In the downlink of an Orthogonal Frequency Division Multiplexing (OFDM) system, pilot symbols are transmitted across the time-frequency plane for the receiving device to estimate the channel’s response. The density of such pilot is often designed to accommodate the devices under most extreme channel conditions. This worst case scenario design results in unnecessary overhead that could have been used for data transmission otherwise. In this paper, we propose an adaptive pilot allocation scheme that, in addition to a low-rate periodically transmitted pilot symbols, sends device associated pilot adaptively only when called for by the channel condition, thereby reducing the overall pilot overhead. Furthermore, the placement of signaling channel and reduced complexity channel estimator required to facilitate this adaptive operation are also addressed.

I. INTRODUCTION

In an Orthogonal Frequency Division Multiplexing (OFDM) system, known symbols referred to as pilot are transmitted across the time-frequency plane for the receiving device to estimate the channel’s time-frequency response in order to perform coherent demodulation of data symbols. In the downlink of a cellular system where these pilot symbols are transmitted continuously for all devices to perform various synchronization tasks, they are also called a common pilot. Since the channel’s time-frequency response is a slow-varying two-dimensional process, the pilot symbols essentially sample this process and therefore need to have a density that is high enough for the receiving device to reconstruct (or interpolate) the full response.

From the sampling theory point of view, in order to guarantee aliasing-free reconstruction of the channel’s response for all devices, the rate at which these pilot symbols are inserted should be greater than or equal to the channel’s Nyquist rate of the device under most extreme channel condition [1]. In fact, it has been suggested [2] and widely accepted that the pilot insertion rate should be twice the Nyquist rate for practical consideration. As an example, the Mobile WiMax [3] air interface has a pilot density of $1/7$, a 14% overhead.

Since the pilot symbols occupy the radio resource that could be used by data transmission otherwise, it is desirable to keep this overhead as low as possible. This is especially critical for future generation cellular system in which large number of base stations with multiple antennas may be visible to a terminal simultaneously. Recently, this has been quantified by the capacity analysis [4] accounting for pilot overhead conducted in the context of a future system being considered in the IST WINNER (World Wireless Initiative New Radio) project [5].

In this paper, we propose an adaptive pilot allocation scheme for the downlink of an OFDM system. The basic concept is to transmit the common pilot symbols at a rate that covers only the majority of devices that are under typical channel conditions. Additional data-associated pilots are allocated to the remaining devices that are under extreme channel conditions.

Although very simple conceptually, other supporting features are required to facilitate such adaptive operation. First of all, the additional pilot disrupt the common pilot’s regularity for which simplified channel estimators are available. A reduced complexity channel estimator that can incorporate the two pilots of distinct characteristics is therefore needed. Secondly, to keep the devices under extreme channel conditions attached to the system before additional pilot symbols are allocated, the control signaling needs to be robust enough to function in all channel conditions. These are the two main issues that will be addressed in this paper.

The remainder of the paper is organized as follows. Sec. II describes the OFDM system model and our basic design principle. The reduced-complexity channel estimator based on the common and data-associated pilots and its error characteristics are given in Sec. III. Sec. IV then introduces a robust distribution of control signaling together with some modem simulation results. A brief summary is given in Sec. V.

II. SYSTEM MODEL AND BASIC DESIGN PRINCIPLE

Fig. 1 shows the basic parameters in an OFDM system on a time-frequency plane. The symbol duration $T_s$ includes a cyclic prefix $T_{cp}$ that ensures sub-carrier orthogonality in dispersive channel. The channel’s maximum delay-Doppler spread $(\tau_{\text{max}}, \nu_{\text{max}})$, normally a fraction of the OFDM symbol’s time duration and sub-carrier spacing, is shown in a larger scale. In order to meet the time domain Nyquist criterion of sampling the channel response, there must be at least one pilot symbol every $1/\nu_{\text{max}}$ second. Or equivalently, there can be at most $1/(\nu_{\text{max}} T_s)$ OFDM symbols between two consecutive pilot symbols. Similarly in the frequency domain, the number of sub-carriers between two consecutive pilot symbols can be at most $(T_s - T_{cp})/\tau_{\text{max}}$. Therefore, the minimum pilot density is given by

$$\rho = \frac{\tau_{\text{max}} \nu_{\text{max}}}{1 - \frac{T_{cp}}{T_s}}. \quad (1)$$
Table I gives an example and summarizes the notations related to the pilot pattern design to be considered in the remaining of this paper. For a mobile system operating at a carrier frequency of 5 GHz, the maximal Doppler and delay spreads of $ν_{\text{max}} = 2000(±1000)$ Hz and $τ_{\text{max}} = 5$ μsec correspond to a vehicle speed of 200 kmph and a maximum scatterer spread over 1500 meters. From Eq. (1), with a 20% cyclic prefix overhead, the theoretical minimum pilot density required to cover a high speed vehicle in a large macro cell is $1/80$. A two-time over-sampling in both time and frequency as suggested in [2] further increases the density to $1/20$. The situation worsens as the carrier frequency and cell size increase. For a terminal that is in the coverage of 4 base stations, 20% of the radio resource is occupied by the pilot. This has not even taken into account the multiple antennas that may also transmit pilots for the spatial multiplexing Multiple Input Multiple Output (MIMO) system that will likely be deployed in the future base stations.

The basic design principle for reducing this large overhead is to use a pilot density only slightly above the Nyquist rate. As will be shown later, this low density pilot is equivalent in performance to a pilot with higher density given that sufficient observations are made. In the event that the time-frequency observation window does not contain enough pilot observation for achieving desired estimation accuracy, additional pilot associated with the data will be transmitted. These situations may arise in certain applications when delay requirement is stringent or in extremely rare cases when the channel’s maximum delay-Doppler spread exceeds the design limit.

The adaptive variation of local pilot density can be facilitated by either control signaling or blind detection. When detected by the transmit device or signaled by the receiving device the need for extra pilot, the transmit device will puncture the data and replace it with the extra pilot. Similarly, the receiving device may be signaled by the transmit device or blindly detect the presence of extra pilot and then perform channel estimation upon the data reception without having to wait for the next occurrence of common pilot.

The transmission of additional data associated pilot to increase local pilot density applies to both time domain and frequency domain. For simplicity, we will consider time domain pilot density increase as example. The extension to frequency domain is straightforward. Moreover, the common pilot used for illustration purpose in this paper has a regularly spaced pattern in both time and frequency. However, the same principle also applies to other variations of hopping patterns [1].

The left hand side of Fig. 2 shows an example of a regularly spaced common pilot with a time insertion frequency of twice the Nyquist rate. In addition to a common pilot of 1.25 times the Nyquist rate in time, the figure to the right also shows a set of data associated pilot (data pilot) and a set of estimated pilot that will be explained later in Sec. III. Upon the reception of the data associated pilot, the intended receive device may proceed to perform channel estimate based on the past observation of the common pilot and the newly acquired data pilot without having to wait until the next occurrence of common pilot. However, this requires a channel estimator that can efficiently combine the two pilot observations and the ability of the receive device to remain attached to the network to receive relevant control information such as the presence, location and density of the data pilot. These two issues will be addressed in Sec. III and Sec. IV respectively.

### III. Reduced Complexity Channel Estimator

The optimal Minimum Mean Square Error (MMSE) channel estimator based on both the common and data pilots is too complicated since it involves the inversion of a large kernel matrix which lacks the structure of the regularity of the common pilot. In this section, we will describe an alternative that essentially combines the estimate from the common pilot with the observation of the data associated pilot to form a sub-optimal, but less complex joint estimate.

#### A. General Formulation for MMSE Channel Estimator

In an OFDM system, the discrete frequency domain received samples can be expressed as

$$X[t, f] = H[t, f]Λ[t, f] + Z[t, f],$$  \hspace{1cm} (2)

where the index $(t, f)$ corresponds to the $f$th sub-carrier in the $t$th OFDM symbol, $H[t, f]$ is the channel’s time-frequency
response at that point, $\Lambda[t, f]$ is the transmitted symbol and $Z[t, f]$ is the Additive White Gaussian Noise (AWGN).

The subset of these samples corresponding to the pilot observations can be concisely expressed in a matrix form if we arrange them into a column vector:

$$X_c = \Lambda_c H_c + Z_c,$$

(3)

where $\Lambda_c$ is a diagonal matrix containing the pilot symbols as its diagonal elements and $X_c$, $H_c$ and $Z_c$ are column vectors of the same size corresponding to the pilot observation, the channel and the noise respectively. The subscript $c$ is used to denote the observation corresponding to the common pilot that is regularly spaced in time and frequency. We will further assume that the number of pilot symbols is $K$, i.e., $\Lambda_c$ is an $K \times K$ matrix and the three vectors are of dimension $K \times 1$.

Similarly, using the subscript $d$, we can write the observation corresponding to the data associated pilot as

$$X_d = \Lambda_d H_d + Z_d.$$

(4)

By stacking both the common pilot and data pilot observations in a column vector,

$$
\begin{bmatrix}
X_c \\
X_d
\end{bmatrix}
= 
\begin{bmatrix}
\Lambda_c & 0 \\
0 & \Lambda_d
\end{bmatrix}
\begin{bmatrix}
H_c \\
H_d
\end{bmatrix} + 
\begin{bmatrix}
Z_c \\
Z_d
\end{bmatrix},
$$

(5)

and introducing the subscript $t$, we can express the total pilot observation in a similar matrix form given by

$$X_t = \Lambda_t H_t + Z_t.$$

(6)

The channel $H[t, f]$ is sometimes modeled as a two-dimensional zero-mean Wide Sense Stationary (WSS) Gaussian random process. Its correlation defined as

$$\Gamma[t_1 - t_2, f_1 - f_2] \equiv E\{H[t_1, f_1]H^*[t_2, f_2]\}$$

(7)

is often assumed known or can be estimated from past observation. With the matrix representation of Eqs. (3,4,6) and the knowledge of the channel’s statistics, an MMSE estimator can be easily formulated for each of the three pilot observations.

As an example, we now consider the MMSE estimator for the common pilot only. Let $H$ denote the $L \times 1$ column vector containing the channel response at the locations where it needs to be estimated in the time-frequency plane. In Fig. 2, this can be the $L = 4 \times 120$ points between 152 < $t$ ≤ 156 and 0 ≤ $f$ < 120. The MMSE estimate of $H$ is simply its mean conditioned on the observation $X_c$ [6]:

$$H(X_c) = E\{H|X_c\} = \Pi_{HH_c}H_{X_c}^{-1}X_c$$

$$= \Pi_{HH_c}A_c^{-1}\Sigma^2 + \sigma^2 I)^{-1}X_c,$$

(8)

where $\Pi_{HH_c} = E\{H_c H^*_c\}$ denotes the $K \times K$ auto covariance matrix of $H_c$, $\Pi_{HH}$ is the $L \times K$ covariance matrix between $H$ and $H_c$, and $\Sigma_c^2$ is the similarly defined $L \times K$ covariance matrix between $H$ and $X_c$.

Since the elements in the involved covariance matrices can be derived from the channel’s correlation function given in Eq. (7), the estimator becomes an $L \times K$ linear operator that transforms the $K \times 1$ observation vector $X_c$ into the $L \times 1$ vector containing the estimates at the desired locations. The MMSE estimate of $H$ based on either the data pilot only or on the total pilots has the same matrix expression as Eq. (8) with corresponding subscript.

B. Reduced Complexity Sub-optimal Estimator

The MMSE estimate given in Eq. (8) is based only on the common pilot observation and therefore suffers from performance degradation as the time index becomes farther away from the last pilot symbols transmitted in time. With the presence of the data pilot, the receiver may choose to use an MMSE estimator based on $X_t$, the combination of both the common and data pilot observations, to reduce the estimation error if waiting for the occurrence of the next pilot symbols in time is not an option.

However, the MMSE estimator based on $X_t$ involves the inversion of a large kernel matrix which lacks the regular structure of the common pilot. One solution is to form an MMSE estimator based on the data pilot and a few neighboring common pilot symbols to reduce the size of the kernel matrix. This fails to exploit the contribution from the common pilot symbols that are not included. Instead, we will now describe a novel method that formulates an MMSE estimator based on the data pilot observation and the estimated channel response of a few neighboring locations based on the common pilot. These are the estimated pilot positions shown in Fig. 2. Such estimate is often available since the receiver needs to continuously estimate the channel in order to demodulate control information.

Let $H_n$ denote the channel response at the estimated pilot location, then the MMSE estimate of $H_n$ is given by

$$\hat{H}_n(X_c) = E\{H_n|X_c\}$$

$$= \Pi_{H_n H_c}A_c^{-1}\Sigma^2 + \sigma^2 I)^{-1}X_c$$

$$\equiv W_n X_c,$$

(9)
where $W_n$ is the corresponding linear operator. Note that in the following derivation $W_n$ is not restricted to be an MMSE estimator. It can be any linear operator on an arbitrary subset of past observation of the common pilot. The two-dimensional DFT-based Maximum Likelihood (ML) estimator [7], [8] is an example with much less complexity.

By stacking the estimated pilot $\hat{H}_n$ and the data pilot observation $X_d$ in a column vector

$$X_m \equiv \begin{bmatrix} \hat{H}_n \\ X_d \end{bmatrix},$$

(10)

we can express the corresponding MMSE estimator as

$$\hat{H}(X_m) = E \{ H|X_m \} = \Pi_{HX_m} \Pi_{X_m}^{-1} X_m,$$

where

$$\Pi_{HX_m} = \begin{bmatrix} \Pi_{HH} \Lambda_{c}^H W_n \Pi_{HH} \Lambda_{d}^H \\ \Pi_{X_m} \end{bmatrix}$$

(12)

$$\Pi_{X_m} = \begin{bmatrix} W_n (\Lambda_c \Pi_{HH} \Lambda_c^H + \sigma_n^2 I)^{-1} W_n \Lambda_c \Pi_{HH} \Lambda_d^H + \sigma_n^2 I \\ \Lambda_d \Pi_{HH} \Lambda_d^H \end{bmatrix}. $$

Since the dimension of the kernel matrix $\Pi_{HX_m}$ is only a fraction of the dimension of $\Pi_{X_m}$, the kernel matrix of the optimal MMSE estimator based on both pilot observations, it is much easier to invert. The performance loss with respect to the optimal estimator, as will be shown in the following section, is minimal.

### C. Estimation Error Characteristics

The covariance matrix of the estimation error vector based on the common pilot observation is given by [6]

$$\Pi_{\tilde{H}(X_c)} = E \{ (H - \hat{H}(X_c)) (H - \hat{H}(X_c))^H \} = \Pi_{H} - \Pi_{HX_c} \Pi_{X_c}^{-1} \Pi_{HX_c} \equiv \Pi_{H} - \Pi_{HH} \Lambda_c^H \cdot (\Lambda_c \Pi_{HH} \Lambda_c^H + \sigma_n^2 I)^{-1} \Lambda_c \Pi_{HH} \Lambda_d^H,$$

(13)

and its diagonal is the average error of the estimator. By substituting the respective matrices into Eq. (13), we can evaluate the error distribution of the various estimators in the time-frequency plane.

Fig. 3 shows the error distribution of the estimator based on the common pilot. The various pilots location is as shown in Fig. 2. All the common pilot symbols starting from $t = 0$ are used in this case. It is clear that the estimation error increases gradually as the OFDM symbol index passes beyond the last pilot symbols transmitted at $t = 152$. The distribution chart for the estimator based on the common pilot with twice the Nyquist rate is shown in Fig. 4. The error does not begin to increase until the OFDM symbol index passes $t = 155$ when the last set of common pilot symbols are transmitted.

Similar charts for the estimator based on the total pilot observations and the low complexity sub-optimal estimator are omitted. To compare the proposed low complexity estimator, we plot the error distributions at $t = 156$ when the data pilot is transmitted for the various estimators in Fig. 5. The curve marked “common+data” corresponds to the optimal MMSE estimator based on both common and data pilots whereas the curve marked “data+estimated” corresponds to the sub-optimal approach, which is only 1 dB worse.

### IV. CONTROL SIGNAL DISTRIBUTION FOR ROBUST ADAPTIVE OPERATION

In an OFDM system, the modulation symbols of a channel-coded information block can be distributed in many ways such as the Time Division Multiple Access (TDMA) or Frequency Division Multiple Access (FDMA), as shown by the circles in the top two sub-plots of Fig. 6. If the pilot density is high enough, the distribution of the data symbols is irrelevant in terms of demodulation performance. However, if the pilot density is lower than the channel’s Nyquist rate, as in the case of the extreme...
conditions a few devices may experience before the additional pilots are transmitted, the demodulation performance may degrade substantially for those data symbols that are far away from the pilot symbols. This is illustrated in Fig. 7 by the normalized Mean Square Error (MSE) distribution in the time-frequency plane over a rectangular area formed by 4 closest pilot symbols. When the pilot density is at the Nyquist rate, the normalized MSE remains fairly flat over the entire region. However, when the pilot density is at $1/1.25$ Nyquist rate, only the positions near the 4 pilot symbols at the corners experience satisfactory estimation accuracy.

This phenomenon suggests that the distribution of the control signaling data symbols around the pilot symbols, as shown in mapping 3 of Fig. 6, is most robust for our proposed adaptive scheme. The placement should preferably start with the first tier of 4 adjacent symbol locations that are either one symbol or one sub-carrier away from the pilot symbol, and then continue next to the second tier of 4 locations that are one symbol and one sub-carrier away from the pilot. In total, 8 adjacent positions around the pilot symbols are available for assignment. Additional tiers may be added if necessary.

### A. Reduced Complexity Sub-optimal Estimator

Since the control signaling is intended for all devices in the system, it should be designed with enough margin for the worst case scenario in which the Signal to Noise Ratio (SNR) is lowest and the channel’s delay-Doppler spread is highest. Under these conditions, the pilot observations other than that of the closest one do not have any correlation with the channel response of the desired location. Therefore, the MMSE channel estimator only needs to include the closest pilot observation and Eq. (8) is reduced to a scalar operator

$$
\hat{H}(X_c) = E\{|H|X_c\} = \Pi_{H_c}X_c
$$

where $\lambda_c$, $X_c$ and $\Pi_{H_c}$ are scalars corresponding to the pilot symbol value, the pilot observation and the correlation between the channel response at the desired location and the observation, respectively. Since the data symbol and pilot symbol are adjacent to each other, we can further assume that $\Pi_{H_c} \approx \Pi_{H_c} = 1$ and Eq. (14) becomes

$$
\hat{H}(X_c) \approx \frac{\lambda_c^HX_c}{|\lambda_c|^2 + \sigma_Z^2},
$$

which is independent of the data symbol’s location.

### B. Simulation Results

Figs. 8 to 10 show the Frame Error Rate (FER) results for the three mappings in Fig. 6 as a function of $E_b/N_0$, ignoring
cyclic prefix. The simulation parameters are summarized in Table II. The advantage of the proposed mapping is clearly demonstrated by the robust performance even when the pilot density is at 1/2 the channel’s maximal delay-Doppler spread. An additional benefit of the proposed mapping is the increased diversity gain resulting from the spreading of the data symbols across the entire frequency band.

V. CONCLUSION

We have proposed an adaptive pilot allocation scheme for OFDM on the downlink. The adaptive transmission of data associated pilot in addition to a continuously broadcasted low density common pilot allows for lower overall pilot overhead while maintaining the system’s ability to serve devices with stringent delay requirement or in extreme but rare channel conditions. The practical implementation of the design is further supported by a reduced-complexity channel estimator and a robust distribution of control signaling. The two-stage channel estimator combines the estimate derived from the common pilot with the data pilot observation to form a simplified linear estimator. The bundling of control signaling with the common pilot guarantees the delivery of critical information to devices in most extreme conditions. Finally, numerical results from error analysis and modem simulation verify the proposition’s feasibility.

REFERENCES