral training for all features and regressors. The error bars indicate standard deviation and the lines with stars above the bars mark cases that are significantly different ($p = 0.05$). In cases when the line ends in between two bars both are meant.
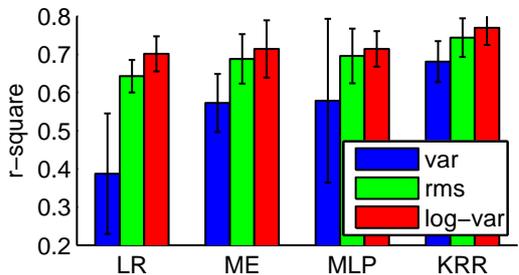
Fig. 5. Cross-validation performance for a subject with congenital deficiency, trained with contra-lateral motion data. Error bars indicate inter-fold standard deviation. The effect of feature transformations is the same as for able-bodied subjects: rms lead to better and log-var to best results for all regressors and the effect was stronger for linear methods.

training trajectories had no strong influence. Even if only single DoF were active in training (1/4 radial trajectories), the regressors performed still very good on the testing data which included many combined movements. This shows that the algorithm is able to generalize also to regions of y for which no training data was provided. The models can generalize from a small set of co-activations to various mixtures of independently combined DoFs. This indicates that the feature space is also linear with respect to the DoFs.

3) Processing time: As an indication of the computational load of the algorithms the processing time for training was measured (Fig. 9a). All processing was done in MATLAB 64 bit, running on a system with a 2.67 GHz processor and 8 GB of memory. Evidently the LR is exceedingly fast (100 ms with all data included) thus permitting potential real-time adaptation. In contrast, the MLP can take substantial amount

of time for training (up to 5 min).

The computational cost for applying the methods is shown in figure 9b. The time to apply LR, ME and MLP does not depend on the amount of training data and is very fast (LR:5 ms, ME:40 ms and MLP:100 ms for the entire test data of 3000 samples or 600 seconds of EMG data). KRR is a non-parametric model and needs to access all training data samples during testing. Testing time for KRR thus increases with increased training-set and reaches around 2.5 s for the largest training set (10ms per sample). This, together with the memory requirement KRR may make embedded processing prohibitive.

E. Reduced channel-set

For this study data was recorded with 192 channels. Cost and power consumption will set limits on the number of channels that can be used in a clinical prosthetic system. Therefore we investigate the performance of the algorithms with reduced sets of 96, 48, 24, 16, 12 and 6 channels (Fig. 10a) with regular spacing (Fig. 10c). For all methods the performance increases with increasing number of channels and saturates at half of the available channels. When the number of channels is reduced below approximately 12 to 16, the performance drops abruptly. KRR performs best in all cases and achieves an r-square value of 0.8 with only 12 monopolar channels. However, the differences between the methods are rather small, e.g. the computational cheaper method ME has with the same number of channels still a median performance of 0.73.

Similar results were obtained for the subject with congenital deficiency (Fig. 10 b). The number of channels differs from figure 10c because the electrode array had to be cut to fit the size of the residual limb without overlap. Again, KRR performance was best and a drop in performance below a certain number of channels was observed (22 channels in this case).
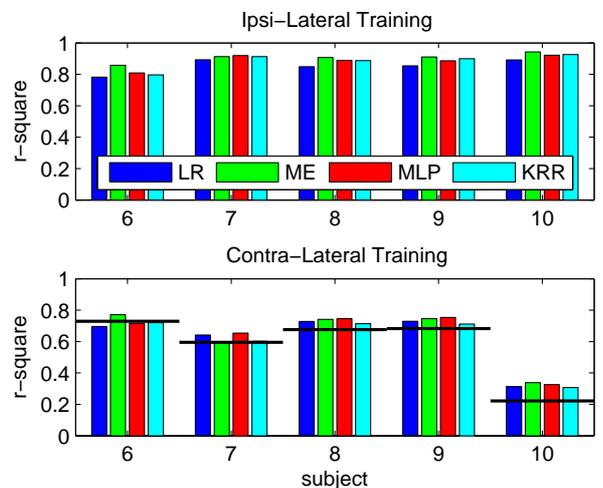
Fig. 6. Cross-validation performance for ipsi- (u.) and contra-lateral training (d.); the decrease in performance form ipsi- to contra-lateral training is approximately proportional to the ability of the subjects to copy the movements from left and right wrist as indicated by the black horizontal lines.
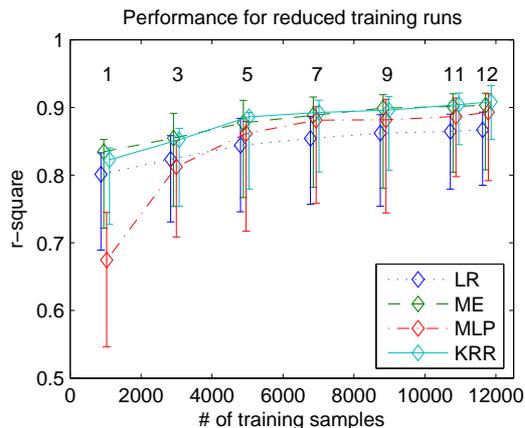
Fig. 7. Cross-validation results with reduced number of training runs. Curves indicate median across subjects and whiskers show 25/75 percentiles. The numbers above the curves indicate the number of runs used. 1000 feature-samples correspond to 200 s of data. ME, KRR and LR are less influenced by the reduction of training data as compared to the MLP.
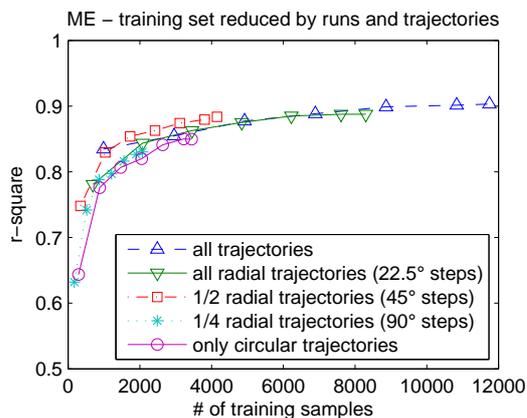
Fig. 8. Cross-validation performance for training sets reduced by number of training runs and training trajectories. Shown is the median across subjects for the ME. Performance increases with increasing number of training samples nearly independently of the specific choice of trajectories.

IV. DISCUSSION

This study presents a systematic comparison of EMG features and control algorithms for simultaneous and proportional control of a hand prostheses with multiple DoFs. The evaluation scenarios in which the methods were compared, have targeted aspects that are important for clinical applications.

A. Feature representation

Previous studies have often used variance to capture EMG activity [12], [11], [10]. However, power (variance) of EMG increases disproportionately as force increases to achieve extreme wrist inclinations. A simple non-linear transformation (square root or logarithm) can account for this non linearity and thus improves performance for all methods tested. This is particularly true for the linear methods (LR and ME) which attained with this simple modification a performance closer to the more complex non-linear algorithms. Opposite direction of movement engages different muscles. This leads to an additional non-linearity of the problem as stated here

(a) training times   (b) testing time
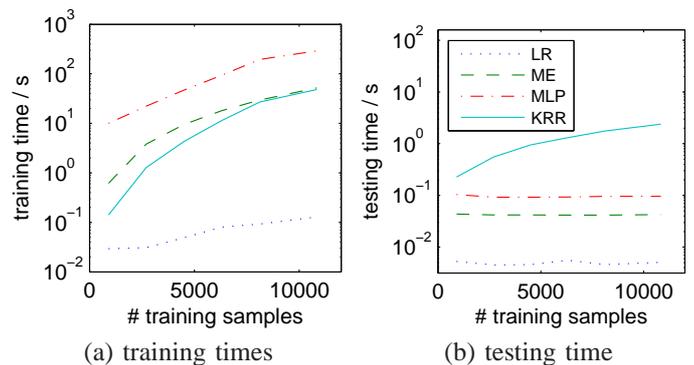
Fig. 9. Training and testing time as functions of the training set size

(where direction is indicated by changing sign). The goal of the mixture of expert technique proposed here was to break the linear trajectory into two regressors, each specializing into positive or negative displacements. With this modification the remaining non-linearity is largely addressed and performance increases to levels comparable to state-of-the-art non-linear regression algorithms (ME, see section II-D2).

B. Clinical applicability

1) Amount of calibration data: In clinical practice it is desirable that the controller requires as little calibration data as possible. Importantly, it should be able to generalize to movements for which exhaustive training data is not available. This is particularly important for simultaneous proportional control with many DoFs, because the amount of data and recording time increases exponentially if the space of movement is to be uniformly and densely sampled. We found that dense sampling of all movement directions is not as important as overall number of training samples. This indicates that the feature space is also linear with respect to the DoFs. In practice this means that not all possible combinations of DoFs are required for calibration, which can reduce the complexity of the training protocol and thus alleviate the effort for the user. With approximately 2000 feature samples (less than seven minutes training data) the ME algorithm performs already reasonably well. Increasing the recording time beyond this point provides diminishing returns. With the current implementation of MLP about 5000 training samples (more than 15 minutes) are needed to avoid a substantial drop in performance. However, there exist techniques that could increase the performance for small training sets [28].

2) Computational costs: The current clinical standard for fitting the prosthetic device involves a computer to visualize the EMG signals and configure the parameter settings. Thus, the computational cost of training is of lesser concern. However, future devices may aim to adaptively calibrate the device in real-time in which case efficient learning algorithms are a key requirement. The training times for LR is negligible and the algorithm is readily converted into a real-time setting. For the full data set ME and KRR needed almost a minute. But assuming a reduced data set of 2000 samples which would still lead to a reasonable performance, ME and KRR could be
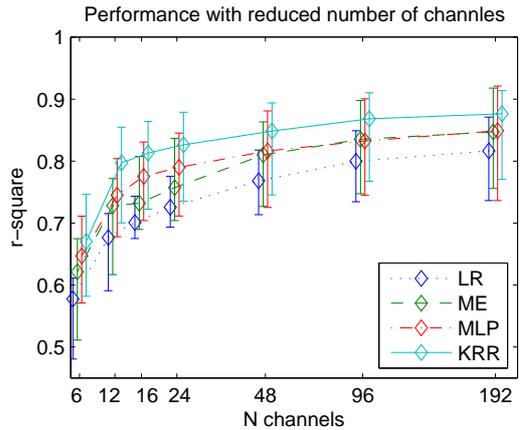
(a) performance of able bodied subjects
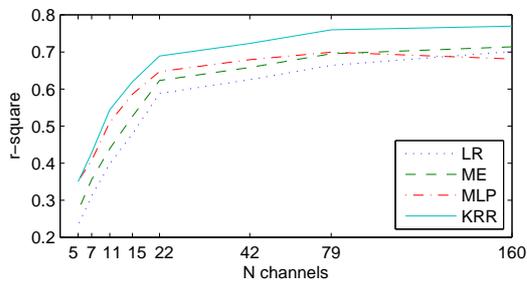
(b) performance for subject with congenital deficiency
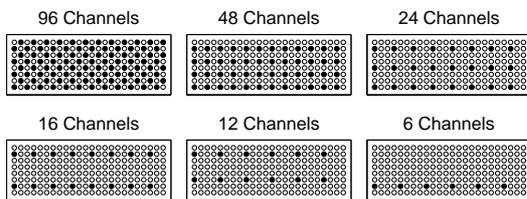
(c) channel-sets

Fig. 10. Reduced channel-sets: (a) Cross-validation performance with reduced channel-sets for able-bodied subjects (median and 25/75 percentiles across subjects). The performance decreased with the decrease in the number of channels and dropped abruptly when fewer than 12 to 16 channels were used. (b) Cross-validation performance with reduced channel-sets for the subject with congenital deficiency. The results were similar to those from the able-bodied subjects, but the performance drop occurred when fewer than 22 channels were used. (c) The definition of the channel subsets

trained in less than 5 seconds. To train MLP with 5000 samples requires approximately 60 seconds which would preclude real-time adaptation. This could perhaps be mitigated by reducing the number of channels and more efficient implementations.

The computational costs during execution is critical because they need to fulfill real-time requirements on an embedded system with little computational power. The time to evaluate one test sample must not exceed a few milliseconds. Therefore the processing times measured on the machine described in section III-D3 can only give a rough assessment. The processing for LR consists only of a single matrix-vector multiplication and is negligible. ME and MLP consist of several matrix-vector multiplications and evaluations of sigmoid functions. This is also possible on a relatively simple system. The application of KRR involves evaluating the kernel-function

for the test sample with all training data points and a matrix-vector multiplication with the kernel matrix. Since the kernel matrix is growing quadratically with the number of training points, the processing costs and the memory requirements are very high already for medium training data sets. (e.g. for 2000 data points the kernel matrix has 4×10^6) entries. This makes the use of KRR prohibitive with currently available prostheses hardware. Note that there exist techniques to reduce the memory requirements and computational costs of KRR (see e.g. [29],[30],[31]).

3) Number of channels: Because of costs, power consumption and reliability, the number of electrodes for a clinical application should be as small as possible. Reducing the number of channels leads to a reduced performance for all investigated methods. But even with 12 channels the regressors were still able to estimate the wrist position with an r-square value of 0.7 to 0.8. For the subject with congenital deficiency, 22 channels were sufficient to reach an r-square value of 0.6 to 0.7. The number of needed channels may vary significantly for subjects with limb deficiency depending on the individual anatomy and capabilities. The channels were selected arbitrarily with a regular spacing. It is expected that with automatic channel-selection methods a higher performance can be reached with even fewer channels. This is important particularly for potential users of myo-prostheses.

4) Transfer to amputees / training strategy: Contra-lateral training is one possibility to apply the methods to uni-lateral amputees. The performance in this case depends on the amount of residual muscles, the ability of the user to execute the contractions with his disabled side and the ability to copy the movements from the intact side. The last factor has been evaluated in this study with five able-bodied subjects. Our results suggest that even for able-bodied subjects there is a large variability in how precise bilateral mirrored movements can be executed. These results indicate that user training and feedback will be essential for a successful application of regression techniques for a simultaneous proportional control of multiple DoF prostheses. Given good mirror movement performance, all other results of this study apply to the case of contra-lateral training. This was shown for one subject with congenital limb deficiency, whose performance was only slightly below that of able-bodied subjects. Moreover, the main findings of our study, including the positive effect of the feature transformations, were valid also for this subject. This indicates that our findings may transfer to potential users of myoelectric prostheses and emphasizes the relevance of this work.

The experiments in this study are based on 2 DoF, namely, flexion/extension and radial/ulnar deviation of the wrist. The latter is not available in current prosthesis hardware. The muscles for those movements are located close to the skin leading to good EMG signals and less problems due to skin-muscle-shifts are expected compared to pronation/supination. These problems might be a minor issue when applied to amputees because of different anatomy. However, the control signals from radial/ulnar deviation can also be used to control the rotation unit of the prosthesis if this leads to more stable results.

C. Linear vs. non-linear methods

Performance comparisons indicate that linear methods can achieve very good results comparable to state-of-the-art non-linear regression algorithms. In fact, when using an appropriate EMG feature representation and a proper regularization the results with ME are almost indistinguishable from those of non-linear methods. A major advantage of linear methods is the dramatically reduced computational demand for training and evaluation; both LR and the ME model are convex problems that can be solved very efficiently. Moreover, linear methods are less prone to over-fitting than non-linear methods. LR and ME can be easily realized on a very simple and cheap micro-controller with little power consumption and are readily modified for real-time adaptation. In contrast to linear methods non-parametric models like KRR suffer from large memory requirements and significantly longer evaluation times for large calibration data sets. Parametric non-linear models such as artificial neural networks on the other hand do not require as much memory and are relatively fast during evaluation, but training can be slow and they required longer calibration sessions.

V. Conclusion

We systematically compared state-of-the-art regression techniques for independent simultaneous and proportional myoelectric control. Linear and non-linear methods were compared under carefully designed experimental paradigms in order to assess their performance in terms of accuracy and robustness targeting clinical requirements.

We identified that a logarithmic transformation of the well established variance feature linearized the relationship between EMG and wrist angles. This allows to apply very simple and computationally cheap linear methods.

The models generalized very well to DoF-combinations for which no training data was provided. This indicates that the log-var feature space is also linear with respect to DoFs and that it is not necessary to record training data for all possible combinations of DoFs.

An additional linearization was achieved by separating movement in opposing directions, which is motivated by the fact that opposing movements are controlled by different sets of muscles. The resulting ME algorithm represents a promising candidate for the next generation of prosthetic devices. If adequately regularized, it performs similarly to, or better than more complex non-linear methods, even when only little training data is available. It is superior in terms of computational cost during both calibration and prediction phase and can be implemented on a very simple hardware. By including one subject with congenital limb deficiency we have shown that our findings transfer well to potential users of myo-prostheses. Future studies will explore the case of co-adaptive learning strategies [32].

References

[1] J. M. Miguelez, "Clinical experiences with the michelangelo hand, a four-year review," in *MyoElectric Controls/Powered Prosthetics Symposium, Fredericton, New Brunswick, Canada*, 2011.

[2] V. D. Niet Otrand, O, A. Heleen, R. M. Bongers, H. Bouwsema, and C. K. V. D. Sluis, "The i-LIMB hand and the DMC plus hand compared: A case report," *Prosthetics and Orthotics International*, vol. 34, no. 2, pp. 216–220, Jan. 2010.

[3] E. Scheme and K. Englehart, "Electromyogram pattern recognition for control of powered upper-limb prostheses: state of the art and challenges for clinical use," *Journal of Rehabilitation Research and Development*, vol. 48, no. 6, pp. 643–659, Sep. 2011.

[4] A. Fougner, E. Scheme, A. Chan, K. Englehart, and O. Stavdahl, "Resolving the limb position effect in myoelectric pattern recognition," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 9, no. 6, pp. 644–651, dec 2011.

[5] J. Hahne, B. Graimann, and K.-R. Müller, "Spatial filtering for robust myoelectric control," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 5, pp. 1436–1443, May 2012.

[6] Y. Geng, D. Tao, L. Chen, and G. Li, "Recognition of combined arm motions using support vector machine," in *Informatics in Control, Automation and Robotics*, Jan. 2012, no. 133, pp. 807–814.

[7] A. Young, L. Smith, E. Rouse, and L. Hargrove, "Classification of simultaneous movements using surface EMG pattern recognition," *IEEE Transactions on Bio-Medical Engineering*, vol. 60, no. 5, pp. 1250–1258, May 2013.

[8] N. Jiang, S. Dosen, K.-R. Muller, and D. Farina, "Myoelectric control of artificial limbs; is there a need to change focus?" *IEEE Signal Processing Magazine*, vol. 29, no. 5, pp. 152–150, Sep. 2012.

[9] J. L. Nielsen, S. Holmgaard, N. Jiang, K. B. Englehart, D. Farina, and P. A. Parker, "Simultaneous and proportional force estimation for multifunction myoelectric prostheses using mirrored bilateral training," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 3, pp. 681–688, Mar. 2011.

[10] S. Muceli and D. Farina, "Simultaneous and proportional estimation of handkinematics from emg during mirrored movements at multipledegrees-of-freedom," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 20, no. 3, pp. 371–378, May 2012.

[11] N. Jiang, J. L. Vest-Nielsen, S. Muceli, and D. Farina, "EMG-based simultaneous and proportional estimation of wrist/hand dynamics in uni-lateral trans-radial amputees," *Journal of NeuroEngineering and Rehabilitation*, vol. 9, no. 1, p. 42, Jun. 2012.

[12] N. Jiang, K. B. Englehart, and P. A. Parker, "Extracting simultaneous and proportional neural control information for multiple-DOF prostheses from the surface electromyographic signal," *IEEE Transactions on Bio-Medical Engineering*, vol. 56, no. 4, pp. 1070–1080, Apr. 2009.

[13] N. Jiang, S. Muceli, B. Graimann, and D. Farina, "Effect of arm position on the prediction of kinematics from emg in amputees," *Medical and Biological Engineering and Computing*, vol. 51, no. 1-2, pp. 143–151, Feb. 2013.

[14] K. Englehart and B. Hudgins, "A robust, real-time control scheme for multifunction myoelectric control," *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 7, pp. 848–854, Jul. 2003.

[15] T. R. Farrell and R. F. Weir, "The optimal controller delay for myoelectric prostheses," *IEEE Trans Neural Syst Rehabil Eng*, vol. 15, no. 1, pp. 111–118, Mar. 2007.

[16] C. F. Gauß, "Theoria motus corporum coelestium in sectionibus conicis solem ambientium," *Göttingen*, 1809.

[17] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer, Oct 2007.

[18] L. C. Parra, C. D. Spence, A. D. Gerson, and P. Sajda, "Recipes for the linear analysis of EEG," *NeuroImage*, vol. 28, no. 2, pp. 326–341, Nov. 2005.

[19] J. Hahne, H. Rehbaum, F. Biessmann, F. Meinecke, K.-R. Müller, N. Jiang, D. Farina, and L. Parra, "Simultaneous and proportional control of 2D wrist movements with myoelectric signals," in *2012 IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, Sep. 2012.

[20] A. Aizerman, E. Braverman, and L. Rozonoer, "Theoretical foundations of the potential function method in pattern recognition learning," *Automation and Remote Control*, vol. 25, pp. 821–837, Jan 1964.

[21] K.-R. Müller, S. Mika, G. Ratsch, K. Tsuda, and B. B. Schölkopf, "An introduction to kernel-based learning algorithms," *IEEE Transactions on Neural Networks*, vol. 12, no. 2, pp. 181–201, Jan 2001.

[22] B. Schölkopf, A. Smola, and K.-R. Müller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural computation*, vol. 10, no. 5, pp. 1299–1319, Jul. 1998.

[23] K.-R. Müller, C. W. Anderson, and G. E. Birch, "Linear and non-linear methods for brain-computer interfaces," *IEEE transactions on neural systems and rehabilitation engineering*, vol. 11, no. 2, pp. 165–169, Jun. 2003.

[24] B. Schölkopf and A. J. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. MIT Press, 2002.

[25] J. Shawe-Taylor and N. Cristianini, *Kernel methods for pattern analysis*. Cambridge University Press, 2004.

[26] S. Lemm, B. Blankertz, T. Dickhaus, and K.-R. Müller, "Introduction to machine learning for brain imaging," *Neuroimage*, vol. 56, no. 2, pp. 387–399, May 2011.

[27] A. d. Avella, A. Portone, L. Fernandez, and F. Lacquaniti, "Control of fast-reaching movements by muscle synergy combinations," *Journal of Neuroscience*, vol. 26, no. 30, pp. 7791–7810, Jul. 2006.

[28] G. Montavon, G. B. Orr, and K.-R. Müller, Eds., *Neural Networks: Tricks of the Trade, Reloaded*, 2nd ed., ser. Lecture Notes in Computer Science (LNCS). Springer, 2012, vol. 7700.

[29] L. Shpigelman, K. Crammer, R. Paz, E. Vaadia, and Y. Singer, "A temporal kernel-based model for tracking hand-movements from neural activities," *Advances in Neural Information Processing Systems*, 2005.

[30] M. Seeger, "Cross-validation optimization for large scale structured classification kernel methods," *Journal of Machine Learning Research*, vol. 9, pp. 1147–1178, 2008.

[31] A. Rahimi and B. Recht, "Weighted Sums of Random Kitchen Sinks: Replacing minimization with randomization in learning," in *NIPS*, 2008, pp. 1313–1320.

[32] C. Vidaurre, C. Sannelli, K.-R. Müller, and B. Blankertz, "Machine-learning based co-adaptive calibration," *Neural computation*, vol. 23, no. 3, pp. 791–816, Mar. 2011.

Janne Hahne received the Diploma degree in electrical engineering in 2008 from Berlin Institute of Technology (TU-Berlin), Germany. In 2009 he became member of the Department of Strategic Technology Management at Otto Bock HC, Duderstadt, Germany, where he was working on control systems for myoelectric hand prostheses. In 2010 he joint the Machine Learning Laboratory at TU-Berlin, where he is currently working toward the Ph.D. degree. His research interests include myoelectric control, biomedical signal processing and machine learning.

Felix Bießmann obtained a BSc in Cognitive Science in 2005 from the University of Osnabrück, a MSc in Neuroscience at the International Max-Planck Research School, Tübingen, and a PhD in Machine Learning from Berlin Institute of Technology. His research interests include statistical learning methods for neuroscientific data with a focus on multimodal data. Since 2013 he is Assistent Professor at Korea University, Seoul.

Ning Jiang Ning Jiang (S'02-M'09) received the B.S. degree in electrical engineering from Xi'an Jiaotong University, Xi'an, China, in 1998, and the M.Sc. and Ph.D. degrees in engineering from the University of New Brunswick, Fredericton, NB, Canada, in 2004 and 2009, respectively. He was a Marie Curie Fellow at the Strategic Technology Management, Otto Bock Healthcare GmbH, Germany from 2010 to 2012, and is currently a research scientist with Department of Neurorehabilitation Engineering, University Medical Center Göttingen, Georg-August University, Göttingen, Germany. His research interests include signal processing of electromyography, advanced prosthetic control, neuromuscular modeling, and BCI for neurorehabilitation.

Hubertus Rehbaum (S'12) received his Dipl.-Ing. Degree in Electrical Engineering and Information Technology from the RWTH Aachen University in 2011. He is currently a Marie Curie Fellow and PhD candidate at the Department of NeuroRehabilitation Engineering, University Medical Center, University of Göttingen, Germany. His research interests include signal processing of electromyography, advanced prosthetic control, and adaptive control algorithms.

Frank C. Meinecke graduated in Physics at University of Potsdam in 2003 and received his Ph.D. degree at TU Berlin in 2011. From 2001 to 2007, he worked at the Intelligent Data Analysis and since 2007 at the Machine Learning Group at TU Berlin. His research is focused on the development of methods for time series analysis and matrix factorization with contributions in the fields of blind source separation, synchronization analysis, nonstationarities and multimodal integration. He uses the developed methods to tackle problems in various application domains, most notably in neuroscience and Brain-Computer-Interfacing.

Dario Farina (M'01-SM'09) received the M.Sc. degree in electronics engineering from Politecnico di Torino (PDT), Italy, in 1998, and the Ph.D. degrees in automatic control and computer science and in electronics and communications engineering from the Ecole Centrale de Nantes, France, and PDT, respectively, in 2002. During 2002-2004, he was Research Assistant Professor at PDT and in 2004-2008 Associate Professor in Biomedical Engineering at Aalborg University (AAU), Denmark. From 2008 to 2010 he was Full Professor in Motor Control and Biomedical Signal Processing and Head of the Research Group on Neural Engineering and Neurophysiology of Movement at AAU. In 2010 he was appointed Full Professor and Founding Chair of the Department of Neurorehabilitation Engineering at the University Medical Center Göttingen, Georg-August University, Germany, within the Bernstein Center for Computational Neuroscience. He is also the Chair for NeuroInformatics of the Bernstein Focus Neurotechnology Göttingen. He is an Associate Editor of Medical & Biological Engineering & Computing and the Journal of Electromyography and Kinesiology, as well as a member of the Editorial Board or reviewer for several international Journals. His research focuses on biomedical signal processing, neurorehabilitation technology, and neural control of movement. Dr. Farina is the President of the International Society of Electrophysiology and Kinesiology (ISEK) since 2012. He is the recipient of the 2010 IEEE Engineering in Medicine and Biology Society Early Career Achievement Award for his contributions to biomedical signal processing and to electrophysiology. He is an Associate Editor of IEEE Transactions on Biomedical Engineering.

Klaus-Robert Müller has been Professor for Computer Science at TU Berlin since 2006; at the same time he is directing the Bernstein Focus on Neurotechnology Berlin. Since 2012 he also holds a distinguished Professorship for Neurotechnology at Korea University, Seoul. He studied physics in Karlsruhe from 1984-89 and obtained his PhD in Computer Science at TU Karlsruhe in 1992. After a PostDoc at GMD-FIRST, Berlin, he was a Research Fellow at University of Tokyo from 1994-1995. From 1995 he built up the Intelligent Data Analysis (IDA) group at GMD-FIRST (later Fraunhofer FIRST) and directed it until 2008. 1999-2006 he was a Professor at University of Potsdam. In 1999, he was awarded the Olympus Prize by the German Pattern Recognition Society, DAGM and in 2006 he received the SEL Alcatel Communication Award. In 2012 he was elected to be a member of the German National Academy of Sciences - Leopoldina. His research interests are intelligent data analysis, machine learning, signal processing and Brain Computer Interfaces.

Lucas C. Parra is Professor of Biomedical Engineering at the City College of the City University of New York. Previously he was head of the adaptive image and signal processing group at Sarnoff Corporation (1997-2003) and member of the machine learning and the imaging departments at Siemens Corporate Research (1995-1997). His background is in machine learning, signal processing, and medical imaging. In recent years his interests have shifted to systems neuroscience. Current projects focus on non-invasive "reading and writing the brain", i.e. interpreting brain signals on a single-trial basis and electrically stimulating the brain to boost neuronal plasticity. Dr. Parra received his Ph.D. in Physics from the Ludwig-Maximilian University in Germany in 1996.