Virtual Pilot Signal for Massive MIMO-OFDM Systems

Sunho Park*, Jun Won Choi†, Ji-Yun Seol‡, and Byonghyo Shim*

*School of Information and Communication, Korea University, Seoul, Korea
†Electrical-Bio Engineering Department, Hanyang University, Seoul, Korea
‡Communications Research Team, DMC R&D Center, Samsung Electronics Co., Ltd., Suwon, Korea

Abstract—A technology exploiting large number of antennas into the wireless transceiver, often called massive MIMO technology, has been received great attention in recent years. The benefits of this approach, however, can be realized only when the quality of estimated channel is ensured at both transmitter and receiver. In this paper, we propose a new channel estimation technique dealing with the pilot shortage in the massive MIMO systems. In a nutshell, we employ reliable data tones obtained from the iterative detection and decoding (IDD) as well as pilot signals in the estimation of channels. We show that the proposed method achieves substantial performance gain over conventional approaches employing pilot signals exclusively.

I. INTRODUCTION

Recently, there have been growing interests on large-scale multi-input multi-output orthogonal frequency division multiplexing (MIMO-OFDM) systems to further improve throughput and reliability of next generation wireless communications. One of major concerns in realizing the massive MIMO-OFDM systems is that it requires significant amount of pilot signals to estimate channels for each antenna link. Clearly, this situation is undesirable since the pilot signals occupy significant portion of resources, eating out the data throughput severely. Note that the reduction of pilot signals density is also undesirable, since this will cause a degradation of the channel estimation quality, affecting the link performance and eventually the data throughput.

Our aim in this work is to demonstrate that the data signals, in conjunction with the state of the art iterative detection and decoding (IDD) technique [9], is effective means to improve the channel estimation quality of massive MIMO-OFDM systems. By employing deliberately chosen ‘reliable’ data tones (henceforth referred to as virtual pilot signals (VPS)) as well as the pilot signals, channel estimation quality and subsequent detection and decoding quality can be improved substantially.

There have been number of studies on decision-directed channel estimation to improve the performance of MIMO-OFDM systems. In [1], [2], approaches suppressing the interference from other transmit antenna have been proposed. In essence, channel estimates predicted from adjacent symbols

III. SYSTEM DESCRIPTION

A. MIMO-OFDM Systems

The input-output relationship of a data tone at the subcarrier index $k$ and the symbol index $n$ of the MIMO-OFDM system is given by

$$y^{(r)}_{k,n} = \sqrt{P_d} \sum_{i=0}^{N_t-1} h^{(r,i)}_{k,n} x^{(i)}_{k,n} + n^{(r)}_{k,n},$$

where $y^{(r)}_{k,n}$ and $n^{(r)}_{k,n}$ are the received signal and noise at the receive antenna $r$, $x^{(i)}_{k,n}$ are the symbol from the transmit antenna $i$, $h^{(r,i)}_{k,n}$ is the channel response from the transmit antenna $i$ to the receive antenna $r$, and $P_d$ is the transmission power per each data tone and transmit antenna.

In a similar way, the received signal for the pilot tone (a.k.a, reference signal) is given by

$$z^{(r)}_{k,n} = \sqrt{P_p} h^{(r,t)}_{k,n} r^{(t)}_{k,n} + n^{(r)}_{k,n},$$

where $z^{(r)}_{k,n}$ is the received data at the receive antenna $r$, $r^{(t)}_{k,n}$ is the pilot signal transmitted at the antenna $t$, and $P_p$ is the pilot transmission power.

Typically, conventional receiver estimates the channel of the data tone via the pilot channel estimation followed by the interpolation. Then, detection of data tones (i.e., generation of log-likelihood ratio (LLR)) is performed using the estimated channel $h^{(r,i)}_{k,n}$ and the received signal $y^{(r)}_{k,n}$ [5], [6]
where \( c_m \) is the \( m \)-th information bit in the symbol vector \( \mathbf{x} = [x_{k,0}, \ldots, x_{k,N_r-1}]^T, \mathbf{y} = [y_{k,0}, \ldots, y_{k,N_r-1}]^T \),

\[
\mathbf{h} = \begin{bmatrix}
\hat{h}_{k,0}^{(0)} & \ldots & \hat{h}_{k,0}^{(N_r-1)} \\
\vdots & \ddots & \vdots \\
\hat{h}_{k,N_r-1,0} & \ldots & \hat{h}_{k,N_r-1,N_r-1}
\end{bmatrix},
\]

\[P(c_m) = P(c_m = +1) \quad P(c_m = -1) \text{ is defined similarly}. \quad \mathbf{c}_m = [c_1, \ldots, c_{m-1}, c_{m+1}, \ldots, c_{r}]^T, \quad \mathbf{L}_A = [L_A(c_1), \ldots, L_A(c_{m-1}), L_A(c_{m+1}), \ldots, L_A(c_r)]^T, \text{ and } \mathbf{L}_A(c_m) = \ln \frac{P(c_m = +1)}{P(c_m = -1)}. \]

In the detection process, soft symbols in a form of LLRs are generated and then delivered to the channel decoder. Further, when the IDD is employed in the receiver, the detector and the channel decoder exchange the extrinsic LLRs in an iterative fashion until the desired performance is satisfied or suitably chosen termination condition is satisfied.

### B. Conventional Channel Estimation

The conventional algorithm estimates the channel using the pilot signals. Typically, received pilot signals in an appropriately chosen local window are used (see Fig. 1). The received pilot signals in the local window are expressed as

\[z_p^{(r)} = \sqrt{P_p} \mathbf{h}_p^{(r)} r_p(t) + n_p, \quad 0 \leq p \leq N_p - 1 \] (4)

for \( 0 \leq p \leq N_p - 1 \) where \( N_p \) is the total number of received pilot signals within the window. Vector form of (4) is

\[\mathbf{z}_r = \sqrt{P_p} \mathbf{R}_r \mathbf{h}_r + \mathbf{n}_r, \quad \text{where } \mathbf{z}_r = [z_0^{(r)}, \ldots, z_{N_r-1}^{(r)}]^T, \mathbf{h}_r = [h_0^{(r)}, \ldots, h_{N_r-1}^{(r)}]^T, \mathbf{n}_r = [n_0^{(r)}, \ldots, n_{N_r-1}^{(r)}]^T, \text{ and } \mathbf{R}_r = \text{diag}(r_0^2, \ldots, r_{N_r-1}^2). \]

Under the assumption that the pilot sequence satisfies unitary property (i.e., \( \mathbf{R}_r \mathbf{R}_r^H = 1 \)), the MMSE estimate \( \hat{\mathbf{h}}_{r,t} \) of the channel \( \mathbf{h}_{r,t} \) becomes

\[\hat{\mathbf{h}}_{r,t} = \text{Cov}(\mathbf{h}_{r,t}, \mathbf{z}_r) \text{ Cov}(\mathbf{z}_r, \mathbf{z}_r)^{-1} \mathbf{z}_r \quad \text{for } 0 \leq t \leq T_s - 1 \] (5)

where \( \text{Cov}(\mathbf{a}, \mathbf{b}) = E[\mathbf{a} \mathbf{b}^H] - E[\mathbf{a}]E[\mathbf{b}^H] \).

In essence, the MMSE channel estimation consists of two steps: 1) descrambling of the received signal using the pilot sequence (i.e., \( 1/\sqrt{P_p} \mathbf{R}_t^H \mathbf{z}_{r,t} \)) and 2) application of the filter \( \mathbf{w}_r^H \) to the descrambled signal. In generating \( \mathbf{w}_r^H \), the covariance matrix \( \text{Cov}(\mathbf{h}_{r,t}, \mathbf{h}_{r,t}) \) is needed and this is computed parametrically by considering the relative location among pilot signals in frequency and symbol domain. In the well-known Jake’s model, for example, correlation of channel gains spaced by \( \Delta f \) frequency tones and \( \Delta \tau \) symbol times is given by

\[E[\mathbf{h}_{k,n}^{(r,t)} \mathbf{h}_{k+\Delta k,n+\Delta \tau}^{(r,t)}] = \left( \sum_{i=0}^{P-1} P_{d}(2\pi f_d \Delta \tau + i \Delta f \Delta \tau) \right), \quad 0 \leq d \leq N_d - 1 \] (6)

where \( f_d \) and \( f_d \) are the spacing between adjacent frequency tones and Doppler frequency respectively, \( T_s \) is the period for one OFDM symbol, and \( J_0(x) \) is the 0-th order Bessel function.

### III. CHANNEL ESTIMATION WITH VIRTUAL PILOT SIGNALS

In this section, we introduce the proposed channel estimation algorithm that exploits soft symbols of data tones in the channel re-estimation. Fig. 2 depicts the block diagram of the proposed scheme. First, a posteriori LLRs, generated by adding an extrinsic LLRs of the MIMO detector and a priori LLRs of the decoder is converted to the soft symbols. After selecting VPSs, channel is re-estimated with chosen VPSs as well as the pilot signals. VPSs used for the channel estimation should meet two conditions. First and foremost, they should be reliable enough to be used for the channel estimation. To do so, clearly, the magnitude of a posteriori LLRs should be large. Secondly, VPSs should be highly correlated with the pilot tones since otherwise VPSs will not be helpful in the channel estimation of pilot position. When the enhanced channel estimates are obtained, they are used for the MIMO detector in the next IDD iteration.

#### A. Channel Re-Estimation Using VPS

As mentioned, other than the received pilot signals \( \mathbf{z}_r \) (see (5)), the proposed scheme employs received VPS signals

\[y_d^{(r)} = \sqrt{P_d} \sum_{i=0}^{N_y-1} g_d^{(i)} x_d^{(i)} + v_d^{(r)}, \quad 0 \leq d \leq N_d - 1 \] (7)
where \( N_d \) is the number of data tones used and \( v_d^{(r)} \) and \( g_d^{(r,i)} \) are noise and channel gain at the \( d \)-th VPS, respectively. Now, combining (5) with (9), we have

\[
\begin{bmatrix}
z_r \\
y_r
\end{bmatrix} = \sqrt{T_p R_t} \begin{bmatrix}
h_{r,t} & 0 & \cdots & 0 \\
0 & \sqrt{T_d X_0} & \cdots & 0 \\
0 & 0 & \cdots & \sqrt{T_d X_{N_r-1}}
\end{bmatrix} \begin{bmatrix}
h_{r,t} \\
g_{r,t} \end{bmatrix} + \begin{bmatrix}
n_r \\
v_r
\end{bmatrix}
\]

(10)

where \( X_i = \text{diag}(x_i^{(1)}, \ldots, x_i^{(N_d-1)})^T \), \( v_r = [v_0^{(r)}; \cdots; v_{N_d-1}^{(r)}]^T \), \( y_r = [y_0^{(r)}; \cdots; y_{N_d-1}^{(r)}]^T \), and \( g_{r,i} = [g_0^{(r,i)}; \cdots; g_{N_d-1}^{(r,i)}]^T. \)

Note that, in contrast to the received pilot signals, received VPS consists of signals \( X_i \) from multiple transmit antennas. Note also that \( X_i \) is not deterministic quantity (random variable) and its statistics can be inferred from LLRs. Suppose \( L(c_1^{(i)}), \ldots, L(c_Q^{(i)}) \) are the LLRs of \( Q \) coded bits mapped to a data symbol \( x_i^{(i)} \). Then the mean and variance of \( x_i^{(i)} \) is

\[
E[x_i^{(i)}] = \sum_{\theta \in \Theta} \sum_{k=1}^{Q} \theta_k \left( 1 + c_{i,k}^{(i)} \tanh \left( L(c_{i,k}^{(i)})/2 \right) \right) / p_r(c_{i,k}^{(i)})
\]

(11)

\[
E[|x_i^{(i)}|^2] = \sum_{\theta \in \Theta} |\theta|^2 \sum_{k=1}^{Q} \left( 1 + c_{i,k}^{(i)} \tanh \left( L(c_{i,k}^{(i)})/2 \right) \right) / p_r(c_{i,k}^{(i)})
\]

(12)

where the set \( \Theta \) includes all possible constellation points.

Denoting \( X_i = \text{diag}(E[x_i^{(i)}], \ldots, E[x_i^{(N_d-1)}]) \), the MMSE estimate of \( h_{r,t} \) is given by

\[
\hat{h}_{r,t} = \text{Cov} \left( \begin{bmatrix}
z_r \\
y_r
\end{bmatrix} \right) \text{Cov} \left( \begin{bmatrix}
z_r \\
y_r
\end{bmatrix}, \begin{bmatrix}
z_r \\
y_r
\end{bmatrix} \right)^{-1} \begin{bmatrix}
z_r \\
y_r
\end{bmatrix}
\]

(13)

where

\[
\text{Cov} \left( \begin{bmatrix}
z_r \\
y_r
\end{bmatrix}, \begin{bmatrix}
z_r \\
y_r
\end{bmatrix} \right) = \left( \sqrt{T_p E[h_{r,t}^H R_t^H]} \right)^T \left( \sqrt{T_d \sum_{i=0}^{N_d-1} E[h_{r,t}^H g_{r,i}^H]} X_i \right)
\]

(14)

and

\[
\text{Cov} \left( \begin{bmatrix}
z_r \\
y_r
\end{bmatrix}, \begin{bmatrix}
z_r \\
y_r
\end{bmatrix} \right) = \begin{bmatrix}
A_{11} & A_{12} \\
A_{21} & A_{22}
\end{bmatrix}
\]

(15)

where

\[
A_{11} = P_p R_t E[h_{r,t}^H h_{r,t}^H] R_t^H + \sigma_n^2 I
\]

(16)

\[
A_{12} = \sqrt{P_d} \sqrt{T_p} R_t \sum_{i=0}^{N_d-1} E[h_{r,t}^H g_{r,i}^H] X_i^H
\]

(17)

\[
A_{21} = \sqrt{P_d} \sqrt{T_p} \sum_{i=0}^{N_d-1} X_i E[g_{r,i}^H h_{r,t}^H] R_t^H
\]

(18)

\[
A_{22} = P_d \sum_{i=0}^{N_d-1} \sum_{j=0}^{N_d-1} E[X_i E[g_{r,i}^H g_{r,j}^H]] X_j^H + I
\]

(19)

In computing the statistics in (16)-(19), correlations of channel among spatial antenna pairs should be required. For simplicity, we assume that the channel correlation of different antenna pairs are negligible. That is,

\[
\sum_{i=0}^{N_d-1} E[h_{r,t} g_{r,i}^H] X_i^H \approx E[h_{r,t}^H g_{r,t}^H] X_t^H
\]

and

\[
\sum_{i=0}^{N_d-1} \sum_{j=0}^{N_d-1} E[X_i E[g_{r,i}^H g_{r,j}^H] X_j^H] \approx \sum_{i=0}^{N_d-1} E[X_i R_{gg} X_i^H].
\]

Then the corresponding channel estimate \( \hat{h}_{r,t} \) becomes

\[
\hat{h}_{r,t} = \left[ \begin{array}{c} C_{hh} \\
C_{hg} X_i^H \end{array} \right]
\]

\[
\left[ C_{hh} + \frac{1}{P_d} I \right]^{-1} \left[ \begin{array}{c} \sqrt{P_t} R_t^H z_r \\
\sqrt{P_d} y_r \end{array} \right]
\]

(20)

where \( C_{hh} = \text{Cov}(h_{r,t}, h_{r,t}) \), \( C_{hg} = \text{Cov}(h_{r,t}, g_{r,i}) \), \( C_{gg} = \text{Cov}(g_{r,i}, g_{r,j}) \) and these can be computed using (8). Also, \( x_i = [x_i^{(r)}; \cdots; x_i^{(N_d-1)}]^T \) and \( X \otimes Y \) denotes Hadamard (element-by-element) product of \( X \) and \( Y \). It is worth mentioning that if the soft information is unavailable (\( X_i = 0 \) and \( X | X_i^H = I \)), then the channel estimate in (20) degenerates to one without VPS in (7).

### B. Data Tone Selection

One important issue to be addressed in the proposed method is that among all data tones in the local window which one should be chosen as the VPS. In order to choose the best possible data tones, we need to evaluate the MSE \( \|h_{r,t} - \hat{h}_{r,t}\|^2 \) of all possible \( N_d \) data tone combinations and then choose one generating the minimum MSE. Since such procedure is overly complicated, we instead compute the MSE with a single data tone and the existing pilot signals and then choose \( N_d \)-best data tones as VPSs.

Under the assumption that channel correlations between different antenna pair are negligible, one can show that

\[
\text{MSE}(\ell) \triangleq E \left[ \|h_{r,t} - \hat{h}_{r,t}\|^2 \right]
\]

\[
= \frac{1}{P_t} tr \left[ C_{hh} \left( C_{hh} + \frac{1}{P_d} I \right)^{-1} \right] - \left( \frac{1}{P_d} \right)^2 \frac{|E[x_i^T]^H|}{A} C_{gh} \left( C_{hh} + \frac{1}{P_d} I \right)^{-2} C_{gh} \left( C_{hh} + \frac{1}{P_d} I \right)^{-1} C_{gh}
\]

(22)

where \( \ell \) is an index of data tone and \( A = \sum_{i=0}^{N_d-1} E[|x_i|^2] \). Note that \( N_d \) data tones minimizing MSE(\( \ell \)) is those maximizing the second term in (22).

It is worth mentioning that the VPS selection process is mainly affected by the correlations between the pilot signals and data tones rather than the correlations within the pilot signals. Also, the correlations between the pilot signals is typically small since the pilot signals are allocated evenly and sparsely (\( C_{hh} \approx I \)). Considering these, we obtain the simplified cost metric

\[
\phi(\ell) = |E[x_i^T]| |C_{gh}|.
\]

(23)
Interestingly, the cost function to choose the VPS is expressed as a product of the reliability of soft decision ($|E[x_t]|$) and the correlation between the data tone and the pilot ($|C_{gh_i}|$). Hence, our task is simplified to choose $N_d$-best data tones maximizing the cost metric $\phi(\ell)$. Note that this process is performed for each transmit antenna.

**IV. SIMULATION RESULTS**

The simulation setup is based on single user MIMO-OFDM system with QPSK modulation. The signal is transmitted over $12 \times 12$ MIMO system with extended vehicular A (so called EVA) channels [7]. Since the pilot signal mapping of MIMO systems larger than $8 \times 8$ is not yet specified in the 3GPP standard [8], we used the mapping that only one pilot (reference signal) $R_i$ is mapped for $i$-th transmit antenna in each slot. As a metric for measuring performance, we consider the mean square error (MSE) and the bit error rate (BER). For comprehensive view, we perform simulations on the proposed method and the conventional LS and MMSE algorithms. Note that the proposed algorithm employs 32 VPSs.

In Fig. 3(a), we plot the MSE performance of channel estimation algorithms. Since the properly chosen VPSs improve the quality of channel estimation, the resulting MSE of proposed method is substantially smaller than that of conventional method in all SNR regime. In Fig. 3(b), we show the BER performance of conventional and proposed method at first and 7th iterations. Since the improved channel estimates provide a positive effect on the IDD process, we see that the proposed method provides a substantial gain (around 1 dB gain at waterfall regime) over conventional methods.

**V. CONCLUDING REMARKS**

In this paper, we proposed a new soft decision-directed channel estimation algorithm for the massive MIMO-OFDM systems. By deliberately choosing the reliable data tones located near the pilot position, the proposed method provides better channel estimates and accurate LLR detection. When combined with IDD, the proposed method achieves substantial performance gain over the conventional channel estimation schemes.

**REFERENCES**


Fig. 3. Performance of conventional and proposed scheme as a function of $E_b/N_0$ : (a) MSE and (b) BER.