Joint Channel Tracking and Data Detection for Massive MIMO Uplink Over Time-Varying Channel

Yuya Akiba and Shinya Sugiura, Senior Member, IEEE

Abstract—In this letter, we propose a novel joint channel estimation and data detection scheme for massive multiple-input multiple-output (MIMO) uplink, which operates in a time-varying fading channel. More specifically, at the receiver of the proposed scheme, a data frame is divided into multiple blocks, and in each block, a demodulated data block is used for updating channel state information (CSI) in an iterative manner. Furthermore, the initial CSI in the block of interest is given by the estimated CSI in the previous block, hence allowing accurate tracking of CSI in a time-varying channel without imposing additional pilot insertion inside the data frame. Since the length of the divided blocks affects both the achievable channel tracking and data detection performances, it is optimized so as to maximize the discrete-input continuous-output memoryless channel’s (DCMC) capacity derived in this letter. It is demonstrated that the DCMC capacity of the proposed scheme is capable of nearly achieving those of the perfect CSI counterpart without imposing any substantial pilot overhead.

Index Terms—Constrained capacity, channel estimation, channel tracking, iterative detection, matched filter, massive MIMO, time-varying channel.

I. INTRODUCTION

M ASSIVE multiple-input multiple-output (MIMO) has attracted much attention owing to its high spectral efficiency and energy efficiency [1–7]. By assuming the employment of a large number of antenna elements at a base station, MIMO channels are decomposed into parallel streams with the aid of low-complexity matched-filter (MF)-based detection for uplink and conjugate beamforming for downlink.

Channel estimation (CE) is the key to achieving the expected high performance of massive MIMO systems [8]. By relying on channel reciprocity in a time-division duplex (TDD) mode, the base station can reuse channel state information (CSI) estimated in the uplink for downlink beamforming. Typically, a receiver estimates CSI based on pilot symbols inserted in a data frame [9], which naturally imposes unignorable overhead. Furthermore, the pilot overhead increases upon increasing the channel’s coherence time and the Doppler frequency in order to maintain the required CE performance. In order to improve the accuracy of CE while avoiding the increase of pilot overhead, joint CE and data detection (DD) schemes have been developed [10–16], where a demodulated data frame is exploited for updating CSI estimated as a longer pilot sequence than the original pilot sequence. In [10], Cozzo and Hughes introduced a semi-blind iterative joint CE and DD scheme for a point-to-point MIMO scenario. In [11, 12], the same idea was exploited for a single-RF spatial modulation (SM) scheme, exploiting antenna switching mechanism at the transmitter. Furthermore, in [13], data-aided CE was proposed for massive MIMO, where partially decoded soft-output information is used to update the channel. Also, space-alternating generalized expectation-maximization based joint CE and DD were proposed to deal with pilot contamination in [16]. However, these schemes assume ideal time-invariant channels over a frame, and hence the achievable performance may severely deteriorate in the presence of unignorable Doppler effects. Moreover, block-by-block joint channel tracking and DD with a fixed block length were proposed for orthogonal space-time-block codes in [17]. In [18], a variable-block-length joint channel tracking and DD scheme was proposed for massive MIMO downlink based on a single-RF-aided SM transmitter.

Against this background, the novel contributions of this letter are that we propose variable-block-length joint channel tracking and DD for time-varying multiuser massive MIMO uplink. Each data frame is divided into multiple short blocks, and the joint CE and DD are performed in an iterative manner by assuming that the channel remains constant over each block. Here, the initial CSI in a block of interest is given by CSI estimated in the previous block, and hence CSI is tracked even in the presence of Doppler effects. Also, we derive the discrete-input continuous-output memoryless channel’s (DCMC) capacity of the proposed scheme. Furthermore, it is demonstrated that the proposed scheme achieves a DCMC capacity close to that of the ideal perfect CSI (PCSI) scenario with both low complexity and a low pilot overhead.

The remainder of this letter is organized as follows. In Section II, we provide the system model and joint CE and DD of the proposed scheme. In Section III, the DCMC capacity and the block length optimization are described, while in Section IV, our performance results are provided. Finally, Section V concludes the present letter.

II. SYSTEM MODEL

In this letter, we consider a massive MIMO uplink with $K$ single-antenna user terminals and a base station equipped with $N$ receive antennas. It is assumed that the number of receive antennas $N$ is higher than the number of users $K$. For simplicity, we consider that there is no inter-cell interference, as well as no pilot contamination. Users simultaneously transmit independent symbols of an $M$-size phase-shift keying (PSK) or quadrature amplitude modulation (QAM) constellation, where the sum rate is $R = K \log_2 M$ bps/Hz.
The frame format is constituted by a $K$-length pilot block and a $T$-length data block, where the pilot block $S_p \in \mathbb{C}^{K \times K}$ is used for estimating channel coefficients. The data signals received at the base station in the $t$th time slot are represented by

$$r(t) = \mathbf{H}(t)s(t) + \mathbf{v}(t) \in \mathbb{C}^N \quad (1 \leq t \leq T), \quad (1)$$

where $\mathbf{H}(t) \in \mathbb{C}^{N \times K}$ is the complex-valued channel matrix, whose $n$th-row and $k$th-column element is a channel coefficient between the $k$th user’s transmit antenna and the $n$th receive antenna at the base station. $s(t) = [s_1(t), \ldots, s_N(t)]^T \in \mathbb{C}^N$ are the $K$ complex-valued modulated symbols. Furthermore, $\mathbf{v}(t) \in \mathbb{C}^N$ represent the additive white Gaussian noises (AWGN), which are independent random variables obeying the complex-valued Gaussian distribution with a zero of mean and a variance of $K\rho^{-1}\mathbf{I}_N$, where $\rho$ is the average signal-to-noise ratio (SNR) and $\mathbf{I}_N$ represents the $N$-sized identity matrix.

In this letter, we assume that the channel coefficients are modeled as independent Rayleigh fading, where the $n$th-row and $k$th-column element of $\mathbf{H}(t)$ is generated according to the Jakes model as follows [18]:

$$h_{n,k}(t) = \frac{1}{\sqrt{N_s}} \sum_{\nu=1}^{N_s} \cos [2\pi (\cos \xi_k n) F_d T_s t + \zeta_{k,n}]$$

$$+ j \frac{1}{\sqrt{N_s}} \sum_{\nu=1}^{N_s} \sin [2\pi (\cos \xi_k n) F_d T_s t + \tilde{\zeta}_{k,n}], \quad (2)$$

in which $N_s$ is the number of scatters, $F_d T_s$ is the normalized Doppler frequency, and $\xi_k n$, $\zeta_{k,n}$ and $\tilde{\zeta}_{k,n}$ are random variables which obey the uniform distribution over the range $[0, 2\pi]$.

In this letter, we assume that the base station employs linear MF-based detection [1]. Assuming also that the estimated channel matrix is $\hat{\mathbf{H}}(t) = [\hat{\mathbf{h}}_1(t), \ldots, \hat{\mathbf{h}}_K(t)] \in \mathbb{C}^{N \times K}$, the MF weights for detecting the $k$th user’s symbol are given by

$$\mathbf{w}^H_k(t) = \frac{\hat{\mathbf{h}}^H_k(t)}{||\hat{\mathbf{h}}_k(t)||^2}, \quad (3)$$

where $(\cdot)^H$ indicates Hermitian transpose. Then, the signal weighted by the MF is represented by

$$y_k(t) = \mathbf{w}^H_k(t)r(t)$$

$$= \mathbf{w}^H_k(t)\mathbf{h}_k(t)s_k(t) + \mathbf{w}^H_k(t)\sum_{l \neq k} \mathbf{h}_l(t)s_l(t)$$

$$+ \mathbf{w}^H_k(t)v(t), \quad (4)$$

where the first term corresponds to the $k$th user’s symbol to be detected, and the second term represents inter-user interference. Furthermore, the third term is the associated additive noise component. When considering the limit of an infinite number of receive antennas $N \rightarrow \infty$, the effects of inter-user interference and additive noises vanish, hence arriving at $y_k(t) \xrightarrow{N \rightarrow \infty} \mathbf{w}^H_k(t)\mathbf{h}_k(t)s_k(t)$, similar to [1], except for the CE errors. Finally, the data can be demodulated based on maximum likelihood (ML) detection with

$$\hat{s}_k(t) = \arg\min_s ||y_k(t) - s||^2. \quad (5)$$

Note that computational complexities required for detecting

![Fig. 1. Frame structure and joint CE and DD process of the proposed scheme.](image)

$K$ users’ symbols are $4KN$ and $2KM$ real-valued multiplications from (5) and (6). By contrast, that of ML detection is $M^K(4KN + 2N)$ real-valued multiplications.

A. Conventional joint CE and DD [10]

As a means to accurately estimate channel coefficients while maintaining a low pilot overhead, the conventional joint CE and DD scheme [10] works under the assumption that the channels are static over an entire frame, i.e., $\mathbf{H}(t) = \mathbf{H}$. Prior to DD, the receiver carries out iterative CE. More specifically, initial CE results are calculated based on the MMSE criterion as follows [19]:

$$\hat{\mathbf{H}}_0 = \mathbf{R}_p\mathbf{S}_p^H \left(\mathbf{S}_p\mathbf{S}_p^H + \frac{K}{\rho}\mathbf{I}_K\right)^{-1}, \quad (7)$$

where $\mathbf{R}_p \in \mathbb{C}^{N \times K}$ is the received signals corresponding to the pilot symbols $\mathbf{S}_p$. In this letter, we assume that the pilot sequence $\mathbf{S}_p$ is given by a $(K \times K)$-sized discrete Fourier transform (DFT) matrix whose $k$th-row and $t$th-column element is $\exp(j2\pi kt/K)$ [9]. Then, in order to improve the accuracy of CE, an iterative joint CE and DD algorithm is applied. In the $i$th iteration, the data frame is demodulated from the received signals $\mathbf{r}(t)$ and the estimated channels $\hat{\mathbf{H}}_{i-1}$ in the $(i-1)$st iteration, based on (3), (4), and (6), in order to obtain the $i$th estimated data frame $\hat{\mathbf{S}}_i = [\hat{s}_i(1), \ldots, \hat{s}_i(T)] \in \mathbb{C}^{K \times T}$.

Then, by regarding the demodulated data frame $\hat{\mathbf{S}}_i$ in the $i$th iteration as a long pilot sequence, CSI is updated as follows:

$$\hat{\mathbf{H}}_i = \mathbf{R}\hat{\mathbf{S}}_i^H \left(\hat{\mathbf{S}}_i\hat{\mathbf{S}}_i^H + \frac{K}{\rho}\mathbf{I}_K\right)^{-1}, \quad (8)$$

where the data frame matrix is $\mathbf{R} = [r(1), \ldots, r(T)] \in \mathbb{C}^{N \times T}$. The maximum number of iterations $I_{itr}$ is set sufficiently high that the estimated CSI converges. Finally, we obtain the demodulated data symbols as $\mathbf{S}_{i\text{itr}}$. Since the conventional joint CE and DD scheme assumes that CSI remains constant over the entire frame length $K + T$, it suffers from severe performance deterioration in a rapidly time-varying scenario.
B. Proposed variable-block-length joint CE and DD scheme

Fig. 1 depicts the frame structure and the iterative joint CE and DD process of the proposed scheme. In the proposed scheme, the data frame is divided into Q blocks, each having length \( L \), and hence, the number of blocks is given by \( Q = \lceil T/L \rceil \). We consider that the \( q \)th data block and the received block are respectively given by

\[
\begin{align*}
S_q &= [s(L(q-1)+1), \ldots, s(Lq)]^T \in \mathbb{C}^{K \times L} \quad (9) \\
R_q &= [r(L(q-1)+1), \ldots, r(Lq)]^T \in \mathbb{C}^{N \times L}. \quad (10)
\end{align*}
\]

In each received block, CSI and block data are iteratively calculated based on the joint CE and DD principal by assuming the channels are static over \( L \) symbols, where CSI in the previous block is used as the initial condition of the estimated CSI in the block of interest. Hence, CSI can be tracked even in a time-varying channel if the length of \( L \) is sufficiently small. We assume the estimated CSI in the \( q \)th block and \( i \)th iteration as \( \hat{H}_{q,i} = [\hat{h}_{1,q,i}, \ldots, \hat{h}_{K,q,i}]^T \in \mathbb{C}^{N \times K} \). The initial CSI of the first block is calculated using pilot symbols in the same manner as the conventional joint CE and DD by (7), which is represented by \( \hat{H}_{1,0} = \mathbb{H}_0 \). The demodulated data block in the \( q \)th (\( 1 \leq q \leq Q \)) block and the \( i \)th (\( 0 \leq i \leq \text{itr} \)) iteration is given by

\[
\hat{S}_{q,i} = [\hat{s}_1(L(q-1)+1), \ldots, \hat{s}_L(Lq)]^T \in \mathbb{C}^{K \times L}, \quad (11)
\]

where \( \hat{s}_{k,i} \)th-row and \( k \)th-column element is calculated by

\[
\hat{s}_{k,i}(L(q-1)+l) = \arg\min_s[\hat{y}_{k,i-1}(L(q-1)+l) - s]^2 \quad (12)
\]

where

\[
\hat{y}_{k,i-1}(L(q-1)+l) = \hat{w}_{k,q,i-1}^H \mathbb{R}_q \hat{S}_{q,i}^H \hat{S}_{q,i} \hat{S}_{q,i}^H + \frac{K}{\rho} \mathbb{I}_K
\]

Then, the CSI in the \( i \)th iteration is updated according to

\[
\hat{H}_{q,i} = \mathbb{R}_q \hat{S}_{q,i}^H \left( \hat{S}_{q,i} \hat{S}_{q,i}^H + \frac{K}{\rho} \mathbb{I}_K \right)^{-1} \quad (15)
\]

After the predefined number of iterations \( \text{itr} \), the initial CSI of the \((q+1)\)th block is set to \( \hat{H}_{q+1,0} = \hat{H}_{q,i_{\text{itr}}} \), while \( \hat{S}_{q,i_{\text{itr}}} \) is the estimated \( q \)th data block.

In terms of CE performance, there is a tradeoff relationship with respect to the parameter \( L \). More specifically, a low \( L \) value allows us to accurately track the time-varying channel. By contrast, a high \( L \) value makes it possible to regard the estimated symbols as a long pilot sequence in the iterative joint CE and DD process. Hence, block length \( L \) has to be optimized depending on the Doppler frequency. Note that the proposed scheme with the parameters of \( L = T \) (\( Q = 1 \)) is identical to the conventional joint CE and DD scheme.

C. Computational Complexity

In this section, the computational complexity imposed by the proposed joint CE and DD scheme is evaluated in terms of the number of real-valued multiplications per symbol time. As for CE of (15), the calculations of \( \mathbb{R}_q \hat{S}_{q,i}^H \) and \( \hat{S}_{q,i} \hat{S}_{q,i}^H \) require \( 4KN \) and \( 4K^2 \) real-valued multiplications, respectively. Also, the inversion of a \( K \times K \)-sized matrix is \( O(K^3/L) \), while a matrix multiplication requires \( 4K^2N/L \) real-valued multiplications. Moreover, the MF-based detection of (12) and (13) requires \( 4KN+2KM \) real-valued multiplications in each symbol time. Finally, considering \( \text{itr} \) iterations, we arrive at the total complexity of the proposed receiver as

\[
(I_{\text{itr}}+1)(4KN+2KM) + I_{\text{itr}} \left[ 4KN+4K^2+O \left( \frac{K^3}{L} \right) + \frac{4K^2}{L} \right]. \quad (16)
\]

III. Capacity Derivation and Block-Length Optimization

In this section, we derive the DCMC capacity of the proposed scheme, which is used for optimizing the block length \( L \). Since our iterative CE and DD scheme, the demodulated data, having specific constellations, are used for improving CE. Hence, the continuous-input continuous-output memoryless channel (CMMC) capacity [20], assuming Gaussian symbols, is not suitable to metric for the proposed scheme.

Similar to (5), the \( k \)th user’s symbol after MF demodulation in the \( I_{\text{itr}} \)th iteration is represented by

\[
\begin{align*}
\tilde{y}_{k,i_{\text{itr}}} &= \mathbb{w}_{k,q,i_{\text{itr}}}^H \hat{h}_k(t)s_k(t) + \mathbb{w}_{k,q,i_{\text{itr}}}^H \sum_{i \neq k} \hat{h}_i(t)s_i(t) + \mathbb{w}_{k,q,i_{\text{itr}}}^H v(t) \\
&\quad \text{interference plus noise } N_k(t)
\end{align*}
\]

Here, let us define the sum of the second and third terms of (17) as \( N_k(t) \). Then, the DCMC capacity of the proposed scheme is defined as follows [21]:

\[
\begin{align*}
\text{CDMC} &= \mathbb{E} \left[ \sum_{k=1}^{K} \left\{ \log_2 M - \frac{1}{M} \sum_{m=1}^{M} \log_2 \left( \sum_{m'=1}^{M} \exp \Psi_{m,m'} \right) \right\} \right],
\end{align*}
\]

where \( \mathbb{E}[\cdot] \) represents the expectation operation. Moreover,

\[
\begin{align*}
\Psi_{m,m'} &= - \frac{\mathbb{E}[|N_k(t)|^2]}{|\mathbb{E}[|N_k(t)|^2]|^2} \\
&\approx - \frac{\mathbb{E}[|N_k(t)|^2]}{|\mathbb{E}[|N_k(t)|^2]|^2} \mathbb{E} \left[ \left( 1 + \frac{1}{\rho} \right) \right],
\end{align*}
\]

where \( s_m \) is the \( m \)th PSK or QAM symbol (\( m = 1, \ldots, M \)), and the approximation from (19) to (20) is valid for large \( N \).

Fig. 2 shows the DCMC capacity of the proposed scheme, where the block length was varied from \( L = 25 \) to 1600 and the system parameters were \((K,N) = (16,128)\) and \( F_0 T_s = 10^{-3} \). The average SNR was given by \( \rho = -10 \, \text{dB}, 0 \, \text{dB}, \) and \( 10 \, \text{dB} \). The DCMC capacity assuming PCSI at the receiver was also plotted as a dotted line for each SNR scenario. As seen in Fig. 2, the DCMC capacity was maximized by \( L = 200, 100, \) and \( 100 \) for the average SNRs of \( \rho = -10 \, \text{dB}, 0 \, \text{dB}, \) and \( 10 \, \text{dB} \), respectively. The proposed scheme attained the DCMC
The normalized Doppler frequency was set to 0 dB. Observe in Fig. 3 that the MSE converged in as few as one iteration for $\rho = 10$ dB, while five iterations were needed for the convergence in the scenario with $\rho = 0$ dB. Here, the MSE of CE was defined as

$$\mathbb{E}\left[\frac{1}{K N L Q} \sum_{q=1}^{Q} \sum_{t=L(q-1)+1}^{L_q} \left\|H(t) - \hat{H}_{q,t} \right\|^2\right]. \quad (21)$$

Observe in Fig. 3 that the MSE converged in as few as one iteration for $\rho = 10$ dB and 20 dB, while five iterations were needed for the convergence in the scenario with $\rho = 0$ dB.

Fig. 4 shows the DCMC capacity, where the $L$ value was optimized for each SNR. The normalized Doppler frequency was set to $F_d T_s = 10^{-3}$ and the data length was given by $T = 1600$. Observe in Fig. 4 that the DCMC capacity of the proposed scheme approached the PCSI limit upon increasing the SNR, while that of the conventional CE and DD scheme did not.

In Fig. 5, the normalized Doppler frequency was varied from $-10$ dB to 10 dB with steps of 5 dB. The optimized block lengths were $L = 100, 100, 50$ for the average SNRs of $\rho = 0$ dB, 10 dB, and 20 dB, respectively. The DCMC capacity nearly the same as the PCSI bound for $\rho = 10$ dB, and the DCMC capacity that is 2.6 dB apart from the PSCI bound for $\rho = 0$ dB.

**IV. Simulation Results**

In this section, we provide our performance results in order to characterize the proposed scheme. We considered the scenarios of $K = 16$ users and $N = 128$ receive antennas while employing QPSK modulation, which corresponds to $R = 32$ bps/Hz for the channel-uncoded system. Moreover, the proposed scheme with idealistic PCSI and the conventional joint CE and DD scheme [11] were considered as benchmarks.

Fig. 3 shows the convergence behavior of MSE of CE for the proposed scheme. We considered $(K, N) = (16, 128)$ and QPSK modulation, hence a sum rate of $R = 32$ bps/Hz. Furthermore, the normalized Doppler frequency was set to $F_d T_s = 10^{-3}$ and the frame length was $T = 1600$. The
$F_d T_s = 1 \times 10^{-3}$ to $5 \times 10^{-3}$ with the step of $1 \times 10^{-3}$. The other system parameters were the same as those used in Fig. 4.

As seen in Fig. 5 that the DCMC capacity of the proposed scheme reached to the PCSI limit upon increasing the SNR.

Fig. 6 illustrates the BER of the proposed scheme with the three-stage turbo-coded architecture. The block length $L$ was optimized for each SNR, and $T = 1600$ frame length was considered. The information bits were channel encoded by the half-rate recursive systematic convolutional (RSC) code, and the RSC-coded bits were interleaved. Then, the interleaved bits were encoded by unity-rate convolutional (URC) code, and the URC-coded bits were interleaved. Hence, the sum rate was $R = 16$ bps/Hz. We employed RSC(2,1,2) code having the generator polynomials of $(G_r, G) = (3, 2)$ in octal format [22]. The number of outer iterations and that of inner iterations were set to 40 and 2, respectively. The normalized Doppler frequency was set to $F_d T_s = 10^{-3}$. As shown in Fig. 6, the BER of the conventional joint CE and DD scheme remained at 0.5 in the SNR range of $-10 \leq \rho \leq 10$ dB, whereas the BER of the proposed scheme becomes lower than $10^{-4}$ at 0 dB, which is expected from the result in Fig. 4.

V. CONCLUSIONS

In this letter, we proposed a novel joint CE and DD scheme with the aid of frame division for time-variant massive MIMO uplink channels. By optimizing the block length for the data frame division, the proposed scheme’s DCMC capacity approached the PCSI counterparts. Moreover, only a single iteration was needed for convergence in the reasonable SNR range of $\rho \geq 10$ dB.