An Efficient SAGE-based Data Detection Algorithm for OFDM Systems in the Presence of Very Fast Fading Channels

Zhicheng Dong
Provincial Key Lab of Information
Southwest Jiaotong University
Chengdu, China
School of Engineering
Tibet University, Lhasa, China
Email: dongzc666@163.com

Pingzhi Fan
Provincial Key Lab of Information
Southwest Jiaotong University
Chengdu, China
Email: pzfan@swjtu.edu.cn

Erdal Panayırcı
Dept of Electronics Engineering
Kadir Has University
Istanbul, Turkey
eepanay@khas.edu.tr

***Abstract*—In this paper, an iterative and computationally efficient data detection algorithm is proposed based on the space alternating generalized expectation maximization (SAGE) technique for orthogonal division multiplexing (OFDM) systems under fast fading channels. The proposed detector includes the original detector presented in [1] as one of its special cases. With a proper choice of its parameters, simulations show that the new detector has negligible performance loss than original one in [1] with smaller number of iterations.**

I. Introduction

Orthogonal frequency-division multiplexing (OFDM) has been shown to be an effective technique to overcome the inter-symbol interference (ISI) caused by frequency-selective fading with a simple transceiver structure. It has emerged as the leading transmission technique for a wide range of wireless communication standards[2] such as the IEEE's 802.16 family - better known as Mobile Worldwide Interoperability Microwave Systems for Next-Generation Wireless Communication Systems (WiMAX) - and the Third-Generation Partnership Project (3GPP) in the form of its Long-Term Evolution (LTE) project. Both systems employ orthogonal frequency division multiplexing/multiple access (OFDMA) as well as a new single-carrier frequency-division multiple access (SC-FDMA) format. To promote the IEEE 802.16 and LTE standards, recently, a high mobility feature has been introduced (IEEE 802.16m, LTE Advanced (LTE-A)) to enable mobile broadband services at vehicular speeds beyond 120 km/h.

The specific structure of the Doppler-induced ICI in OFDM systems operating over highly mobile channels presents a distinctive feature of limited support of the Doppler spread that can be exploited by the receiver. In the presence of very high mobility, the channel cannot be assumed constant even during one OFDM symbol period. This will results in a loss of subchannel orthogonality which leads to inter-carrier interference (ICI). ICI will degrade the spectral efficiency of the system. For detection of data in an OFDM system in the presence of ICI, the maximum likelihood (ML) detector is optimum [3]. However, the it demands the highest computational load and cannot be realized in practice. In [4], the performance of an minimum mean square error successive interference cancellation(MMSE-SIC) with optimal ordering, namely the MMSE-SIC algorithm (VBLAST) is investigated. However the higher computational load inhibits it to realize in practice. In [5], a low complexity algorithm based on MMSE-SIC algorithm is presented. In [1], the joint channel estimation, equalization and data detection for OFDM systems in the presence of very high mobility is studied.

In this paper, a new computationally feasible data detection algorithm is proposed based on the generalized expectation maximization (SAGE) algorithm. It is shown that the convergency rate of the algorithm much faster than that of [1] with a negligible performance loss.

Notations: A boldface large and small letter mean a matrix or a vector, respectively. $\mathbf{A}_{m,n}$ denotes the (m,n)th element of $\mathbf{A}$. Also, $\mathbf{A}_{m,:}$ and $\mathbf{A}_{:,n}$ denote the mth row vector and the nth column vector of $\mathbf{A}$, respectively. $\|\bullet\|$ denotes the norm of a vector. 0_N and 1_N represent the $N \times N$ zero and identity matrices, respectively. F_N and F_N^H denote the $N \times N$ fast Fourier transform (FFT) matrix and inverse Fourier transform (IFFT) matrix, respectively. $F_N = (1/\sqrt{N})[\exp^{-j2\pi(m-1)(n-1)/N}] \quad m,n = 1,\cdots,N$. Also, $(k)_N$ denotes k modulo N, $[\ \ ]^T$ and $[\ \ ]^H$ stand for transpose and Hermitian, respectively.

II. System Model

Let us consider an OFDM system with N subcarriers and available bandwidth B=1/T_s, where T_s is the sampling period. A given sampling period is divided into N subchannels by equal frequency spacing $\Delta f = B/N$. At the transmitter, information symbols are mapped into possibly complex-valued transmitted symbols according to the modulation format employed. The symbols are processed by an $N-$length Inverse Fast Fourier Transform (IFFT) block that transforms the data symbol sequence into the time domain. The time-domain signal is extended by a guard interval containing G samples whose length is chosen to be longer than the expected delay spread to avoid ISI. The guard interval includes a cyclically extended part of the OFDM block to avoid ICI. Hence, the complete OFDM block duration is $P = N + G$ samples. The resulting signal is converted to an analog signal by a digital-to-analog (D/A) converter. After shaping with a low-pass filter (e.g. a raised-cosine filter) with bandwidth B, it is transmitted through the transmit antenna with the overall symbol duration of PT_s.

Let $h(m, l)$ represent the lth path (multipath component) of the time-varying channel impulse response at time instant $t = mT_s$. The discrete-time received signal can then be expressed as follows:

$$y(m) = \sum_{l=0}^{L-1} h(m, l) d(m - l) + w(m), \tag{1}$$

where the transmitted signal $d(m)$ at discrete sampling time mT_s is given by

$$d(m) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X(k) e^{j2\pi mk/N}, \tag{2}$$

L is the total number of paths of the frequency selective fading channel, and $w(m)$ is additive white Gaussian noise (AWGN) with zero mean and variance $E\{|w(m)|^2\} = \sigma_w^2$. The sequence $X(k), k = 0, 1, \cdots, N - 1$, in (2) represent either quadrature-amplitude modulation (QAM) or phase-sift-keying (PSK) modulated data symbols with $E\{|X(k)|^2\} = 1$.

At the receiver, after passing through the analog-to-digital converter (A/D) and removing the cyclic prefix (CP), a fast Fourier transform (FFT) is used to transform the data back into the frequency domain. Lastly, the binary data is obtained after demodulation and channel decoding.

The fading channel coefficients $h(m, l)$ can be modeled as zero-mean complex Gaussian random variables. Based on the wide-sense stationary uncorrelated scattering (WSSUS) assumption, the fading channel coefficients in different paths are uncorrelated with each other. However, these coefficients are correlated within each individual path and have a Jakes Doppler power spectral density [6] having an autocorrelation function given by

$$E\{h(m, l)h^*(n, l)\} = \sigma_{h_l}^2 J_0(2\pi f_d T_s(m - n)), \tag{3}$$

where $\sigma_{h_l}^2$ denotes the power of the channel coefficients of the lth path. f_d is the Doppler frequency in Hertz so that the term $f_d T_s$ represents the normalized Doppler frequency of the channel coefficients. $J_0(.)$ is the zeroth order Bessel function of the first kind.

By using (1) in (2), the received signal can be written as

$$y(m) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X(k) \sum_{l=0}^{L-1} h(m, l) e^{j\frac{2\pi k(m-l)}{N}} + w(m), \tag{4}$$

which upon defining the time-varying channel transfer function

$$H(k, m) = \sum\nolimits_{l=0}^{L-1} h(m, l) e^{-j2\pi lk/N}, \tag{5}$$

becomes

$$y(m) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X(k) H(k, m) e^{j2\pi mk/N} + w(m). \tag{6}$$

The FFT output at the k^{th} subcarrier, after excluding the guard interval, can be expressed as

$$\begin{aligned} Y(k) &= \frac{1}{\sqrt{N}} \sum_{m=0}^{N-1} y(m) e^{-j2\pi mk/N} \\ &= X(k)G(k, k) + I(k) + W(k), \end{aligned} \tag{7}$$

where $I(k)$ is ICI caused by the time-varying nature of the channel given as

$$I(k) = \sum_{i=0, i \neq k}^{N-1} X(i) G(k, i). \tag{8}$$

$G(k, i)$ in (8) represents the average frequency domain time-varying channel response, defined as

$$G(k, i) = (1/N) \sum_{m=0}^{N-1} H(i, m) e^{j2\pi m(i-k)/N}. \tag{9}$$

Similarly, the term $G(k, k) = \frac{1}{N} \sum_{m=0}^{N-1} H(k, m)$ in (7) represent the portion of the average frequency domain channel response at the kth subcarrier and $W(k)$ denotes discrete Fourier transform of the white Gaussian noise $w(m)$:

$$W(k) = \frac{1}{\sqrt{N}} \sum_{m=0}^{N-1} w(m) e^{-j2\pi mk/N}. \tag{10}$$

Because of the term $I(k)$ in (7), there is an irreducible error floor even in the training sequences since pilot symbols are also corrupted by ICI, arising from the fact that the time-varying channel destroys the orthogonality between sub-carriers. Therefore, channel estimation should be performed either jointly with data or before the FFT block in order to compensate for the ICI.

From (6) and (7), the FFT output received signal can be expressed in vector form as

$$\mathbf{Y} = \mathbf{G}\mathbf{X} + \mathbf{W}, \tag{11}$$

where $\mathbf{Y} = [Y(0), Y(1), ..., Y(N - 1)]^T$, $\mathbf{X} = [X(0), X(1), ..., X(N - 1)]^T$ and $\mathbf{W} = [W(0), W(1), ..., W(N - 1)]^T$. For $k, i = 0, 1, \cdots, N - 1$, the (k, i)th element of the matrix $\mathbf{G} = [G(k, i)] \in \mathcal{C}^{N \times N}$ representing the time-varying channel is given by (9).

Eq.(11) can be expressed as

$$\mathbf{Y} = \sum_{n=1}^{N} \mathbf{z}(n), \tag{12}$$

where $\mathbf{z}(n) = \mathbf{G}(:, n) X(n) + \mathbf{W}(n)$. $\mathbf{W}(n)$ is the decomposed version of $\mathbf{W}$, that is, $\mathbf{W} = \sum_{n=0}^{N-1} \mathbf{W}(n)$. Thus, it is also a complex Gaussian noise with zero mean and the variance σ_n^2 with $\sum\limits_{n=1}^{N} \sigma_n^2 = {\sigma_w}^2$.

III. Data Detection Based on SAGE Algorithm

The data detection scheme is based on SAGE algorithm. The SAGE algorithm is a twofold generalization of the so-called ”expectation maximization” (EM) algorithm that provides updated estimates for an unknown parameter set [7]. Rather than updating all parameters simultaneously as done in the EM algorithm, at each iteration step, only a subset of the parameter set is updated while keeping the parameters in the complement set set fixed. The convergence rate of the SAGE algorithm is usually higher than that of the EM algorithm, because the conditional Fisher information matrix for each set of parameters is likely smaller than that of the complete data, given for the entire space.

At the ith iteration, the expectation-step (E-step) of the SAGE algorithm is defined as follows:

A. Expectation-Step (E-Step):

The E-Step computes the the average log-likelihood function of the observed value. The conditional expectation is then taken over $X(n)$ given the observation $\mathbf{Y}$ and that $\mathbf{X}$ equals its estimate calculated at ith iteration

$$U_n\left(X(n), \mathbf{X}^i\right) = E\left\{\ln p\left(\mathbf{z}(n)\left|\mathbf{X}^i\right.\right)\left|\mathbf{Y}, \mathbf{X}^i\right.\right\}. \tag{13}$$

By neglecting the terms independent of X, $\ln p\left(\mathbf{z}(n)\left|\mathbf{X}^i\right.\right)$ can be calculated from (13) as

$$\ln p\left(\mathbf{z}(n)\left|\mathbf{X}^i\right.\right) \sim \tfrac{1}{\sigma_n^2}\Re\left\{X(n)^*\mathbf{G}(:,n)^H\mathbf{z}(n)\right\}. \tag{14}$$

The details are shown in appendix A and $\Re\{\bullet\}$ denotes that the real part of its argument. Inserting (14) into (13), we have for $U_n\left(X(n), \mathbf{X}^i\right)$

$$U_n\left(X(n), \mathbf{X}^i\right) = \tfrac{1}{\sigma_n^2}\Re\left\{X_n^*\mathbf{G}(:,n)^H E\left\{\mathbf{z}(n)\left|\mathbf{Y}, \mathbf{X}^i\right.\right\}\right\} \tag{15}$$

where the conditional distribution of $\mathbf{y}(n)$ given $\mathbf{Y}$ and $\mathbf{X}^i$ is Gaussian with mean

$$\begin{aligned} E\left\{\mathbf{z}(n)\left|\mathbf{Y}, \mathbf{X}^i\right.\right\} =& X^i(n)\mathbf{G}(:,n) \\ &+ \frac{\sigma_n^2}{\sigma_w^2}\left(\mathbf{Y} - \sum_{j=1}^{N}\xi_j^i\right), \end{aligned} \tag{16}$$

where $\xi_j^i = X_j^i\mathbf{G}(:,j)$. The details are shown in appendix B.

B. Maximization-Step (M-Step):

In the maximization step (M-step) of the SAGE algorithm the estimates of the data sequence are updated at the $i+1$th iteration as follows:

$$X^{i+1}(n) = \arg\max_{X(n)} U_n\left(X(n), \mathbf{X}^i\right), \tag{17}$$

substituting (15) into (17)

$$\begin{aligned} X^{i+1}(n) = \arg\max_{X(n)} \Re\Bigg\{X^*(n)\Bigg\{\tfrac{W_n}{\sigma_n^2}\Bigg(\left(1 - \tfrac{\sigma_n^2}{\sigma_w^2}\right)X^i(n) \\ + \tfrac{\sigma_n^2}{\sigma_w^2 W_n}\Bigg(\mathbf{G}(:,n)^H\mathbf{Y} - \sum_{\substack{j=1\\j\neq k}}^{N}\psi_j^i\Bigg)\Bigg)\Bigg\}\Bigg\}, \end{aligned} \tag{18}$$

where $W_n = \mathbf{G}(:,n)^H\mathbf{G}(:,n)$, $\psi_j^i = \mathbf{G}(:,n)^H X^i(j)\mathbf{G}(:,j)$. Since W_n/σ_n^2 is a positive real number, we can obtain the new expressions as follows

$$\begin{aligned} X^{i+1}(n) = \arg\max_{X(n)} \Re\Bigg\{X^*(n)\Bigg\{\Bigg(\left(1 - \frac{\sigma_n^2}{\sigma_w^2}\right)X^i(n) \\ + \frac{\sigma_n^2}{\sigma_w^2 W_n}\Bigg(\mathbf{G}(:,n)^H\mathbf{Y} - \sum_{\substack{j=1\\j\neq k}}^{N}\psi_j^i\Bigg)\Bigg)\Bigg\}\Bigg\}. \end{aligned} \tag{19}$$

Decison on $X^{i+1}(n)$ depends on the two parts in (19). For $\frac{\sigma_n^2}{\sigma_w^2} = 1$, the detector is the same as [1], and for $\frac{\sigma_n^2}{\sigma_w^2} = 0$, the detector is reduced to the one used in the last iteration. The proper value of $\frac{\sigma_n^2}{\sigma_w^2}$, for a fast convergence of the iterative algorithm or a good performance in a limited number of iteration, depends on the energy of the signal and the ICI.

Here we consider the following simple functions for choosing these parameters [8]:

$$\tfrac{\sigma_n^2}{\sigma_w^2} = \tfrac{\mathrm{INSR}(n)}{c + \mathrm{INSR}(n)}, \tag{20}$$

where $c = 0.5$, $\mathrm{INSR}(n)$ denotes the interference-plus-noise to signal ratio of nth subcarrier, calculated as follows:

$$\mathrm{INSR}(n) = \tfrac{1}{\phi(n,n)}\sum_{q\neq n}^{N}\left(|\phi(n,q)| + \sigma_w^2\phi(n,n)\right), \tag{21}$$

where $\phi = \mathbf{G}^H\mathbf{G}$.

In order to detect the initial data symbols X^0. From the observations (11), we use a MMSE data detection techniques, expressed as

$$\mathbf{X}^0 = \mathbf{G}^\dagger\left(\mathbf{G}\mathbf{G}^\dagger + \sigma_w^2 I_N\right)^{-1}\mathbf{Y}. \tag{22}$$

As we know, time-varying channels produce a nearly-banded channel matrix whose only main and few subdiagonal terms are significant [1], [9]. The banded property of the channel in (22) can be exploited to reduce the computational complexity by means of low complexity decompositions such as the Cholsky or $\mathrm{LL}^\dagger$ factorization of Hermitian banded matrices. In this scheme, we choose the $\mathrm{LL}^\dagger$ factorization to obtain inverse of matrix. The reader is referred to [1] for details.

In summary, the algorithm proceeds as follows.

1) Initialization: Compute the initialization using (22). Set the iteration counter to $i = 0$.

2) Set the sub-carrier index to $n = 1$.

3) Based on the current values of X, compute the $X^{i+1}(n)$ using (19), which will replace the value of the corresponding element of $X^i(n)$.

4) If $n \leq N$, let $n = n + 1$, and go to step 3). Otherwise, go to step 5).

5) If X has converged or the iteration index has reached its maximum, it is terminated. Otherwise, let $i = i + 1$, and return to step 2).

Note that, Fig.1 shows the flow chart to describe the proposed algorithm.

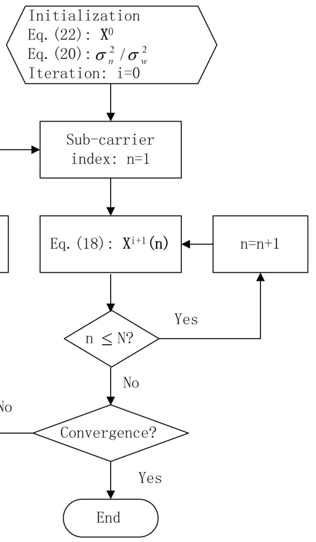

Fig. 1. Flow Chart of the Proposed Algorithm

IV. Numerical Results And Discussions

In this section, the simulation results about the performance of OFDM systems based on the proposed detector is presented. The system operates over a 15KHz subcarrier space with 128 subcarriers with carrier frequency of 5GHz and with M-ary phase shift keying signaling. A multipath wireless channel having an exponentially decaying power delay profile with normalized power [1], $\sigma_0^2 = 0.448$, $\sigma_1^2 = 0.321$, and $\sigma^2 = 0.230$ is chosen.

In Fig. 2, the BER performance of the proposed algorithm is presented as a function of SNR for the normalized Doppler frequency $f_D T = 0.1543$ corresponding to a mobile terminal moving at speeds of 500 km/h. The Fig. 2 shows that the performance of proposed SAGE scheme is closed to the SAGE scheme in [1]. The ICI power increases with transmitted SNR when speed is fixed. The figure also indicates that all detection schemes have similar BER when SNR is below 10 dB due to fact that the noise gets larger than ICI power. However, the different schemes have different BER performance when SNR is above 10 dB (the interference mainly comes from ICI). Therefore, the proposed algorithm can efficiently solves the ICI problem. Consequently, the performance based on SAGE is better than MMSE scheme, but poorer than MMSE-SIC in [4].

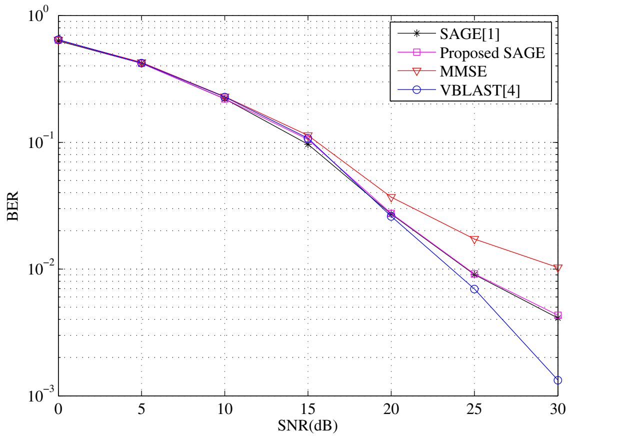

Fig. 2. BER Comparison of Different Detection Scheme for $v = 500km/h$, 16PSK Signaling.

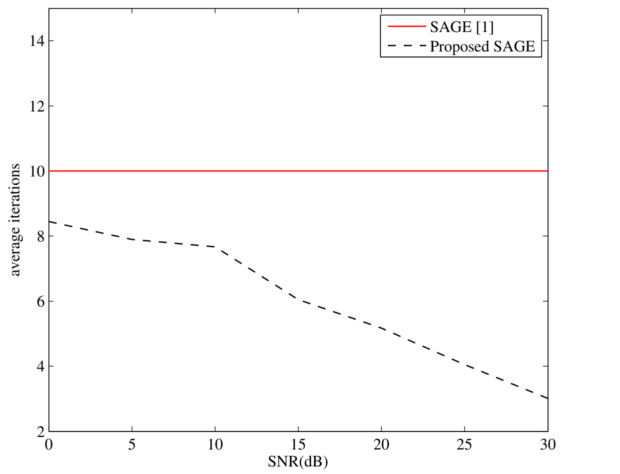

Fig. 3. Average Number Iterations Comparison of Two Different SAGE Scheme

Fig. 3 shows the average number iterations comparison of two different SAGE scheme. In order to prove the convergence speed of the algorithm, the number of iterations is chosen as 10 in our computer simulations. However, as can be seen from Fig. 3, the average number of iterations needed for convergence of our algorithm is less than 10 and gets smaller as the SNR becomes takes larger values, due to fact that the proposed algorithm efficiently solves the ICI problem rather than Gassuian noise. The Fig. 3 clearly shows that the proposed SAGE scheme has less smaller computational complexity than the original SAGE scheme presented in [1] with almost the same performance.

V. Conclusion

In this paper, an efficient data detection for OFDM systems under fast fading channel based on SAGE is presented. The proposed scheme includes the original scheme of [1]. The proposed SAGE scheme was shown to substantially decrease computational complexity compared to the original scheme. Extension of the proposed algorithm to a joint channel estimation and data detection algorithm is straightforward.

ACKNOWLEDGMENT

This work was supported by the National Basic Research Program of China (973 Program No. 2012CB316100), the National Science Foundation of China (NSFC, No. 61032002), the 111 Project (No.111-2-14), and the 2013 Doctoral Innovation Funds of Southwest Jiaotong University and the Fundamental Research Funds for the Central Universities.

APPENDIX A

It is easily shown that the maximization of the log-likelihood function [10].

$$L(X) = \frac{-1}{2\sigma_w^2}\left(\left|\mathbf{Y} - \sum_{n=1}^{N}\mathbf{G}(:,n)\,X(n)\right|^2\right) \tag{23}$$

is equivalent to the maximization of the function $U\left(X, \mathbf{X}'\right)$, which is defined as follows:

$$U\left(X, \mathbf{X}'\right) = E\left\{\ln p\left(\mathbf{z}(n)\left|\mathbf{X}^i\right.\right)\left|\mathbf{Y}, X = \mathbf{X}'\right.\right\}, \tag{24}$$

where

$$\begin{aligned}\ln p\left(\mathbf{z}(n)\left|\mathbf{X}^i\right.\right) = &\frac{1}{(2\pi)^{N/2}\prod_{n=1}^{N}\sigma_n}\times\\ &\exp\left(\sum_{n=1}^{N}\frac{-1}{2\sigma_n^2}\left|\mathbf{y}(n) - \mathbf{G}(:,n)\,X(n)\right|^2\right).\end{aligned} \tag{25}$$

It can be easily observed that

$$\begin{aligned}\ln p\left(\mathbf{z}(n)\left|\mathbf{X}^i\right.\right) = A + &\sum_{n=1}^{N}\frac{1}{\sigma_n^2}\Re\left(\mathbf{z}(n)^H\mathbf{G}(:,n)\,X(n)\right.\\ &\left.-\frac{1}{2}\mathbf{G}(:,n)^H\,\mathbf{G}(:,n)\,X(n)X(n)^H\right),\end{aligned} \tag{26}$$

where A is a constant. Since $X(n)X(n)^H = |X(n)|^2 = B$ (for MPSK), the second term does not depend on X, and (26) can be simplified to

$$\ln p\left(\mathbf{z}(n)\left|\mathbf{X}^i\right.\right) = C + \sum_{n=1}^{N}\frac{1}{\sigma_n^2}\Re\left(\mathbf{z}(n)^H\mathbf{G}(:,n)\,X(n)\right). \tag{27}$$

In the SAGE algorithm, the N-dimensional maximization problem of (27) can be reduced into N one- dimensional maximization problems. The above equation can be written as

$$\ln p\left(\mathbf{z}(n)\left|\mathbf{X}^i\right.\right) \sim \frac{1}{\sigma_n^2}\Re\left\{X^*(n)\mathbf{G}(:,n)^H\mathbf{z}(n)\right\}. \tag{28}$$

APPENDIX B

In this Appendix, the (16) is proved. Since both $\mathbf{z}(n)$ and $\mathbf{Y}$ given X are Gaussian,

$$\begin{aligned}E\left\{\mathbf{z}(n)\left|\mathbf{Y}, \mathbf{X}^i\right.\right\} = &E\left\{\mathbf{z}(n)\left|\mathbf{X}^i\right.\right\}\\ &+ \mathrm{C}_{zY}\mathrm{C}_{YY}^{-1}\left[\mathbf{Y} - E\left\{\mathbf{Y}\left|\mathbf{X}^i\right.\right\}\right],\end{aligned} \tag{29}$$

where

$$\mathrm{C}_z = E\left\{\left[\mathbf{z}(n) - E\left\{\mathbf{z}(n)\left|\mathbf{X}^i\right.\right\}\right]\times\left[\mathbf{Y} - E\left\{\mathbf{Y}\left|\mathbf{X}^i\right.\right\}\right]\left|\mathbf{X}^i\right.\right\},$$

$$\mathrm{C}_{YY} = E\left\{\left(\mathbf{Y} - E\left\{\mathbf{Y}\left|\mathbf{X}^i\right.\right\}\right)^2\left|\mathbf{X}^i\right.\right\},$$

$$E\left\{\mathbf{Y}\left|\mathbf{X}^i\right.\right\} = \sum_{n=1}^{N}\mathbf{G}(:,n)\,X(n),$$

$$E\left\{\mathbf{z}(n)\left|\mathbf{X}^i\right.\right\} = \left[\mathbf{G}(:,1)\,X^i(1), \mathbf{G}(:,2)\,X^i(2), ..., \mathbf{G}(:,n)\,X^i(n)\right]^T$$

$$\mathrm{C}_{zY} = \left[\sigma_1^2, \sigma_2^2, ..., \sigma_n^2\right]^T,$$

It can be easily shown that [11]

$$\begin{aligned}E\left\{\mathbf{z}(n)\left|\mathbf{Y}, \mathbf{X}^i\right.\right\} = &X^i(n)\mathbf{G}(:,n)\\ &+ \frac{\sigma_n^2}{\sigma_w^2}\left(\mathbf{Y} - \sum_{j=1}^{N}X^i(j)\mathbf{G}(:,j)\right).\end{aligned} \tag{30}$$

REFERENCES

[1] E. Panayirci, H. Senol, and H. Poor, "Joint channel estimation, equalization, and data detection for ofdm systems in the presence of very high mobility," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4225–4238, 2010.

[2] R. Prasad, *OFDM for wireless communications systems*. Artech House Publishers, 2004.

[3] Y. Liang, Q. Zhang, Z. Liu, F. Shu, and S. Berber, "Low-complexity sphere decoding for detection of ofdm systems in doubly-selective fading channels," *Electr. lett.*, vol. 45, no. 15, pp. 797–798, 2009.

[4] Y. Choi, P. Voltz, and F. Cassara, "On channel estimation and detection for multicarrier signals in fast and selective rayleigh fading channels," *IEEE Trans. Commun.*, vol. 49, no. 8, pp. 1375–1387, 2001.

[5] S. Chen and Y. Yang, "Low-complexity mmse-sic equalizer employing ldlh factorization for ofdm systems over time-varying channels," *IEEE Trans. Veh. Technol.*, vol. 59, no. 8, pp. 4128–4131, 2010.

[6] W. Jakes and D. C. Cox, *Microwave Mobile Communications*. New York: Wiley-IEEE, 1994.

[7] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 1–38, 1977.

[8] M. Borran and M. Nasiri-Kenari, "An efficient detection technique for synchronous cdma communication systems based on the expectation maximization algorithm," *IEEE Trans. Veh. Technol.,*, vol. 49, no. 5, pp. 1663–1668, 2000.

[9] E. Panayirci, H. Dogan, and V. Poor, "Low-complexity map-based successive data detection for coded ofdm systems over highly mobile wireless channels," *IEEE Trans. Veh. Technol.*, vol. 60, no. 6, pp. 2849–2857, 2011.

[10] S. Kay, "Fundamentals of statistical signal processing, volume i: Estimation theory (v. 1)," 1993.

[11] G. H. Golub and C. F. Van Loan, *Matrix computations*. Johns Hopkins University Press, 1996, vol. 3.