# On-line Differentiation Of Neuroelectric Activities: Algorithms And Applications.

M. Krauledat[1,2], B. Blankertz[1], G. Dornhege[1], M. Schröder[1], G. Curio[3] and K.-R. Müller[1,4]

[1]Fraunhofer FIRST.IDA, Kekuléstr. 7, 12 489 Berlin, Germany
[2]Technical University Berlin, Str. d. 17. Juni 135, 10 623 Berlin, Germany
[3]Dept. of Neurology, Campus Benjamin Franklin, Charité University Medicine Berlin, Hindenburgdamm 30, 12 203 Berlin, Germany
[4]University of Potsdam, August-Bebel-Str. 89, 14 482 Potsdam, Germany.

***Abstract*— This paper discusses machine learning methods and their application to Brain-Computer Interfacing. A particular focus is placed on linear classification methods which can be applied in the BCI context. Finally, we provide an overview of the Berlin-Brain Computer Interface (BBCI).**

## I. INTRODUCTION

Brain-Computer Interfacing is an interesting, active and highly interdisciplinary research topic ([2], [3], [4], [5]) at the interface between medicine, psychology, neurology, rehabilitation engineering, man-machine interaction, machine learning and signal processing. A BCI could, e.g., allow a paralyzed patient to convey her/his intentions to a computer application. From the perspective of man-machine interaction research, the communication channel from a healthy human's brain to a computer has not yet been subject to intensive exploration, however it has potential, e.g., to speed up reaction times, cf. [6] or to supply a better understanding of a human operator's mental states.

Classical BCI technology has been mainly relying on the adaptability of the human brain to biofeedback, i.e., a subject learns the mental states required to be understood by the machines, an endeavour that can take months until it reliably works [7], [8].

The Berlin Brain-Computer Interface (BBCI) pursues another objective in this respect, i.e., to impose the main load of the learning task on the 'learning machine', which also holds the potential of adapting to specific tasks and changing environments given that suitable machine learning (e.g. [9]) and adaptive signal processing (e.g. [10]) algorithms are used. Short training times, however, imply the challenge that only few data samples are available for learning to characterize the individual brain states to be distinguished. In particular when dealing with few samples of data (trials of the training session) in a high-dimensional feature space (multi-channel EEG, typically several features per channel), overfitting needs to be avoided. It is in this high dimensional – small sample statistics scenario where modern machine learning can prove its strength.

The studies were partly supported by the *Bundesministerium für Bildung und Forschung* (BMBF), FKZ 01IBB02A and FKZ 01IBB02B and the PASCAL Network of Excellence, EU # 506778. This paper is based on excerpts of [1].

The present paper introduces basic concepts of linear classification (for a discussion of nonlinear methods in the context of BCI, see [11], [9], [12], [13], [14]). Finally, we briefly describe our BBCI activities where some of the discussed machine learning ideas come to an application and conclude.

## II. LINEAR METHODS FOR CLASSIFICATION

In BCI research it is very common to use linear classifiers, but although linear classification already uses a very simple model, things *can* still go terribly wrong if the underlying assumptions do not hold, e.g. in the presence of outliers or strong noise which are situations very typically encountered in BCI data analysis. We will discuss these pitfalls and point out ways around them.

Let us first fix the notation and introduce the linear hyperplane classification model upon which we will rely mostly in the following (cf. Fig. 1, see e.g. [15]). In a BCI set-up we measure $k = 1 \dots K$ samples $\mathbf{x}_k$, where $\mathbf{x}$ are some appropriate feature vectors in $n$ dimensional space. In the training data we have a class label, e.g. $y_k \in \{-1, +1\}$ for each sample point $\mathbf{x}_k$. To obtain a linear hyperplane classifier

$$\mathbf{y} = \operatorname{sign}\left(\mathbf{w}^\top \mathbf{x} + b\right) \tag{1}$$

we need to estimate the normal vector of the hyperplane $\mathbf{w}$ and a threshold $b$ from the training data by some optimization technique [15]. On unseen data $\mathbf{x}$, i.e. in a BCI feedback session we compute the projection of the new data sample onto the direction of the normal $\mathbf{w}$ via Eq.(1), thus determining what class label $\mathbf{y}$ should be given to $\mathbf{x}$ according to our linear model.

### *A. Optimal linear classification: large margins versus Fisher's discriminant*

Linear methods assume a linear separability of the data. We will see in the following that the optimal separating hyperplane from last section maximizes the *minimal* margin (minmax). In contrast, Fisher's discriminant maximizes the *average* margin, i.e., the margin between the class means.

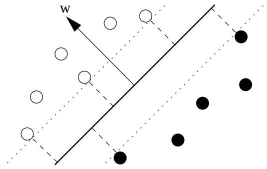


Fig. 1. Linear classifier and margins: A linear classifier is defined by a hyperplane's normal vector $\mathbf{w}$ and an offset $b$, i.e. the decision boundary is $\{\mathbf{x} \,|\, \mathbf{w}^\top \mathbf{x} + b = 0\}$ (thick line). Each of the two halfspaces defined by this hyperplane corresponds to one class, i.e. $f(\mathbf{x}) = \mathrm{sign}(\mathbf{w}^\top \mathbf{x} + b)$. The *margin* of a linear classifier is the minimal distance of any training point to the hyperplane. In this case it is the distance between the dotted lines and the thick line. From [9].

*1) Large margin classification:* For linearly separable data there is a vast number of possibilities to determine $(\mathbf{w}, b)$, that all classify correctly on the training set, however that vary in quality on the unseen data (test set). An advantage of the simple hyperplane classifier (in canonical form cf. [16]) is that literature (see e.g. [15], [16]) tells us how to select the *optimal* classifier $\mathbf{w}$ for unseen data: it is the classifier with the largest margin $\rho = 1/\|\mathbf{w}\|_2^2$, i.e. of minimal (euclidean) norm $\|\mathbf{w}\|_2$ [16] (see also Fig. 1). Linear Support Vector Machines (SVMs) realize the large margin by determining the normal vector $\mathbf{w}$ according to

$$\begin{aligned} \min_{\mathbf{w},b,\xi} \ & {}^1\!/_2 \|\mathbf{w}\|_2^2 + {}^C\!/_K \|\xi\|_1 \quad && \text{subject to} \\ & y_k(\mathbf{w}^\top \mathbf{x}_k + b) \geqslant 1 - \xi_k && \text{and} \\ & \xi_k \geqslant 0 && \text{for } k = 1, \ldots, K, \end{aligned} \tag{2}$$

where $\|\cdot\|_1$ denotes the $\ell_1$-norm: $\|\xi\|_1 = \sum |\xi_k|$. Here the elements of vector $\xi$ are slack variables and parameter $C$ controls the size of the margin vs. the complexity of the separation. While the user has not to care about the slack variables, it is essential to select an approppriate value for the free parameter $C$ for each specific data set. The process of choosing $C$ is called model selection, see e.g. [9]. One particular strength of SVMs is that they can be turned in nonlinear classifiers in an elegant and effective way (see e.g. [16], [9], [12]).

*2) Fisher's discriminant:* Fisher's discriminant computes the projection $\mathbf{w}$ differently. Under the restrictive assumption that the class distributions are (identically distributed) Gaussians of equal covariance, it can be shown to be Bayes optimal. The separability of the data is measured by two quantities: How far are the projected class means apart (should be large) and how big is the variance of the data in this direction (should be small). This can be achieved by maximizing the so-called Rayleigh coefficient of between and within class variance with respect to $\mathbf{w}$ [17], [18]. The slightly stronger assumptions have been fulfilled in several of our BCI experiments e.g. in [13], [14]. When the optimization to obtain (regularized) Fisher's discriminant is formulated as a mathematical programm, cf. [19], [9], [20], it resembles the SVM:

$$\begin{aligned} \min_{\mathbf{w},b,\xi} \ & {}^1\!/_2 \|\mathbf{w}\|_2^2 + {}^C\!/_K \|\xi\|_2^2 \quad \text{subject to} \\ & y_k(\mathbf{w}^\top \mathbf{x}_k + b) = 1 - \xi_k \quad \text{for } k = 1, \ldots, K. \end{aligned}$$

### *B. Some remarks about regularization and non-robust classifiers*

Linear classifiers are generally more robust than their non-linear counterparts, since they have only limited flexibility (less free parameters to tune) and are thus less prone to overfitting. Note however that in the presence of strong noise and outliers *even* linear systems can fail. In the cartoon of Fig.2 one can clearly observe that one outlier or strong noise event can change the decision surface drastically, if the influence of single data points on learning is not limited. Although this effect can yield strongly decreased classification results for linear learning machines, it can be even more devastating for nonlinear methods. A more formal way to control one's mistrust in the available training data, is to use regularization (e.g. [21], [22], [15]). Regularization helps to limit (a) the influence of outliers or strong noise (e.g. to avoid Fig.2 middle), (b) the complexity of the classifier (e.g. to avoid Fig.2 right) and (c) the raggedness of the decision surface (e.g. to avoid Fig.2 right). No matter whether linear or nonlinear methods are used, one should *always* regularize, – in particular for BCI data!

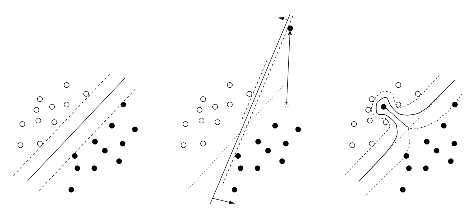

Fig. 2. The problem of finding a maximum margin "hyper-plane" on reliable data (left), data with an outlier (middle) and with a mislabeled pattern (right). The solid line shows the resulting decision line, whereas the dashed line marks the margin area. In the middle and on the right the original decision line is plotted with dots. Illustrated is the noise sensitivity: only one strong noise/outlier pattern can spoil the whole estimation of the decision line. From [23].

## III. The Berlin Brain-Computer Interface

The Berlin Brain-Computer Interface is driven by the idea to shift the main burden of the learning task from the human subject to the computer under the motto 'let the machines learn'. To this end, the machine learning methods presented in the previous sections are applied to EEG data from selected BBCI paradigms: selfpaced [13], [14] and imagined [24], [25], [26] experiments.

### *A. Self-paced Finger Tapping Experiments*

In preparation of motor tasks, a negative readiness potential precedes the actual execution. Using multi-channel EEG recordings it has been demonstrated that several brain areas contribute to this negative shift (cf. [27], [28]). In unilateral finger or hand movements the negative shift is mainly focussed on the frontal lobe in the area of the corresponding motor cortex, i.e., contralateral to the performing hand. Based on the laterality of the pre-movement potentials it is possible to discriminate multi-channel EEG recordings of upcoming left from right hand movements. Fig. 3 shows the lateralized

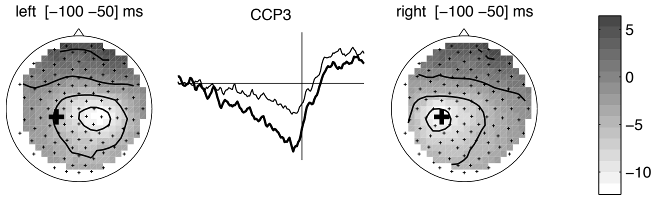


Fig. 3. The scalp plots show the topography of the electrical potentials prior to keypress with the left resp. right index finger. The plot in the middle depicts the event-related potential (ERP) for left (thin line) vs. right (thick line) index finger in the time interval -1000 to -500 ms relative to keypress at electrode position CCP3, which is marked by a bigger cross in the scalp plots. The contralateral negativation (lateralized readiness potential, LRP) is clearly observable. Approx. 260 trials per class have been averaged.

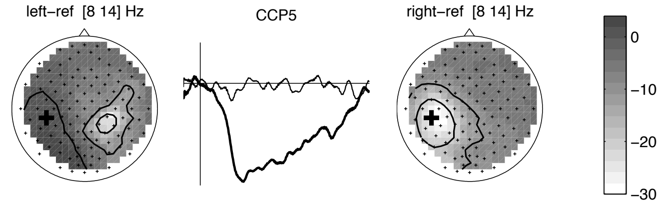


Fig. 4. This scalp plots show the topography of the band power in the frequency band 8–14 Hz relative to a reference period. The plot in the middle shows ERD curves (temporal evolution of band power) at channel CCP5 (mark by a bigger cross in the scalp plots) for left (thin line) and right (thick line) hand motor imagery. The contralateral attenuation of the $\mu$-rhythm during motor imagery is clearly observable. For details on ERD, see [30].

readiness potential during a 'self-paced' experiment, as it can be revealed here by averaging over 260 trials in one subject.

In the 'self-paced' experiments, subjects were sitting in a normal chair with fingers resting in the typing position at the computer keyboard. In a deliberate order and on their own free will (but instructed to keep a pace of approximately 2 seconds), they were pressing keys with their index and little fingers.

EEG data was recorded with 27 up to 120 electrodes, arranged in the positions of the extended 10-20 system, referenced to nasion and sampled at 1000 Hz. The data were downsampled to 100 Hz for further offline analyses. Surface EMG at both forearms was recorded to determine EMG onset. In addition, horizontal and vertical electrooculograms (EOG) were recorded to check for correlated eye movements.

In [6], it has been demonstrated that when analyzing LRP data offline with the methods detailed in the previous sections, classification accuracies of more than 90% can be reached at 110 ms before the keypress, i.e. a point in time where classification on EMG is still at chance level. These findings suggest that it is possible to use a BCI in time critical applications for an early classification and a rapid response.

Table I shows the classification results for one subject when comparing different machine learning methods. Clearly regularization and careful model selection are mandatory which can, e.g., be seen by comparing LDA and RLDA. Of course, regularization is of more importance the higher the dimensionality of features is. The reason of the very bad performance of $k$-NN is that the underlying Euclidean metric

TABLE I

TEST SET ERROR ($\pm$ STD) FOR CLASSIFICATION AT 110 MS BEFORE KEYSTROKE; ›MC‹ REFERS TO THE 56 CHANNELS OVER (SENSORI) MOTOR CORTEX, ›ALL‹ REFERS TO ALL 105 CHANNELS. THE ALGORITHMS IN QUESTION ARE LINEAR DISCRIMINANT ANALYSIS (LDA), REGULARIZED LINEAR DISCRIMINANT ANALYSIS (RLDA), LINEAR PROGRAMMING MACHINE (LPM), SUPPORT VECTOR MACHINE WITH GAUSSIAN RBF KERNEL (SVMRBF) AND $k$-NEAREST NEIGHBOR ($k$-NN).

| channels | LDA | RLDA | LPM | SVMrbf | $k$-NN |
|---|---|---|---|---|---|
| all | 16.9±1.3 | 8.4±0.6 | 7.7±0.6 | 8.6±0.6 | 28.4±0.9 |
| mc | 9.3±0.6 | 6.3±0.5 | 7.4±0.7 | 6.7±0.7 | 22.0±0.9 |

is not appropriate for the bad signal-to-noise ratio found in EEG trials. For further details refer to [13], [14]. Note that the accuracy of 90% can be maintained in recent realtime feedback experiments [29]. Here, as no trigger information is available beforehand, the classification decision is split into one classifier that decides whether a movement is being prepared and a second classifier that decides between left and right movement to come.

### B. Motor Imagery Experiments

During imagination of a movement, a lateralized attenuation of the $\mu$- and/or central $\beta$-rhythm can be observed localized in the corresponding motor and somatosensory cortex. Besides a usual spectral analysis, this effect can be visualized by plotting event-related desynchronization (ERD) curves [30] which show the temporal evolution of the band-power in a specified frequency band. A typical averaged ERD is shown in Fig. 4.

We performed experiments with 6 healthy subjects performing motor imagery. The subjects were sitting comfortably in a chair with their arms in a relaxed position on an arm rest. Two different sessions of data collection were provided: In both a target "L", "R" and "F" (for left, right hand and foot movement) is presented for the duration of 3.5 seconds to the subject on a computer screen. In the first session type this is done by visualizing the letter on the middle of the screen. In the second session type the left, right or lower triangle of a moving gray rhomb is colored red. For the whole length of this period, the subjects were instructed to imagine a sensorimotor sensation/movement in left hand, right hand resp. one foot. After stimulus presentation, the screen was blank for 1.5 to 2 seconds. In this manner, 35 trials per class per session were recorded. After 25 trials, there was a short break for relaxation. Four sessions (two of each training type) were performed. EEG data was recorded with 128 electrodes together with EMG from both arms and the involved foot, and EOG as described above.

An offline machine learning analysis of the "imagined"-experiments yields again high classification rates (up to 98.9% with the feature combination algorithm PROB [25], [24]), which predicts the feasibility of this paradigm for online feedback situations (see also [26]). In fact, our recent online experiments have confirmed this prediction by showing high bitrates for several subjects. These subjects were untrained and had to play video games like 'brain pong', 'basket' (a spelling task) and 'controlled 1-D cursor movement' [31]. Depending on the 'game' scenario the best subjects could achieve information transfer rates of up to 37 Bits/min. A detailed description of this on-line feedback study is given in [32].

### C. *Hex-O-Spell Application*

The major challenge in designing a BCI text input system like the novel "Hex-o-Spell" is to map the small number of detectable mental states into a high number of symbols (26 letters plus punctuation characters) while accounting for the low signal-to-noise ratio in the control signal. Fig. 5 shows how this problem is solved in the Hex-o-Spell system: using two different mental states, the BCI subject controls the movement of an arrow in the middle of the circle. Adjoining to this circle, six hexagons are located which contain groups of letters and symbols. By imagining left hand movement, the subject can turn the arrow clockwise at constant speed. Once the arrow points into the direction of the hexagon which contains the desired letter, an imagined foot movement increases the size of the arrow. If the arrow touches one of the hexagons, this field is selected and all other hexagons are cleared. The letters of the selected hexagon are distributed over the other fields, and the arrow can once more be used to select a single letter via turning and extending. At

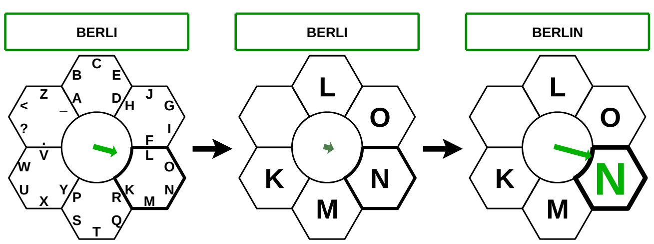


Fig. 5. The mental typewriter "Hex-o-Spell". The two states of the BCI system control rotation and extension of the arrow. After selecting a group of letters, a single letter can be chosen by selecting the hexagonal field containing the letter. See text for details.

the CeBIT computer fair 2006 in Hannover, Germany, this system was operated by two users in live demonstrations on two consecutive days. Despite the long period of use (starting at 9 AM, the subjects performed until 5 PM with only a short break for lunch), the BBCI proved very robust against external influences such as electric and acoustic noise and the slowly deteriorating signal-to-noise ratio caused by drying electrode gel. The typing speed for one subject was between 2.3 and 5 letters/min for one subject and between 4.6 and 7.6 letters/min for the other subject. More details on Hex-o-Spell can be found in [33], [34].

## IV. Conclusion

After a brief review of general linear machine learning techniques, this paper demonstrated their application in the context of real BCI-experiments. Using these techniques, it can be seen that the paradigm shift away from subject training to individualization and adaptation ('let the machines learn') of the signal processing and classification algorithm to the specific brain 'under study' holds the key to the success of the BBCI. Only this interplay of machine learning algorithms with specifically tailored feedback applications enables even untrained subjects to use the BCI and therefore dramatically enhances and broadens the spectrum of practical applications in human-computer interfacing.

**Acknowledgments**: We thank our co-authors from previous publications for letting us use the figures and joint results [12], [9], [11], [14], [23].

## References

[1] Klaus-Robert Müller, Matthias Krauledat, Guido Dornhege, Stefan Jähnichen, Gabriel Curio, and Benjamin Blankertz, "A note on the Berlin Brain-Computer Interface," in *Human Interaction with Machines: Proceedings of the 6th International Workshop held at the Shanghai Jiao Tong University*, Günter Hommel and Sheng Huanye, Eds., 2006, pp. 51–60.
[2] Jonathan R. Wolpaw, Niels Birbaumer, Dennis J. McFarland, Gert Pfurtscheller, and Theresa M. Vaughan, "Brain-computer interfaces for communication and control," *Clin. Neurophysiol.*, vol. 113, pp. 767–791, 2002.
[3] J. R. Wolpaw, N. Birbaumer, William J. Heetderks, D. J. McFarland, P. Hunter Peckham, G. Schalk, E. Donchin, Louis A. Quatrano, C. J. Robinson, and T. M. Vaughan, "Brain-computer interface technology: A review of the first international meeting," *IEEE Trans. Rehab. Eng.*, vol. 8, no. 2, pp. 164–173, 2000.
[4] Andrea Kübler, Boris Kotchoubey, Jochen Kaiser, Jonathan Wolpaw, and Niels Birbaumer, "Brain-computer communication: Unlocking the locked in," *Psychol. Bull.*, vol. 127, no. 3, pp. 358–375, 2001.
[5] Eleanor A. Curran and Maria J. Stokes, "Learning to control brain activity: A review of the production and control of EEG components for driving brain-computer interface (BCI) systems," *Brain Cogn.*, vol. 51, pp. 326–336, 2003.
[6] Matthias Krauledat, Guido Dornhege, Benjamin Blankertz, Gabriel Curio, and Klaus-Robert Müller, "The Berlin brain-computer interface for rapid response," *Biomed. Technik*, vol. 49, no. 1, pp. 61–62, 2004.
[7] J. R. Wolpaw, D. J. McFarland, and T. M. Vaughan, "Brain-computer interface research at the Wadsworth Center," *IEEE Trans. Rehab. Eng.*, vol. 8, no. 2, pp. 222–226, 2000.
[8] N. Birbaumer, N. Ghanayim, T. Hinterberger, I. Iversen, B. Kotchoubey, A. Kübler, J. Perelmouter, E. Taub, and H. Flor, "A spelling device for the paralysed," *Nature*, vol. 398, pp. 297–298, 1999.
[9] K.-R. Müller, S. Mika, G. Rätsch, K. Tsuda, and B. Schölkopf, "An introduction to kernel-based learning algorithms," *IEEE Neural Networks*, vol. 12, no. 2, pp. 181–201, May 2001.
[10] S.S. Haykin, *Adaptive Filter Theory*, Prentice Hall, 1995.
[11] Klaus-Robert Müller, Charles W. Anderson, and Gary E. Birch, "Linear and non-linear methods for brain-computer interfaces," *IEEE Trans. Neural Sys. Rehab. Eng.*, vol. 11, no. 2, pp. 165–169, 2003.
[12] Klaus-Robert Müller, Matthias Krauledat, Guido Dornhege, Gabriel Curio, and Benjamin Blankertz, "Machine learning techniques for brain-computer interfaces," *Biomed. Technik*, vol. 49, no. 1, pp. 11–22, 2004.
[13] Benjamin Blankertz, Gabriel Curio, and Klaus-Robert Müller, "Classifying single trial EEG: Towards brain computer interfacing," in *Advances in Neural Inf. Proc. Systems (NIPS 01)*, T. G. Diettrich, S. Becker, and Z. Ghahramani, Eds., 2002, vol. 14, pp. 157–164.
[14] Benjamin Blankertz, Guido Dornhege, Christin Schäfer, Roman Krepki, Jens Kohlmorgen, Klaus-Robert Müller, Volker Kunzmann, Florian Losch, and Gabriel Curio, "Boosting bit rates and error detection for the classification of fast-paced motor commands based on single-trial EEG analysis," *IEEE Trans. Neural Sys. Rehab. Eng.*, vol. 11, no. 2, pp. 127–131, 2003.
[15] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*, John Wiley & Sons, second edition, 2001.
[16] V.N. Vapnik, *The nature of statistical learning theory*, Springer Verlag, New York, 1995.
[17] R.A. Fisher, "The use of multiple measurements in taxonomic problems," *Annals of Eugenics*, vol. 7, pp. 179–188, 1936.

[18] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, Academic Press, San Diego, 2nd edition, 1990.

[19] S. Mika, G. Rätsch, and K.-R. Müller, "A mathematical programming approach to the kernel Fisher algorithm," in *Advances in Neural Information Processing Systems*, T.K. Leen, T.G. Dietterich, and V. Tresp, Eds. 2001, vol. 13, pp. 591–597, MIT Press.

[20] S. Mika, G. Rätsch, J Weston, B. Schölkopf, A. Smola, and K.-R. Müller, "Constructing descriptive and discriminative non-linear features: Rayleigh coefficients in kernel feature spaces," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 25, no. 5, pp. 623–628, May 2003.

[21] T. Poggio and F. Girosi, "Regularization algorithms for learning that are equivalent to multilayer networks," *Science*, vol. 247, pp. 978–982, 1990.

[22] G. Orr and K.-R. Müller, Eds., *Neural Networks: Tricks of the Trade*, vol. 1524, Springer LNCS, 1998.

[23] G. Rätsch, T. Onoda, and K.-R. Müller, "Soft margins for AdaBoost," *Machine Learning*, vol. 42, no. 3, pp. 287–320, Mar. 2001, also NeuroCOLT Technical Report NC-TR-1998-021.

[24] Guido Dornhege, Benjamin Blankertz, Gabriel Curio, and Klaus-Robert Müller, "Increase information transfer rates in BCI by CSP extension to multi-class," in *Advances in Neural Information Processing Systems*, Sebastian Thrun, Lawrence Saul, and Bernhard Schölkopf, Eds., vol. 16, pp. 733–740. MIT Press, Cambridge, MA, 2004.

[25] Guido Dornhege, Benjamin Blankertz, Gabriel Curio, and Klaus-Robert Müller, "Boosting bit rates in non-invasive EEG single-trial classifications by feature combination and multi-class paradigms," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 993–1002, June 2004.

[26] Matthias Krauledat, Guido Dornhege, Benjamin Blankertz, Florian Losch, Gabriel Curio, and Klaus-Robert Müller, "Improving speed and accuracy of brain-computer interfaces using readiness potential features," in *Proceedings of the 26th Annual International Conference IEEE EMBS on Biomedicine, San Francisco*, 2004.

[27] R. Q. Cui, D. Huter, W. Lang, and L. Deecke, "Neuroimage of voluntary movement: topography of the Bereitschaftspotential, a 64-channel DC current source density study," *Neuroimage*, vol. 9, no. 1, pp. 124–134, 1999.

[28] W. Lang, O. Zilch, C. Koska, G. Lindinger, and L. Deecke, "Negative cortical DC shifts preceding and accompanying simple and complex sequential movements," *Exp. Brain Res.*, vol. 74, no. 1, pp. 99–104, 1989.

[29] Roman Krepki, Benjamin Blankertz, Gabriel Curio, and Klaus-Robert Müller, "The Berlin Brain-Computer Interface (BBCI): towards a new communication channel for online control in gaming applications," *Journal of Multimedia Tools and Applications*, 2004, invited contribution.

[30] Gert Pfurtscheller and F. H. Lopes da Silva, "Event-related EEG/MEG synchronization and desynchronization: basic principles," *Clin. Neurophysiol.*, vol. 110, no. 11, pp. 1842–1857, Nov 1999.

[31] Roman Krepki, Benjamin Blankertz, Gabriel Curio, and Klaus-Robert Müller, "The Berlin Brain-Computer Interface (BBCI): towards a new communication channel for online control of multimedia applications and computer games," in *9th International Conference on Distributed Multimedia Systems (DMS'03)*, 2003, pp. 237–244.

[32] Benjamin Blankertz, Guido Dornhege, Matthias Krauledat, Klaus-Robert Müller, Volker Kunzmann, Florian Losch, and Gabriel Curio, "The Berlin Brain-Computer Interface: EEG-based communication without subject training," *IEEE Trans. Neural Sys. Rehab. Eng.*, vol. 14, no. 2, 2006, in press.

[33] Klaus-Robert Müller and Benjamin Blankertz, "Toward non-invasive brain-computer interfaces," 2006, accepted.

[34] Benjamin Blankertz, Guido Dornhege, Matthias Krauledat, Michael Schröder, John Williamson, Roderick Murray-Smith, and Klaus-Robert Müller, "The berlin brain-computer interface presents the novel mental typewriter hex-o-spell," *biomed*, 2006, accepted.