Bit loading using imperfect CSIT for prediction based resource allocation in mobile OFDMA

Jorge F. Schmidt, Student Member, IEEE, Juan E. Cousseau, Senior Member, IEEE, Risto Wichman, and Stefan Werner, Senior Member, IEEE

Abstract—We present a prediction-based resource allocation algorithm (RA) for orthogonal frequency division multiple access (OFDMA) downlink, where inaccuracies in the wireless channel predictions are accounted for in the problem formulation. As the prediction quality degrades significantly with the prediction horizon, we propose a solution based on the histogram of the prediction error. This characterization also enables different mobile stations (MSs) to use different channel predictors as it does not rely on a specific prediction scheme. Using this characterization of the prediction error and based on classical resource allocation strategies we derive an algorithm that incorporates imperfect channel prediction information of future time slots. We evaluate the proposed algorithm using a practical low complexity channel predictor, suitable for implementation at the MSs. Simulation results show that the proposed algorithm outperforms previous prediction-based RA strategies without the characterization of the prediction error, and the system throughput is comparable to the case with perfect channel state information in the transmitter.

I. INTRODUCTION

OFDMA is a multiple access technique capable of exploiting multiuser diversity in a frequency selective fading scenario. Because of its orthogonal structure, OFDMA allows multiple users to transmit simultaneously on the different subcarriers of one OFDM symbol. Considering time-varying environments, the problem of allocating specific subcarriers to users over successive time slots, taking into account overall data throughput as well as fairness constraints, has received a lot of attention over the last years [1]-[8].

Predicted channel state information in the transmitter (CSIT) plays an important role on efficient sharing of channel resources in time varying wireless channels. For example, in long-term evolution (LTE) downlink [9], the physical layer scheduler allocates channel resources between mobile stations (MSs) in 1 ms resolution. The RA is based on achievable rate values reported by the MSs through a feedback channel. In this context, channel prediction has already been considered for compensating feedback latency in rapidly varying scenarios, see, e.g., [1], [2]. However, the use of channel predictors to improve RA algorithms has not been fully explored yet.

With the development of long range channel predictors [10]-[14], it becomes feasible to include information about the channel state on future time slots in the RA algorithm. However, in order to exploit this new information, prediction accuracy has to be taken into account. Specifically, in high mobility scenarios, channel prediction degrades appreciably as the prediction horizon increases [10], and the assumptions of perfect CSIT or constant prediction error are no longer valid. Several RA algorithms for OFDMA downlink considering imperfect CSIT have been proposed recently [1]-[4]. A modification of the proportional fair scheduler is proposed in [4] to predict the allocation for several time slots. This method is shown to improve fairness among users when comparing to the case without prediction. However, in high mobility scenarios and with practical implementation of channel predictors, two main issues related to prediction-based RA in rapidly varying channels can be identified: 1) The trade-off between prediction horizon and prediction accuracy, and; 2) Characterization of the prediction error to improve system performance.

In this paper, we propose a prediction-based RA algorithm for high mobility scenarios. The main concern is to characterize the prediction error associated to practical channel predictors. Imperfect CSIT due to prediction error has an important impact on the overall system performance as achievable rates of MSs are computed from CSIT by evaluating the channel power on each subchannel.

The main contributions of this paper are: 1) The derivation of a prediction-based RA algorithm, which is aware of prediction error and capable of providing close to perfect CSIT system throughput at the system BER constraint, and 2) A statistical description of the prediction error that does not assume a specific error model (usually associated to a specific prediction technique), and is thus feasible to be applied to general long range channel predictors. To the best of our knowledge, general theoretical models for long range predictors have not been developed yet.

Specifically, it is shown that based on the typical parameters of practical OFDMA systems, a set of histograms can be calculated and updated periodically to approximate the statistics of the prediction error. To corroborate the effectiveness of the proposed algorithm in practical OFDMA systems, we consider realistic system parameters as well as a realistic channel model for each MS. Also, as OFDMA downlink channel prediction is performed at mobile stations (MS), we consider in this work a computationally inexpensive channel predictor, practical for
implementation at the MSs [10].

The outline of the paper is as follows. In Section II we define the system model, introduce the notation, review the prediction-based RA literature and describe a long range channel predictor suitable for application on prediction based RA schemes. The error aware prediction-based RA is derived in Section III, where the effect of imperfect predicted CSIT is analyzed and an empirical characterization is proposed to compensate the prediction error. An analysis of the computational cost associated with the proposed scheme is presented in Section IV and the performance of the proposed prediction correction scheme is evaluated in Section V. Finally, Section VI provides our conclusions.

II. PROBLEM STATEMENT

A. System model

We consider an OFDMA downlink transmission where resources are allocated on subcarrier basis. On the other hand, in LTE and mobile WiMAX, resource blocks consist of several subcarriers, either contiguous or distributed within the OFDMA band, and the signaling of rate values corresponding to the resource blocks is vendor specific. Nevertheless, we chose to minimize the system assumptions and study a generic case where a resource block consists of one subcarrier. However, the concept can be adapted to other types of resource blocks as well.

The system under consideration has $N$ available subchannels, $K$ active MSs, and resources are allocated on a time slot basis, where a time slot consists of $M$ consecutive OFDM symbols. For each time slot $s$, the base station (BS) scheduler decides which MS is assigned to a particular subchannel $n$. More than one subchannel might be assigned to a MS depending on its rate requirement or channel state.

$P_f$ symbols carrying pilot subcarriers are evenly distributed over a time slot and $P_f$ pilot subcarriers are evenly distributed in frequency over subcarriers on the pilot symbols to aid channel estimation and prediction at the MSs. Pilot symbols are QPSK while data symbols can be taken from a set of available QAM constellations $[\beta_1, ..., \beta_K]$. Furthermore, convolutional coding of the data symbols is used to allow a finer grid of possible bit rates. The base station transmits on each subchannel using one of the modulation and coding schemes subject to the system BER constraint. MSs are assumed to experience independent channel fading with the same statistics. Regarding the time selectivity of the channel, we assume that the channel varies significantly from one time slot to the next. As the RA is performed on a time slot basis, we assume that the channel variation over a time slot can be neglected. Thus, the channel frequency response for time slot $s$ is given by

$$H_k(s, n) \approx \sum_{l=1}^{L} h_k(s, l)e^{-j2\pi \frac{2n}{W}}.$$  \hspace{1cm} (1)

which is a complex Gaussian random variable with zero mean and $\sigma_h^2$ variance. For $l = 1, ..., L$, $h_k(s, l)$ denotes the $l$-th channel tap for user $k$ over time slot $s$, being $L$ the length of the channel impulse response. CSIT is estimated for current time slot and predicted for the following $W$ time slots. There are two alternatives to acquire CSIT at the BS. The prediction can be performed either in BS or in mobile stations. The latter case requires a larger overhead in the feedback channel as $W$ estimates are fed back for each slot. In the former case, the overhead is less at the expense of a high computational load at the BS. The first alternative is implemented in this paper, while the derivations apply in both cases. In the following, transmitted power is assumed to be the same in all subchannels to emphasize the bit loading. We denote the rate achieved by user $k$ on time slot $s$ by $R_k(s)$, which is given as the sum rate over all subchannels assigned to user $k$ as $R_k(s) = M \sum_{n \in I_k} r_k(m, n)$, being $I_k$ the index set for the subchannels assigned to user $k$ and $r_k(m, n)$ the rate for user $k$ at symbol time $m$ on subchannel $n$. This way, the overall system throughput for time slot $s$ is given by

$$R(s) = \sum_{k=1}^{K} R_k(s).$$  \hspace{1cm} (2)

The design goal is to maximize the system throughput defined in (2) with constraints on total transmit power, system BER and fairness among users.

B. Prediction-based RA

A fair and efficient sharing of radio resources among users is an important design factor in wireless networks. Proportional fair scheduler (PFS) is a popular solution to provide a fair distribution of resources among users when CSIT is available. Denoting the 4th MS average data rate by $\bar{R}_k$, based on a fairness reasoning, it aims to maximize $\bar{R}_k$ over time, based on allocations on previous time slots. That is, $U(s+1) = \arg \max_{U} \sum_{k=1}^{K} \log(\bar{R}_k)$, where $U(s+1)$ is the utility function to be maximized for allocating time slot $s+1$ using information up to time slot $s$ and $U$ is the set of all possible allocations for the considered time slot. The algorithm keeps track of the average throughput $\bar{R}_k(s)$ of each MS in an exponentially weighted window of length $\tau$.

The performance of this scheduling algorithm can be improved, if channel predictions from MSs are available for future time slots. However, long-range channel predictors have not been fully exploited within resource allocation algorithms. An extension of the PFS to include CSIT prediction has been derived in [4]. This prediction-based PFS (P-PFS) calculates the achievable rates in a prediction window from slot $s+1$ to $s+W$, and the average rates at the end of this window are maximized to allocate time slot $s+1$. The prediction-based allocation scheme can be expressed as

$$U(s+W) = \arg \max_{U} \sum_{k=1}^{K} \log(\bar{R}_k^W),$$  \hspace{1cm} (3)

where $\bar{R}_k^W$ denotes the average rate of user $k$ at the end of the next $W$ time slots and is computed as

$$\bar{R}_k^W = \left(1 - \frac{1}{\tau}\right)^{W} \bar{R}_k(s) + \frac{1}{\tau} \sum_{w=1}^{W} \left(1 - \frac{1}{\tau}\right)^{W-w} \bar{R}_k(s+w).$$  \hspace{1cm} (4)
where $\hat{R}_k$ indicates a “virtual” allocated rate, as time slots $s+2$ to $s+W$ have not been allocated yet. A practical algorithm to implement (3) is also given in [4], which will be used in this paper.

C. Channel prediction

To be able to use information on future time slots, achievable rates must be estimated within the prediction window. In what follows we briefly describe a low-complexity channel predictor [10] that will be used both as a motivating example and for testing the proposed prediction-based RA scheme.

To simplify the notation, we focus in this subsection on only one MS and drop MS index $k$. It is shown in [10] that the time variation over $M$ (several times larger than $M$) OFDM symbols of the channel coefficient corresponding to one subcarrier can be well described in terms of a size $M$ discrete cosine transform (DCT) truncated to its first $G \ll M$ basis functions. The basis dimension $G$ is determined from the maximum expected channel Doppler shift and the DCT energy compression characteristics. By interpreting the DCT basis functions as the impulse responses of ideal bandpass filters centered at the cosines frequencies it is also shown that an approximation $\hat{H}_{FB}(m,n)$ of $H(m,n)$ for $m = 1, \ldots, M$ can be obtained by recursive filtering of $H(m,n)$ with a second order IIR filter bank based on a normalized allpass lattice realization, which frequency response is given by

$$H_{FB}(e^{j\omega}) = \frac{0.5(1 - s_{20})(1 + e^{-j2\omega})}{1 - s_{20}e^{-j2\omega}} + \sum_{i=1}^{G-1} \frac{0.5(1 - s_{2i})(1 - e^{-j2\omega})}{1 + (s_{2i} + 1)s_{1i}e^{-j\omega} + s_{2i}e^{-j2\omega}},$$

(5)

where $s_{1i} = -\cos\left(\frac{i\pi}{M}\right)$ is the lattice parameter defining the filters central frequencies, and $s_{2i}$ ($0 < s_{2i} < 1$) is related to the $3$ dB bandwidth of each narrow band filter. By defining $x_i(m,n)$, $0 \leq i \leq G-1$, the state vector for each passband filter on subcarrier $n$, the filter bank structure of (5) can be expressed in state space form as

$$x_i(m+1,n) = A_i x_i(m,n) + b_i H(m,n)$$
$$\hat{H}_{FB}(m,n) = c_i x_i(m,n) + d_i H(m,n)$$

(6)

where, $A_i$, $b_i$, $c_i$, and $d_i$ are the state space description of the filters in (5). Using (6), the optimum algorithm for estimating $H(m,n)$ in AWGN is a set of $G$ scalar Kalman filters for the filter bank of (6). Further, as the DCT basis functions are fixed for the whole transmission, the Kalman gains have a steady state solution resulting in a low complexity limiting Kalman filter. Having the Kalman estimate $\hat{H}(m,n)$ as input, a long range channel predictor can be obtained by decimating on $T$ (determined according to the maximum channel Doppler shift such that the decimated channel is sampled above the Nyquist frequency) the input estimates, scaling up in frequency the passband filters by the same factor $T$ and using $L$ step extrapolation on the limiting Kalman filter. In this manner, the following limiting Kalman predictor can be obtained as

$$e_p(m+\ell,n) = \hat{H}(m+\ell-T,n) - \hat{H}(m+\ell-T,n)$$
$$\mathbf{x}_p(m+\ell,n) = (A_p)^{\ell} \mathbf{x}_p(m+\ell-\ell L,n)$$

$$+ k_p e_p(m+\ell,n)$$

$$\hat{H}_p(m+\ell,n) = e_p^T \mathbf{x}_p(m+\ell,n)$$

$$\hat{H}_p(m+\ell+T(L-1),n) = \sum_{i=0}^{L-1} \hat{H}_p(m+\ell,n),$$

(7)

where $\ell = 0, \ldots, (T \times L) - 1$ are the state vector samples used at each iteration, $\mathbf{x}_p$ is the predictor state vector, $A_p$ and $c_p$ are the transition matrix and output vector of the passband filters scaled up in frequency by $T$, and $k_p$ are the corresponding steady state Kalman gains. Finally, to obtain the predicted channel over all subcarriers, one limiting Kalman filter predictor is run over each of the $P_f$ available pilot subcarriers and interpolated for all other subcarriers using DCT interpolation in frequency.

III. ERROR AWARE PREDICTION-BASED RA

In this section we evaluate achievable rates $\hat{r}_k(s+w,n)$ for future time slots (input to the P-PFS of Section II-B) when imperfect CSIT due to prediction errors is available. When perfect CSIT is available, extending [17] for coded modulation, the instantaneous BER for subchannel $n$ in symbol time $m$ can be approximated by

$$P_e(s,n) \approx c_1 \exp \left\{-c_2 \gamma \beta(s,n) |H(s,n)|^2 \frac{2^{\beta(s,n)} - 1}{2^{\beta(s,n)} - 1}\right\},$$

(8)

where $c_1 = 0.2$, $c_2 = 1.6$, $\beta(s,n)$ is the number of bits per symbol of the QAM constellation used for a subchannel and time slot and $\delta(s,n)$ is the coding gain, with respect to the uncoded case, for the corresponding QAM constellation. SNR value $\gamma$ accounts for the path loss while small scale fading effects are represented by $H(s,n)$. By evaluating (8) for the possible pairs of $\beta(s,n)$ and $\delta(s,n)$, the one that satisfies the system BER constraint is selected and determines the number of bits per symbol $\hat{r}(s,n)$ for transmission.

In a practical scenario, it is impossible to have perfect CSIT and the determination of $\hat{r}(s,n)$ is based on an estimated $\hat{H}(s,n)$ of $H(s,n)$. In this case, the use of (8) no longer guarantees that the BER constraint is satisfied. In case of imperfect CSIT [17], an average BER $P_e(s,n) = E_{|H(s,n)|}[\hat{H}(s,n)] \{P_e(s,n)\}$ is considered instead of the instantaneous BER of (8), where the expectation is evaluated over $|H(s,n)|[\hat{H}(s,n)]$. By defining the random variable $\alpha = |H(s,n)|[\hat{H}(s,n)]$ it results in

$$P_e(s,n) = \int_0^\infty P_e(s,n) f(\alpha) d\alpha$$

$$= \int_0^{c_1} \exp \left\{-c_2 \gamma \beta(s,n) |\alpha|^2 \frac{2^{\beta(s,n)} - 1}{2^{\beta(s,n)} - 1}\right\} f(\alpha) d\alpha,$$

(9)

where $f(\alpha)$ is the probability density function (pdf) of $|H(s,n)|[\hat{H}(s,n)]$. When $\hat{H}(s,n)$ is an estimate or slightly delayed estimate of the true channel coefficient $H(s,n)$, several assumptions hold regarding $f(\alpha)$ [17], [18], [2] and closed form solutions can be derived for evaluating (9).
Unfortunately, for the prediction-based RA application considered here, $H(s, n)$ is a prediction of $H(s, n)$ based on significantly older CSIT and a derivation of $f(\alpha)$, besides complex to obtain, depends on the characteristics of the channel predictor used to obtain $H(s, n)$. In the following paragraphs we reformulate (9) in terms of the prediction error instead of $\alpha$, as the former is the available observation in case of the prediction-based RA.

For a prediction-based RA algorithm considering CSIT up to $W$ time slots in the future, $W$ different pdf should be evaluated. Let us define for each considered prediction horizon the random variables $e_w = (\alpha_w - [\hat{H}(s+w, n)]_{|H(s, n)}) / \sigma_H^2$ such that the pdfs of $\alpha_w$ and $e_w$ are related by

$$f_{\alpha_w}(\alpha_w) = \frac{1}{\sigma_H^2} f_{e_w}(e_w).$$

Replacing (10) in (9), the estimated BER for each prediction horizon can be evaluated in terms of the defined normalized prediction error $e_w$ using

$$P_{ew}(s+w, n) = \hat{P}_{ew}(s+w, n) \times \rho_w(n),$$

where $\hat{P}_{ew}(s+w, n)$ denotes the evaluation of (8) for the available predicted value of the channel coefficient

$$\hat{P}_{ew}(s+w, n) = c_1 \exp \left\{ -\frac{c_2 \gamma \delta(s+w, n) H(s+w, n)_{|H(s, n)}^2}{2(\delta(s+w, n) - 1)} \right\},$$

and $\rho_w(n)$ is the correction factor compensating the imperfect CSIT through the normalized prediction error $e_w$ using

$$\rho_w(n) = \frac{1}{\sigma_H^2} \int_{-\infty}^{\infty} \exp \left\{ -\frac{c_2 \gamma \delta(s+w, n) \theta(s+w, n)}{2(\delta(s+w, n) - 1)} \right\} f_{e_w}(e_w) \, de_w,$$

where $\theta(s+w, n) = 2 \sigma_H^2 \hat{H}(s+w, n)_{|H(s, n)} e_w + \sigma_H^2 \hat{e}_w$. It is clear that for perfect CSIT $\rho_w(n) = 1$. We derive next an estimator of $f_{ew}(e_w)$ in (13) to evaluate this correction factor.

Assuming that the prediction error $e_w$ is identically distributed over the system subchannels and is independent of $H(s, n)$, the number $Q$ of prediction error samples for the different subchannels can be used to construct a histogram to estimate $f_{ew}(e_w)$. Following [19, Ch. 5.5], to find $\hat{f}_{ew}(e_w)$ we use the empirical pdf. We expect that as $Q \to \infty$, $\hat{f}_{ew}(e_w)$ approaches the true density $f_{ew}(e_w)$.

To derive the estimator, let the $i$th histogram interval be centered on $\epsilon_i$, and let $\omega$ be its width. Hence the $i$th interval will be denoted by $\Delta_i = \{e_w : \epsilon_i - \frac{\omega}{2} < e_w \leq \epsilon_i + \frac{\omega}{2}\}$, so that the histogram $\hat{f}_{ew}(e_w)$ is defined by

$$\hat{f}_{ew}(e_w) = \frac{Q_i}{Q \omega} \text{ for } e_w \in \Delta_i,$$

being $Q_i$ the number of observations for which $e_w \in \Delta_i$. The interval width $\omega$, together with the setting of a maximum error threshold $\epsilon_Q$, determine the number of probability mass points of the estimator. The bias of $f_{ew}(e_w)$ as an estimator of $f_{ew}(e_w)$ can be shown to be for each interval $b(\hat{f}_{ew}(e_w)) \approx f_{ew}(\epsilon_i) \frac{\omega}{2} \frac{Q_i}{Q \omega}$, where $f_{ew}(\epsilon_i)$ is the second derivative of $f_{ew}(\epsilon_i)$, and the normalized standard error $\sigma_\epsilon$ for estimating $f_{ew}(\epsilon_i)$ is

$$\sigma_\epsilon = \sqrt{\frac{1 - \omega f_{ew}(\epsilon_i)}{Q \omega f_{ew}(\epsilon_i)}}.$$

It can be noted that the effect of the cell width $\omega$ on $\sigma_\epsilon$ is the opposite to that on the bias such that an appropriate choice of $\omega$ is to set $\omega \propto 1/\sqrt{Q}$ [19].

Using the estimated pdf for the prediction error, (13) can be rewritten in terms of (14) as

$$\hat{P}_{ew}(s+w, n) = \frac{1}{\sigma_H^2} \sum_{i=1}^{Q} \exp \left\{ -\frac{c_2 \gamma \delta(s+w, n) \theta(s+w, n)}{2(\delta(s+w, n) - 1)} \right\} \hat{f}_{e_i}(\epsilon_i),$$

being $\theta(s+w, n) = 2 \sigma_H^2 \hat{H}(s+w, n)_{|H(s, n)} \epsilon_i + \sigma_H^2 \hat{e}_w$. This last expression is then inserted into (11) replacing $\rho_w(n)$ and used to evaluate $P_{ew}(s+w, n)$ on each subchannel such that

$$P_{ew}(s+w, n) = \hat{P}_{ew}(s+w, n) \times \hat{\rho}_w(n).$$

The pair $\beta^*(s+w, n)$, $\delta^*(s+w, n)$ satisfying (16) determines the achievable rate $F(s+w, n)$ on subchannel $n$. It is worth noting that as the number of subcarriers in typical OFDMA systems is usually large (above 512 subcarriers), the bias and standard error will be small, and the $W$ histograms built for each MS will describe the statistics of the prediction error accurately. Further, the histograms for estimating $f_{ew}(\epsilon_i)$ are constructed assuming perfect CSIT at the MS for current time slot. In a practical implementation $H(s, n)$ is estimated at the MS for symbol detection. Assuming that the estimation error is much lower than the prediction error in the prediction range, which is a reasonable assumption for practical estimators/predictors (as will be shown in Section V), the former can be used as a reference for the characterization of the latter without affecting much the resulting $f_{ew}(\epsilon_i)$. Also, if we further assume that the statistics of $e_w$ are slowly time variant, then the histograms can be periodically improved with the incorporation of new data points. That is

$$\hat{f}_{e_i}^{(s)}(\epsilon_i) = \xi \hat{f}_{e_i}^{(s)}(\epsilon_i) + (1 - \xi) \hat{f}_{e_i}^{(s-W)}(\epsilon_i),$$

where $0.5 \leq \xi \leq 1$ is a tuning parameter selected to compensate for the possibly time-varying statistics of $e_w$. Finally, the resulting values of $f_{ew}(s+w, n)$ corresponding to each MS are fed to the prediction-based RA algorithm.

IV. COMPLEXITY ANALYSIS

The proposed prediction error characterization/compensation scheme involves an additional computational cost when compared with the baseline scenario of uncompensated imperfect CSIT. For the uncompensated case, the computational cost for the bit loading is given by the evaluation of (8) for the possible pairs of $\beta$ and $\delta$. Using the binary search algorithm to this end, the average number of required evaluations of (8) for this search is $\log_2(K) - 1$ and at most $\log_2\hat{K}$ probes in the worst case, being $\hat{K}$ the number of available $\beta-\delta$ pairs, which is a small number for practical systems. On the other hand, for the proposed scheme, the computational cost involved in the histogram construction of (14) and in the evaluation of the correction factor $\rho_w(n)$ of (15) should be considered besides the evaluation of $P_{ew}$ similar to (8).
The constructed histograms have $2Q + 1$ bins (probability mass points). The value of $Q_i$ in (14) for each probability mass point can also be evaluated using the binary search algorithm requiring on average $\log_2(2Q + 1)$ trials for each sample. Thus, the computational cost associated to the $W$ histograms construction is evaluated as $O(WQ \log_2(2Q + 1))$, and assuming that the histograms are updated every $W$ time slots it reduces to $O(Q \log_2(2Q + 1))$. Regarding the computation of the correction factor $\hat{\rho}_w(n)$ of (15), it must be evaluated for all possible pairs of $\beta$ and $\delta$ to be then inserted into (16). The evaluation of (15) involves $O(2Q + 1)$ complex operations, which leads to a computational cost of $O(\bar{K}(2Q + 1))$ complex operations.

Summarizing the discussion from the previous paragraphs, the additional computational cost of the proposed scheme, when compared to the baseline of the uncompensated case, is $O(Q \log_2(2Q + 1) + \bar{K}(2Q + 1))$. Thus, in practice the complexity is dominated by the number of error samples $Q$, which is much larger than the number of available modulation/coding pairs $\bar{K}$.

V. NUMERICAL EVALUATION

In this section we evaluate the impact of the prediction error aware bit loading presented in Section III to the performance of the prediction-based RA algorithm of [4]. We seek to quantify the performance improvement achieved with the proposed empirical characterization of the prediction error.

We consider the downlink of an OFDMA system based on 3GPP specification [9] operating at a carrier frequency $f_C = 2$ GHz with 10 MHz bandwidth, 15 KHz subcarrier spacing and $5 \mu$s cyclic prefix. FFT size is 1024 and $N = 600$ subcarriers are in use occupying 90% of the bandwidth. The length of the time slot is $M = 15$ OFDMA symbols, equivalent to 1 ms and the system uses 5% pilot ratio. 4, 16 and 64-QAM constellations are available for data transmission. Also, the system employs convolutional coding for data transmission using the (133,171) rate 1/2 code with puncturing to obtain 2/3 and 3/4 code rates. A bit loading function similar to [8] is used corresponding to a system target BER of $1 \cdot 10^{-3}$. All the simulation results are obtained by averaging over time in order to average the channel fading statistics.

Based on the derivations on Section III, the parameters for the construction of the histograms are chosen as $Q = 600$ (all system subchannels used), $\omega = 0.078$ and the limits for the histogram set to $\pm 0.8$. Finally, parameter $\xi$ for updating of the histograms is set to $\xi = 0.7$. To evaluate a realistic propagation environment, we consider independent wireless fading channel for each MS following the ITU-Vehicular A model, which results in a 27-tap frequency selective channel for the given system parameters. Each channel tap varies in time according to Jakes Doppler spectrum, and we employ the practical channel estimator/predictor structure of [10]. Following [10], the parameters for this channel predictor are $M = 12M$ such that $s_{11} = -\cos(\frac{\pi}{M})$. It is shown in [10] that the performance is robust to the typical values of the selectivity parameter $s_{21}$ such that its value is set to 0.998. Also, for the MSs speeds of interest a value of $G = 8$ results for the basis expansion dimension. Finally, the extrapolation factor for the Kalman predictor is fixed to $\bar{L} = 3$ and the decimation factor $T$ adjusted to meet the different prediction horizons.

Before addressing the performance of P-PFS, Fig. 1 illustrates the virtual rates $\bar{R}_k(s+w)$ in (4) (input to P-PFS) when imperfect CSIT is available. The figure shows the average computed $\bar{R}_k(s+w)$ for a single MS, over 3 successive time slots based on the predicted CSIT with a SNR of 25 dB. The input to the P-PFS of (3) is the average rate $\bar{R}_k^W$ in (4) of user $k$ at the end of the $W = 3$ time slots considered computed based on these virtual rates $\bar{R}_k(s+w)$. In general, it noticed that imperfect CSIT degrades the virtual rates. However, even with imperfect CSIT the prediction increases system throughput as will be seen later on. The proposed bit loading based on (16) is compared against the uncompensated case (8), and the case of perfect CSIT is also shown for reference. As expected, the average bit loading is independent on the prediction horizon when perfect CSIT is available. Two bars for imperfect CSIT are shown for the proposed bit loading when the true channel (Perf. Ref.) and channel estimates (Est. Ref.) are employed for the histograms construction. It can be noted that the proposed scheme effectively reduces the loading gap between the perfect and imperfect cases obtaining a reduction from 0.35, 0.75 and 1.1 to 0.1, 0.35 and 0.6 bits/subchannel for $w = 1, 2$ and 3 time slots prediction respectively. Although not used by P-PFS, bit loading without prediction $(w = 0)$ is also plotted for reference.

![Fig. 1. Average bit loading per subcarrier for no prediction and](image-url)
the same time the impact of the degraded virtual rates due to imperfect CSIT $R_k(s + w)$ (illustrated in Figure 1) on the performance of P-PFS is evident.

In Figure 3, system throughput is evaluated for a prediction window of length $W = 3$ and up to 20 MSs moving at 25km/h in a high and low SNR scenarios. Channel power for MSs is uniformly distributed between -3 and 0dB. The 25dB and 10dB SNR settings result in channel estimate references with MSE of -60.15dB and -33.64dB respectively for the construction of the histograms. In both scenarios, the results for Perfect CSIT are plotted as a best performance bound. As the RA for the next time slot is based on the computation of the average rates achieved by MSs at the end of the prediction window, if the computed achievable rates for future time slots differ too much from the true achievable rates, the prediction-based RA performance will degrade in terms of system throughput or in terms of the system BER constraint. P-PFS with imperfect CSIT is not able to exploit the available resources to full extent and the system throughput is reduced, in particular for the case of few MSs. On the other hand, it can be seen that the proposed correction scheme gives close to perfect CSIT performance, which demonstrates its effectiveness. This happens because, as the histograms constructed in (14) are not restricted to have zero mean, the correction factor of (15) has the capability to compensate a bias on the predicted CSIT. As expected, in this low SNR regime, system throughput is reduced as channel quality for the MSs is degraded. However, the proposed error correction scheme still reduces effectively the throughput gap for both references used in the histograms construction. It is worth noting that the performance improvement obtained with the proposed scheme is larger in this last case, as the uncompensated bit loading cannot approach the perfect CSIT case even for a large number of users (where the probability of having a user with good channel state at each time slot increases). This latter conclusion justifies the added complexity of the proposed scheme.

Figure 4 shows again the system throughput for the high SNR setting of Figure 3, this time evaluated for a linear predictor [12] and a sum of sinusoids predictor [11], which we refer to as MMSE and ESPRIT predictors respectively. It is worth noting that the computational cost of these predictors is significantly larger than that of the predictor of [10] so that their implementation at the MSs may not be feasible. It is observed for both predictors, the proposed scheme also gives close to perfect CSIT performance showing that it can be applied to different prediction schemes as the characterization of the prediction error is independent of the channel prediction scheme.

Regarding BER behavior using the proposed error correction scheme, Figure 5 shows the BER results for the high SNR setting of Figure 3 and all channel predictors tested. As the bit loading function used is discrete, all curves are below the BER constraint as expected and thus, the Perfect CSIT case is used as a reference for performance comparison. When the channel power is underestimated, the uncompensated bit loading algorithm assigns modulation/coding schemes that meet the BER constraint for this underestimated power. If the channel state is in reality better, the BER attained with the assigned more robust modulation/coding pair will be lower, indicating that the system resources are not fully exploited. On the other hand, if the channel power is overestimated, the assigned modulation/coding scheme will not be supported by the real channel state, thus leading to a higher BER, indicating a loss in transmission reliability. It is easy to see in Figure 5 that while the RBEM predictor of Section II-C underestimates the channel power, the MMSE and ESPRIT predictors overestimate the channel power. However, in all cases the proposed compensation scheme effectively reduces the BER gap relative to the Perfect CSIT case. As observed
in Fig. 3, the constructed histograms take into account a possible under/overestimation of the channel power and thus can compensate this effect in the bit loading. In the high SNR scenarios of Figures 3 and 4, for a large number of users all schemes converge to the same total system throughput. This happens because as the number of users increases, the probability of having a user with good channel conditions at all time slots increases, and so does the system throughput. However, achievable rates computed by the different schemes are not necessarily the same. Thus, MSs scheduled by P-PFS might not be the same in all cases leading to different BER performance for each scheme. This figure shows that the proposed technique approaches the Perfect CSIT case for all schemes, demonstrating that it does not rely on a specific prediction structure.

A last simulation setting is shown in Figure 6 where we further investigate the performance dependence on the MSs speed. We consider again a prediction window of $W=3$ time slots and 10 MSs moving at speeds from 50 to 200km/h with 25dB SNR. This figure shows how the bit loading gap increase for higher speeds (prediction error increases for fixed prediction range) results in a throughput loss for high MSs speeds. In all cases the proposed scheme reduces this loss significantly compared to the uncompensated case.

VI. Conclusions

We proposed a characterization of the prediction error for prediction-based resource allocation for OFDMA downlink over mobile wireless channel when imperfect channel state information is available. Based on the large amount of frequency data samples available in a typical OFDMA system, we derived an empirical approach based on histograms for the characterization of the prediction error for the different prediction horizons considered in the prediction window.

We evaluated the proposed scheme under realistic channel conditions, system parameters and a practical channel predictor feasible for implementation at mobile stations. Simulation results indicate that the proposed scheme outperforms similar prediction-based resource allocation algorithms that disregard the prediction error.

REFERENCES