International Journal of Electrical, Electronic and Communication Sciences
ISSN: 2517-9438
Vol:10, No:1, 2016

Sparsity-Aware and Noise-Robust Subband Adaptive Filter

Young-Seok Choi

Abstract—This paper presents a subband adaptive filter (SAF) for a system identification where an impulse response is sparse and disturbed with an impulsive noise. Benefiting from the uses of l_1-norm optimization and l_0-norm penalty of the weight vector in the cost function, the proposed l_0-norm sign SAF (l_0-SSAF) achieves both robustness against impulsive noise and much improved convergence behavior than the classical adaptive filters. Simulation results in the system identification scenario confirm that the proposed l_0-norm SSAF is not only more robust but also faster and more accurate than its counterparts in the sparse system identification in the presence of impulsive noise.

Keywords—Subband adaptive filter, l_0-norm, sparse system, robustness, impulsive interference.

I. Introduction

THE normalized least mean square (NLMS) algorithm has become one of the most popular and widely used adaptive filtering algorithms because of its simplicity and robustness. Despite these advantages, the use of NLMS has been limited since it converges poorly for correlated input signals [1]–[4]. To address this problem, various approaches have been presented, such as the recursive least squares algorithm [1], [2], the affine projection algorithm [2], [3], and subband adaptive filtering (SAF) [5]–[9]. Among these, the SAF algorithm allocates the input signals and desired response into almost mutually exclusive subbands. This prewhitening characteristic of SAF allows each subband to converge almost separately so that the subband algorithms obtain faster convergence behavior. On the basis of these characteristics, Lee and Gan proposed a normalized SAF (NSAF) algorithm in [8], [9]. This work improves the convergence speed, while using almost the same computational complexity as the NLMS algorithm. However, the NSAF still suffers from the degradation of convergence performance in cases when an underlying system to be identified is sparse such as network echo path [10], underwater channel [11], and digital TV transmission channel [12]. Motivated by the proportionate step-size adaptive filtering [13], [14], the proportionate NSAF (PNSAF) has been presented to combat poor convergence in a sparse system identification [15]. However, it does not exploit the sparsity condition itself. Moreover, the NSAF and PNSAF algorithms are highly sensitive to impulsive interference, leading to deteriorated convergence behavior. Impulsive interference exits in various applications such as acoustic echo cancellation [17], network cancellation [18], subspace tracking [19], and so on.

Y.-S. Choi is with the Department of Electronic Engineering, Gangneung-Wonju National University, 7 Jukheon-gil, Gangneung, 210-702 Republic of Korea (phone: +82-33-640-2429; fax: +82-33-646-0740; e-mail: yschoi@gwnu.ac.kr).

To address the robustness issue, the sign SAF (SSAF) [16] has been developed based on the l_1-norm optimization, making it robust against impulsive interference. However, its use is limited in case of a sparse system identification. Moreover, the SSAF converges poorly and fails to accelerate the convergence rate with the number of subbands.

Motivated by compressive sensing framework [20], [21] and the least absolute shrinkage and selection operator (LASSO) [22], a variety of adaptive filtering algorithms which incorporates the sparsity of a system has been developed unlike the proportionate adaptive filtering approach [23]–[26]. Especially, the l_0-norm of a system is able to represent the actual sparsity [24]–[26]. In this paper, a l_0-norm constraint SSAF (l_0-SSAF) is presented, aiming at developing a sparsity-aware SSAF. With this in mind, by integrating the l_0-norm penalty of the current weight vector into the l_1-norm optimization criterion, the l_0-SSAF benefits both much improved convergence for a sparse system identification and robustness against impulsive noise. In addition, the l_0-SSAF is derived from a l_1-norm optimization of the *a priori* error instead of the *a posteriori* error used in the SSAF. Thus, there is no need to approximate the *a posteriori* error with the *a priori* error to derive the update recursion of the l_0-SSAF. Simulation results show that the l_0-SSAF is superior to the conventional SAFs in identifying a sparse system in the presence of severe impulsive noise.

II. Conventional SAFs

Consider a desired signal $d(n)$ that arise from the system identification model

$$d(n) = \mathbf{u}(n)\mathbf{w}^{\circ} + v(n), \tag{1}$$

where $\mathbf{w}^{\circ}$ is a column vector for the impulse response of an unknown system that we wish to estimate, $v(n)$ accounts for measurement noise with zero mean and variance σ_v^2 and $\mathbf{u}(n) = [u(n)\ \ u(n-1)\ \ \cdots u(n-M+1)]$ is a $1 \times M$ input vector.

Fig. 1 shows the structure of the NSAF, where the desired signal $d(n)$ and output signal $u(n)$ are partitioned into N subbands by the analysis filters $H_0(z), H_1(z), \ldots, H_{N-1}(z)$. The resultant subband signals, $d_i(n)$ and $y_i(n)$ for $i = 0, 1, \ldots, N-1$, are critically decimated to a lower sampling rate commensurate with their bandwidth. Here, the variable n to index the original sequences, and k to index the decimated sequences are used for all signals. Then, the decimated desired

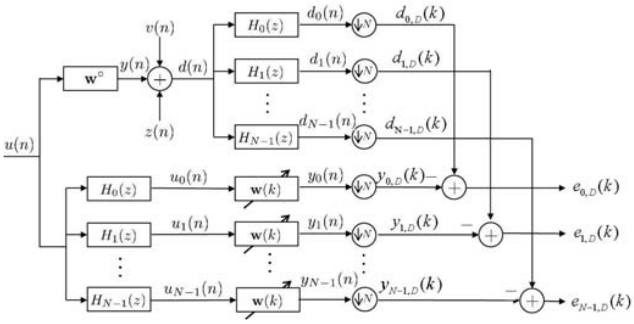

Fig. 1 Subband structure used in the proposed SAF

signal and the decimated filter output signal at each subband are defined as $d_{i,D}(k) = d_i(kN)$ and $y_{i,D}(k) = \mathbf{u}_i(k)\mathbf{w}(k)$, where $\mathbf{u}_i(k)$ is the input data vector for the *i*th subband such that

$$\mathbf{u}_i(k) = [u_i(kN), u_i(kN-1), \ldots, u_i(kN-M+1)]$$

and $\mathbf{w}(k) = [w_0(k), w_1(k), \ldots, w_{M-1}(k)]^T$ denotes an estimate for $\mathbf{w}^\circ$. Then, the decimated subband error vector is given by

$$e_{i,D}(k) = d_{i,D}(k) - y_{i,D}(k) = d_{i,D}(k) - \mathbf{u}_i(k)\mathbf{w}(k). \quad (2)$$

In [8], the authors have presented the update recursion of the NSAF algorithm is given by

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu \sum_{i=0}^{N-1} \frac{\mathbf{u}_i^T(k)}{\|\mathbf{u}_i(k)\|^2} e_{i,D}(k), \quad (3)$$

where μ is a step-size parameter. Then, the estimation error in all the *N* subbands, i.e., $\mathbf{e}_D(k) = [e_{0,D}(k), \ldots, e_{N-1,D}(k)]^T$ can be written in a compact form as

$$\mathbf{e}_D(k) = \mathbf{d}_D(k) - \mathbf{U}(k)\mathbf{w}(k), \quad (4)$$

where the $N \times M$ subband data matrix $\mathbf{U}(k)$ and the $N \times 1$ desired response vector $\mathbf{d}(k)$ are given by

$$\mathbf{U}(k) = [\mathbf{u}_0(k),\, \mathbf{u}_1(k)\, \ldots,\, \mathbf{u}_{N-1}(k)]^T, \quad (5)$$

and

$$\mathbf{d}_D(k) = [d_{0,D}(k),\, d_{1,D}(k), \ldots, d_{N-1,D}(k)]^T. \quad (6)$$

More recently, the SSAF [16] has been obtained from the following optimization criterion:

$$\min_{\mathbf{w}(k+1)} \|\mathbf{d}_D(k) - \mathbf{U}(k)\mathbf{w}(k+1)\|_1$$
$$\text{subject to } \|\mathbf{w}(k+1) - \mathbf{w}(k)\|_2^2 \le \mu^2, \quad (7)$$

where $\|\cdot\|_1$ denotes the l_1-norm and μ^2 is a parameter which prevents the weight coefficient vectors from abrupt change. Using Lagrange multipliers to solve the constrained optimization problem and utilizing the accessible $\mathbf{e}_D(k)$ instead of unavailable *a posteriori* error, i.e., $\mathbf{d}_D(k) - \mathbf{U}(k)\mathbf{w}(k+1)$, the update recursion of the SSAF is formulated as

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu \frac{\mathbf{U}^T(k)\mathrm{sgn}[\mathbf{e}_D(k)]}{\sqrt{\sum_{i=0}^{N-1} \mathbf{u}_i(k)\mathbf{u}_i^T(k) + \delta}}, \quad (8)$$

where δ is a regularization parameter and $\mathrm{sgn}(\cdot)$ denotes the sign function, where $\mathrm{sgn}[\mathbf{e}_D(k)] = [\mathrm{sgn}(e_{0,D}(k)), \ldots, \mathrm{sgn}(e_{N-1,D}(k))]^T$.

III. Proposed l_0-Norm Constraint SSAF (l_0-SSAF)

Our objective is to cope with the sparsity of an underlying system while inheriting robustness from the l_1-norm optimization criterion. Our approach is to find a new weight vector, $\mathbf{w}(k+1)$, that minimizes the l_1-norm of the *a priori* error vector with the l_0-norm penalty of the current weight vector $\mathbf{w}(k)$ as follows:

$$\min_{\mathbf{w}} J(k) \triangleq \min_{\mathbf{w}} \left[\|\mathbf{e}_D(k)\|_1 + \gamma\|\mathbf{w}(k)\|_0\right]. \quad (9)$$

where $\|\cdot\|_0$ denotes the l_0-norm and $\gamma(>0)$ is a regularization parameter which governs the comprise between the effect of the l_0-norm penalty term and the error vector related term. Note that the *a priori* error $\mathbf{e}_D(k)$ is used unlike the SSAF, leading to no approximation of the *a posteriori* error with the *a priori* error.

Taking derivative of $J(k)$ with respect to $\mathbf{w}(k)$, it leads to

$$\begin{aligned}\nabla_{\mathbf{w}(k)} J(k) &= -\,\mathbf{U}^T(k)\mathrm{sgn}(\mathbf{e}_D(k)) + \gamma \frac{\partial \|\mathbf{w}(k)\|_0}{\partial \mathbf{w}(k)} \\ &\triangleq -\,\mathbf{U}^T(k)\mathrm{sgn}(\mathbf{e}_D(k)) + \gamma \mathbf{f}_\beta(\mathbf{w}(k)),\end{aligned} \quad (10)$$

where $\mathbf{f}_\beta(\mathbf{w}(k)) \triangleq [f_\beta(w_0(k)),\, f_\beta(w_1(k)) \ldots, f_\beta(w_{M-1}(k))]^T$. To avoid a Non-polynomial hard problem from the l_0-norm minimization, the l_0-norm penalty is often approximated as follows [27]:

$$\|\mathbf{w}(k)\|_0 \approx \sum_{i=0}^{M-1} \left(1 - e^{-\beta|w_i(k)|}\right), \quad (11)$$

where the parameter β plays a role adjusting the degree of zero attraction. A *m*th component of the gradient for (11) is given by

$$\frac{\partial \|\mathbf{w}(k)\|_0}{\partial w_m(k)} = f_\beta(w_m(k)) = \beta \mathrm{sgn}[w_m(k)] e^{-\beta|w_m(k)|} \quad (12)$$

$\forall 0 \le m \le M-1$. To reduce the computational cost in (12), the 1st order Taylor series expansion of the exponential function is employed

$$e^{-\beta|x|} \approx \begin{cases} 1 - \beta|x|, & |x| \le \frac{1}{\beta} \\ 0, & \text{elsewhere} \end{cases} \quad (13)$$

Then, a gradient (12) is computed as

$$f_\beta(w_m(k)) = \begin{cases} \beta^2 w_m(k) + \beta, & -\frac{1}{\beta} \le w_m(k) < 0 \\ \beta^2 w_m(k) - \beta, & 0 < w_m(k) \le \frac{1}{\beta} \\ 0, & \text{elsewhere.} \end{cases} \quad (14)$$

Finally, the update recursion of the l_0-SSAF is given by

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu \mathbf{U}^T(k)\mathrm{sgn}(\mathbf{e}_D(k)) - \kappa \mathbf{f}_\beta(\mathbf{w}(k)), \quad (15)$$

where μ is the step-size parameter and $\kappa = \mu\gamma$.

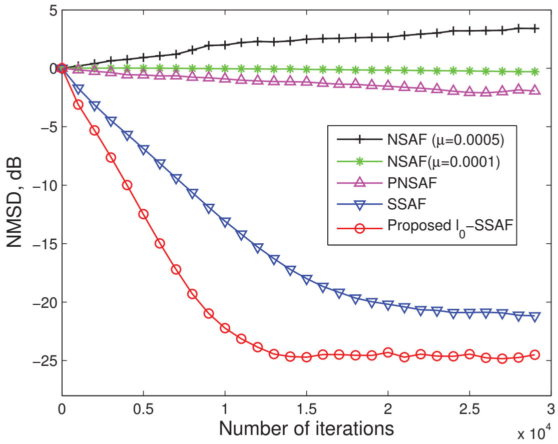

Fig. 2 NMSD Learning curves of the NSAF, PNSAF, SSAF, and l_0-SSAF algorithms [N=4, SIR=-30 dB, Input: Gaussian AR(1) with pole at 0.9]

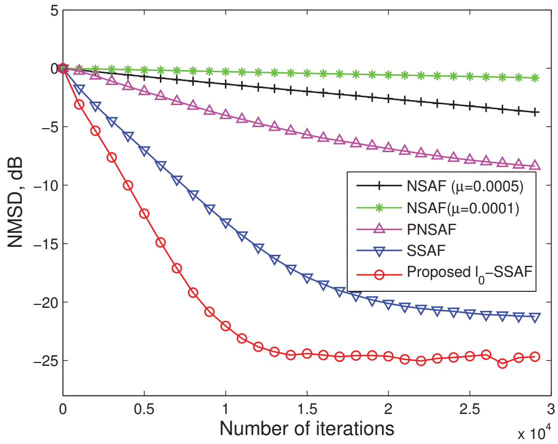

Fig. 4 NMSD Learning curves of the NSAF, PNSAF, SSAF, and l_0-SSAF algorithms [N=4, SIR=-10 dB, Input: Gaussian AR(1) with pole at 0.9]

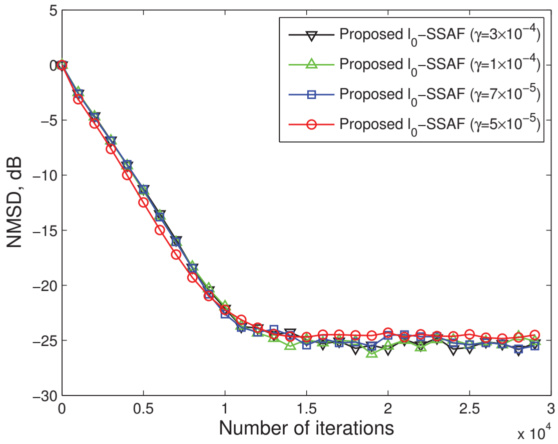

Fig. 3 NMSD Learning curves of the l_0-SSAF algorithm with various γ values [N=4, SIR=-30 dB, Input: Gaussian AR(1) with pole at 0.9]

IV. Simulation Results

To validate the performance of the proposed l_0-SSAF, computer simulations are carried out in a system identification scenario in which the unknown system is randomly generated. The length of the unknown system is $M = 128$, where S of them are non-zero. The non-zero filter weights are positioned randomly and their values are taken from a Gaussian distribution $\mathcal{N}(0, 1/S)$. Here, the sparse systems of the sparsity $S = 4, 8, 16, 32$ are considered. The adaptive filter and the unknown system are assumed to have the same number of taps. The input signals $u(n)$ are obtained by filtering a white, zero-mean, Gaussian random sequence through a first-order system

$$G_1(z) = 1/(1 - 0.9z^{-1}), \tag{16}$$

or a second-order system

$$G_2(z) = (1 + 0.5z^{-1} + 0.8z^{-2})/(1 - 0.9z^{-1}). \tag{17}$$

A measurement noise $v(n)$ with white Gaussian distribution is added to the system output $y(n)$ such that the signal-to-noise ratio (SNR) is 20 dB, where the SNR is defined as

$$\text{SNR} = 10\log_{10}\left(\frac{E[y^2(n)]}{E[v^2(n)]}\right),$$

where $y(n) = \mathbf{u}(n)\mathbf{w}^\circ$.

An impulsive noise $z(n)$ is added to the system output $y(n)$ with the signal-to-interference ratio (SIR) of -30 or -10 dB correspondingly. The impulsive noise is modeled by a Bernoulli-Gaussian (BG) distribution [17], which is obtained as the product of a Bernoulli distribution and a Gaussian one, i.e., $z(n) = \omega(n)\eta(n)$, where $\omega(n)$ is a Bernoulli process with a probability mass function given by $P(\omega) = 1 - Pr$ for $\omega = 0$ and $P(\omega) = Pr$ for $\omega = 1$. In addition, $\eta(n)$ is an additive white Gaussian noise with zero mean and variance $\sigma_\eta^2 = 1000\sigma_y^2$. Here, $Pr = 0.01$ is used. In order to compare the convergence performance, the normalized mean square deviation (NMSD),

$$\text{Normalized MSD} = \text{E}\left[\frac{\|\mathbf{w}^\circ - \mathbf{w}(k)\|^2}{\|\mathbf{w}^\circ\|^2}\right],$$

is taken and averaged over 50 independent trials. The cosine-modulated filter banks [28] with the subband numbers of $N = 4$ is used in the simulations. The prototype filter of length $L = 32$ is used. The parameters used in simulations are as follows: NSAF ($\mu = 0.0005$ or 0.0001), SSAF ($\mu = 0.0005$, $\delta = 0.001$), PNSAF ($\mu = 0.0005$, $\rho = 0.01$), l_0-SSAF ($\mu = 0.0003$, $\beta = 20$). The γ of the l_0-norm SSAF is obtained by repeated trials to minimize the steady-state NMSD.

Fig. 2 shows the NMSD learning curves of the NSAF, PNSAF, SSAF, and l_0-norm SSAF algorithms in the case of SIR $= -30$ dB. For the l_0-SSAF, $\gamma = 5 \times 10^{-5}$ is chosen. Compared to the conventional SAF algorithms, the proposed l_0-SSAF yields remarkably improved convergence performance in terms of the convergence rate and the steady-state misalignment.

In Fig. 3, to verify the effect of γ on convergence performance, the NMSD curves of the l_0-SSAF for different

γ values are illustrated in the case of $\mathrm{SIR} = -30$ dB. With different γ values ($\gamma = 3 \times 10^{-4}$, 1×10^{-4}, 7×10^{-5}, and 5×10^{-5}), the l_0-SSAF is not excessively sensitive to γ. The analysis for an optimal γ value remains a future work.

Fig. 4 illustrates the NMSD learning curves of the NSAF, PNSAF, SSAF, and l_0-norm SSAF algorithms under $\mathrm{SIR} = -10$ dB. The same γ value with Fig. 2 is chosen. In figure, the similar results shown in Fig. 2 are observed.

V. Conclusion

This paper has proposed a robust and sparse-aware SSAF algorithm which incorporates the sparsity condition of a system into the l_1-norm optimization criterion of the *a priori* error vector. By utilizing the l_0-norm penalty of the current weight vector and approximating it to avoid a Non-polynomial hard problem, the update recursion of the proposed l_0-norm SSAF is obtained while reducing the computational cost using Taylor series expansion. The simulation results indicate that the proposed l_0-SSAF achieves highly improved convergence performance over the conventional SAF algorithms where a system is not only sparse but also disturbed with impulsive noise.

References

[1] S. Haykin, *Adaptive Filter Theory*, 4th ed., Upper Saddle River, NJ: Prentice Hall, 2002.

[2] A. H. Sayed, *Fundamentals of Adaptive Filtering*, New York: Wiley, 2003.

[3] P. S. R. Diniz, *Adaptive Filtering: Algorithms and Practical Implementations*, 3rd ed. Boston: Kluwer, 2008.

[4] M. M. Sondhi, "The History of Echo Cancellation," *IEEE Transactions on Signal Processing Magazine*, vol. 23, no. 5, pp. 1730–1740, 2006.

[5] A. Gilloire and M. Vetterli, "Adaptive filtering in subbands with critical sampling: analysis, experiments, and application to acoustic echo cancellation," *IEEE Transactions on Signal Processing*, vol. 40, no. 8, pp. 1862–875, 1992.

[6] M. D. Couriville and P. Duhamel, "Adaptive filtering in subbands using a weighted criterion," *IEEE Transactions on Signal Processing*, vol. 46, no. 9, pp. 2359–2371, 1998.

[7] S. S. Pradhan and V. U. Reddy, "A new approach to subband adaptive filtering," *IEEE Transactions on Signal Processing*, vol. 47, no. 3, pp. 655–664, 1999.

[8] K. A. Lee and W. S. Gan, "Improving convergence of the NLMS algorithm using constrained subband updates," *IEEE Signal Processing Letters*, vol. 11, no. 9, pp. 736–739, 2004.

[9] K. A. Lee and W. S. Gan, "Inherent decorrelating and least perturbation properties of the normalized subband adaptive filter," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4475–4480, 2006.

[10] D. L. Duttweiler, "Proportionate normalized least-mean-squares adaptation in echo cancellers," *IEEE Transactions on Speech Audio Processing*, vol. 8, no. 5, pp. 508–518, 2000.

[11] W. Li and J. C. Preisig, "Estimation of rapidly time-varying sparse channels," *IEEE Journal of Ocean Engineering*, vol. 32, no. 4, pp. 929–939, 2007.

[12] W. F. Schreiber, "Advanced television systems for terrestrial broad-casting: some problems and some proposed solutions," *Proceedings of IEEE*, vol. 83, no. 6, pp. 958–981, 1995.

[13] S. L. Gay, "An efficient, fast converging adaptive filter for network echo cancellation," in *Proceedings of the Asilomar Conference on Signals, Systems, Computer*, pp. 394–398, Nov. 1998.

[14] H. Deng and M. Doroslovacki, "Improving convergence of the PNLMS algorithm for sparse impulse response identification," *IEEE Signal Processing Letters*, vol. 12, no. 3, pp. 181–184, 2005.

[15] M. S. E. Abadi, "Proportionate normalized subband adaptive filter algorithms for sparse system identification," *Signal Processing*, vol. 89, no. 7, pp. 1467–1474, 2009.

[16] J. Ni, and F. Li, "Variable regulaisation parameter sign subband adaptive filter," *Electronis Letters*, vol. 89, no. 7, pp. 1605–1607, 2010.

[17] L. R. Vega, H. Ray, J. Benesty, and S. Tressens, "A new robust variable step-size NLMS algorithm," *IEEE Transions on Signal Processing*, vol. 56, no. 5, pp. 1878–1893, 2008.

[18] Z. Yang, Y. R. Zheng, and S. L. Grant, "Proportionate affine projection sign algorithms for network echo cancellation," *IEEE Transactions on Audio, Speech, Language Processing*, vol. 19, no. 8, pp. 2273–2284, 2011.

[19] B. Liao, Z. G. Zhang, and S. C. Chan, "A new robust Kalman filter-based subspace tracking algorithm in an impulsive noise environment," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 57, no. 9, pp. 740744, 2010.

[20] D. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[21] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[22] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of Royal statistical Society Series B*, vol. 58, pp. 267–288, 1996.

[23] Y. Chen, Y. Gu, and A. O. Hero, "Sparse LMS for system identification," in *Proceeding of the International Conference on Acoustic, Speech, and Signal Process (ICASSP '09)*, pp. 3125–3128, April 2009.

[24] Y. Gu, J. Jin, and S. Mei, "l_0 norm constraint LMS algorithm for sparse system identification," *IEEE Signal Processing Letters*, vol. 16, no. 9, pp. 774–777, 2009.

[25] J. Jin, Y. Gu, and S.Mei, "A stochastic gradient approach on compressive sensing signal reconstruction based on adaptive filtering framework," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 409–420, 2010.

[26] E. M. Eksioglu and A. K. Tanc, "RLS Algorithm with convex regularization," *IEEE Signal Processing Letters*, vol. 18, no. 8, pp. 470–473, 2011.

[27] P. S. Bradley and O. L. Mangasarian, "Feature selection via concave minimization and support vector machines," in *Proceeding of International Conference on Machine Learning (ICML '98)*, pp. 82–90, 1998.

[28] P. P. Vaidyanathan, *Multirate Systems and Filterbanks*, Englewood Cliffs, NJ: Prentice-Hall, 1993.