# BLIND SOURCE SEPARATION IN AN ARRAY OF ISFET SENSORS: I. MATHEMATICAL FOUNDATIONS

**S. BERMEJO**
UNIVERSITAT POLITÈCNICA DE CATALUNYA (UPC), SPAIN

**KEYWORDS: Blind Source Separation, Independent Component Analysis, Array Processing, ISFETs, Smart Sensors.**

**ABSTRACT**: Blind source separation (BSS) is a recent statistical technique that addresses the problem of separating several sources from a linear mixture of signals in array processing. This paper presents the underlying principles of BSS and other related methods like Independent Component Analysis (ICA) and shows how to apply them in the context of an array of ISFET sensors.

## INTRODUCTION

The incorporation of artificial intelligence processing techniques into traditional sensor systems has allowed the development of the so-called smart sensors [1]. The underlying idea of these systems is to overcome the inherent limitations of sensors by introducing signal processing techniques that enhance their output signals in order to facilitate the extraction of relevant information in further stages.

Smart sensors are typically involved in the monitoring and control of complex real-world processes which require the processing of signals from multiple sources, consequently, an array of sensors is employed. In this context, some kind of array processing of the sensors' output signals must be performed. One of the emerging methods for array processing in the last few years is blind source separation (BSS) [2] that can be very effective in separating sources from a mixture of observed signals. See e.g. [3][4] for recent applications of BSS techniques in smart sensors.

The contents of this paper are organized as follows. Section 2 presents the basics of BSS and other related techniques like independent component analysis (ICA). Section 3 introduces an ISFET-based array processing architecture for detecting concentrations of ions in water and shows that the overall system can be reduced to a BSS problem. Finally, preliminary conclusions are given and some further work is suggested.

## BLIND SOURCE SEPARATION AND OTHER RELATED TECHNIQUES

### Blind Source Separation as a Statistical Learning Method

Blind source separation (BSS) attempts to reconstruct a set of hidden signals $\{s_i\}$ from several observed signals $\{x_j\}$ that (presumably) have been generated from a linear mixture of the originals signals. The term "blind" refers to the fact that we must recover the unseen signals from the observed ones and also that there is no (or little) prior information about how the mixture has been produced. However, this deficiency is compensated to some degree by the existence of a set of (empirical) samples of the observed signals $\mathbf{D}=\{\mathbf{x}_r\}$, which allow "learning from data" in order to recover the original signals. Independent component analysis (ICA) [5][6] is a method related to BSS that follows similar goals.

BBS methods are a class of statistical learning machines [7] that aim to construct a verisimilar (or reliable) model of the mixing process through the information extracted from the set of samples **D**. Learning machines achieve this goal using a cost (or objective) function that depends on **D** and the parameters of the model. This function is minimized during a learning (or training) phase in order to compute the desired reliable estimation of these parameters. Hence, model's learning process can be viewed as a search that takes place within the space of parameters. In order to find a solution within a feasible time and to ensure that the computed solution is reliable or generalizable to new data that has not been seen in the training phase, this search must be biased or constrained to some degree. From a practical point of view what this fact implies is that the learning machine must make several (implicit or explicit) statistical assumptions about the problem.

A BSS method, as a statistical learning procedure, mainly consists of three parts:

1) A probabilistic model of the data (i.e. sources) that denotes in which way the data is distributed and how the original sources are related to the mixing signals. There are two main approaches for determining the data distribution: parametric and non-parametric models. In parametric modeling a particular distribution is assumed (e.g. uniform, exponential, etc.) while non-parametric approaches attempt to produce consistent estimates of any distribution given enough training samples to construct the non-parametric model. On the other hand,

the simplest model assumed between sources and mixing signals is linear. However extensions to non-linear mappings can also be possible.

2) An objective function in which the minimum (or maximal) point assures the achievement of a good solution for the BSS problem.

3) An optimization (or learning) algorithm which minimizes (or maximizes) the objective function in order to compute the solution. Many standard algorithms in optimization can be employed for this purpose. See [8] for a review of the basic methods of unconstrained optimization including steepest descent, which is the technique most widely employed.

The limitations of BSS methods are the same as those that arise in other techniques that learn from data [9]:

1) The use of an inappropriate model for the problem (the so-called bias error). If the model is too simple it will be impossible to compute a solution that copes with the problem. In the simplest forms of BSS models the dependence between the observed signals and the sources is linear. Clearly, this model can be insufficient in some domains. In this case, a non-linear extension could be used. Furthermore, the (implicitly or explicitly) assumed probabilistic distribution of the sources affects the degree to which reconstruction is possible. As the model departs from the real distribution, the more difficult reconstruction can be. See II.C[2] for an extended discussion about the problem of using a wrong model for the probability density function (pdf) of sources.

2) An excessive dependence of the estimated parameters of the model on the training samples (the so-called variance error). Generally speaking, as the complexity (e.g. the number of free parameters) of the model grows it is more probable that the estimated parameters are particularly tuned ("over-fitted") to the seen samples in the training phase. In BSS the use of a simple model such as a linear one allows a good estimation of the parameters with few training samples, as was reported in [2] . And finally:

3) Optimization errors produced by the use of on-line learning algorithms since they only converge to the same points as their batch counterparts under certain conditions (see e.g. [10][11]). However the use of on-line algorithms would allow the system to track the evolution of statistics of the source signals over time [12], i.e. non-stationary statistics.

## The BSS Model

In the simplest BSS model [2], we observe m discrete-time signals $x_1[n],\ldots, x_m[n]$ that correspond to a linear mixture of a p source signal $s_1[n],\ldots, s_p[n]$, i.e.

$$\begin{aligned} x_1[n] &= a_{11}s_1[n] + \ldots + a_{1p}s_p[n] \\ &\vdots \\ x_m[n] &= a_{m1}s_1[n] + \ldots + a_{mp}s_p[n] \end{aligned} \quad \textbf{(1)}$$

or expressed in a vector form

$$\mathbf{x}[n] = (x_1[n]\ldots x_m[n])^T = \mathbf{A}\,\mathbf{s}[n] \quad \textbf{(2)}$$

where **A** is known as the mxp "mixing matrix" and T denotes transpose. Equation **(2)** is a form of what is known in statistics as latent variable modeling [13]. In latent variable modeling, we have the observable vector **x** which is being generated from hidden (or latent) variables **s** through a linear mapping **A**. Then, the goal of this modeling is to create an inverse of **A** (within a multiplicative factor) from a set of examples **D** extracted from **x**[n] in order to recover the latent variables **s**[n]. To compute a feasible solution, latent variable modeling assumes some probabilistic model of the hidden variables **s** and then constructs a generative model that explains the observable vectors **x**. Examples of latent variable modeling, other than BSS, include e.g. mixture models [14] in which **x** is assumed to be generated from a mixture of (typically) gaussian random vectors where the latent variables are binary vectors that denote from which of the distributions the data has been generated. Other well-known forms of latent variable modeling are hidden Markov models.

A simple extension of Equation **(2)** includes the presence of a noise vector **n** as an additive presence in the BSS model, see e.g.[15],

$$\mathbf{x}[n] = \mathbf{A}\,\mathbf{s}[n] + \mathbf{n}[n] \quad \textbf{(3)}$$

Equation **(3)** is also used in factor analysis [16] where the noise components $\{n_i\}$ represent the part of the observed vectors that are not explained by the factors **s**. In factor analysis, these components are assumed to be distributed independently of **s** and to have no correlation between them.

Other extensions of the basic model denoted in Equation **(2)** include the generalization to a non-linear mapping [17][18][19]. A nonlinear BSS (or ICA) can be expressed as the estimation of the following generative model for the data,

$$\mathbf{x}[n] = \mathbf{f}(\mathbf{s}[n]) \quad \textbf{(4)}$$

where **f** is an unknown function from $\Re^p$ to $\Re^m$. Note that for a linear **f**, Equation **(4)** gives the linear BSS model.

According to the BSS model in Equation **(2)** and given the observable vector **x**, we must compute a pxm separating matrix B which allows an estimation of the source signals $\{s_i[n],\ i=1,\ldots,p\}$ using the following reconstruction algorithm,

$$\mathbf{y}[n] = (y_1[n] \ldots y_p[n])^T = \mathbf{B}\mathbf{x}[n] \approx \mathbf{s}[n] \qquad \textbf{(5)}$$

where clearly **B** must tend to $\mathbf{A}^{-1}$ to obtain a good estimation of **s** with **y** if it is assumed there is no noise. It is worth noting that **y** can be considered as a linear projection from a high-dimensional space X down to a low-dimensional one when p<m. This fact relates BSS to other linear approaches like principal component analysis (PCA) [20] or projection pursuit (PP) [21] both of which also compute their transformations using Equation **(5)**. However, it is also possible to have more sources than observed signals (p>m), which implies the inverse problem. See [22] for an introduction to the p>m case.

The classical approach to BSS and ICA is to use a non-Gaussian identically and independently distributed (i.i.d.) model in a stationary environment. However, current research directions focus modeling on the sources as; 1) non-Gaussian i.i.d., 2) Gaussian non-stationary, and 3) Gaussian, stationary correlated in time [23]. In the following sub-sections, we basically deal with the traditional approach pointing out some straightforward extensions.

**Independence of the components of s[n]**

The main assumption in the BSS model is the independence of the source signals $\mathbf{s}=(s_1 \ldots s_n)^T$, which is realistic in most of the cases in a sensor array context as we can sense a lot of different physical information which is (in principle) unrelated.

If **s** is formed by independent random variables, then its pdf can be expressed as the product of the pdf of its components ([24] p.132),

$$f_S(\mathbf{s}) = \prod_{i=1}^{p} f_{S_i}(s_i) \qquad \textbf{(6)}$$

Since $s_i$ and $s_j$ are independent, it can be shown that $g(s_i)$ and $h(s_j)$ are also independent where h and g are any arbitrary function([24] p.133). Then,

$$\begin{aligned} E[g(s_i)h(s_j)] &= \int g(s_i)h(s_j)f_{S_i}(s_i)f_{S_j}(s_j)ds_i\, ds_j = \\ &= \int g(s_i)f_{S_i}(s_i)ds_i\, h(s_j)f_{S_j}(s_j)ds_j = E[g(s_i)]E[h(s_j)] \end{aligned} \qquad \textbf{(7)}$$

Note that Equation **(7)** when g(x)=h(x)=x is reduced to the condition that must fulfill two uncorrelated random variables, which is a well-known result since independence implies uncorrelatedness. Many ICA methods constrain the form of computing **B** so that it always gives uncorrelated estimates of the independent components in order to decrease the number of free parameters, which allows simplification of the learning algorithm (p.9 [6]).

**x[n] as a discrete-time ergodic or stationary random process**

The basic approach in BSS and ICA [25] is to consider $\{x_i[n],\ n=1,\ldots,T,\ i=1,\ldots,m\}$ as a set of T realizations of m random variables. In other words, each set $\{x_i[n],\ n=1,\ldots,T\}$ is considered as a set of samples extracted from one random variable. The underlying assumption of this method is that the statistics of T samples extracted from a particular realization of the random process $x_i[n]$, i.e. a random variables $x_i[n_j]$, converges to the statistics of T samples collected from the extraction of one sample from T different realizations of $x_i[n]$ over time, when $T \rightarrow \infty$. This property is called ergodicity in the context of random processes.

On the other hand, time is not explicit in the simplest BSS model since **A** does not depend on the discrete-time n. Thus, the model assumes that the discrete-time random process **s**[n] is stationary, or more precisely, **strict-sense stationary** since the processes **s**[n] and **s**[n+c] must have the same statistics for any c. However, less restrictive approaches have been proposed recently that rely on stationary second-order statistics [26]. Conversely, extensions to non-stationary environments have also been considered in recent times [27][28][29][32].

**Gaussians are not allowed in the model**

Given that $s_i$ and $s_j$ are independent it is easy to observe that

$$\begin{aligned} &E\{(s_i - E[s_i])(s_j - E[s_j])\} = E\{(s_i - E[s_i])\}E\{(s_j - E[s_j])\} = \\ &= 0\ \forall i \neq j \end{aligned} \qquad \textbf{(8)}$$

The imposition of Equation **(8)** in the computation of the reconstruction matrix **B** allows p(p-1)/2 constraints over a total of mxp that are needed to compute the mxp free parameters of **B**.

On the other hand, the covariance matrix of the sources **s** using Equation **(8)** gives

$$\mathbf{K_s} = E\{(\mathbf{s} - E[\mathbf{s}])(\mathbf{s} - E[\mathbf{s}])^T\} = \mathrm{diag}(\sigma_{s_1}^2\ \sigma_{s_2}^2 \ldots \sigma_{s_p}^2) \qquad \textbf{(9)}$$

where diag(•) denotes a diagonal matrix and $\{\sigma_{s_i}^2, i=1,\ldots,p\}$ are the variances of the source signals $\{s_i\}$. Additional p constrains can be obtained if variances are forced into unity, i.e.

$$\mathbf{K_s} = \mathbf{I} \qquad \textbf{(10)}$$

The constrains imposed by Equation **(9)** are called spatial whiteness (or sphering) and leave mxp-[p(p+1)/2] to be determined by information other than second-order information. However, note that second-order information does at least half the job in BSS when p≥m-1 (i.e. the same number of sources as mixed signals or more sources than mixed signals). Then, the additional p(2m-p-1)/2 constrains could be extracted from high-order moments. Consequently, a simple solution to finding a numerical solution to the computation of **B** is to force **s** to be non-gaussian, since the gaussian (or normal) random variables restrict their statistical information to first and second-order moments. The search of hidden signals from the observed signals departing from gaussianity is not new. Projection pursuit [21], a technique introduced by Friedman and Tukey in the seventies [30], seeks to minimize measures for computing a transformation **B** that allows **y** to be non-gaussian with the purpose of finding clusters of data.

Finally, it is worth observing that searching non-gaussian distribution for **s** is not the only possible solution. For instance, factor analysis, in spite of using Equation **(9)** in order to eliminate some of the indeterminacy of its model, can extract independent gaussians (see [16] for further details).

**Parametric vs. non-parametric estimation of the model**

BSS techniques differ broadly in the assumptions made for the individual distributions of sources (Equation **(6)**). In principle, we are interested in assuming as little as possible in the distributions in order to cope with a wide range of problems. A possible solution is to estimate these distributions using non-parametric models like e.g. kernel density estimation [31]. See [32] for the use of non-parametric estimation of pdf based on kernels to compute **B** using the EM algorithm [33] in a maximum likelihood approach. However, simpler non-parametric approaches based on high-order moments are widely employed (see for a review [2][5][6]). Finally, a hybrid approach is also possible using semi-parametric statistics [34].

## Objective Functions

A natural function that can detect non-gaussianity is the differential entropy,

$$H(\mathbf{y}) = -\int f_y(\mathbf{y}) \log f_y(\mathbf{y}) d\mathbf{y} \tag{11}$$

which is the highest for a gaussian variable among all the random variables of the same variance (p.14 [6]). Suppose that we model this departure from normality in the 1-dimensional D as follows

$$f_y(y) = \varphi(y)[1 + \varepsilon(y)] \tag{12}$$

where $\varphi(y)$ is a gaussian variable and $\varepsilon(y)$ is a function that quantifies the degree of departure. According to [21], Equation **(11)** can be approximated by

$$H(y) = -\int f_y(y) \log f_y(y) dy \approx -\int \varphi(y) \varepsilon^2(y) dy \approx \approx \left( \frac{k_3^2 + \frac{1}{2} k_4^2}{12} \right) \tag{13}$$

where $k_3$ and $k_4$ are the third- and fourth-order cumulants. Extensions to the p-dimensional case can be done using cumulants for higher dimensions [35], which is a very common approximation in BSS and ICA. Equation **(13)** is an example of a non-parametric modeling of the pdf of the sources since no explicit distribution is assumed. See [6] for an extended discussion about the use of differential entropy and other related measures based on cumulants like Kurtosis. An alternative to the above approach is based on measuring the degree of independence between the components of **y**, which can be done with the Kullback-Leibler divergence [5].

On the other extreme, we can maximize the log likelihood of $\{x_i[n], n=1,\ldots,T, i=1,\ldots,m\}$ given a parametric form of the pdf of **s** (II.C[2]) One striking feature of the ML approach is that when T→∞ the log likelihood converges to a function which is formed by the differential entropy and the Kullback-Leibler divergence. Thus, non-parametric and parametric approaches are in fact related. A comprehensive overview of objectives functions for BSS and ICA can be found in [37].

## Learning Algorithms

Different optimization techniques can be used depending on which kind of objective must be minimized. If the objective function is based on non-parametric statistics, standard optimization methods like steepest descent [8] (e.g. gradient descent) are typically employed. On the other hand, if a ML approach is chosen, procedures like the EM algorithm [33] can be utilized. See [36] for an ICA algorithm using this approach.

# SOURCE SEPARATION IN AN ARRAY OF ISFETS

## ISFETS

The ion-sensitive FET (ISFET), introduced over 25 years ago [38], is a field effect transistor that is sensitive to chemical particles. See [39][40] for models of ISFETS.

## A Linear Model of Mixed Sources in an Array of ISFETS

According to [41], the drain current of an ISFET i which is active in ions of class k can be expressed in a linear range as

$$Id_i = a_i\left(V_{gs} - V_{T_i} - 0.5\,Vds\right)Vds = \tag{14}$$
$$= a_i\left(V_{gs} - \left(E_{oi} + b_i \ln\left(a_k + \sum_j K_{kj} a_j^{\frac{Zk}{Zj}}\right)\right) - 0.5\,Vds\right)Vds$$

where $V_{gs}$ is the gate-source voltage, $V_{ds}$ is the drain-source voltage, $E_{oi}$ is the membrane potential referring to the bulk solution consisting of a single type of ions of ISFET i, $a_k$ is the activity of the main ion k, $K_{kj}$ is the selectivity coefficient which relates the response to the interfering ions $a_j$, $Z_k$ is the valence of the main ion k and $Z_j$ is the valence of the disturbing ion j in the solution.

If we use the first-order approximation of the natural log function around a working point q,

$$\ln(x) = \ln(q) + \left(\frac{x}{q} - 1\right) + O(x^2) \tag{15}$$

allows the transformation of Equation **(14)** after some simple algebraic manipulations into the following expression

$$Id_i = A_i + B_i\left(a_k + \sum_j K_{kj} a_j^{\frac{Zk}{Zj}}\right) \quad i = 1,\ldots,m \tag{16}$$

where m denotes the number of sensors in the ISFETs array. Since the sum of interference ions can be considered as noise, Equation **(16)** can be reduced to a particular form of Equation **(3)**. Consequently, we have a problem of source separation that is suitable to be solved by BSS techniques. Note that if we only have one class of ISFETs, which responds to one kind of ion, in the array, the BSS problem is reduced to one source signal (1-D case).

## CONCLUSIONS AND FURTHER WORK

BSS is a mature statistical method that allows the recovery of hidden signals or sources from observed mixed signals. Consequently, it can be employed in array processing problems in the context of smart sensors since it can easily enhance the output signals of the sensors eliminating undesired information like noise or simple constants that depend on each specific sensor device for the recovery of original source signals from the observed mixtures. As there are many extensions to the original BSS approach, a free-noise linear and stationary model are available, a great variety of complex real-world problems can be addressed.

In our work, we have shown how the drain current of an array of ISFETS can be reduced to a linear model in which concentrations of ions appear mixed in the array. Hence, the BSS method could be employed to recover the original concentrations, which would simplify the computation of later processing stages. Since noise (originated from the interference of other ions with the principal ion) and non-linearities (caused by the log term in ISFET equations) appear in the ISFET array equations, further research will be concentrated on advanced BSS models.

## ACKNOWLEDGEMENTS

This work was supported by the IST Programme No. 2000-28084 (SEWING) of the EU.

## THE AUTHOR

Dr. Sergio Bermejo is with the Advanced Hardware Architectures group in the Department of Electronic Engineering, Universitat Politècnica de Catalunya, C4 building, Jordi Girona 1-3, 08034 Barcelona.
E-mail: sbermejo@eel.upc.es

## REFERENCES

[1] J.W. Gardner, "Microsensors: Principles and Applications", Chichester: John Wiley & Sons, 1994, chapters 11 & 12

[2] J.-F.Cardoso, "Blind Signal Separation: Statistical Principles", Proceedings of the IEEE, Vol. 86, No. 10, 1998, pp. 2009-2025

[3] A. Paraschiv-Ionescu, C. Jutten, G. Bouvier, "Source Separation Based Processing for Smart Sensor Arrays", Submitted, 2002

[4] M. Kermit, O. Tomicy, "Independent Components of Odour Signals", 3rd International Conference on Independent Component Analysis and Blind Source Separation, San Diego, California, USA, 2001.

[5] A. Hyvärinen, "Survey on Independent Component Analysis", Neural Computation Surveys, 1999, Available electronically in http://www.cis.hut.fi/_aapo/

[6] A. Hyvärinen, E. Oja, "Independent Component Analysis: A Tutorial", International Joint Confrenece on Neural Networks, ijcnn99, 1999, Available electronically in http://www.cis.hut.fi/projects/ica/

[7] V. Vapnik, "Statistical Learning Theory", New York: Wiley, 2000

[8] J.E. Dennis Jr, R. B. Schnabel, "A View of Unconstrained Optimization", In G.L. Nemhauser, A.H.G. Rinnoy Kan, M.J. Todd, Optimization, Amsterdam: North-Holland, 1989

[9] S. Geman, S., E. Bienenstock, R. Doursat, "Neural Networks and the Bias/Variance Dilemma", Neural Computation, Vol.4, 1992, pp. 1-58.

[10] S. Bermejo, J. Cabestany, "Finite-Sample Convergence of the LVQ1 Algorithm and the BLVQ1 algorithm", Neural Processing Letters, Vol. 13, 2001, pp.135-157.

[11] S. Bermejo, J. Cabestany, "The Effect of Finite Sample in the Online K-means Algorithm", Accepted for publication in Neurocomputing, 2001.

[12] A. Benveniste, M. Métivier, P. Priouret, "Adaptive Algorithms and Stochastic Approximations", Berlin: Springer-Verlag, 1990.

[13] B. S. Everitt, An Introduction to Latent Variable Models. London, U.K.:Chapman and Hall, 1984

[14] G. J. McLachlan, K.E. Basford, "Mixture Models. Inference and Applications to Clustering", New York: Marcel Dekker, 1988.

[15] A. Belouchrani, K. Abed-Meraim, J.F. Cardoso, E. Moulines, "A Blind Source Separation Technique Using Second-Order Statistics", IEEE Transactions on Signal Processing, 1997, Vol.45, No.2, pp.434-444.

[16] T.W. Anderson, "An Introduction to Multivariate Statistical Analysis", $2^{nd}$ Edition, New York: John Wiley & Sons, 1984, chapter 19

[17] P. Pajunen, A. Hyvärinen and J. Karhunen, "Non-Linear Blind Source Separation by Self-Organizing Maps", Proc. Int. Conf. on Neural Information Processing, Hong Kong, 1996, pp.1207-1210

[18] P. Pajunen, "A Maximum Likelihood Approach to Nonlinear Blind Source Separation", Proceedings of the 1997 Int. Conf. on Artificial Neural Networks (ICANN'97), Lausanne, Switzerland, October 1997, pp. 541-546

[19] A. Hyvärinen, P. Pajunen, "Nonlinear Independent Component Analysis: Existence and Uniqueness results, Neural Networks, Vol. 12, No. 3, 1999, pp. 429-439

[20] I.T. Jolliffe, "Principal Components Analysis", New York: Springer Verlag, 1986.

[21] M.C. Jones, R. Sibson, "What is Projection Pursuit?", Journal of the Royal Statistical Society A, Vol. 150, 1987, pp. 1-36

[22] J.-F. Cardoso, P. Common, "Independent Component Analysis, A Survey of Some Algebraic Methods", 1994.

[23] J.-F. Cardoso, "The Three Easy Routes to Independent Component Analysis; Contrast and Geometry", 3rd International Conference on Independent Component Analysis and Blind Source Separation, San Diego, California, USA, 2001

[24] A. Papoulis, "Probability, Random Variables and Stochastic Processes", 3rd Edition, Singapore: McGraw-Hill, 1991.

[25] A. Hyvärinen, K. Karhunen, E. Oja, "Independent Component Analysis", New York: Wiley, 2001, p.2

[26] A. Belouchrani, K. Abed-Meraim, J.F. Cardoso, "A Blind Source Separation Technique Using Second-Order Statistics", IEEE Transactions on Signal Processing, Vol. 45, No. 2, 1997, pp. 434-444

[27] D.T. Pham, J.F. Cardoso, "Blind separation of instantaneous mixtures of non stationary sources", IEEE Transactions on Signal Processing, Vol. 49, No. 9, 2001, pp. 1837-1848

[28] A. Ziehe, K. R. Muller "TDSEP: an efficient algorithm for blind separation using time structure", Proceedings of ICANN, 1998, pp. 675-680

[29] N. Murata, K.-R. Müller, A. Ziehe, S. Amari, "Adaptive on-line learning in changing environments", Advances in Neural Information Processing Systems 9 (NIPS'96), D.S. Touretzky, M.C. Mozer and M.E. Hasselmo (eds.), MIT Press: Cambridge, MA, 1997, pp. 599-605

[30] J. H. Friedman, J.W. Tukey, "A Projection Pursuit Algorithm for Exploratory Data Analysis", IEEE Transaction on Computers, Vol. 23, No. 9, 1974, pp. 881-890

[31] W. Scott, "Multivariate Density Estimation: Theory, Practice and Visualization", New York: John Wiley& Sons, 1992.

[32] N. Vlassis, Y. Motomura, "Efficient Source Adaptivity in Independent Component Analysis", IEEE Transactions on Neural Networks, Vol. 12, No. 3, 2001, pp. 559-566

[33] A.P. Dempster, N. M. Laird, D.B. Rubin, "Maximum Likelihood from incomplete data via the EM algorithm", Journal of the Royal Statistical Society B, Vol. 39, pp. 1-38

[34] S. Amari, J.F. Cardoso, "Blind Source Separation-Semiparametric approach", IEEE Transaction on Signal Processing, Vol. 45, No.11, 1997, pp. 2692-2700

[35] G. Deco, D. Obradovic, "An Information-Theoretic Approach to Neural Computing", New York: Springer-Verlag, 1996

[36] M. Welling, M. Weber, "A Constrained EM Algorithm for Indepnednet Component Analysis", Neural Computation, Vol. 13, pp. 677-689

[37] J.-F. Cardoso, "Higher-order contrasts for Independent Component Analysis", Neural Computation, Vol. 11, 1999, pp. 157-192

[38] P. Bergveld, "Development of an ion-sensitive solid state device for neurophysiological measurements", IEEE Transaction on Biomedical Engineering, Vol. 17, 1970, pp. 70–71

[39] M. Grattarola, G. Massobrio, S. Martinoia, " Modeling H/sup +/-sensitive FETs with SPICE" IEEE Transactions on Electron Devices, Vol. 39, 1992, pp. 813–819

[40] S. Jamasb, S.D. Collins, R.L. Smith," A physical model for threshold voltage instability in Si/sub 3/N/sub 4/-gate H/sup +/-sensitive FET's (pH ISFET's)", IEEE Transactions on Electron Devices, Vol. 45, 1998, pp. 1239 –1245

[41] A. Napieralski, M. Daniel, Michal Szermer, K. Slusarczyk, "Library of CHEMFET for the ELDO environment", SEWING Project Report, 2001, http://www.sewing.mixdes.org.