# ADAPTIVE SIGNAL PROCESSING USING MAXIMUM ENTROPY ON THE MEAN METHOD AND MONTE CARLO ANALYSIS

*Pavla Holejšovská, Ing.* [*)], *Z. Peroutka, Ing.* [**)], *J.-F. Bercher, Prof. Assist.* [***)]

Západočeská Univerzita v Plzni, Czech Republic and Groupe ESIEE Paris, France

**Abstract:** *Designed signal and image processing method enables to avoid impact of measurement system on experimental evidence, mitigates noises and increases precision of obtained data. The simulations in Matlab were used to verify the proposed algorithm consisted of two methods. One of them is the Maximum Entropy on the Mean Method (MEMM) that deals with the solving of the linear and noisy inverse problem of the form **y=Ax+b.** The second one is the Expectation-Minimization Algorithm (EM algorithm) consisting of two steps. The expectation step (E step) computes estimation $\hat{\boldsymbol{x}}$ and it is followed by the maximization step (M step) that provides new estimation. Both steps have to be iterated until convergence. In the exponential family, the E step gives similar results as the MEMM that is why we combine both methods in order to explore the advantages of both. The measured data (**y**) processed by proposed algorithm allows to eliminate the noise **b** and the influence of the measurement system properties (presented by degradation matrix **A**). Thus, this algorithm enables to reconstruct real data **x** that object produces. Importance of this method in the signal and image processing and diagnostics is obvious.*

## 1. MAXIMUM ENTROPY ON THE MEAN METHOD (MEMM)

The Maximum Entropy on the Mean Method (MEMM) solve the linear and noisy inverse problem of the form of **y=Ax+b**, where **y** are the observations, **A** is the supposed degradation matrix, vector **x** is the measured object (typically a signal or image) and **b** is the noise which has to be estimated. The measured data processed by this algorithm allows to eliminate the noise **b** and the influence of the measurement system properties (presented by degradation matrix **A**) to the measured data **y** and so to obtain real data that the object produces. Given reference measure **μ** defined on the object **x** and noise **b**, the MEMM consists of selecting the distribution $\hat{\boldsymbol{p}}$ which is the closest to **μ** according to the Kullback distance and which satisfy a given constraint, in this case the observation equation. The MEMM estimation $\hat{\boldsymbol{x}}$ is the mean of the selected distribution $\hat{\boldsymbol{p}}$. So, it is the minimizer of a convex cost function defined on **x** and **b.**

Let have a linear inverse problem **y=Ax+b**. The observation matrix **A** is supposed to be known and some statistical characteristics of the noise **b** too. When the observation matrix **A** is not regular or ill-conditioned, the problem is ill-posed. It means the convex constraint

$$\boldsymbol{x} \in \boldsymbol{C} \ , \tag{1}$$

where C is a convex set, is necessary. In some specific problems where the lower and upper bound $(a_k,b_k)$ are known, the convex set can be defined as

$$\boldsymbol{C} = \left\{ \boldsymbol{x} \in \boldsymbol{R}^N / \boldsymbol{x}_k \in \left\langle \boldsymbol{a}_k, \boldsymbol{b}_k \right\rangle, \boldsymbol{k} = \mathbf{1}..\boldsymbol{N} \right\}. \tag{2}$$

For the estimation $\hat{\boldsymbol{x}}$ we select the distribution $\hat{\boldsymbol{p}}$ which is the closest to reference measure **μ** according to the Kullback distance. The Kullback distance **D(p||μ)** is defined for a reference measure **μ** and probability measure **P** by

$$\mathbf{D}(\mathbf{p} \,||\, \mu) = \int \log \frac{d\boldsymbol{P}}{d\mu} d\boldsymbol{P} \,. \tag{3}$$

Thus, the MMEM method begins by the specification of the convex set **C** and the reference measure **dμ(x)** over it. The actual observations **y** are considered as the mean $\mathbf{E_p(X)}$ of the probability distribution **P** defined on **C**.

The distribution **P** is selected as the minimizer of the μ-entropy submitted on the constraints of the mean $\mathbf{AE_p(x) = y}$ in the noiseless case. It means the **P** is the nearest distribution respect to the Kullback distance $\mathbf{D(p||\mu)}$ to the reference measure $\boldsymbol{\mu}$ satisfying equation $\mathbf{AE_p(x) = y}$. MEMM problem in the noiseless case is given by:

$$\begin{cases} \hat{p} = \arg\min_{P} \int \log \frac{dP}{d\mu}(x) dP(x), \\ such\ that\ \ y = A\int x dP(x) \\ and\ \ x = E_p(x) \end{cases} \tag{4}$$

The solution exists if it is belongs the exponential family. For each $x \in C$, the $\mathbf{AE_p(x) = y}$. We define the function $\mathbf{F(x)}$ as the optimum value of the Kullback distance.

$$F(x) = \operatorname*{Inf}_{P \in P_X} \mathrm{D}(\mathrm{p} \,||\, \mu),\ \ where\ \ P_X = \{P : E_p(x) = x\}\,. \tag{5}$$

Than the problem can be posed as:

$$\begin{cases} \hat{x} = \arg\min_{x} F(x), \\ such\ that\ \ y = Ax. \end{cases} \tag{6}$$

This method amounts to minimizing the convex criterion and so admits the dual formulation. The dual function is defined as

$$D(\lambda) = \lambda^t y - F^*(\lambda^t A), \tag{7}$$

and allows to calculate numerically the expectation $\mathbf{E_p(x)}$. The function $F^*(\lambda^t A)$ is the convex conjugate of the function $\mathbf{F(x)}$. It means that dual formulation of the problem that have to be solved is to find estimation $\hat{\lambda}$ by maximizing the dual function $\mathbf{D(\lambda)}$:

$$\left\{ \hat{\lambda} = \sup_{\lambda}(\lambda^t y - F^*(\lambda^t A)), \right. \tag{8}$$

The problem number one of the MEMM is to define the function $F^*(\lambda^t A)$ that presents the optimum value of the dual function and to maximize it. The problem is that in the noisy case the function $F^*(\lambda^t A)$ has to be separated to two functions:

$$F^*(\lambda^t A) = F_x^*(\lambda^t A) + F_b^*(\lambda^t). \tag{9}$$

The first one is the signal function $F_x^*(\lambda^t A)$ and the second one is the noise function $F_b^*(\lambda^t)$. As you see the problem start to be more complicated because both parts have to be defined and solved separately. The solution of this problem is really complicated. In the exponential family, the MEMM gives similar results as the E step of the EM algorithm. That is why we tried to use the EM algorithm to eliminate the need of the function $F^*(\lambda^t A)$ separation in the noisy case.

## 2. E-M Algorithm with Maximum Entropy on the Mean Method

Expectation-Minimization Algorithm (EM algorithm) consists of two steps: the expectation step followed by the maximization step. The E step computes an estimation $\hat{x}$ satisfying the conditions upon the observations.

Let have **x** the data, the **y** observations, $\mathbf{f(x|\theta)}$ probability density function and $\boldsymbol{\theta}$ set of parameters of the density, then for E step we compute:

$$Q(\theta | \theta^{[k]}) = E[\log f(x \mid \theta) \mid y, \theta^k]. \tag{10}$$

Then the M step provides new estimation:

$$Q(\lambda, \lambda^k, y) = \lambda^t AE[x \mid y, \theta^k] \tag{11}$$

These two steps must be iterated until convergence occurs - it may be determined as:

$$\left\| \theta^{[k]} - \theta^{[k-1]} \right\| < \varepsilon \; . \tag{12}$$

In the exponential family, the MEMM gives similar results as the E step of the EM algorithm. That is why the new method of the iterative algorithm was used to avoid the difficulties with the function $F^*(\lambda^t A)$ separation in the noisy case of MEMM and the both methods were combined. We compute the estimation of the **x** by applying the Maximum Entropy on the Mean Method (MEMM) into the EM algorithm. The first step is computes the estimation $\hat{\lambda}$ by maximizing the dual function **D(λ)** using the MEMM and then the EM algorithm starts by implementation of the estimation $\hat{\lambda}$ - it means we compute:

$$Q(\lambda, \lambda^k, y) = \lambda^t AE[x \mid y, \lambda^k]. \tag{13}$$

$$\lambda^{k+1} = \arg\max_{\lambda} \lambda^t AE[x \mid y, \lambda^k] - F^*(\lambda^t A). \tag{14}$$

And these two steps are iterated until the convergence condition occurs.

$$\left\| \lambda^{[k]} - \lambda^{[k-1]} \right\| < \varepsilon \; . \tag{15}$$

## 3. RESULTS

The simulations were made to reconstruct real signal **x** ($x_{true}$) from the measured observations **y** ($y_{measured}$). Results of the data reconstruction are well seen in the simulations below. Here you can see the results of the simulations of the signal **x** that is given by:

```
x=abs((.9*cos(2*pi*[.5:1:N-.5]/N)+.05))'.
```
(16)

The noise is supposed to be Gaussian. Various values of the $\sigma^2$ have been considered. In this paper, $\sigma^2$=0.01, $\sigma^2$=0.2 a $\sigma^2$=0.5 are presented in figures 1, 2 and 3.

On the left side, there is the comparison between the real signal **x** (normal) and signal $\mathbf{x_{ME}}$ (bold) computed by the MEMM with EM algorithm. Experimental data **y** are presented by points.

On the right side of each figure you can see the iteration process of estimation $\hat{x}$. The real signal **x** is presented by dotted line, last iteration of the computed signal $\mathbf{x_{ME}}$ by bold.

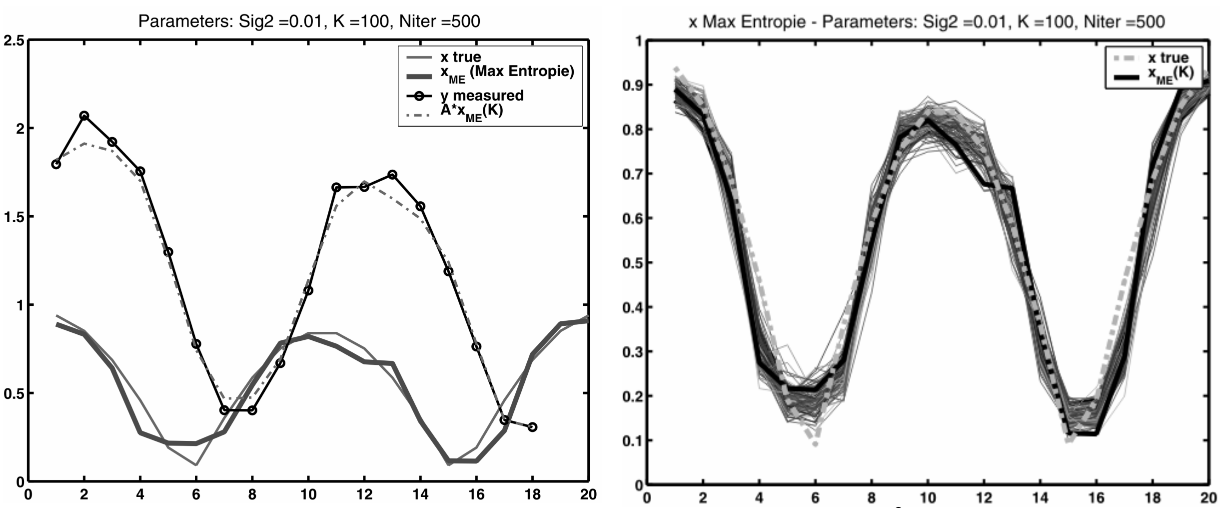


Fig. 1 Simulation of the MEMM with EM algorithm (Parameters: $\sigma^2$=0.01, 100 iter. cycles)

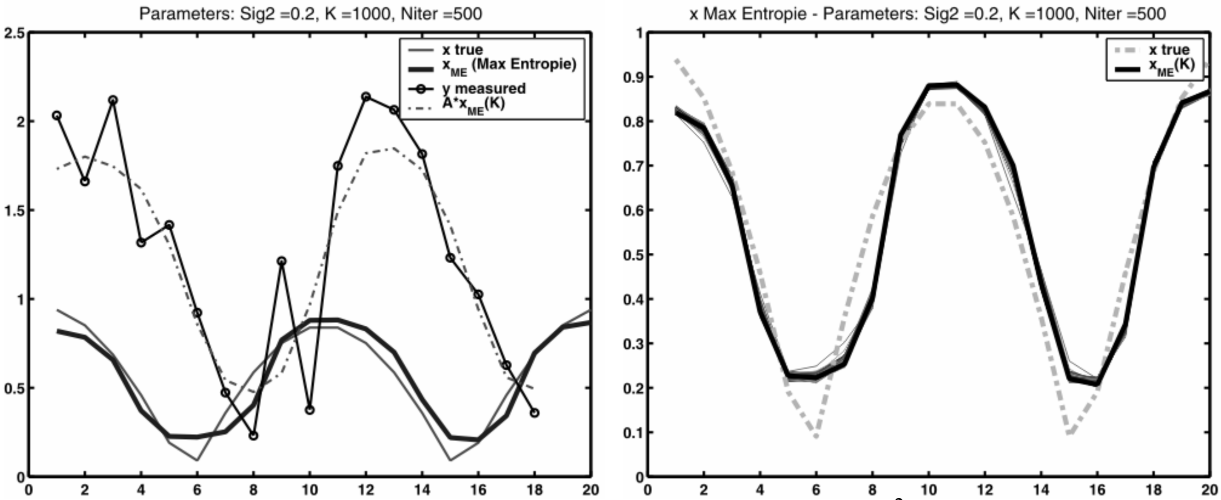


Fig. 2 Simulation of the MEMM with EM algorithm (Parameters: $\sigma^2$=0.2, 1000 iter. cycles)

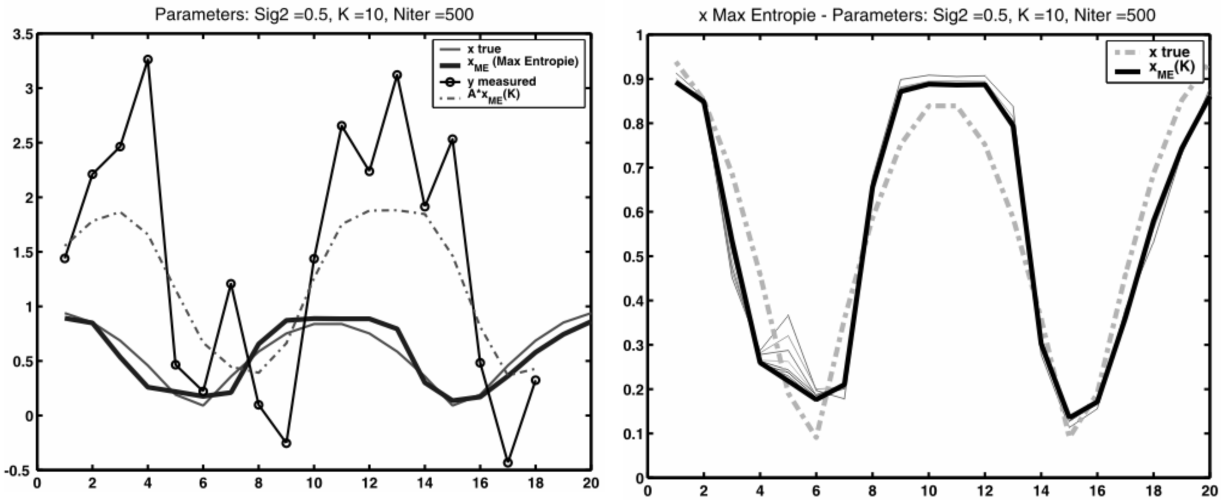


Fig. 3 Simulation of the MEMM with EM algorithm (Parameters: $\sigma^2$=0.5, 10 iter. cycles)

## 4. MONTE CARLO ANALYSIS

Various simulations in Matlab were made to reconstruct real signal **x** from the measured observations **y** computed by the MEMM with EM algorithm. If the signal is too noisy, the results will be less satisfactory as you can see in the figure 4a),b).

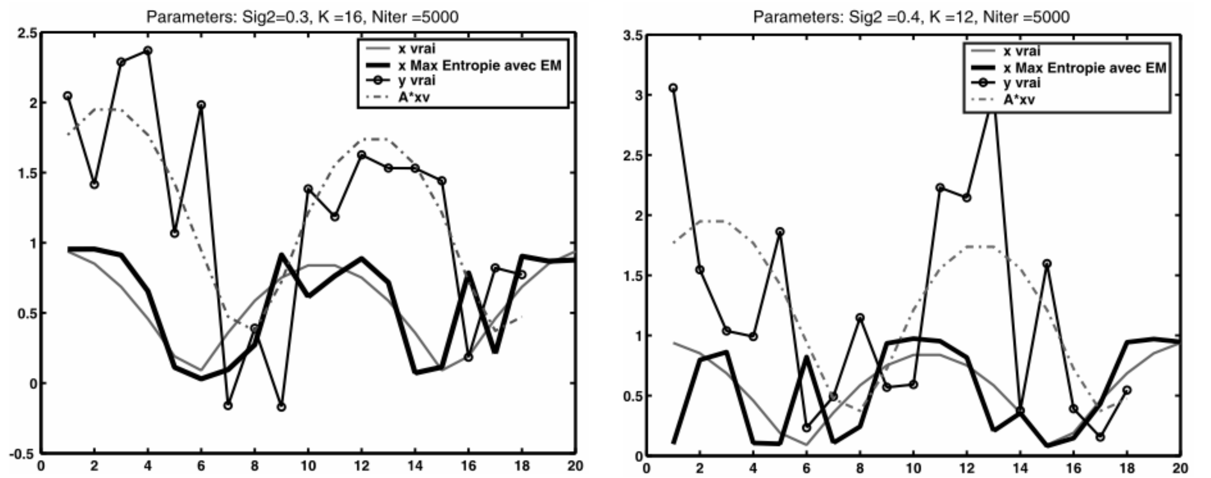


Fig. 4 Simulations of the MEMM with EM algorithm

a) Parameters: $\sigma^2$=0.3, 16 iter. cycles b) Parameters: $\sigma^2$=0.4, 13 iter. cycles

The noise and its distribution can lead to the reconstruction faults that can cause the estimated signal distortion. In this case, the Monte Carlo analysis is applied to avoid the method mistakes and computing faults. It allows to reconstruct the estimation $\hat{x}$ of the real signal **x** from measured, very noisy observations **y** with high precision.

The Monte Carlo analysis was applied to 30 simulations with the same parameters and the obtained results are shown in the figures 5 and 6.

On the left side you can see all the simulations and the result of the Monte Carlo analysis (black line) with its fault range (black dotted line). You can see that even if the simulations can provide quite great reconstruction faults, the result of the Monte Carlo analysis corresponds well with the real data **x** (gray bold line).

On the right side, the comparison of the real signal **x** (gray bold line) and the mean of the computed estimation $\hat{x}$ of all the simulations (black thin line) with its fault range (gray thin dotted line) is presented. From the figures is evident that quit great estimation error is very satisfactorily eliminated by the Monte Carlo analysis (black line).

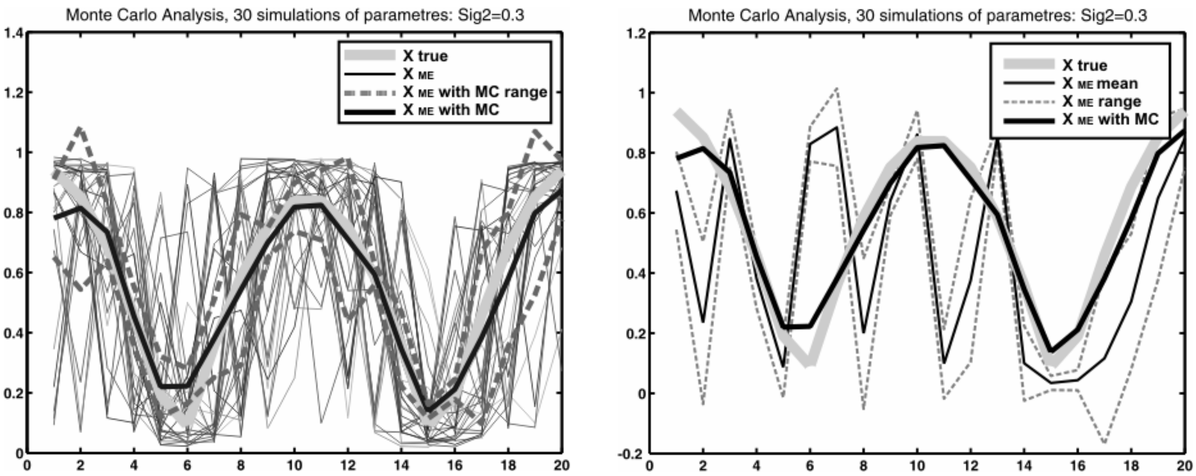


Fig. 5 Monte Carlo Analysis of the MEMM with EM algorithm
(Parameters: $\sigma^2$=0.3, 30 simulations)

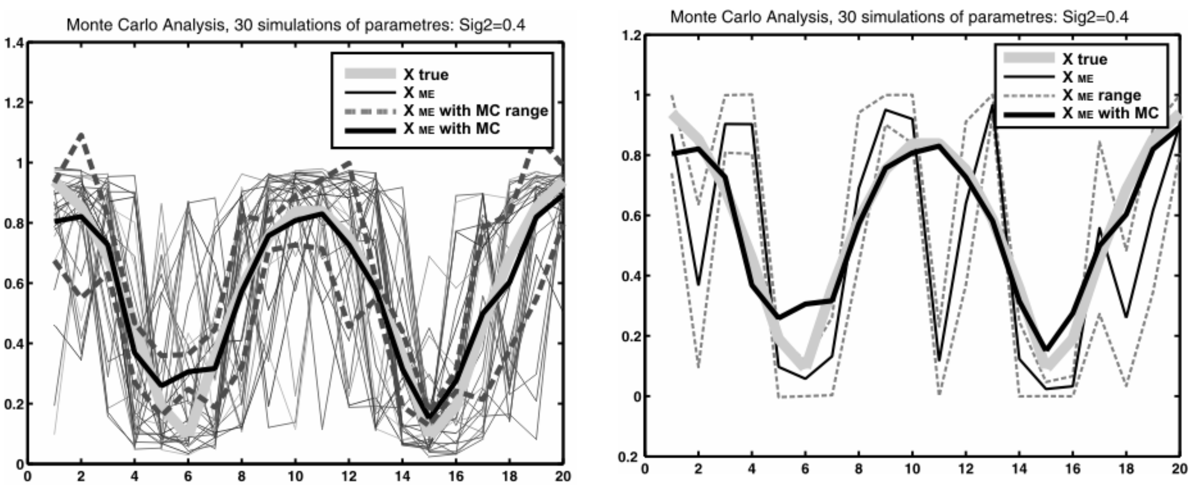


Fig. 6 Monte Carlo Analysis of the MEMM with EM algorithm
(Parameters: $\sigma^2$=0.4, 30 simulations)

## 5. CONCLUSIONS

The Maximum Entropy on the Mean Method (MEMM) is used to solve the linear and noisy inverse problem. Because of difficulties in the computation process in the noisy case, the MEMM is used as the first step of the Expectation Maximization algorithm that allows to converge to the real signal **x** by the iteration process. The problem is to find the best estimation $\hat{x}$ of the real data and satisfy the given constraints. We combine two methods in order to gain the advantages of both. The proposed algorithm reconstructs the real data **x** from measured, noisy observations **y**. This method enables to mitigate the noise **b** and the measurement system impact.

The simulations in Matlab were made to verify this new reconstruction method based on the Maximum Entropy on the Mean Method combined with EM algorithm. The noise is supposed to be Gaussian. Various values of the $\sigma^2$ have been considered. If the signal is too noisy, the results will be less satisfactory. The noise and its distribution can lead to the reconstruction faults that can cause the estimated signal distortion. In this case, the Monte Carlo analysis is applied to avoid the method mistakes and computing faults. It allows to reconstruct the estimation $\hat{x}$ of the real signal x from measured, very noisy observations y with favorable precision. Importance of this method in the signal and image processing and diagnostics is obvious.

The method is still in progress.


## ACKNOWLEDGEMENT

The research was performed at Groupe ESIEE Paris, France and is partly supported by project FRVŠ 2328/2003/G1.


## REFERENCES


[1] J.-F.Bercher, G.Le Besnerais, G.Demoment: *The Maximum Entropy on the Mean Method, Noise and Sensitivity.* Maximum Entropy and Bayesian Methods, Kluwer Academic Publishers, Cambridge,U.K.,1995.

[2] C.Heinrich, J.-F.Bercher, G.Demoment: *The Maximum Entropy on the Mean Method, Correlations and Implementation Issues.* In: IEEE Transactions on Information Theory,1995.

[3] Todd K. Moon: *The Expectation-Maximization Algorithme..* In: IEEE Signal Processing Magazine. November 1996, IEEE1996. p.47-60. 1053-5888/96.



*) University of West Bohemia in Pilsen, Faculty of Electrical Engineering
Department of Applied Electronics
Univerzitní 8, 306 14 Plzeň, Czech Republic
e-mail: holejsov@kae.zcu.cz

**) University of West Bohemia in Pilsen, Faculty of Electrical Engineering
Department of Electromechanics and Power Electronics
Univerzitní 8, 306 14 Plzeň, Czech Republic
e-mail: peroutka@ieee.org

***) École Supérieure d'Ingénieurs en Électrotechnique et Électronique (Groupe ESIEE),
Département Traitement du signal et Télécommunications et Laboratoire Parole Signal Image (LPSI)
Cité Descartes, BP 99, 93 162 Noisy le Grand Cedex