Improved Adaptive Sparse Channel Estimation Based on the Least Mean Square Algorithm

Guan Gui, Wei Peng and Fumiyuki Adachi
Department of Communication Engineering
Graduate School of Engineering,
Tohoku University
Sendai, Japan
{gui,peng}@mobile.ecei.tohoku.ac.jp, adachi@ecei.tohoku.ac.jp

***Abstract*—Least mean square (LMS) based adaptive algorithms have been attracted much attention since their low computational complexity and robust recovery capability. To exploit the channel sparsity, LMS-based adaptive sparse channel estimation methods, e.g., zero-attracting LMS (ZA-LMS), reweighted zero-attracting LMS (RZA-LMS) and L_p-norm sparse LMS (LP-LMS), have also been proposed. To take full advantage of channel sparsity, in this paper, we propose several improved adaptive sparse channel estimation methods using L_p-norm normalized LMS (LP-NLMS) and L_0-norm normalized LMS (L0-NLMS). Comparing with previous methods, effectiveness of the proposed methods is confirmed by computer simulations.**

Keywords—least mean square (LMS); adaptive sparse channel estimation; normalized LMS (NLMS); sparse penalty; compressive sensing (CS).

I. INTRODUCTION

The demand for high-speed data services is getting more insatiable due to the number of wireless subscribers roaring increase. Various portable wireless devices, e.g., smart phones and laptops, have generated rising massive data traffic [1]. It is well known that the broadband transmission is an indispensable technique in the next generation communication systems [2-7]. However, the broadband signal is susceptible to interference by frequency-selective fading. In the sequel, the broadband channel is described by a sparse channel model in which multipath taps are widely separated in time, thereby create a large delay spread [7-12]. In other words, unknown channel impulse response (CIR) in broadband wireless communication system is often described by sparse channel model, supporting by a few large coefficients. That is to say, most of channel coefficients are zero or close to zero while only a few channel coefficients are dominant (large value) to support the channel. A typical example of sparse channel is shown in Fig. 1, where the number of dominant channel taps is 4 while the length of channel is 16.

Traditional least mean square (LMS) algorithm is one of the most popular methods for adaptive system identification [13], e.g. channel estimation. Indeed, LMS-based adaptive channel estimation can be easily implemented by LMS-based filter due to its low computational complexity or fast convergence speed. However, the LMS-based method never takes advantage of channel sparse structure as prior information and then it may loss some estimation performance.

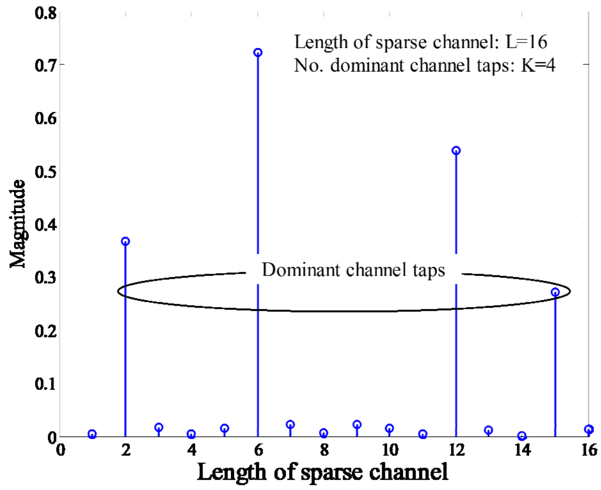

Fig. 1. A typical example of sparse multipath channel.

Recently, many algorithms have been proposed to take advantage of sparse nature of the channel. For example, based on the theory of compressive sensing (CS) [14], [15], various sparse channel estimation methods have been proposed in [16-19]. For one thing, these CS-based sparse channel estimation methods require that the training signal matrices satisfy the restricted isometry property (RIP) [20]. However, design these kinds of training matrices is non-deterministic polynomial-time (NP) hard problem [21]. For another thing, some of these methods achieve robust estimation at the cost of high computational complexity, e.g., sparse channel estimation using least-absolute shrinkage and selection operator (LASSO) [22]. To avoid the high computational complexity on sparse channel estimation, a variation of the LMS algorithm with L_1-norm penalty term in the LMS cost function has also been developed in [23]. The L_1-norm penalty was incorporated into the cost function of conventional LMS algorithm, which resulted in two sparse LMS algorithms, namely zero-attracting LMS (ZA-LMS) and reweighted-zero-attracting LMS (RZALMS) [23]. Following this idea, adaptive sparse channel estimation method using L_p-norm sparse penalty LMS (LP-LMS) was also proposed in order to further improve estimation performance [24].

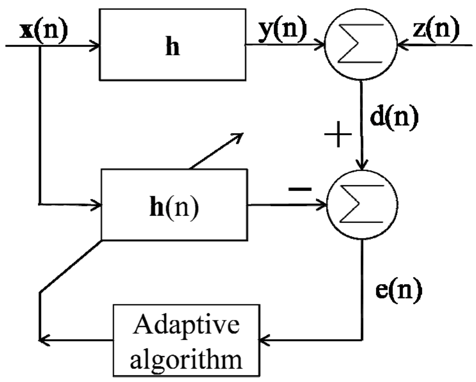

Fig.2. A sparse multipath communication system.

In this paper, we propose an improved sparse channel estimation method by introducing L_0-norm LMS algorithm (L0-LMS) which was proposed in [25]. In the following, based on above mentioned sparse LMS algorithms in [24-25], we propose two kinds of improved adaptive sparse channel estimation methods using L_p-norm normalized LMS (LP-NLMS) and L_0-norm normalized LMS (L0-NLMS), respectively. Effectiveness of propose approaches will be evaluated by computer simulations.

Section II introduces sparse system model and problem formulation. Section III discusses adaptive sparse channel estimation methods using different LMS-based algorithms. In section V, computer simulation results are given and their performance comparisons are also discussed. Concluding remarks are resented in Section V.

II. System Model

Consider a sparse multipath communication system, as shown in Fig. 2, the input signal $\mathbf{x}(n)$ and output signal $d(n)$ are related by

$$d(n) = \mathbf{h}^T\mathbf{x}(n) + z(n), \tag{1}$$

where $\mathbf{h} = [h_0, h_1, \dots, h_{N-1}]^T$ is a N-length unknown sparse channel vector which is supported only by K dominant channel taps, $\mathbf{x}(n) = [x(n), x(n-1), \dots, x(n-N+1)]^T$ is N-length input signal vector and $z(n)$ is an additive noise variable at time n. The object of LMS adaptive filter is to estimate the unknown sparse channel coefficients $\mathbf{h}$ using the input signal $\mathbf{x}(n)$ and output signal $d(n)$. According to Eq. (1), channel estimation error $e(n)$ is written as

$$e(n) = d(n) - \mathbf{h}^T(n)\mathbf{x}(n), \tag{2}$$

where $\mathbf{h}(n)$ is the LMS adaptive channel estimator. Based on Eq. (2), LMS cost function can be given by

$$L(n) = \frac{1}{2}e^2(n). \tag{3}$$

Hence, the update equation of LMS adaptive channel estimation is derived by

$$\mathbf{h}(n+1) = \mathbf{h}(n) + \mu e(n)\mathbf{x}(n)\,, \tag{4}$$

where $\mu \in (0, 2/\gamma_{\max})$ is a step size of gradient descend step-size and $\gamma_{\max}$ is the maximum eigenvalue of the covariance matrix of $\mathbf{x}(n)$.

III. LMS-based Adaptive Sparse Channel Estimation

From the above Eq. (4), we can find that the LMS-based channel estimation method never take advantage of sparse structure in $\mathbf{h}$. To a better understood, the standard LMS-based channel estimation can be concluded as

$$\mathbf{h}(n+1) = \mathbf{h}(n) + \text{adaptive update}. \tag{5}$$

Unlike the standard LMS method, we exploit the channel sparsity by introducing several L_p-norm ($0 \le p \le 1$) penalties to LMS-based cost function. Hence, the LMS-based adaptive sparse channel estimation can be written as

$$\mathbf{h}(n+1) = \mathbf{h}(n) + \text{adaptive update} + \text{sparse penalty}. \tag{6}$$

From above update Eq. (6), the objective of this paper is introducing different sparse penalties to take the advantage of sparse structure as for prior information.

A. ZA-LMS algorithm

To exploit the channel sparsity in CIR, the cost function of ZA-LMS [23] is given by

$$L_{ZA}(n) = \frac{1}{2}e^2(n) + \lambda_{ZA}\|\mathbf{h}(n)\|_1, \tag{7}$$

where λ_{ZA} is a regularization parameter which balances the adaptive estimation error and sparse penalty of $\mathbf{h}(n)$. The corresponding update equation of ZA-LMS is

$$\begin{aligned}\mathbf{h}(n+1) &= \mathbf{h}(n) - \mu\frac{\partial L_{ZA}(n)}{\partial \mathbf{h}(n)} \\ &= \mathbf{h}(n) + \mu e(n)\mathbf{x}(n) - \rho_{ZA}\mathrm{sgn}\{\mathbf{h}(n)\}\,,\end{aligned} \tag{8}$$

where $\rho_{ZA} = \mu\lambda_{ZA}$ and $\mathrm{sgn}\{\cdot\}$ is a component-wise function which is defined as

$$\mathrm{sgn}(h) = \begin{cases} h/|h|, & \text{when } h \neq 0 \\ 0, & \text{when } h = 0\text{'} \end{cases} \tag{9}$$

where the h is one of taps of $\mathbf{h}$. From the update equation in Eq. (8), the second term attracts the small filter coefficients to zero, which speed up convergence when the most of the channel coefficients $\mathbf{h}$ are zeros.

B. RZA-LMS algorithm

The ZA-LMS cannot distinguish between zero taps and non-zero taps since all the taps are forced to zero uniformly; therefore, its performance will degrade in less sparse systems. Motivated by reweighted L_1-minimization sparse recovery algorithm [26], Chen et. al. proposed a heuristic approach to zero-attracting LMS (RZA-LMS) [23]. The cost function of RZA-LMS is given by

$$L_{RZA}(n) = \frac{1}{2}e^2(n) + \lambda_{RZA}\sum_{i=1}^{N}\log(1+\varepsilon_{RZA}|h_i|), \quad (10)$$

where $\lambda_{RZA} > 0$ is a regularization parameter which trades off the estimation error and channel sparsity. The corresponding update equation is

$$\begin{aligned}\mathbf{h}(n+1) &= \mathbf{h}(n) - \mu\frac{\partial L_{RZA}(n)}{\partial \mathbf{h}(n)} \\ &= \mathbf{h}(n) + \mu e(n)\mathbf{x}(n) \\ &\quad -\mu\lambda_{RZA}\varepsilon_{RZA}\sum_{i=1}^{N}\frac{\mathrm{sgn}(|h_i(n)|)}{1+\varepsilon_{RZA}|h_i(n)|} \\ &= \mathbf{h}(n) + \mu e(n)\mathbf{x}(n) - \rho_{RZA}\frac{\mathrm{sgn}(\mathbf{h}(n))}{1+\varepsilon_{RZA}|\mathbf{h}(n)|}, \quad (11)\end{aligned}$$

where $\rho_{RZA} = \mu\lambda_{RZA}\varepsilon_{RZA}$ is a parameter which depends on step-size μ, regularization parameter λ_{RZA} and threshold ε_{RZA}, respectively. In the second term of Eq. (11), if magnitudes of $h_i(n)$, $i = 1,2,\dots,N$ are smaller than $1/\varepsilon_{RZA}$, then these channel coefficients will be replaced by zeros.

C. LP-LMS and LP-NLMS algorithms

Following the idea in Eq. (11), LP-LMS based adaptive sparse channel estimation method has been proposed in [24].The cost function of LP-LMS is given by

$$L_{LP}(n) = \frac{1}{2}e^2(n) + \lambda_{LP}\|\mathbf{h}(n)\|_p, \quad (12)$$

where $\lambda_{LP} > 0$ is a regularization parameter which balances the estimation error and channel sparsity. The corresponding update equation of LP-LMS is

$$\begin{aligned}\mathbf{h}(n+1) &= \mathbf{h}(n) - \mu\frac{\partial L_{LP}(n)}{\partial \mathbf{h}(n)} \\ &= \mathbf{h}(n) + \mu e(n)\mathbf{x}(n) - \rho_{LP}\frac{\|\mathbf{h}(n)\|_p^{1-p}\mathrm{sgn}(\mathbf{h}(n))}{\varepsilon_{LP}+|\mathbf{h}(n)|^{1-p}}, \quad (13)\end{aligned}$$

where $\rho_{LP} = \mu\lambda_{LP}$ which is decided by step-size μ and regularization parameter λ_{LP}, and $\varepsilon_{LP} > 0$. According to the updating equation of LP-LMS in Eq. (13), the update equation of LP-NLMS can be derived as

$$\begin{aligned}\mathbf{h}(n+1) &= \mathbf{h}(n) + \mu_N\frac{e(n)\mathbf{x}(n)}{\mathbf{x}^H(n)\mathbf{x}(n)} \\ &\quad -\rho_{LPN}\frac{\|\mathbf{h}(n)\|_p^{1-p}\mathrm{sgn}(\mathbf{h}(n))}{\varepsilon_{LPN}+|\mathbf{h}(n)|^{1-p}}, \quad (14)\end{aligned}$$

where $\varepsilon_{LPN} > 0$ and μ_N is a step size which controls the gradient descend speed and $\rho_{LPN} = \mu_N\lambda_{LPN}$ is a parameter which depends on step-size and regularization parameter.

D. L0-LMS and L0-NLMS algorithms

Consider the L_0-norm penalty on the cost function of LMS so that it can produce sparse channel estimator since this penalty term forces the channel taps values of $\mathbf{h}(n)$ to approach zero. Then, the cost function of L0-LMS is given by

$$L_{L0}(n) = \frac{1}{2}e^2(n) + \lambda_{L0}\|\mathbf{h}(n)\|_0, \quad (15)$$

where $\lambda_{L0} > 0$ is a regularization parameter. Since solving the L_0-norm minimization is a Non-Polynomial (NP) hard problem, we replace it with approximate continuous function

$$\|\mathbf{h}\|_0 \approx \sum_{i=1}^{N}(1-e^{-\beta|h_i|}), \quad (16)$$

According to the approximate function in Eq. (16), L0-LMS cost function can be changed as

$$L_{L0}(n) = \frac{1}{2}e^2(n) + \lambda_{L0}\sum_{i=1}^{N}(1-e^{-\beta|h_i|}). \quad (17)$$

Then, the update equation of L0-LMS based adaptive sparse channel estimation can be derived as

$$\begin{aligned}\mathbf{h}(n+1) &= \mathbf{h}(n) - \mu\frac{\partial L_{L0}(n)}{\partial \mathbf{h}(n)} \\ &= \mathbf{h}(n) + \mu e(n)\mathbf{x}(n) \\ &\quad -\rho_{L0}\beta\,\mathrm{sgn}(\mathbf{h}(n))e^{-\beta|\mathbf{h}(n)|}, \quad (18)\end{aligned}$$

where $\rho_{L0} = \mu\lambda_{L0}$. It is worth mention that the exponential function in Eq. (18) causes high computational complexity. To reduce the computational complexity, the first-order Taylor series expansion of exponential functions is taken into consideration

$$e^{-\beta|h|} \approx \begin{cases}1-\beta|h|, & \text{when } |h| \le 1/\beta \\ 0, & \text{others}\end{cases}. \quad (19)$$

Then, the update equation of L0-LMS based adaptive sparse channel estimation can be derived as

$$\mathbf{h}(n+1) = \mathbf{h}(n) + \mu e(n)\mathbf{x}(n) - \rho_{L0}J(\mathbf{h}(n)), \quad (20)$$

where $J(h)$ is defined as

$$J(h) \approx \begin{cases}2\beta^2 h - 2\beta\,\mathrm{sgn}(h), & \text{when } |h| \le 1/\beta \\ 0, & \text{others}\end{cases}. \quad (21)$$

Based on this algorithm in Eq. (20), we further propose an improved adaptive sparse channel estimation method by using L0-NLMS algorithm

$$\mathbf{h}(n+1) = \mathbf{h}(n) + \mu_N\frac{e(n)\mathbf{x}(n)}{\mathbf{x}^H(n)\mathbf{x}(n)} - \rho_{L0}J(\mathbf{h}(n)), \quad (22)$$

where μ_N is the step size of gradient descend which is same as in Eq. (14).

IV. NUMERICAL SIMULATIONS

In this section, we compare the performance of proposed channel estimators using 10000 independent Monte-Carlo runs for averaging. The length of sparse multipath channel $\mathbf{h}$ is set as $N = 16$ and its number of dominant taps is set as $K = 1$ and 4 respectively. The values of dominant channel taps follow random Gaussian distribution and the positions of dominant taps are randomly allocated within the length of $\mathbf{h}$ which is subjected to $E\{\|\mathbf{h}\|_2^2\} = 1$. The signal-to-noise ratio (SNR) is defined as $10\log(E_0/\sigma_n^2)$, where E_0 is transmitted power. Here, we set the SNR as 10dB, 20dB and 30dB, respectively. All of the step sizes of gradient descend and regularization parameters are listed in Tab. I.

TABLE I. SIMULATION PARAMETERS FOR LMS-BASED ADAPTIVE SPARSE CHANNEL ESTIMATION.

Type of parameters	Value
μ	5e-2
μ_N	5e-1
λ_{ZA}	4e-2
λ_{RZA}	8e-2
λ_{LP}	1e-2
λ_{LPN}	7e-2
λ_{L0}	8e-2
λ_{L0N}	4e-3

The estimation performance is evaluated by mean square error (MSE) standard which is defined as

$$MSE(\mathbf{h}(n)) = E\{\|\mathbf{h}(n) - \mathbf{h}\|_2^2\}, \quad (23)$$

where $E[\cdot]$ denotes expectation operator, $\mathbf{h}$ and $\mathbf{h}(n)$ are the actual channel vector and its estimator, respectively.

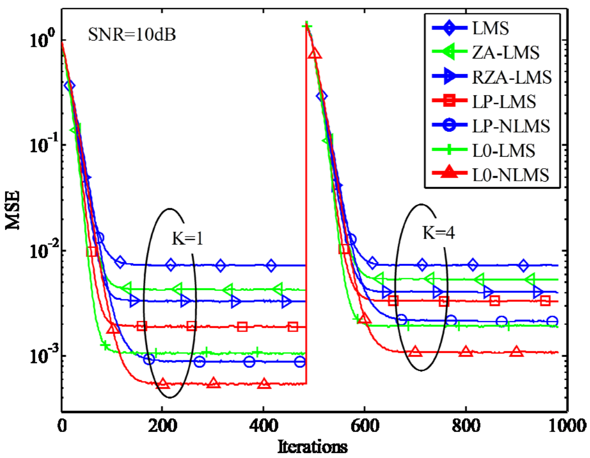

Fig. 3. MSE of different methods versus the number of iterations SNR = 10dB.

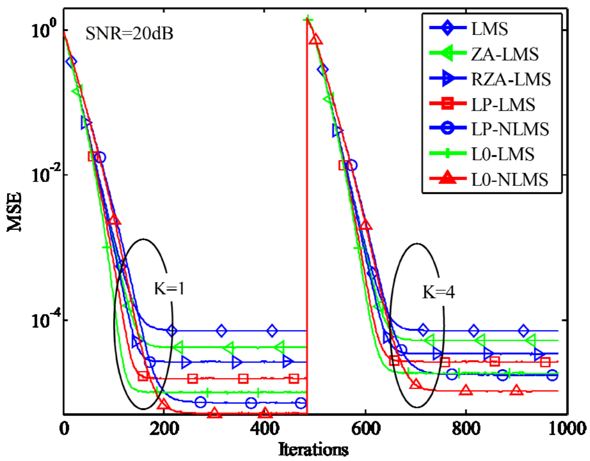

Fig. 4. MSE of different methods versus the number of iterations (SNR = 20dB).

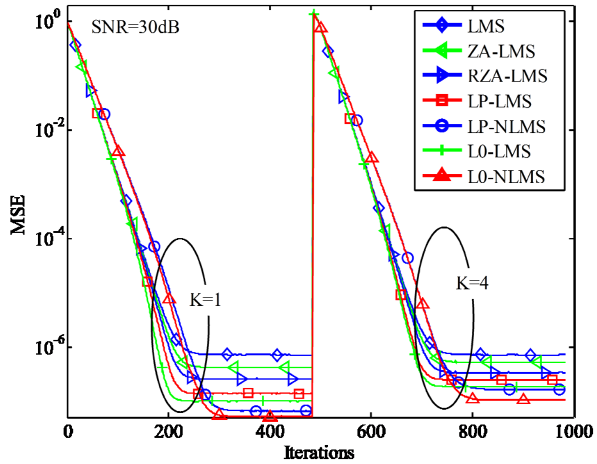

Fig. 5. MSE of different methods versus the number of iterations (SNR = 30dB).

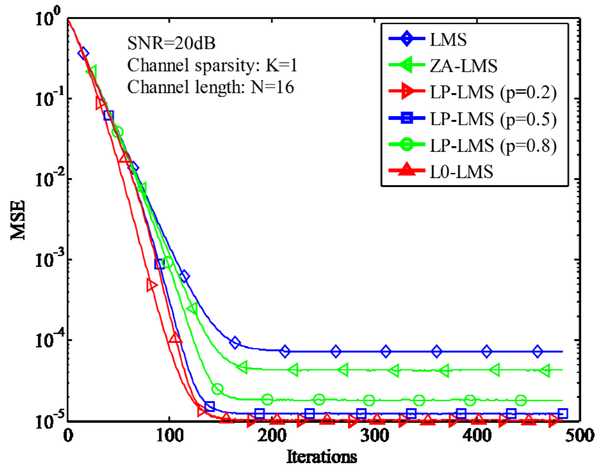

Fig. 6. MSE of LP-LMS versus the number of iterations (SNR = 20dB), where $p = 0.2, 0.5$ and 0.8.

At first, we compare all the LMS-based adaptive sparse channel estimation methods with different SNRs. When $SNR = 10\text{dB}$, the proposed methods can achieve better estimation performance than previous methods as shown in Fig. 3. Note that the performance of LP-NLMS and L0-NLMS based adaptive sparse channel estimation methods are better than LP-LMS and L0-LMS based ones. In addition, we can also find that the convergence speed of LMS-based estimation methods (LMS, ZA-LMS, RZA-LMS, LP-LMS and L0-LMS) is faster than the NLMS-based ones (LP-NLMS and L0- NLMS). However, as the iteration times increase, NLMS based adaptive sparse channel estimation can achieve better estimation performance than LMS-based method by using same sparse penalty. To further confirm the advantage of our proposed methods in different SNR region, e.g., SNR = 20dB and 30dB, they are evaluated and simulation results are shown in Fig. 4 and Fig. 5, respectively. According to the above computer simulations, we can find our proposed methods can work well in different SNRs. Also, effectiveness of proposed adaptive sparse channel estimation methods has been verified.

Secondly, we evaluate the estimation performance of LP-LMS as a function of p as shown in Fig. 6. From the figure, estimation performance of LP-LMS increases as the value of p decrease. In other words, smaller L_p-norm sparse penalty on cost function of LMS can achieve better estimation performance. For example, when $p = 0.2$, the performance curve of LP-LMS is closes to L0-LMS and vice versa.

V. Conclusion

In this paper, we have investigated LMS-based adaptive sparse channel estimation methods by enforcing different sparse penalties. According to the CS theory, we have proposed an improved adaptive sparse channel estimation method by using ℓp-norm zero attracting LMS algorithm, where $0 \leq p \leq 1$. In addition, to further improve the estimation performance, L0-NLMS based adaptive sparse channel estimation methods have been proposed. Compared to the existing methods, the proposed methods exhibit better convergence and their performance advantages are demonstrated by several representative numerical simulations.

Acknowledment

This work was supported in part by the Japan Society for the Promotion of Science (JSPS) postdoctoral fellowship.

References

[1] D. Raychaudhuri, and N. Mandayam, "Frontiers of Wireless and Mobile Communications," Proceedings of the IEEE, vol. 100, no. 4, pp. 824-840, Apr. 2012.

[2] F. Adachi, D. Grag, S. Takaoka, and K. Takeda. "Broadband CDMA techniques," IEEE Wireless Communications, vol. 12, no. 2, pp. 8-18, Apr. 2005.

[3] J. Wang and J. Chen, "Performance of wideband CDMA systems with complex spreading and imperfect channel estimation," IEEE Journal on Selected Areas in Communications, vol. 19, no. 1, pp. 152-163, Jan. 2001.

[4] Y. Q. Zhou, J. Wang and M. Sawahashi, "Downlink transmission of broadband OFCDM systems Part I: hybrid detection," IEEE Transactions on Communications, vol. 53, no. 4, pp. 718-729, April 2005.

[5] H. Zhu and J. Wang, "Chunk-based resource allocation in OFDMA systems - Part I: chunk allocation," IEEE Transactions on Communications, vol. 57, no. 9, pp. 2734-2744, Sept. 2009.

[6] H. Zhu and J. Wang, "Chunk-based resource allocation in OFDMA systems - Part II: joint chunk, power and bit allocation," IEEE Transactions on Communications, vol. 60, pp. 499-509, no. 2, Feb. 2012.

[7] F. Adachi and E. Kudoh, "New direction of broadband wireless technology," Wireless Communications and Mobile Computing, vol. 7, no. 8, pp. 969-983, May 2007.

[8] W. F. Schreiber, "Advanced television systems for terrestrial broadcasting: Some problems and some proposed solutions," Proceedings of the IEEE, vol. 83, pp. 958-981, Jun. 1995.

[9] A. F. Molisch, "Ultra wideband propagation channels Theory, measurement, and modeling," IEEE Transactions on Vehicular Technology, vol. 54, no. 5, pp. 1528-1545, Sept. 2005.

[10] Z. Yan, M. Herdin, A. M. Sayeed, and E. Bonek, "Experimental study of MIMO channel statistics and capacity via the virtual channel representation," University of Wisconsin-Madison, Madison, Technique Report, Feb. 2007.

[11] N. Czink, X. Yin, H. Ozcelik, M. Herdin, E. Bonek, and B. H. Fleury, "Cluster characteristics in a MIMO indoor propagation environment," IEEE Transactions on Wireless Communication, vol. 6, no. 4, pp. 1465-1475, Apr. 2007.

[12] L.Vuokko, V.-M. Kolmonen, J. Salo, and P. Vainikainen, "Measurement of large-scale cluster power characteristics for geometric channel models," IEEE Transactions on Antennas Propagation, vol. 55, no. 11, pp. 3361-3365, Nov. 2007.

[13] Widrow and S. D. Stearns, "Adaptive signal processing," New Jersey: Prentice Hall, 1985.

[14] Candes, J. Romberg and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," IEEE Transaction on Information Theory, vol. 52, no. 2, pp. 489-509, Feb. 2006.

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

[16] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressed Channel Sensing: A New Approach to Estimating Sparse Multipath Channels," Proceedings of the IEEE, vol. 98, no. 6, pp. 1058-1076, Jun. 2010.

[17] G. Taubock, F. Hlawatsch, D. Eiwen, and H. Rauhut, "Compressive Estimation of Doubly Selective Channels in Multicarrier Systems: Leakage Effects and Sparsity-Enhancing Processing," IEEE Journal of Selected Topics in Signal Processing, vol. 4, no. 2, pp. 255-271, Apr. 2010.

[18] G. Gui, Q. Wan, W. Peng, and F. Adachi, "Sparse Multipath Channel Estimation Using Compressive Sampling Matching Pursuit Algorithm," in Prof. IEEE VTS APWCS, Kaohsiung, Taiwan, 20-21 May 2010.

[19] G. Gui, P. Wei, A. Mehbodniya, and F. Adachi, "Compressed Channel Estimation for Sparse Multipath Non-orthogonal Amplify-and-forward Cooperative Networks," in IEEE 75th Vehicular Technology Conference (VTC2012-Spring), Yokohama, Japan, 2012.

[20] E. J. Candes, "The restricted isometry property and its implications for compressed sensing," Comptes Rendus Mathematique, vol. 346, no. 9- 10, pp. 589-592, May 2008.

[21] Michael R. Garey and David S. Johnson, "Computers and Intractability: A Guide to the Theory of NP-Completeness," W. H. Freeman; First Edition edition, 1979.

[22] R. Tibshirani, "Regression Shrinkage and Selection via the Lasso," Journal of the Royal Statistical Society (B), vol. 58, no. 1, pp. 267-288, 1996.

[23] Y. Chen, Y. Gu, and A. O. H. III, "Sparse LMS for System Identification," in Prof. IEEE ICASSP, Taibei, Taiwan, Apr. 2009.

[24] O. Taheri, Sergiy A. Vorobyov, "Sparse channel estimation with Lp-norm and reweighted L1-norm penalized least mean squares," in Proc. IEEE ICASSP, Prague, Czech Republic, 22-27 May 2011.

[25] Y. Gu, J. Jin, and S. Mei, "L0-Norm Constraint LMS Algorithm for Sparse System Identification," IEEE Signal Processing Letters, vol. 16, no. 9, pp. 774-777, Sept. 2009.

[26] E. J. Candes, Michael B. Wakin and S. P. Boyd, "Enhancing Sparsity by Reweighted L1 Minimization," Journal of Fourier Analysis Applications, vol. 14, no. 5-6, pp. 87-905, Dec. 2008.