A Fast Converging and Self-Adjusting SHARF Algorithm

Omer Sezer, M.S. and Mohammed Ferdjallah, Ph.D.

Department of Electrical & Computer Engineering, The University of Tennessee, Knoxville, TN

ABSTRACT

Adaptive infinite impulse response (IIR), or recursive, filters are less attractive mainly because of the stability and the difficulties associated with their adaptive algorithms. Hyperstability is a concept from nonlinear stability theory and its convergence is directly related to strictly positive real (SPR) transfer functions. The simple hyperstable adaptive recursive filters (SHARF) is a simplified version of hyperstable recursive filters designed for real time applications. In this paper, SHARF is investigated using constraint recursive least-squares (RLS) method with SPR transfer functions designed without any priori knowledge of the parameters of the filter by the pole-zero placement on the unit circle method. To demonstrate its fast convergence and self-adjustment, SHARF algorithm is applied to a pure four-pole autoregressive process.

1. INTRODUCTION

Adaptive infinite impulse response (IIR), or recursive, filters have the advantage over finite impulse response (FIR) filters by requiring fewer coefficients and less computational cost. However, adaptive IIR filters are instable during adaptation due to their nonlinear phase property, which limits their practical implementations [1,2]. The adaptive IIR algorithm is formulated using output error identification method via hyperstability. Hyperstability, proposed by Popov, is a concept from nonlinear stability theory, which is used to develop real time adaptive recursive filter. The hyperstability theory is first applied to applications in control systems, then simplified and adapted into the adaptive digital signal processing applications by Landau and others [3]. The hyperstable adaptive recursive filter (HARF) represents the modification of the hyperstable output error identification method. The simple hyperstable adaptive recursive filter (SHARF), the simplified version of HARF, is more suitable for real time applications with a lower computation but slower rate of adaptation with the least mean squares (LMS) method [1,4,5]. The convergence of the hyperstability theory is directly related to SPR transfer function. The design of SPR requires the knowledge of unknown filter parameters. This undesirable characteristic of SPR initiated efforts to minimize the designer efforts, which lead to new methods with different degree of complications [6]. In this paper, a new algorithm is investigated in order to design SPR transfer functions without any prior knowledge of filter parameters. The pole-zero placement on the unit circle method related the numerator and denominator coefficients of SPR transfer function through fundamental coefficients. The fundamental coefficients are updated using a constrained recursive least-squares (CRLS) method. Section 2 gives a background on SHARF algorithms and SPR transfer functions. The design process of the self-adjusting SHARF algorithm with SPR transfer function model using pole-zero placement on the unit circle method and CRLS method is illustrated in Section 3. In Section 4, self-adjusting SHARF algorithm is applied to a pure four-pole autoregressive process to demonstrate its fast-convergence and self-adjustment.

2. BACKGROUND

The hyperstable adaptive recursive filter (HARF) is an early version of the application of hyperstability to signal processing and suffers many setbacks, which made it very hard to implement. The simple hyperstable adaptive recursive filter (SHARF) is a simplified version of HARF designed to apply for real time applications [4]. The SHARF algorithm consists of an input signal *x(n)* and output signal *y(n),* which closely approximates the desired signal *d(n)* (Fig. 1).

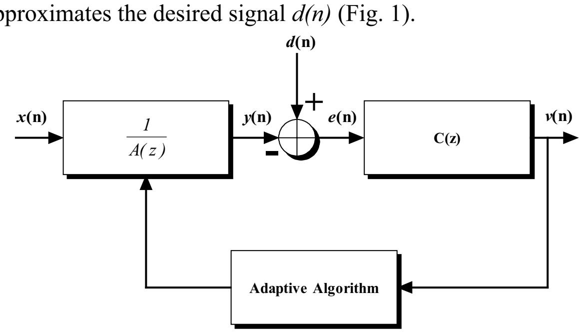

Fig. 1 Simple Hyperstable Adaptive Recursive Filter (SHARF)

The error signal *e(n)* is defined as:

$$e(n) = d(n) - y(n) \tag{1}$$

Hyperstability is directly related to SPR transfer functions. Because the autoregressive form *1/A(z)* fails to be SPR in general, the system is guaranteed to be SPR by adding *C(z)* function [5]. The output of *C(z)* is the moving average of the error signal, which can be defined as [5]:

$$v(n) = e(n) + \sum_{i=1}^{N} c_i e(n-i) \tag{2}$$

The minimum mean square error formulates the performance of the algorithm through the cost function $J_v(n)$. The cost function of the error signal $J_e(n)$ and the cost function of the moving average of the error signal $J_v(n)$ can be defined as:

$$J_e(n) = \frac{1}{2}E\left[e^2(n)\right] \quad and \quad J_v(n) = \frac{1}{2}E\left[v^2(n)\right] \tag{3}$$

The transfer function $G(z)$ in Fig. 1 can be written as:

$$G(z) = \frac{C(z)}{A(z)} = \frac{1 + \sum_{i=1}^{M} c_i z^{-i}}{1 + \sum_{j=1}^{N} a_j z^{-j}} \tag{4}$$

where, a_j are the coefficients and N is the order of the autoregressive process, and the order M and the error smoothing coefficients c_i are chosen by the designer so that $G(z)$ is SPR [5]. The system in Fig. 1 is hyperstable if and only if $G(z)$ is SPR, that is $G(z)$ must have a positive real part, i.e. [5];

$$Re\left[G(z)\right] > 0, \quad for \quad |z| = 1 \tag{5}$$

The problem is that the denominator coefficients of $G(z)$, a required source for adaptation, is unknown.

3. SELF-ADJUSTING SHARF ALGORITHM

3.1 SPR Transfer Function Design

The unknown denominator coefficients can be estimated by designing an SPR transfer function using pole-zero placement on the unit circle method. This method is proven to be robust for IIR band-rejection filter algorithm [7]. The transfer function $G(z)$ can be designed by cascading N second-order SPR transfer functions. The pole-zero placement method locates the zeros on the unit circle while the poles are located inside the circle at a radial distance from the zeros. However, the zeros are located inside and near the unit circle, not on the unit circle, for the reason explained at the end of this section. The zeros and the poles of the SPR transfer function $G(z)$ are expressed as [7]:

$$z_i, z_i^* = \gamma_i\left(cos(\theta_i) \pm j\,sin(\theta_i)\right) \tag{6a}$$

$$p_i, p_i^* = \gamma_i \alpha_i (\cos(\theta_i) \pm j \sin(\theta_i)) \tag{6b}$$

where, γ_i and α_i are the distances of zeros and poles from the origin respectively, and θ_i is the angle of the poles and zeros (Fig. 2).

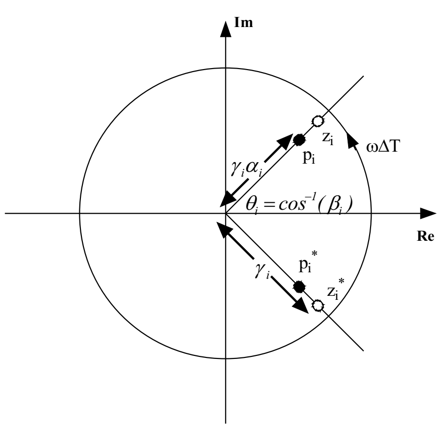

Fig. 2 Pole-zero placement on the unit circle method

The transfer function $G(z)$ can be written as [7]:

$$G(z) = \prod_{i=1}^{N} \frac{(1 - 2\gamma_i \beta_i z^{-1} + \gamma_i^2 z^{-2})}{(1 - 2\gamma_i \alpha_i \beta_i z^{-1} + \gamma_i^2 \alpha_i^2 z^{-2})}, \tag{7a}$$

$$\beta_i = cos(\theta_i), \quad \theta_i = \omega_i \Delta T, \quad \gamma_i > \alpha_i \tag{7b}$$

where, ΔT is the sampling period. The system in Fig. 1 is hyperstable if and only if $G(z)$ is SPR. The real part of a second-order transfer function $G(z)$ is a quadratic equation and can be written in the frequency domain as:

$$\begin{aligned} Re\left(G(j\omega)\right) = {} & 2\gamma^2\left(\alpha^2 + 1\right)cos^2\omega \\ & - 2\gamma\beta\left(\gamma^2\alpha + 1\right)(\alpha + 1)cos\,\omega \\ & + 4\gamma\alpha\beta^2 + \left(1 - \gamma^2\right)\left(1 - \gamma^2\alpha^2\right) \end{aligned} \tag{8a}$$

The transfer function $G(z)$ is SPR if and only if the discriminant of equation (8a) is strictly negative as:

$$\begin{aligned} \Delta = {} & \beta^2\left(\gamma^2\alpha + 1\right)^2(\alpha + 1)^2 - 8\gamma\alpha\beta^2\left(\alpha^2 + 1\right) \\ & - 2\left(1 - \gamma^2\right)\left(1 - \gamma^2\alpha^2\right)\left(\alpha^2 + 1\right) < 0 \end{aligned} \tag{8b}$$

Fig. 3 illustrates that $G(z)$ is SPR for only certain conjugate pole-pairs, where region of SPR is shaded for different γ values in relation to the unit circle. Notice, however, for γ=1, the discriminant of equation (8b) is always positive and $G(z)$ is not an SPR. Thus, the zeros must be located inside and near the unit circle. In reference [5], the region of SPR is illustrated with no smoothing coefficient and with a single smoothing coefficient c_1. In this paper, the region of SPR is illustrated with double smoothing coefficients, c_1 and c_2 (Fig. 3).

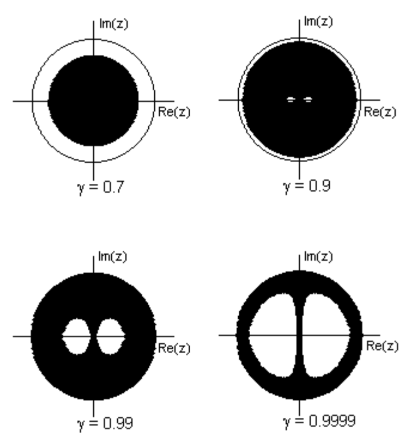

Fig. 3 Illustration of SPR region (shaded area) for second-order system with double smoothing coefficients

3.2 Adaptive Algorithm

The coefficients of the multiple order transfer function $G(z)$ are calculated by convoluting the coefficients of N second order SPRs as follows:

$$\left\{c_i\right\} = \underset{i=1}{\overset{N}{Cv}} \left\{1, -2\gamma_i\beta_i, \gamma_i^2\right\} \tag{9a}$$

$$\left\{a_i\right\} = \underset{j=1}{\overset{N}{Cv}} \left\{1, -2\gamma_i\alpha_i\beta_i, \gamma_i^2\alpha_i^2\right\} \tag{9b}$$

The coefficients c_i and a_i of $G(z)$ are functions of only α_i, β_i, and γ_i, which are the fundamental coefficients to optimize the design of $G(z)$ and are assumed to be independent. The fundamental coefficients are updated using recursive least-squares (RLS) method. A possible objective function for adaptive IIR filtering based on output error is the least-squares function defined as:

$$\xi^d(n) = \sum_{i=0}^{n} \lambda^{n-i} e^2(i) = \sum_{i=0}^{n} \lambda^{n-i} [d(i) - y(i)]^2 \tag{10}$$

where, λ is the forgetting factor and is usually chosen in the range 0 << λ <1. RLS method for adaptive IIR filtering is derived with the following steps:

$$e(n) = d(n) - y(n) = d(n) - \theta^T(n)\phi(n) \tag{11a}$$

$$\theta^T(n) = [\alpha_1(n) \;\; .. \;\; \gamma_1(n) \;\; .. \;\; \beta_1(n) \;\; ..] \tag{11b}$$

$$\phi(n) = [y(n-1) \;\; .. \;\; y(n-N) \;\; x(n) \;\; .. \;\; x(n-M)]^T \tag{11c}$$

$$\psi^T(n) = \left[\frac{\partial v(n)}{\partial \alpha_1(n)} \;\; .. \;\; \frac{\partial v(n)}{\partial \gamma_1(n)} \;\; .. \;\; \frac{\partial v(n)}{\partial \beta_1(n)} \;\; .. \right] \tag{11d}$$

$$S_d(n+1) = \frac{1}{\lambda}\left[S_d(n) - \frac{S_d(n)\psi(n)\psi^T(n)S_d(n)}{\lambda + \psi^T(n)S_d(n)\psi(n)} \right] \tag{11e}$$

where, $S_d(n)$ and $S_d(n+1)$ are $2Nx2N$ matrix and $S_d(n)$ is initially set to unit matrix. Using RLS method, the fundamental coefficients are updated as follows:

$$\theta(n+1) = \theta(n) + v(n)S_d(n+1)\psi(n) \tag{12}$$

The above equations constitute the constrained RLS method for the SHARF algorithm.

4. APPLICATION AND SIMULATION

The self-adjusting SHARF algorithm is applied as an output error identification structure, where filter order matching is assumed. The desired signal $d(n)$ is an autoregressive process generated by a four-pole resonator:

$$\begin{aligned} d(n) &= 2.2618d(n-1) - 2.3341d(n-2) \\ &\quad + 1.2535d(n-3) - 0.3390d(n-4) + \varepsilon(n) \end{aligned} \tag{13}$$

where, $\varepsilon(n)$ is a zero mean unit variance gaussian white process. Using polar coordinates, the poles and the zeros of the resonator are located at:

$$\begin{aligned} p_1, p_1^* &= \alpha_1 \angle \pm 2\pi f_1 \Delta T, \quad f_1 = 10Hz, \quad \alpha_1 = 0.86 \\ p_2, p_2^* &= \alpha_2 \angle \pm 2\pi f_2 \Delta T, \quad f_2 = 20Hz, \quad \alpha_2 = 0.76 \end{aligned} \tag{14a}$$

$$\begin{aligned} z_1, z_1^* &= \gamma_1 \angle \pm 2\pi f_1 \Delta T, \quad f_1 = 10Hz, \quad \gamma_1 = 0.99 \\ z_2, z_2^* &= \gamma_2 \angle \pm 2\pi f_2 \Delta T, \quad f_2 = 20Hz, \quad \gamma_2 = 0.90 \end{aligned} \tag{14b}$$

The poles and zeros correspond to the following values of the fundamental coefficients β_1 and β_2:

$$\beta_1 = 0.8819, \quad \beta_2 = 0.5556 \tag{14c}$$

where, the sampling frequency is $f_s = 128\ Hz$. The coefficients of the transfer function $G(z)$ are calculated from the convolution operations. The fundamental coefficients are updated using the following recursive gradient formulas:

$$\begin{aligned} \frac{\partial v(n)}{\partial \alpha_{1,2}(n)} &= w_{\alpha_{1,2}}(n) + c_2 w_{\alpha_{1,2}}(n-1) + c_3 w_{\alpha_{1,2}}(n-2) \\ &\quad + c_4 w_{\alpha_{1,2}}(n-3) - a_1 \frac{\partial v(n-1)}{\partial \alpha(n)} - a_2 \frac{\partial v(n-2)}{\partial \alpha(n)} \\ &\quad - a_3 \frac{\partial v(n-3)}{\partial \alpha(n)} - a_4 \frac{\partial v(n-4)}{\partial \alpha(n)} \end{aligned} \tag{15a}$$

$$\begin{aligned} \frac{\partial v(n)}{\partial \gamma_{1,2}(n)} &= w_{\gamma_{1,2}}(n) + c_2 w_{\gamma_{1,2}}(n-1) + c_3 w_{\gamma_{1,2}}(n-2) \\ &\quad + c_4 w_{\gamma_{1,2}}(n-3) - a_1 \frac{\partial v(n-1)}{\partial \gamma(n)} - a_2 \frac{\partial v(n-2)}{\partial \gamma(n)} \\ &\quad - a_3 \frac{\partial v(n-3)}{\partial \gamma(n)} - a_4 \frac{\partial v(n-4)}{\partial \gamma(n)} \end{aligned} \tag{15b}$$

$$\begin{aligned} \frac{\partial v(n)}{\partial \beta_{1,2}(n)} &= w_{\beta_{1,2}}(n) + c_2 w_{\beta_{1,2}}(n-1) + c_3 w_{\beta_{1,2}}(n-2) \\ &\quad + c_4 w_{\beta_{1,2}}(n-3) - a_1 \frac{\partial v(n-1)}{\partial \beta(n)} - a_2 \frac{\partial v(n-2)}{\partial \beta(n)} \\ &\quad - a_3 \frac{\partial v(n-3)}{\partial \beta(n)} - a_4 \frac{\partial v(n-4)}{\partial \beta(n)} \end{aligned} \tag{15c}$$

where,

$$\begin{aligned} w_{\alpha_{1,2}}(n) &= -2\gamma_{1,2}\beta_{1,2}y(n-1) \\ &\quad + (2\gamma_{1,2}^2\alpha_{1,2} + 4\gamma_1\gamma_2\alpha_{2,1}\beta_1\beta_2)y(n-2) \\ &\quad - 2\gamma_1\gamma_2\alpha_{2,1}(\gamma_{2,1}\alpha_{2,1}\beta_{1,2} + 2\gamma_{1,2}\alpha_{1,2}\beta_{2,1})y(n-3) \\ &\quad + 2\gamma_1^2\gamma_2^2\alpha_{1,2}\alpha_{2,1}^2 y(n-4) \end{aligned} \tag{16a}$$

$$\begin{aligned} w_{\gamma_{1,2}}(n) &= -2\alpha_{1,2}\beta_{1,2}y(n-1) \\ &\quad + (2\gamma_{1,2}\alpha_{1,2}^2 + 4\gamma_{2,1}\alpha_1\alpha_2\beta_1\beta_2)y(n-2) \\ &\quad - 2\gamma_{2,1}\alpha_1\alpha_2(\gamma_{2,1}\alpha_{2,1}\beta_{1,2} + 2\gamma_{1,2}\alpha_{1,2}\beta_{2,1})y(n-3) \\ &\quad + 2\gamma_{1,2}\gamma_{2,1}^2\alpha_1^2\alpha_2^2 y(n-4) \end{aligned} \tag{16b}$$

$$\begin{aligned} w_{\beta_{1,2}}(n) &= -2\gamma_{1,2}\alpha_{1,2}y(n-1) + 4\gamma_1\gamma_2\alpha_1\alpha_2\beta_{2,1}y(n-2) \\ &\quad - 2\gamma_{1,2}\gamma_{2,1}^2\alpha_{1,2}\alpha_{2,1}^2 y(n-3) \end{aligned} \tag{16c}$$

The convergence of the filter is guaranteed via the desired location of the poles and zeros, which are in the region of SPR with the selected values of $\gamma_1 = 0.99$ and $\gamma_2 = 0.90$ (Fig. 3). After each fundamental coefficient is initially set to a random location within the unit circle, the fourth order output error identification filter converges to its optimum performance after 9k iterations with a forgetting factor λ=0.98. The final locations of the zeros are at $(0.9898\angle \pm 2\pi(9.9966)\Delta T)$ and $(0.8999\angle \pm 2\pi(19.9968)\Delta T)$ while the final locations of the poles are at $(0.8599\angle \pm 2\pi(9.9966)\Delta T)$ and $(0.7601\angle \pm 2\pi(19.9968)\Delta T)$ which are very close to desired locations. The zeros are observed to converge first while poles follow after. The fundamental coefficient β_1 converges rapidly toward its optimum value after 4k iterations while other fundamental coefficients α_1, α_2, γ_1, γ_2, and β_2 experientially converge toward their optimum values after 9k iterations. The adaptation processes for the fundamental coefficients of the fourth-order output error identification filter are illustrated in Fig. 4, 5, and 6. The cost function of the error signal $J_e(n)$ and the cost function of the moving average of the error signal $J_v(n)$ converge asymptotically to the variance of the noise as the algorithm converges to its optimum state (Figure 7). Finally, the output signal $y(n)$ closely approximates the desired signal $d(n)$ as illustrated in Figure 8.

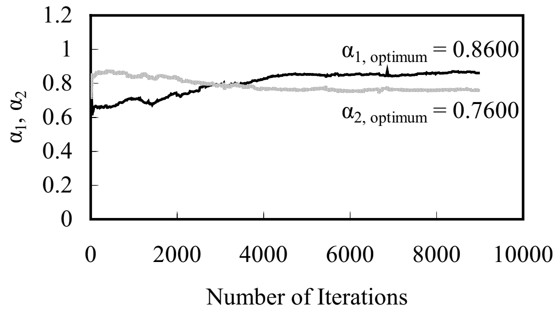

Fig. 4 Adaptive process of α_1 and α_2

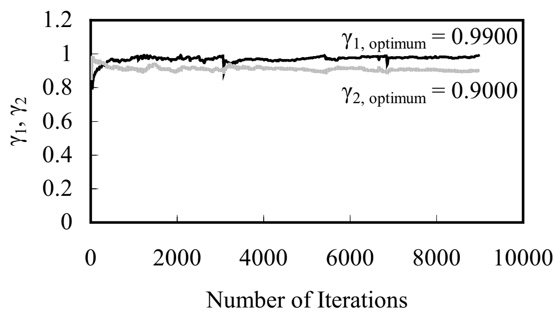

Fig. 5 Adaptive process of γ_1 and γ_2

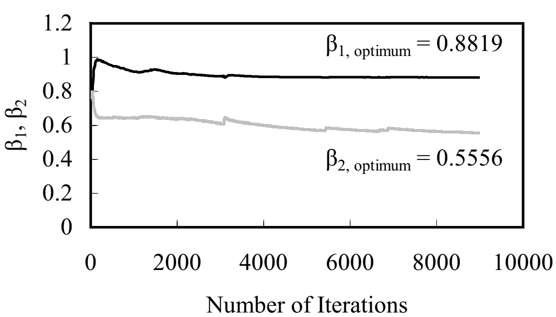

Fig. 6 Adaptive process of β_1 and β_2

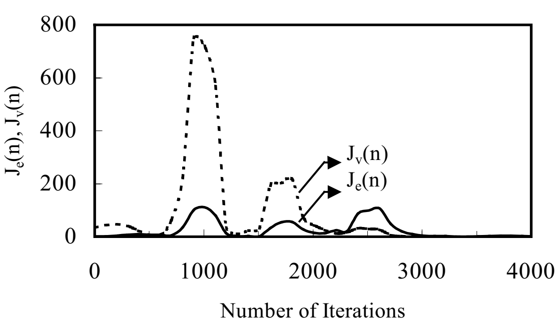

Fig. 7 The cost functions $J_e(n)$ and $J_v(n)$

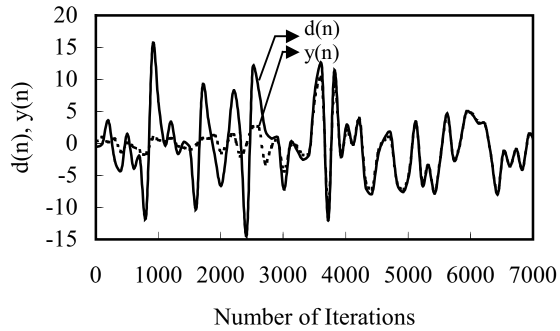

Fig. 8 The desired signal $d(n)$ and the output signal $y(n)$

5. CONCLUSION

When applied without any self-adjusting algorithm, SHARF algorithm presents a burden on the designer in establishing a SPR transfer function. However, when the SPR transfer function is designed using pole-zero placement on the unit circle method, the SHARF algorithm is made self-adjusting through a biased cost function. In addition, the convergence rate of SHARF algorithm is faster with CRLS method. With this unique structure, the adaptive IIR filter has been successfully demonstrated in the output error identification structure.

6. REFERENCES

[1] C. R. Jr. Johnson, and M. G. Larimore, "Comments on and additions to "An adaptive recursive LMS filter," Proceedings of The IEEE, vol. 65, pp. 1399-1401, 1977.

[2] C. R. Johnson, Jr, "A convergence proof for a hyperstable adaptive recursive filter," IEEE Transactions on Information Theory, vol. it-25, no. 6, pp. 745-749, Nov. 1979.

[3] I. D. Landau, "Unbiased recursive identification using model reference adaptive techniques," IEEE Transactions on Automatic Control, vol. 21, pp. 194-202, 1976.

[4] C. R. Johnson, Jr, M. G. Larimore, J. R. Treichler, and B. D. O. Anderson, "SHARF convergence properties," IEEE Transactions on Acoustics, speech, and signal processing, vol. assp-29, no. 3, pp. 659-670, June 1981.

[5] M. G. Larimore, J. R. Treichler, and C. R. Jr. Johnson, "SHARF: An algorithm for adapting IIR digital filters," IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 28, no. 4, pp. 428-440, 1980.

[6] M. Rupp, "Normalization and convergence of gradient-based algorithms for adaptive IIR filters," Signal Processing, vol. 46, no. 1, pp. 15-30, 1995.

[7] O. Sezer and M. Ferdjallah, "Constrained RLS algorithm for narrow band interference rejection from EEG signals during CES", 26th Annual International Conference, IEEE Engineering in Medicine and Biology Society (EMBS), San Francisco, September 1-5, 2004.