Fig. 3 indicates that the computational cost of iterative GBF is significantly lower than those of greedy bit filling [3] and the algorithm in [2], particularly in cases of larger N and smaller $\bar{b}$. For instance, when $N = 512$ and $\bar{b} = 6$ is the case, R_1 varies from 88% to 98% along with η, whereas the value of R_2 is almost steadily 88%. For given values of N and $\bar{b}$, R_1 steadily increases along with η, whereas R_2 slightly decreases. These observations reveal the appealing merits of iterative GBF in comparison with earlier algorithms of $O(N^2)$ complexity.

Fig. 4 presents the improvement of iterative GBF compared with EBF in [1]. As can be seen, for each given pair of N and $\bar{b}$, the corresponding curve of R_3 has a sawtooth shape. It is due to the fact that the number of comparisons saved by iterative GBF, i.e., $d(N-1)$, is proportional to d, which periodically varies from 0 to $\bar{b}-1$ along with η. Numerically, the improvement of iterative GBF can be significant when d is close to $\bar{b}-1$. For instance, R_3 can be up to 11% and 9% for scenarios of $N = 128$ and $N = 512$, respectively, when $\bar{b} = 10$ is the case. It can be also observed that the improvement of iterative GBF is comparatively more significant in scenarios of smaller N and larger $\bar{b}$. These observations confirm that the computational cost of iterative GBF is lower than that of EBF.

Obviously, the same conclusions can be drawn for iterative GBR in comparison with greedy bit removal [5], the algorithm in [2], and EBR in [1].

VI. Concluding Remarks

We have proposed r-GBF and r-GBR schemes for QAM-modulated OFDM systems, based on which a couple of iterative group-by-group bit-loading algorithms, i.e., iterative GBF and iterative GBR, are presented. In comparison with the original EBF and EBR in [1], the newly presented algorithms get rid of the dependence on conventional greedy algorithms thoroughly and converge to the optimal BAP with lower computational costs.

Acknowledgment

The authors would like to thank Prof. J.-Y. Chouinard and the anonymous reviewers for their constructive comments and suggestions.

References

[1] D. Wang, Y. Cao, and L. Zheng, "Efficient two-stage discrete bit-loading algorithms for OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 59, no. 7, pp. 3407–3416, Sep. 2010.

[2] N. Papandreou and T. Antonakopoulos, "A new computationally efficient discrete bit-loading algorithm for DMT applications," *IEEE Trans. Commun.*, vol. 53, no. 5, pp. 785–789, May 2005.

[3] E. Baccarelli and M. Biagi, "Optimal integer bit-loading for multicarrier ADSL systems subject to spectral-compatibility limits," *Signal Process.*, vol. 84, no. 4, pp. 729–741, Apr. 2004.

[4] A. Fasano, G. Di Blasio, E. Baccarelli, and M. Biagi, "Optimal discrete bit loading for DMT based constrained multicarrier systems," in *Proc. IEEE ISIT*, Jul. 2002, p. 243.

[5] A. Fasano, "On the optimal discrete bit loading for multicarrier systems with constraints," in *Proc. IEEE VTC*, Apr. 2003, pp. 915–919.

[6] R. D. Dutton, "Weak-heap sort," *BIT*, vol. 33, pp. 372–381, 1993.

Optimal Linear Channel Prediction for LTE-A Uplink Under Channel Estimation Errors

Jwo-Yuh Wu and Wen-Ming Lee

***Abstract*—This paper proposes an optimal design of the linear channel prediction scheme for LTE-A uplink transmission, taking into account the effect of channel estimation errors. The optimization criterion is to minimize mean square prediction errors, averaged over the distributions of true channel coefficients and estimation errors. We derive a closed-form solution of the optimal channel predictor and provide exact analysis for the achievable mean square error (MSE). The performance gain, in terms of the amount of reduction in MSE, as compared with the conventional channel predictor, is also analytically characterized. Simulation results are used to validate our MSE analyses and to confirm the performance advantage of the proposed solution.**

***Index Terms*—Channel estimation, channel prediction, Long-Term Evolution/Long-Term Evolution Advanced (LTE/LTE-A), training overhead reduction.**

I. Introduction

Low energy consumption has been a critical demand for the design of next-generation wireless communications [1], [2]. This thus spurs the development of new energy-efficient *green communication* techniques from various aspects, e.g., efficient signal processing algorithms at the physical (PHY) layer, intelligent network resource management for improved medium access control, and cognitive capability for better adaptation to the environments, to name just a few. From a PHY-layer perspective, typical approaches toward improved energy efficiency include, e.g., low-complexity algorithm implementation and reduced overheads for various signaling protocols [2]. While the former has been a widely considered design target in the study and design of PHY-layer signal processing techniques, the latter has received relatively little attention.

Reliable acquisition of channel state information (CSI) at the receiver is key to realizing various performance gains in modern wireless communications. The development of low-overhead CSI acquisition techniques is thus of great importance. In this paper, we consider the Long-Term Evolution Advanced (LTE-A) uplink, in which single-carrier frequency-domain multiple-access (SC-FDMA) modulation [3], [4] is employed. To reduce the training overhead for CSI acquisition, we propose to adopt the linear channel prediction technique at the base station (BS). The proposed approach is built on the fact that consecutive samples of real-world wireless channels are correlated in the time domain. Hence, as long as the BS has acquired a set of channel estimates during the past few training phases, it is then plausible to

Manuscript received October 30, 2012; revised February 5, 2013; accepted March 29, 2013. Date of publication April 12, 2013; date of current version October 12, 2013. This work was supported in part by the National Science Council of Taiwan under Grant NSC 97-2221-E-009-101-MY3, Grant NSC 99-2628-E-009-004, Grant NSC 100-2221-E-009-104-MY3, and Grant NSC 100-3113-P-009-001; by the Ministry of Education of Taiwan under the MoE ATU Program; and by the Telecommunication Laboratories, Chunghwa Telecom Co., Ltd., under Grant TL-101-G106. The review of this paper was coordinated by Dr. H. Lin.

The authors are with the Department of Electrical and Computer Engineering, National Chiao Tung University, Hsinchu 30010, Taiwan (e-mail: jywu@cc.nctu.edu.tw; jerry90908@hotmail.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2013.2257909

exploit the channel temporal correlation to predict the latest CSI via certain channel prediction schemes. In this way, mobile users no longer need to frequently send training signals to the BS, and thus, the training overhead can be reduced. The main contributions of this paper can be summarized as follows.

1) Assume that during the past K training phases the channels are estimated at the BS based on the linear minimum mean-square-error (LMMSE) criterion [5]. Associated with each user, we propose to adopt a bank of K-tap linear finite-impulse response filters (one for each channel path) to linearly combine K channel estimates for future CSI prediction. The design criterion of the predictor coefficients is the minimization of the MSE between the predicted and the true channels, whereby the expectation is taken with respect to the statistics of both the true channels and estimation errors. We derive a closed-form solution of the proposed channel predictor. In addition, the achievable MSE of the proposed solution is also analytically derived.
2) In case channel estimation is perfect, the proposed channel predictor then reduces to the conventional solution widely seen in the literature, e.g., [6]. We also derive for this solution the formula of the achievable MSE under channel estimation errors. With the aid of this result, it is analytically shown that, compared with this conventional solution, the proposed channel predictor does yield a smaller MSE; in particular, the amount of reduction in MSE is analytically characterized. The computational complexity of the proposed robust channel prediction scheme and the conventional solution is also provided.
3) The proposed channel prediction technique can be utilized in conjunction with training-based channel estimation to develop a two-phase CSI acquisition protocol for training overhead reduction. The training overhead reduction that is achieved by the proposed two-phase protocol is characterized. Simulation results are used to illustrate the resultant MSE and bit-error-rate (BER) performances.

We would like to remark that linear channel prediction has been considered for improving CSI quality in wireless communications, e.g., [7]–[16], to list just a few. Most of the existing works relied on the idealized assumption that channel estimation is perfect. To the best of the authors' knowledge, our paper provides an original study of algorithm development and exact performance analyses for LMMSE channel prediction in LTE-A uplink under channel estimation errors. The rest of this paper is organized as follows. Section II is the preliminary. Section III shows the main results of this paper, including the development of the proposed channel prediction algorithm, characterization of the resulting MSE performance, complexity analysis, and a discussion on the issue of training overhead reduction. Section IV contains the simulation results. Finally, Section V is the conclusion.

II. Preliminary

A. System Model During the Training Phase

We consider the discrete-time baseband model for SC-FDMA transmission over an L-path frequency-selective fading channel [3], [4]. The communication environment is assumed to be quasi-static fading, in which the channel gains remain constant during one resource block and can independently vary from block to block. Let $\mathbf{p}[k] \in \mathbb{C}^M$ be the frequency-domain training symbol at the kth resource block; $\mathbf{p}[k]$ is mapped onto consecutive M out of a total number of N subcarriers to obtain

$$\tilde{\mathbf{p}}[k] = \mathbf{A}\mathbf{p}[k] \tag{2.1}$$

where, for $0 \leq \mu \leq N - M (N > M)$

$$\mathbf{A} = \begin{bmatrix} \mathbf{0}_{\mu \times M} \\ \mathbf{I}_M \\ \mathbf{0}_{(N-M-\mu)\times M} \end{bmatrix} \in \mathbb{C}^{N\times M} \tag{2.2}$$

is the subcarrier mapping matrix, in which $\mathbf{0}_{m\times n}$ and $\mathbf{I}_M$ denote, respectively, the zero matrix of size $m \times n$ and the $M \times M$ identity matrix.[1] Before transmission, $\tilde{\mathbf{p}}[k]$ is first postmultiplied by the $N \times N$ inverse fast Fourier transform matrix $\mathbf{F}_N^H \in \mathbb{C}^{N\times N}$ to get

$$\mathbf{x}[k] = \mathbf{F}_N^H \tilde{\mathbf{p}}[k] \tag{2.3}$$

which is then appended with a cyclic prefix (CP) of length no less than $L-1$ to combat ISI. Let

$$\mathbf{h}[k] \triangleq [\, h_0(k) \quad h_1(k) \quad \cdots \quad h_{L-1}(k)\,]^T \in \mathbb{C}^L \tag{2.4}$$

be the channel impulse response vector during the kth resource block; the corresponding channel frequency responses are thus

$$\mathbf{g}[k] \triangleq \mathbf{F}_N \left[\mathbf{h}[k]^T \quad \mathbf{0}_{1\times(N-L)}\right]^T \in \mathbb{C}^N. \tag{2.5}$$

At the receiver side, with CP removal, N-point fast Fourier transform (FFT) processing, and subcarrier demapping, the frequency-domain training data read

$$\mathbf{y}[k] \triangleq \mathbf{G}[k]\mathbf{p}[k] + \mathbf{w}[k] \tag{2.6}$$

where

$$\mathbf{G}[k] \triangleq \text{diag}\left\{\mathbf{A}^T\mathbf{g}[k]\right\} \in \mathbb{C}^{M\times M} \tag{2.7}$$

is a diagonal matrix with the M selected channel frequency responses on the main diagonal, and $\mathbf{w}[k] \sim \mathcal{CN}(\mathbf{0}, \sigma_w^2 \mathbf{I}_M)$, the complex Gaussian distribution with zero mean and covariance $\sigma_w^2\mathbf{I}_N$. To facilitate channel estimation, let us rewrite (2.6) into

$$\mathbf{y}[k] = \mathbf{P}[k] \cdot \left(\mathbf{A}^T\mathbf{g}[k]\right) + \mathbf{w}[k] \tag{2.8}$$

where

$$\mathbf{P}[k] = \text{diag}\left\{\mathbf{p}[k]\right\} \in \mathbb{C}^{M\times M} \tag{2.9}$$

is a diagonal pilot signal matrix. By invoking the definition of $\mathbf{g}[k]$ in (2.5), it is straightforward to obtain the following equivalent representation of the training system (2.8):

$$\mathbf{y}[k] = \mathbf{P}[k]\tilde{\mathbf{F}}_N\mathbf{h}[k] + \mathbf{w}[k] \tag{2.10}$$

in which $\tilde{\mathbf{F}}_N \in \mathbb{C}^{M\times L}$ is obtained from $\mathbf{F}_N$ by retaining the first L columns and M rows corresponding to the M selected subcarriers. The following assumptions are made throughout this paper.

Assumption 1: For each fixed k, channel gains $h_i(k)$'s are independent and identically distributed such that $h_i(k) \sim \mathcal{CN}(0, \sigma_{h,i}^2)$, $0 \leq i \leq L-1$.

Assumption 2: For each fixed i, the temporal variation of the channel gains follows Jake's model [17]; thus, the temporal correlation function is given as $E\{h_i(k)h_i^*(m)\} = \sigma_{h,i}^2 J_0(2\pi f_d(k-m))$, where $J_0(\cdot)$ is the zero-order Bessel function, and f_d is the Doppler frequency in hertz.

[1] Application of the proposed approach to the case of distributed subcarrier mapping is straightforward.

B. Time-Domain LMMSE Channel Estimation

Based on (2.10), we consider the LMMSE estimation of the time-domain impulse response vector $\mathbf{h}[k]$. By following the standard procedures, e.g., [5] and [18], the resultant channel estimate can be obtained as

$$\widehat{\mathbf{h}}[k] = \left[\widehat{h}_0(k) \quad \widehat{h}_1(k) \quad \cdots \quad \widehat{h}_{L-1}(k)\right]^T = \sigma_w^{-2}\left[\mathbf{R}_{\mathbf{h}}^{-1} + \sigma_w^{-2}\tilde{\mathbf{F}}_N\tilde{\mathbf{F}}_N^H\right]^{-1}\tilde{\mathbf{F}}_N^H\mathbf{P}[k]^{-1}\mathbf{y}[k] \tag{2.11}$$

where

$$\mathbf{R}_{\mathbf{h}} \triangleq \text{diag}\left\{\left[\sigma_{h,0}^2 \quad \sigma_{h,1}^2 \quad \cdots \quad \sigma_{h,L-1}^2\right]\right\}. \tag{2.12}$$

We assume that the receiver has acquired a sequence of K channel estimates $\widehat{\mathbf{h}}[1], \ldots, \widehat{\mathbf{h}}[K]$ according to (2.11) during K consecutive training-based resource blocks, based on which channel prediction is then employed to predict the future CSI, which is denoted hereafter by $\widehat{\mathbf{h}}_{\mathcal{P}}[K+n]$, for $n \geq 1$. This paper proposes a new linear channel prediction scheme, which takes into account the channel estimation errors, and studies the achievable MSE performance of the proposed solution. Based on the developed channel prediction technique, we further propose a channel acquisition scheme with reduced training overhead.

III. Proposed Channel Prediction Scheme

This section introduces the proposed optimal channel prediction scheme. Section III-A derives the optimal channel predictor. Section III-B shows the achievable MSE performance. Complexity analysis in terms of flop count evaluation is given in Section III-C. Training overhead reduction achieved by the proposed channel prediction scheme is then discussed in Section III-D.

A. Algorithm Development

We assume that the receiver has acquired a sequence of K channel estimates $\widehat{\mathbf{h}}[1], \cdots, \widehat{\mathbf{h}}[K]$ according to (2.11) during K consecutive training resource blocks. Based on the estimated channels, channel prediction is employed to predict the future CSI $\widehat{\mathbf{h}}_{\mathcal{P}}[K+n] = [\widehat{h}_{0,\mathcal{P}}(K+n)\widehat{h}_{1,\mathcal{P}}(K+n)\cdots\widehat{h}_{L-1,\mathcal{P}}(K+n)]^T$, for $n \geq 1$. For analytic simplicity, prediction is done on a tap-by-tap basis. Thus, associated with the ith path, $0 \leq i \leq L-1$, we consider a linear channel predictor of the form

$$\widehat{h}_{i,\mathcal{P}}(K+n) = \sum_{k=1}^{K} a_{i,k}^{(n)}\hat{h}_i(K+1-k), \qquad n \geq 1 \tag{3.1}$$

where $a_{i,k}^{(n)}$'s are the predictor coefficients to be determined. The design of $a_{i,k}^{(n)}$'s is done via minimizing the MSE, i.e.,

$$\text{MSE}_i[n] \triangleq E\left\{\left|\widehat{h}_{i,\mathcal{P}}(K+n) - h_i(K+n)\right|^2\right\} \tag{3.2}$$

in which the expectation is taken with respect to the considered channel and noise statistics. By following the orthogonal principle [6], [18], the optimal solution can be obtained by solving the following equation, for $1 \leq j \leq K$:

$$E\left\{\left[h_i(K+n) - \sum_{k=1}^{K} a_{i,k}^{(n)}\widehat{h}_i(K+1-k)\right]\widehat{h}_i^*(K+1-j)\right\} = 0 \tag{3.3}$$

which can be rearranged as

$$\sum_{k=1}^{K} a_{i,k}^{(n)} E\left\{\widehat{h}_i(K+1-k)\widehat{h}_i^*(K+1-j)\right\} = E\left\{h_i(K+n)\widehat{h}_i^*(K+1-j)\right\}, \qquad 1 \leq j \leq K. \tag{3.4}$$

Equation (3.4) defines a set of K linear equations with $a_{i,k}^{(n)}$, $1 \leq k \leq K$, as unknowns. Based on (3.4) and by invoking the assumptions about the channel and noise statistics, the optimal predictor coefficients are derived in the following theorem.

Theorem 3.1: Define

$$\mathbf{R} = \begin{bmatrix} r[0] & r[1] & \cdots & r[K-1] \\ r[1] & r[0] & \cdots & r[K-2] \\ \vdots & \vdots & \ddots & \vdots \\ r[K-1] & r[K-2] & \cdots & r[0] \end{bmatrix} \tag{3.5}$$

$$\mathbf{r}_K^{(n)} = \left[\, r[n] \quad r[n+1] \quad \cdots \quad r[K+n-1]\,\right]^T \tag{3.6}$$

where $r[k-m] \triangleq J_0(2\pi f_d(k-m))$. The optimal channel predictor $\bar{\mathbf{a}}_i^{(n)} \triangleq [\bar{a}_{i,1}^{(n)}\bar{a}_{i,2}^{(n)}\cdots\bar{a}_{i,K}^{(n)}]^T \in \mathbb{C}^K$ is given by

$$\bar{\mathbf{a}}_i^{(n)} = \alpha_i\mathbf{R}_i^{-1}\mathbf{r}_K^{(n)} \tag{3.7}$$

in which

$$\mathbf{R}_i \triangleq \beta_i \cdot \mathbf{R} + \gamma_i \cdot \mathbf{I}_K \tag{3.8}$$

with

$$\alpha_i \triangleq \sigma_{h,i}^2 \times \mathbf{e}_i^T\tilde{\mathbf{F}}_N^H\tilde{\mathbf{F}}_N\left(\sigma_w^2\mathbf{R}_{\mathbf{h}}^{-1} + \tilde{\mathbf{F}}_N^H\tilde{\mathbf{F}}_N\right)^{-1}\mathbf{e}_i \tag{3.9}$$

$$\beta_i \triangleq \mathbf{e}_i^T\left(\sigma_w^2\mathbf{R}_{\mathbf{h}}^{-1} + \tilde{\mathbf{F}}_N^H\tilde{\mathbf{F}}_N\right)^{-1}\tilde{\mathbf{F}}_N^H\tilde{\mathbf{F}}_N\mathbf{R}_{\mathbf{h}} \times \tilde{\mathbf{F}}_N^H\tilde{\mathbf{F}}_N\left(\sigma_w^2\mathbf{R}_{\mathbf{h}}^{-1} + \tilde{\mathbf{F}}_N^H\tilde{\mathbf{F}}_N\right)^{-1}\mathbf{e}_i \tag{3.10}$$

$$\gamma_i \triangleq \sigma_w^2\mathbf{e}_i^T\left(\sigma_w^2\mathbf{R}_{\mathbf{h}}^{-1} + \tilde{\mathbf{F}}_N^H\tilde{\mathbf{F}}_N\right)^{-1}\tilde{\mathbf{F}}_N^H \times \tilde{\mathbf{F}}_N\left(\sigma_w^2\mathbf{R}_{\mathbf{h}}^{-1} + \tilde{\mathbf{F}}_N^H\tilde{\mathbf{F}}_N\right)^{-1}\mathbf{e}_i. \tag{3.11}$$

Proof: See Appendix A. ■

It is noted that, with perfect channel estimation, viz., $\sigma_w^2 = 0$, it is easy to check that $\alpha_i = \beta_i = \sigma_{h,i}^2$ and $\gamma_i = 0$; as a result, the optimal predictor $\bar{\mathbf{a}}_i^{(n)}$ in (3.7) is reduced to the well-known solution [6]

$$\tilde{\mathbf{a}}_i^{(n)} = \left[\tilde{a}_{i,1}^{(n)} \quad \tilde{a}_{i,2}^{(n)} \quad \cdots \quad \tilde{a}_{i,K}^{(n)}\right]^T \triangleq \mathbf{R}^{-1}\mathbf{r}_K^{(n)}. \tag{3.12}$$

Moreover, while the proposed channel predictor (3.7) is different across the channel paths, it can be seen from (3.12) that $\tilde{\mathbf{a}}_i^{(n)}$ is identical, irrespective of the paths.

B. MSE Analysis

To further validate the performance advantage of the proposed error-resistant predictor (3.7) as compared with (3.12), we will compare the achievable MSE of the two prediction schemes, which are defined respectively as

$$\overline{\text{MSE}}[n] \triangleq \sum_{i=0}^{L-1} E\left\{\left|\sum_{k=1}^{K} \bar{a}_{i,k}^{(n)}\widehat{h}_i(K+1-k) - h_i(K+n)\right|^2\right\} \tag{3.13}$$

$$\widetilde{\mathrm{MSE}}[n] \triangleq \sum_{i=0}^{L-1} E\left\{ \left| \sum_{k=1}^{K} \tilde{a}_{i,k}^{(n)} \hat{h}_i(K+1-k) - h_i(K+n) \right|^2 \right\}. \tag{3.14}$$

The following theorem provides analytic expressions for the two MSE formulas in (3.13) and (3.14).

Theorem 3.2: Let $\overline{\mathrm{MSE}}[n]$ and $\widetilde{\mathrm{MSE}}[n]$, which are defined in (3.13) and (3.14), respectively, be the MSEs achieved by channel predictors (3.7) and (3.12). Under the assumptions that $\sum_{i=1}^{L} \sigma_{h,i}^2 = 1$ and $r[0] = 1$, we have

$$\overline{\mathrm{MSE}}[n] = 1 - \sum_{i=0}^{L-1} \alpha_i^2 \left[\mathbf{r}_K^{(n)T} \left(\beta_i \mathbf{R} + \gamma_i \mathbf{I}_K\right)^{-1} \mathbf{r}_K^{(n)} \right] \tag{3.15}$$

$$\widetilde{\mathrm{MSE}}[n] = \overline{\mathrm{MSE}}[n] + \tilde{\mathbf{a}}_i^{(n)H} \mathbf{R}_i^{-1} \times \left[\left(1 - \frac{\alpha_i}{\beta_i}\right)^2 \cdot \mathbf{R}_i + \frac{\alpha_i \gamma_i}{\beta_i} \cdot \mathbf{I}_K \right]^2 \tilde{\mathbf{a}}_i^{(n)} \tag{3.16}$$

where $\mathbf{R}_i, \alpha_i, \beta_i, \gamma_i$, and $\tilde{\mathbf{a}}_i^{(n)}$ are defined in (3.8)–(3.12), respectively.

Proof: See Appendix B. ■

From (3.16), it is easy to verify that $\widetilde{\mathrm{MSE}}[n] \geq \overline{\mathrm{MSE}}[n]$. Thus, the proposed solution, which takes into account the effect of channel estimation errors, does yield improved channel prediction accuracy; this will be further confirmed via numerical simulation in Section IV.

C. Algorithmic Complexity

We go on to compare the computational complexity (measured by the number of flop counts of real addition and multiplication) of the proposed robust channel prediction scheme (3.7) and the conventional solution (3.12). First of all, the computations involved, as well as the required flop cost, for obtaining (3.7) are detailed as follows.

1) For each $i = 1, \ldots, L$, compute α_i in (3.9), β_i in (3.10), and γ_i in (3.11), $i = 1, \ldots, L$.
   a) To obtain α_i, computation of $[(\sigma_w^2 \mathbf{R}_{\mathbf{h}}^{-1} + \tilde{\mathbf{F}}_N^H \tilde{\mathbf{F}}_N)^{-1} \mathbf{e}_i]$ can be done via LU decomposition together with backward and forward substitutions [19]; this requires $4[(L(L-1)(L+1)/3) + L^2]$ multiplications and $(L(L-1)(4L+1)/3) + 4L^2 - 2L$ additions. Computation of the matrix–vector product $\sigma_{h,i}^2 \cdot [\mathbf{e}_i^T \tilde{\mathbf{F}}_N^H \tilde{\mathbf{F}}_N] \cdot [(\sigma_w^2 \mathbf{R}_{\mathbf{h}}^{-1} + \tilde{\mathbf{F}}_N^H \tilde{\mathbf{F}}_N)^{-1} \mathbf{e}_i]$ requires $4L + 4$ multiplications and $4L$ additions.
   b) To obtain β_i, computation of $\mathbf{R}_{\mathbf{h}} \tilde{\mathbf{F}}_N^H \tilde{\mathbf{F}}_N [(\sigma_w^2 \mathbf{R}_{\mathbf{h}}^{-1} + \tilde{\mathbf{F}}_N^H \tilde{\mathbf{F}}_N)^{-1} \mathbf{e}_i]$ requires $4L^2 + 4L$ multiplications and $4L^2$ additions. To obtain $[\mathbf{e}_i^T (\sigma_w^2 \mathbf{R}_{\mathbf{h}}^{-1} + \tilde{\mathbf{F}}_N^H \tilde{\mathbf{F}}_N)^{-1} \tilde{\mathbf{F}}_N^H \tilde{\mathbf{F}}_N] \times [\mathbf{R}_{\mathbf{h}} \tilde{\mathbf{F}}_N^H \tilde{\mathbf{F}}_N (\sigma_w^2 \mathbf{R}_{\mathbf{h}}^{-1} + \tilde{\mathbf{F}}_N^H \tilde{\mathbf{F}}_N)^{-1} \mathbf{e}_i]$, it requires $4L$ multiplications and $4L - 2$ additions.
   c) To obtain γ_i, it only requires $4L + 4$ multiplications and $4L$ additions, since $\tilde{\mathbf{F}}_N^H \tilde{\mathbf{F}}_N (\sigma_w^2 \mathbf{R}_{\mathbf{h}}^{-1} + \tilde{\mathbf{F}}_N^H \tilde{\mathbf{F}}_N)^{-1} \mathbf{e}_i$ and $(\sigma_w^2 \mathbf{R}_{\mathbf{h}}^{-1} + \tilde{\mathbf{F}}_N^H \tilde{\mathbf{F}}_N)^{-1} \mathbf{e}_i$ have been available.
2) Once α_i, β_i, and γ_i are obtained, the predictor coefficients $\bar{\mathbf{a}}_i^{(n)}$, $i = 1, \ldots, L$, $n \geq 1$, can be computed by using the Levinson–Durbin algorithm [6], which can exploit the Toeplitz structure of the system matrix for complexity reduction. Based on [19, pp. 194–196], a total number of $(2n-1)K^2$ multiplications and $(2n-1)K^2$ additions are needed.
3) Once $\bar{\mathbf{a}}_i^{(n)}$ is available, the computation of the predicted channel taps [using (3.1)] calls for $4K$ multiplications and $4K - 2$ additions.

To summarize the given results, the total flop counts required for obtaining the proposed scheme (3.7) are listed in Table I. By following similar analyses, the total flop counts needed for computing the conventional solution (3.12) are also given in the table. Fig. 1 plots the flop costs of the two channel prediction schemes (3.7) and (3.12) with respect to different lengths of the channel prediction phase (n); the system parameters are set to be $L = 6$ and $K = 3$. (The same setting is also used in the simulation section.) The figure shows that the algorithmic complexity of the proposed robust predictor (3.7) is higher; this is mainly due to the additional matrix inversion operations needed in computing coefficients α_i and β_i. Nevertheless, as will be seen in the following section, our robust scheme (3.7) can largely improve the MSE and BER performances as compared with the conventional solution (3.12).

TABLE I
FLOP COUNT COMPARISON FOR THE PROPOSED ROBUST CHANNEL PREDICTION SCHEME (3.7) AND THE NONROBUST SOLUTION (3.12) (n: THE LENGTH OF THE TRAINING PHASE, L: THE NUMBER OF CHANNEL PATHS, K: PREDICTOR ORDER, M: THE NUMBER OF OCCUPIED SUBCARRIERS)

	Number of flop counts (counting both addition and multiplication)
Proposed Robust channel predictor (3.7)	multiplications : $\left[K^2 + 4K + \frac{4L(L-1)(L+1)}{3} + 8L^2 + 16L + 8\right] \cdot L + (n-1)(2K^2 + 4K)L$ additions: $\left[K^2 + 4K + \frac{L(L-1)(4L+1)}{3} + 8L^2 + 10L - 4\right] \cdot L + (n-1)\left(2K^2 + 4K - 2\right)L$
Non-robust channel predictor (3.12)	multiplications : $K^2 + 4KL + (n-1)(2K^2 + 4KL)$ additions: $K^2 + (4K-2) \cdot L + (n-1)(2K^2 + (4K-2)L)$

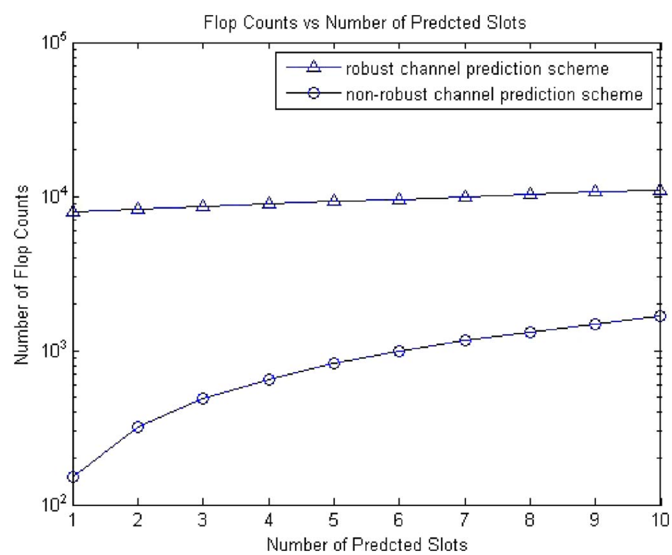

Fig. 1. Comparison of algorithmic complexity of the proposed robust channel predictor (3.7) with the conventional solution (3.12) with respect to different lengths of the channel prediction phase. The system parameters are set to be $K = 3$ and $L = 6$.

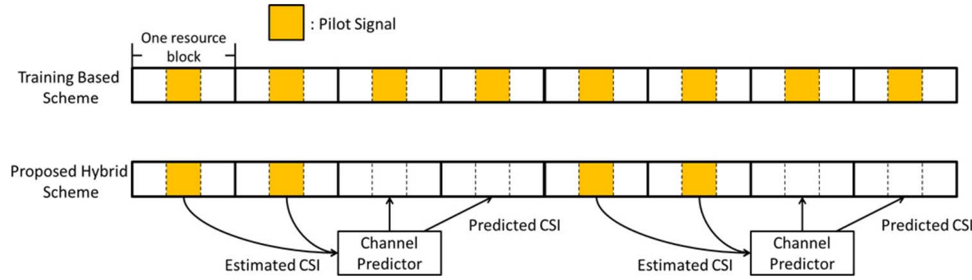

Fig. 2. (a) Schematic depiction of training-based channel estimation. (b) Schematic depiction of the proposed two-phase CSI acquisition protocol ($K = 2$ and $n = 2$).

D. Two-Phase CSI Acquisition Scheme for Training Overhead Reduction

The proposed robust channel prediction technique can be utilized for developing a reduced-overhead CSI acquisition scheme. Toward this end, we propose a two-phase CSI acquisition protocol, as shown in Fig. 2(b). The *training phase* consists of K consecutive training-based resource blocks, during each LMMSE channel estimation [5] is conducted. Once the K channel estimates in the training phase are available, a bank of K-tap channel predictors (one for each path) is employed in the *prediction phase* to predict the CSI for the next n consecutive resource blocks. Notably, since no training symbols are inserted in the resource blocks during the prediction phase, the training overhead is thus reduced.[2] Specifically, for a given n, which is the duration of the prediction phase, the proposed two-phase protocol can achieve

$$\eta = [n/(K+n)] \times 100\% \tag{3.17}$$

reduction in the training overhead, as compared with the training-based protocol. In the simulation section, we will further compare the proposed two-phase protocol with the training-based scheme in terms of the MSE of the acquired CSI and the resultant BER.

IV. Simulation Results

Here, computer simulations are conducted to illustrate the performance of the proposed approach and to validate our theoretical MSE study. We consider an SC-FDMA system, in which the size of FFT is 1024, the number of channel paths is $L = 6$, and the length of CP is equal to $L - 1 = 5$; the sum of the channel delay power profile is normalized so that $\sum_{i=0}^{5} \sigma_{h,i}^2 = 1$. We compare the proposed robust channel prediction scheme (3.7) with two other methods: the conventional nonrobust solution (3.12) and the training-based LMMSE channel estimator [5]. For training-based channel estimation, pilot symbol placement in each time slot follows the LTE-A standard. The prediction-based scheme is implemented by using the two-phase protocol proposed in Section III-D. In our simulations, we set $K = 3$, and the velocity of the mobile user is 10 km/h.

Fig. 3 compares the MSE achieved by the three channel acquisition schemes; for the two-phase protocol, the duration of the prediction phase is set to be $n = 1, 2, 3$, which result in, respectively, 25%, 40%, and 50% reduction in the training overhead [cf. (3.17)]. We can first see in the figure that, for both prediction-based solutions, the theoretical MSE is very close to the simulated outcome. This thus validates our MSE analyses shown in Section III-B. Moreover, it can be seen that the proposed robust channel prediction scheme (3.7) outperforms the conventional nonrobust solution (3.12) and, surprisingly, achieves an even smaller MSE as compared with the training-based approach. Figs. 4 and 5 then show the resultant BER performances for source symbols drawn from quaternary phase-shift keying (QPSK) and 16-quadrature amplitude modulation (QAM), respectively. As expected,

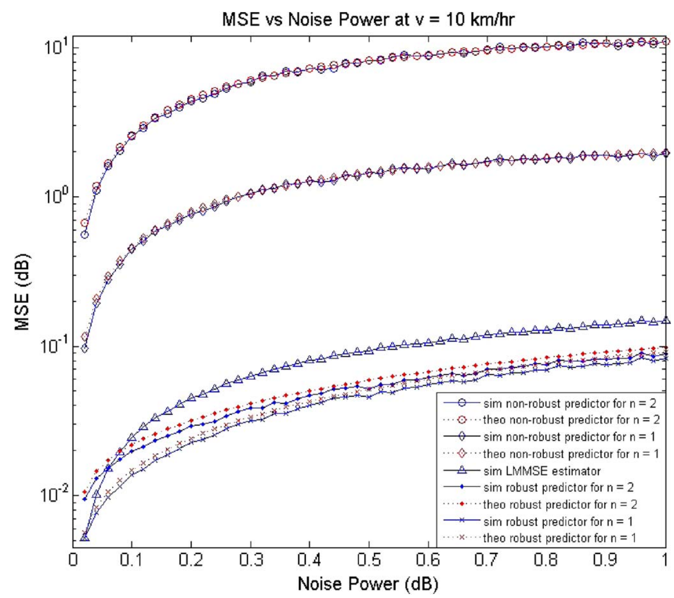

Fig. 3. MSE performances of the proposed robust predictor (3.7), the conventional nonrobust predictor (3.12), and the LMMSE training-based channel estimator [5].

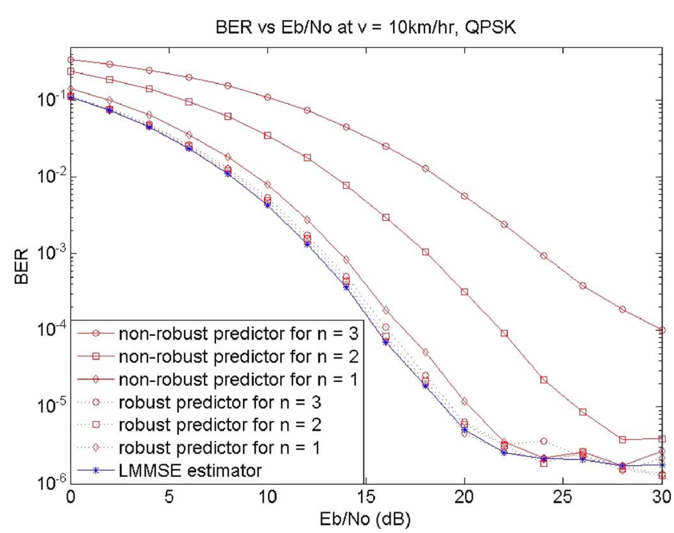

Fig. 4. BER performances of the proposed robust predictor (3.7), the conventional nonrobust predictor (3.12), and the LMMSE training-based channel estimator [5] (QPSK modulation).

[2]It is noted that the amount of transmission energy dedicated to training signals can be computed according to, e.g., [20], which takes into account the modulation types and various conditions of the propagation environment.

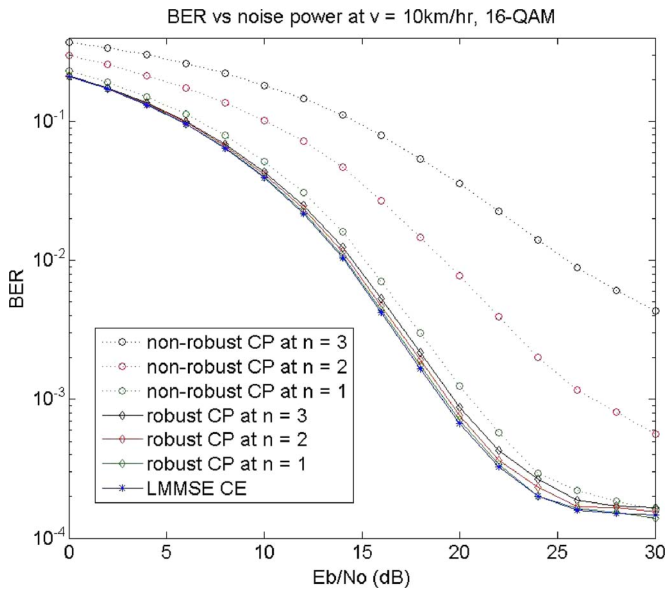

Fig. 5. BER performances of the proposed robust predictor (3.7), the conventional nonrobust predictor (3.12), and the LMMSE training-based channel estimator [5] (16-QAM modulation).

the BER of the proposed solution (3.7) is substantially lower than that of the conventional predictor (3.12) and is very close to the error rate attained by the training-based scheme.

V. Conclusion

We propose a robust linear channel prediction scheme for LTE-A uplink transmission under channel estimation errors. The design criterion is to minimize the mean square prediction errors, averaged over the distributions of the true channel coefficients and estimation errors. We derive a closed-form solution of the optimal channel predictor and provide exact MSE analyses. The reduction in MSE as compared with the conventional channel predictor is also analytically characterized. Simulation results validate our MSE analyses. In addition, the proposed channel prediction scheme combined with the training-based channel estimation can realize a two-phase CSI acquisition protocol for training overhead reduction. Through computer simulations, it is seen that such a two-phase protocol favorably compares with the LMMSE-based training scheme in terms of MSE and BER, at reduced signaling overhead.

Appendix A
Proof of Theorem 3.1

We shall first evaluate the cross-correlation term $E\{h_i(K+n)\widehat{h}_i^*(K+1-j)\}$ and auto-correlation term $E\{\widehat{h}_i(K+1-k)\widehat{h}_i^*(K+1-j)\}$ involved in (3.4). Toward this end, let us first write

$$E\left\{h_i(K+n)\widehat{h}_i^*(K+1-j)\right\} = \mathbf{e}_i^T E\left\{\mathbf{h}[K+n]\widehat{\mathbf{h}}^H[K+1-j]\right\}\mathbf{e}_i. \quad \text{(A.1)}$$

Based on (A.1), we have

$$\begin{aligned}
&\mathbf{e}_i^T E\left\{\mathbf{h}[K+n]\hat{\mathbf{h}}^H[K+1-j]\right\}\mathbf{e}_i \\
&\overset{(a)}{=} \mathbf{e}_i^T E\Big\{\mathbf{h}[K+n]\Big[\left(\sigma_w^2\cdot\mathbf{R}_\mathbf{h}^{-1}+\tilde{\mathbf{F}}_N^H\tilde{\mathbf{F}}_N\right)^{-1}\mathbf{F}_N^H \\
&\qquad\times\left(\tilde{\mathbf{F}}_N\mathbf{h}[K+1-j]+\mathbf{X}_p^{-1}[k]\mathbf{w}[k]\right)\Big]^H\Big\}\mathbf{e}_i \\
&\overset{(b)}{=} \mathbf{e}_i^T E\left\{\mathbf{h}[K+n]\mathbf{h}^H[K+1-j]\right\} \\
&\qquad\times\tilde{\mathbf{F}}_N^H\tilde{\mathbf{F}}_N\left(\sigma_w^2\cdot\mathbf{R}_\mathbf{h}^{-1}+\tilde{\mathbf{F}}_N^H\tilde{\mathbf{F}}_N\right)^{-1}\mathbf{e}_i \\
&\overset{(c)}{=} r[n+j-1]\cdot\mathbf{e}_i^T\mathbf{R}_\mathbf{h}\tilde{\mathbf{F}}_N^H\tilde{\mathbf{F}}_N\left(\sigma_w^2\cdot\mathbf{R}_\mathbf{h}^{-1}+\tilde{\mathbf{F}}_N^H\tilde{\mathbf{F}}_N\right)^{-1}\mathbf{e}_i \\
&\overset{(d)}{=} r[n+j-1]\cdot\sigma_{h,i}^2\cdot\mathbf{e}_i^T\tilde{\mathbf{F}}_N^H\tilde{\mathbf{F}}_N\left(\sigma_w^2\cdot\mathbf{R}_\mathbf{h}^{-1}+\tilde{\mathbf{F}}_N^H\tilde{\mathbf{F}}_N\right)^{-1}\mathbf{e}_i \\
&= r[n+j-1]\cdot\alpha_i \qquad \text{(A.2)}
\end{aligned}$$

where (a) follows by definition of $\hat{\mathbf{h}}[k]$ in (2.11), (b) holds since true channel $\mathbf{h}[k]$ and noise $\mathbf{w}[k]$ are assumed to be uncorrelated, and (c) and (d) are obtained through straightforward manipulations. To determine the auto-correlation term, by following similar procedures, we have

$$\begin{aligned}
&E\left\{\widehat{h}_i(K+1-k)\widehat{h}_i^*(K+1-j)\right\} \\
&= \mathbf{e}_i^T E\left\{\widehat{\mathbf{h}}[K+1-k]\widehat{\mathbf{h}}^H[K+1j]\right\}\mathbf{e}_i \\
&\overset{(e)}{=} \mathbf{e}_i^T\left(\sigma_w^2\mathbf{R}_\mathbf{h}^{-1}+\tilde{\mathbf{F}}_N^H\tilde{\mathbf{F}}_N\right)^{-1}\tilde{\mathbf{F}}_N^H \\
&\qquad\times E\left\{\left(\tilde{\mathbf{F}}_N\mathbf{h}[K+1-k]+\mathbf{X}_p^{-1}\mathbf{w}[K+1-k]\right)\right. \\
&\qquad\qquad\left.\times\left(\mathbf{h}^H[K+1-j]\tilde{\mathbf{F}}_N^H+\mathbf{w}^H[K+1-j]\mathbf{X}_p\right)\right\} \\
&\qquad\times\tilde{\mathbf{F}}_N\left(\sigma_w^2\mathbf{R}_\mathbf{h}^{-1}+\tilde{\mathbf{F}}_N^H\tilde{\mathbf{F}}_N\right)^{-1}\mathbf{e}_i \\
&\overset{(f)}{=} \mathbf{e}_i^T\left(\sigma_w^2\mathbf{R}_\mathbf{h}^{-1}+\tilde{\mathbf{F}}_N^H\tilde{\mathbf{F}}_N\right)^{-1}\tilde{\mathbf{F}}_N^H \\
&\qquad\times\left\{r[j-k]\cdot\tilde{\mathbf{F}}_N\mathbf{R}_\mathbf{h}\tilde{\mathbf{F}}_N^H+\sigma_w^2\cdot\delta[j-k]\cdot\mathbf{I}_M\right\} \\
&\qquad\times\tilde{\mathbf{F}}_N\left(\sigma_w^2\mathbf{R}_\mathbf{h}^{-1}+\tilde{\mathbf{F}}_N^H\tilde{\mathbf{F}}_N\right)^{-1}\mathbf{e}_i \\
&= \Big\{r[j-k]\cdot\mathbf{e}_i^T\left(\sigma_w^2\mathbf{R}_\mathbf{h}^{-1}+\tilde{\mathbf{F}}_N^H\tilde{\mathbf{F}}_N\right)^{-1}\tilde{\mathbf{F}}_N^H\tilde{\mathbf{F}}_N\mathbf{R}_\mathbf{h}\tilde{\mathbf{F}}_N^H\tilde{\mathbf{F}}_N \\
&\qquad\times\left(\sigma_w^2\mathbf{R}_\mathbf{h}^{-1}+\tilde{\mathbf{F}}_N^H\tilde{\mathbf{F}}_N\right)^{-1}\mathbf{e}_i\Big\}+\delta[j-k]\mathbf{e}_i^T \\
&\qquad\times\left(\sigma_w^2\mathbf{R}_\mathbf{h}^{-1}+\tilde{\mathbf{F}}_N^H\tilde{\mathbf{F}}_N\right)^{-1}\tilde{\mathbf{F}}_N^H\tilde{\mathbf{F}}_N\left(\sigma_w^2\mathbf{R}_\mathbf{h}^{-1}+\tilde{\mathbf{F}}_N^H\tilde{\mathbf{F}}_N\right)^{-1}\mathbf{e}_i \\
&= r[j-k]\cdot\beta_i+\delta[j-k]\cdot\gamma_i \qquad \text{(A.3)}
\end{aligned}$$

where (e) follows (2.11), and (d) holds due to the uncorrelated assumption between the channel and noise. Based on (A.2) and (A.3), the K linear equations in (3.4) can be put into a matrix–vector form as

$$(\beta_i\mathbf{R}+\gamma_i\mathbf{I}_K)\cdot\bar{\mathbf{a}}_i^{(n)} = \alpha_i\mathbf{r}_K^{(n)}. \quad \text{(A.4)}$$

The assertion of the theorem directly follows from (A.4). ■

Appendix B
Proof of Theorem 3.2

To derive (3.15), let us expand $\overline{\text{MSE}}[n]$ in (3.13) into

$$\begin{aligned}
&\overline{\text{MSE}}[n] \\
&= \sum_{i=1}^{L}\Bigg\{E\left\{|h_i(K+n)|^2\right\} \\
&\qquad -2\sum_{k=1}^{K}\bar{a}_{i,k}^{(n)}\text{Re}\left\{E\left(h_i(K+n)\widehat{h}_i^*(K+n-k)\right)\right\}
\end{aligned}$$

$$+\sum_{k=1}^{K}\sum_{j=1}^{K}\bar{a}_{i,k}^{(n)}E\left\{\widehat{h}_i(K+n-k)\times\widehat{h}_i^*(K+n-j)\right\}a_{i,j}^{(n)}\Bigg\}. \quad \text{(B.1)}$$

Since $E\{|h_i(K+n)|^2\}=\sigma_{h,i}^2$ and using (A.3), (B.1) can be further rewritten as

$$\begin{aligned}
&\overline{\text{MSE}}[n]\\
&=\sum_{i=1}^{L}\Bigg(\sigma_{h,i}^2-2\sum_{k=1}^{K}\bar{a}_{i,k}^{(n)}\text{Re}\left\{E\left(h_i(K+n)\widehat{h}_i^*(K+n-k)\right)\right\}\\
&\qquad+\sum_{k=1}^{K}\sum_{k=1}^{K}\bar{a}_{i,k}^{(n)}\left(r[j-k]\cdot\beta_i+\delta[j-k]\cdot\gamma_i\right)\bar{a}_{i,j}^{(n)}\Bigg)\\
&\overset{(a)}{=}1-\sum_{i=1}^{L}\Bigg(2\sum_{k=1}^{K}\bar{a}_{i,k}^{(n)}\text{Re}\left\{E\left(h_i(K+n)\widehat{h}_i^*(K+n-k)\right)\right\}\\
&\qquad+\sum_{k=1}^{K}\sum_{k=1}^{K}\bar{a}_{i,k}^{(n)}\left(r[j-k]\cdot\beta_i+\delta[j-k]\cdot\gamma_i\right)\bar{a}_{i,j}^{(n)}\Bigg)\\
&\overset{(b)}{=}1-\sum_{i=0}^{L-1}\left(2\alpha_i\bar{\mathbf{a}}_i^{(n)T}\mathbf{r}_K^{(n)}-\bar{\mathbf{a}}_i^{(n)T}\mathbf{R}_i\bar{\mathbf{a}}_i^{(n)}\right)\\
&\overset{(c)}{=}1-\sum_{i=0}^{L-1}\left[\alpha_i^2\cdot\mathbf{r}_K^{(n)T}\left(\beta_i\cdot\mathbf{R}_K+\gamma_i\cdot\mathbf{I}_K\right)^{-1}\mathbf{r}_K^{(n)}\right]
\end{aligned} \quad \text{(B.2)}$$

where (a) follows from the assumption $\sum_{i=1}^{L}\sigma_{h,i}^2=1$, (b) is obtained through direct manipulations, and (c) holds by invoking $\bar{\mathbf{a}}_i^{(n)}=\alpha_i\mathbf{R}_i^{-1}\mathbf{r}_K^{(n)}$ together with some manipulations. Equation (3.15) is thus proved. To derive (3.16), let us define $\mathbf{u}_i\triangleq[\widehat{h}_i[K]\quad\widehat{h}_i[K-1]\quad\cdots\quad\widehat{h}_i[1]]^T$ and then rewrite $\widetilde{\text{MSE}}[n]$ in (3.14) as

$$\widetilde{\text{MSE}}[n]=\sum_{i=1}^{L}E\left\{\left|\tilde{\mathbf{a}}_i^{(n)T}\mathbf{u}_i-h_i(K+n)\right|^2\right\}. \quad \text{(B.3)}$$

Since, from (3.7)

$$\begin{aligned}
\bar{\mathbf{a}}_i^{(n)}&=\alpha_i\mathbf{R}_i^{-1}\mathbf{r}_K^{(n)}=\alpha_i\left(\beta_i\cdot\mathbf{R}+\gamma_i\cdot\mathbf{I}_K\right)^{-1}\cdot\mathbf{r}_K^{(n)}\\
&=\frac{\alpha_i}{\beta_i}\cdot\left(\mathbf{R}+\frac{\gamma_i}{\beta_i}\cdot\mathbf{I}_K\right)^{-1}\cdot\mathbf{r}_K^{(n)}
\end{aligned} \quad \text{(B.4)}$$

using the matrix inversion lemma [18], we can obtain the following relation between $\bar{\mathbf{a}}_i^{(n)}$ and $\tilde{\mathbf{a}}_i^{(n)}$:

$$\begin{aligned}
\bar{\mathbf{a}}_i^{(n)}&=\frac{\alpha_i}{\beta_i}\cdot\left(\mathbf{R}+\frac{\gamma_i}{\beta_i}\cdot\mathbf{I}_K\right)^{-1}\cdot\mathbf{r}_K^{(n)}\\
&=\frac{\alpha_i}{\beta_i}\cdot\left[\mathbf{R}^{-1}-\mathbf{R}^{-1}\left(\mathbf{R}^{-1}+\frac{\beta_i}{\gamma_i}\mathbf{I}_K\right)^{-1}\mathbf{R}^{-1}\right]\cdot\mathbf{r}_K^{(n)}\\
&=\frac{\alpha_i}{\beta_i}\cdot\left[\mathbf{R}^{-1}-\frac{\gamma_i}{\beta_i}\left(\mathbf{R}+\frac{\gamma_i}{\beta_i}\mathbf{I}_K\right)^{-1}\mathbf{R}^{-1}\right]\cdot\mathbf{r}_K^{(n)}\\
&=\frac{\alpha_i}{\beta_i}\left[\mathbf{I}_K-\frac{\gamma_i}{\beta_i}\left(\mathbf{R}+\frac{\gamma_i}{\beta_i}\mathbf{I}_K\right)^{-1}\right]\cdot\mathbf{R}^{-1}\cdot\mathbf{r}_K^{(n)}\\
&=\frac{\alpha_i}{\beta_i}\left[\mathbf{I}_K-\frac{\gamma_i}{\beta_i}\left(\mathbf{R}_K+\frac{\gamma_i}{\beta_i}\mathbf{I}_K\right)^{-1}\right]\cdot\tilde{\mathbf{a}}_i^{(n)}.
\end{aligned} \quad \text{(B.5)}$$

By defining $\Delta_i^{(n)}\triangleq\tilde{\mathbf{a}}_i^{(n)}-\bar{\mathbf{a}}_i^{(n)}$ and based on (B.5), we have

$$\begin{aligned}
\Delta_i^{(n)}&=\tilde{\mathbf{a}}_i^{(n)}-\bar{\mathbf{a}}_i^{(n)}\\
&=\tilde{\mathbf{a}}_i^{(n)}-\frac{\alpha_i}{\beta_i}\left[\mathbf{I}_K-\frac{\gamma_i}{\beta_i}\left(\mathbf{R}+\frac{\gamma_i}{\beta_i}\mathbf{I}_K\right)^{-1}\right]\cdot\tilde{\mathbf{a}}_i^{(n)}\\
&=\left[\left(1-\frac{\alpha_i}{\beta_i}\right)\cdot\mathbf{I}_K+\frac{\alpha_i\cdot\gamma_i}{\beta_i}\left(\beta_i\cdot\mathbf{R}+\gamma_i\cdot\mathbf{I}_K\right)^{-1}\right]\cdot\tilde{\mathbf{a}}_i^{(n)}\\
&=\left[\left(1-\frac{\alpha_i}{\beta_i}\right)\cdot\mathbf{I}_K+\frac{\alpha_i\cdot\gamma_i}{\beta_i}\mathbf{R}_i^{-1}\right]\cdot\tilde{\mathbf{a}}_i^{(n)}.
\end{aligned} \quad \text{(B.6)}$$

With the aid of (B.6), $\widetilde{\text{MSE}}[n]$ in (3.14) reads

$$\begin{aligned}
&\widetilde{\text{MSE}}[n]\\
&=\sum_{i=1}^{L}E\left\{\left|\tilde{\mathbf{a}}_i^{(n)T}\mathbf{u}_i-h_i[K+n]\right|^2\right\}\\
&=\sum_{i=1}^{L}E\left\{\left|\left(\bar{\mathbf{a}}_i^{(n)}+\Delta_i^{(n)}\right)^T\mathbf{u}_i-h_i[K+n]\right|^2\right\}\\
&=\sum_{i=1}^{L}E\left\{\left|\left(\bar{\mathbf{a}}_i^{(n)T}\mathbf{u}_i-h_i[K+n]\right)+\Delta_i^{(n)T}\mathbf{u}_i\right|^2\right\}\\
&\overset{(d)}{=}\sum_{i=1}^{L}E\left\{\left|\bar{\mathbf{a}}_i^{(n)T}\mathbf{u}_i-h_i[K+n]\right|^2\right\}+\sum_{i=1}^{L}E\left\{\left|\Delta_i^{(n)T}\mathbf{u}_i\right|^2\right\}
\end{aligned} \quad \text{(B.7)}$$

where (d) holds because, with the orthogonality condition (3.3), it follows that

$$\begin{aligned}
&E\left\{\left(\bar{\mathbf{a}}_i^{(n)T}\mathbf{u}_i-h_i[K+n]\right)\left(\Delta_i^{(n)T}\mathbf{u}_i\right)^*\right\}\\
&\qquad=E\left\{\left(\bar{\mathbf{a}}_i^{(n)T}\mathbf{u}_i-h_i[K+n]\right)\cdot\mathbf{u}_i^H\right\}\cdot\Delta_i^{(n)}=0.
\end{aligned} \quad \text{(B.8)}$$

Finally, it remains to specify the term $E\{|\Delta_i^{(n)T}\mathbf{u}_i|^2\}$, which can be expressed as

$$\begin{aligned}
&E\left\{\left|\Delta_i^{(n)T}\mathbf{u}_i\right|^2\right\}\\
&=\Delta_i^{(n)H}\cdot E\left(\mathbf{u}_i\mathbf{u}_i^H\right)\cdot\Delta_i^{(n)}\\
&=\Delta_i^{(n)H}\cdot\mathbf{R}_i\cdot\Delta_i^{(n)}\\
&=\tilde{\mathbf{a}}_i^{(n)H}\cdot\left[\left(1-\frac{\alpha_i}{\beta_i}\right)\cdot\mathbf{I}_K+\frac{\alpha_i\cdot\gamma_i}{\beta_i}\mathbf{R}_i^{-1}\right]\\
&\quad\times\mathbf{R}_i\left[\left(1-\frac{\alpha_i}{\beta_i}\right)\cdot\mathbf{I}_K+\frac{\alpha_i\cdot\gamma_i}{\beta_i}\mathbf{R}_i^{-1}\right]\cdot\tilde{\mathbf{a}}_i^{(n)}\\
&=\tilde{\mathbf{a}}_i^{(n)H}\cdot\left[\left(1-\frac{\alpha_i}{\beta_i}\right)^2\cdot\mathbf{R}_i+2\cdot\left(1-\frac{\alpha_i}{\beta_i}\right)\right.\\
&\qquad\left.\cdot\frac{\alpha_i\cdot\gamma_i}{\beta_i}\cdot\mathbf{I}_K+\frac{\alpha_i^2\cdot\gamma_i^2}{\beta_i^2}\cdot\mathbf{R}_i^{-1}\right]\cdot\tilde{\mathbf{a}}_i^{(n)}
\end{aligned}$$

$$= \tilde{\mathbf{a}}_i^{(n)H} \cdot \mathbf{R}_i^{-1} \cdot \left[\left(1 - \frac{\alpha_i}{\beta_i}\right)^2 \cdot \mathbf{R}_i^2 + 2 \cdot \left(1 - \frac{\alpha_i}{\beta_i}\right) \right.$$

$$\left. \cdot \frac{\alpha_i \cdot \gamma_i}{\beta_i} \cdot \mathbf{R}_i + \frac{\alpha_i^2 \cdot \gamma_i^2}{\beta_i^2} \cdot \mathbf{I}_K \right] \cdot \tilde{\mathbf{a}}_i^{(n)}$$

$$= \tilde{\mathbf{a}}_i^{(n)H} \cdot \mathbf{R}_i^{-1} \cdot \left[\left(1 - \frac{\alpha_i}{\beta_i}\right) \cdot \mathbf{R}_i + \frac{\alpha_i \cdot \gamma_i}{\beta_i} \cdot \mathbf{I}_K \right]^2 \cdot \tilde{\mathbf{a}}_i^{(n)}. \quad \text{(B.9)}$$

Hence, (3.16) is thus proved. ■

REFERENCES

[1] H. Bogucka and A. Conti, "Degrees of freedom for energy savings in practical adaptive wireless systems," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 38–45, Jun. 2011.

[2] G. Y. Li, Z. Xu, C. Xiong, C. Yang, S. Zhang, Y. Chen, and S. Xu, "Energy-efficient wireless communications: Tutorial, survey, and open issues," *IEEE Wireless Commun.*, vol. 18, no. 6, pp. 28–35, Dec. 2011.

[3] H. G. Myung and D. J. Goodman, *Single Carrier FDMA: A New Interface for Long Term Evolution*. Chichester, U.K.: Wiley, 2008.

[4] S. Sesia, I. Toufik, and M. Baker, *LTE, The UMTS Long Term Evolution: From Theory to Practice*, 2nd ed. Chichester, U.K.: Wiley, 2011.

[5] S. C. Huang, J. C. Lin, and K. P. Chou, "Novel channel estimation techniques on SC-FDMA uplink transmission," in *Proc. IEEE VTC-Spring*, 2010, pp. 1–5.

[6] J. G. Proakis, C. M. Rader, F. Ling, C. L. Nikias, M. Moonen, and I. K. Proudler, *Algorithms for Statistical Signal Processing*. Englewood Cliffs, NJ, USA: Prentice-Hall, 2001.

[7] A. Svensson, "An introduction to adaptive QAM modulation schemes for known and predicted channels," *Proc. IEEE*, vol. 95, no. 12, pp. 2322–2336, Dec. 2007.

[8] A. Duel-Hallen, "Fading channel prediction for mobile radio adaptive transmission systems," *Proc. IEEE*, vol. 95, no. 12, pp. 2299–2313, Dec. 2007.

[9] A. Duel-Hallen, H. Hallen, and T. S. Yang, "Long range prediction and reduced feedback for mobile radio adaptive OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 5, no. 10, pp. 2723–2733, Oct. 2006.

[10] J. Akhtman and L. Hanzo, "Channel impulse response tap prediction for time-varying wireless channels," *IEEE Trans. Veh. Technol.*, vol. 56, no. 5, pp. 2767–2769, Sep. 2007.

[11] D. Schafhuber and G. Matz, "MMSE and adaptive prediction of time-varying channels for OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 593–602, Mar. 2005.

[12] J. Hajipour and V. C. M. Leung, "Proportional fair scheduling in multi-carrier networks using channel predictions," in *Proc. IEEE ICC*, 2010, pp. 1–5.

[13] J. F. Schmidt, J. E. Cousseau, R. Wichman, and S. Werner, "Bit loading using imperfect CSIT for prediction-based resource allocation in mobile OFDMA," *IEEE Trans. Veh. Technol.*, vol. 60, no. 8, pp. 4082–4088, Oct. 2011.

[14] V.-H. Pham, X. Wang, M. J. Rahman, and J. Nadeau, "Channel prediction-based adaptive power control for dynamic wireless communications," *Proc. IEEE VTC-Spring*, pp. 1–5, 2011.

[15] V. H. Pham, X. Wang, and J. Nadeau, "Long term cluster-based channel envelop and phase prediction for dynamic link adaptation," *IEEE Commun. Lett.*, vol. 15, no. 7, pp. 713–715, Jul. 2011.

[16] H. M. Tu, J. S. Lin, T. S. Chang, and K. T. Feng, "Prediction-based handover schemes for relay-enhanced LTE-A systems," in *Proc. IEEE WCNC*, 2012, pp. 2879–2884.

[17] G. L. Stuber, *Principles of Mobile Communications*, 2nd ed. Norwell, MA, USA: Kluwer, 2001.

[18] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.

[19] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD, USA: Johns Hopkins Univ. Press, 1996.

[20] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2349–2360, Sep. 2005.

[21] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1998.