Minimum Mean-Squared Error Pilot-Aided Transmission for MIMO Doubly Selective Channels

Arun P. Kannu and Philip Schniter
Dept. ECE, The Ohio State University
Columbus, OH 43210
pachaik@ece.osu.edu, schniter@ece.osu.edu

***Abstract*— This paper considers cyclic-prefixed block-based pilot-aided transmission (PAT), with possibly superimposed pilot and data symbols, over multiple-input multiple-output (MIMO) doubly selective channels (DSCs) that obey a complex-exponential basis expansion model. First, a tight lower bound on the mean-squared error (MSE) of pilot-aided channel estimates is derived, along with necessary and sufficient conditions on pilot/data patterns that achieve this bound. From these conditions, novel minimum-MSE (MMSE) PAT schemes are proposed and upper/lower bounds on their achievable rates are derived. The pilot/data power allocation that maximizes the lower bound is also derived. A high-SNR asymptotic analysis of the achievable rates is then performed, which suggests that the channel spreading parameters should be taken into account when choosing among MMSE-PAT schemes. The number of antennas maximizing the high-SNR achievable rates is also derived.**

I. Introduction

The wireless communication channel is typically modeled as linear transformation and parameterized by a set of time-varying coefficients. Coherent receivers estimate these channel coefficients for subsequent use in data detection. In pilot aided transmission (PAT), a known signal is embedded in the transmitted stream to facilitate channel estimation.

Recently, the information theoretic optimality of PAT schemes which minimize the mean squared error (MSE) of channel estimates has been investigated for multiple-input multiple-output (MIMO) flat fading [1], [2] and frequency selective fading [3] channels. We studied minimum MSE (MMSE) PAT schemes for the single-input single-output (SISO) doubly selective channel in [4]. In this paper, we extend our MMSE-PAT study to the MIMO doubly selective channel (DSC). In the sequel, we present novel MMSE-PAT schemes for the MIMO-DSC and demonstrate an inherent time-frequency duality. We also derive achievable rates (i.e., capacity lower bounds) for these schemes. The high-SNR asymptotic achievable rates are then optimized with respect to the number of active antennas.

The work of Yang et al. [5], while similar in direction, is restricted to (time) non-superimposed pilots. Our work is more general in that we search for MIMO-MMSE-PAT schemes among all linear modulations and allow the possibility of superimposed pilots. In fact, the MIMO-MMSE-PAT scheme found in [5] is but one of several MIMO-MMSE-PAT schemes uncovered here. In addition, we present novel asymptotic achievable-rate results for MMSE-PAT over the MIMO-DSC.

II. System Model

We consider a MIMO system with T transmit and R receive antennas, cyclic-prefix block transmission, and a DSC that satisfies a complex exponential basis expansion model. Details are given below.

A. Cyclic-Prefix Block Transmission Model

The sampled complex-baseband output signal $\{y^{(r)}(n)\}$ at the r^{th} receive antenna is related to the transmitted signals $\{x^{(t)}(n)\}$ from the t^{th} transmit antenna via

$$y^{(r)}(n) = \sum_{t=0}^{T-1}\sum_{\ell=0}^{N_\mathrm{t}-1} h^{(r,t)}(n,\ell)x^{(t)}(n-\ell) + v^{(r)}(n), \quad (1)$$

where $\{v^{(r)}(n)\}$ is zero-mean σ_v^2-variance circular white (spatial and temporal) Gaussian noise (CWGN) and $h^{(r,t)}(n,\ell)$ is the time-n channel response at the r^{th} receive antenna to an impulse applied at time $n-\ell$ on the t^{th} transmit antenna. Here, N_t denotes the channel's maximum time spread normalized to the sampling interval T_s, which is assumed equal for all (r,t) pairs. The length-N transmission block $\{x^{(t)}(n)\}_{n=0}^{N-1}$ is preceded by a cyclic prefix (CP) of length $N_\mathrm{t}-1$, whose contribution is discarded in forming $\boldsymbol{y}^{(r)} = [y^{(r)}(0),\dots,y^{(r)}(N-1)]^\top$. Note that, by making N large compared to N_t, the CP overhead can be made insignificant. Throughout this paper, we assume modulo-N indexing, i.e., $z(i) = z(\langle i\rangle_N)$. With the definitions

$$\begin{aligned}
\boldsymbol{X} &= \mathrm{diag}(\boldsymbol{X}^{(0)},\dots,\boldsymbol{X}^{(T-1)})\\
\boldsymbol{X}^{(t)} &= [\boldsymbol{X}_0^{(t)} \cdots \boldsymbol{X}_{-N_\mathrm{t}+1}^{(t)}]\\
\boldsymbol{X}_k^{(t)} &= \mathrm{diag}(x^{(t)}(k),\dots,x^{(t)}(k+N-1))\\
\boldsymbol{h}^{(r)} &= [\boldsymbol{h}^{(r,0)\top} \cdots \boldsymbol{h}^{(r,T-1)\top}]^\top\\
\boldsymbol{h}^{(r,t)} &= [\boldsymbol{h}_0^{(r,t)\top} \cdots \boldsymbol{h}_{N_\mathrm{t}-1}^{(r,t)\top}]^\top\\
\boldsymbol{h}_k^{(r,t)} &= [h^{(r,t)}(0,k),\dots,h^{(r,t)}(N-1,k)]^\top\\
\boldsymbol{v}^{(r)} &= [v^{(r)}(0),\dots,v^{(r)}(N-1)]^\top,
\end{aligned}$$

the DSC model (1) can be rewritten as

$$\boldsymbol{y}^{(r)} = \boldsymbol{X}\boldsymbol{h}^{(r)} + \boldsymbol{v}^{(r)}. \quad (2)$$

Collecting the observations from different receive antennas as $\bar{\boldsymbol{y}} = [\boldsymbol{y}^{(0)\top},\dots,\boldsymbol{y}^{(R-1)\top}]^\top$, we have

$$\bar{\boldsymbol{y}} = \bar{\boldsymbol{X}}\bar{\boldsymbol{h}} + \bar{\boldsymbol{v}}, \quad (3)$$

where $\bar{\boldsymbol{X}} = \boldsymbol{I}_R \otimes \boldsymbol{X}$, $\bar{\boldsymbol{h}} = [\boldsymbol{h}^{(0)\top}, ..., \boldsymbol{h}^{(R-1)\top}]^\top$ and $\bar{\boldsymbol{v}} = [\boldsymbol{v}^{(0)\top}, ..., \boldsymbol{v}^{(R-1)\top}]^\top$ and $\otimes$ denotes Kronecker product.

The transmit signal is constructed as $x^{(t)}(n) = p^{(t)}(n) + d^{(t)}(n)$, where $\{p^{(t)}(n)\}$ is the pilot sequence and $\{d^{(t)}(n)\}$ is the zero-mean data sequence. Note the superposition of pilots and data. Using $\{p^{(t)}(n)\}$ and $\{d^{(t)}(n)\}$ to construct $\boldsymbol{P}$ and $\boldsymbol{D}$, respectively, in the manner of $\boldsymbol{X}$, we see that

$$\boldsymbol{X} = \boldsymbol{P} + \boldsymbol{D}, \tag{4}$$

which again shows the superposition of pilots and data. Similarly, we have $\bar{\boldsymbol{X}} = \bar{\boldsymbol{P}} + \bar{\boldsymbol{D}}$. Defining the pilot vector $\boldsymbol{p}^{(t)} = [p^{(t)}(0), \dots, p^{(t)}(N-1)]^\top$, the pilot energy is constrained as

$$\sum_{t=0}^{T-1} \|\boldsymbol{p}^{(t)}\|^2 = E_p. \tag{5}$$

The data vector $\boldsymbol{d}^{(t)} = [d^{(t)}(0), \dots, d^{(t)}(N-1)]^\top$ is obtained by linear modulation of N_s information bearing symbols $\boldsymbol{s}^{(t)} = [s^{(t)}(0), \dots, s^{(t)}(N_s-1)]^\top$ according to

$$\boldsymbol{d}^{(t)} = \boldsymbol{B}^{(t)} \boldsymbol{s}^{(t)}, \tag{6}$$

where $\boldsymbol{B}^{(t)}$ is the t^{th} transmit antenna's "data modulation matrix." We require that the columns of $\boldsymbol{B}^{(t)}$ are orthonormal.

B. Doubly Selective Channel Model

We assume that the channel coefficients between different antenna pairs are independent with same second-order statistics. The following CE-BEM [6] describes the channel response between r^{th} receive and t^{th} transmit antenna over the N-length block duration. For $n \in \{0, \dots, N-1\}$ and $\ell \in \{0, \dots, N_\mathsf{t} - 1\}$,

$$h^{(r,t)}(n, \ell) = \frac{1}{\sqrt{N}} \sum_{k=-(N_\mathsf{f}-1)/2}^{(N_\mathsf{f}-1)/2} \lambda^{(r,t)}(k, \ell) e^{j\frac{2\pi}{N}kn}, \tag{7}$$

where CE-BEM coefficients $\{\lambda^{(r,t)}(k, \ell)\}$ are assumed to be zero-mean uncorrelated Gaussian with positive variance. With $f_\mathsf{d} T_s$ denoting the one-sided Doppler spread normalized to the sampling frequency, we refer to $N_\mathsf{f} = 2\lceil f_\mathsf{d} T_s N \rceil + 1$ as the discrete "frequency spread." We assume an underspread channel, i.e., a time-frequency spreading index $\gamma = 2 f_\mathsf{d} T_s N_\mathsf{t} \approx N_\mathsf{t} N_\mathsf{f}/N$ such that $\gamma < 1$. We define the $N \times N_\mathsf{f}$ matrix $\bar{\boldsymbol{F}}$ element-wise as $[\bar{\boldsymbol{F}}]_{n,m} = \frac{1}{\sqrt{N}} e^{j\frac{2\pi}{N} n(m - \frac{N_\mathsf{f}-1}{2})}$ and notice that $\bar{\boldsymbol{F}}^\mathsf{H} \bar{\boldsymbol{F}} = \boldsymbol{I}_{N_\mathsf{f}}$. With the definitions

$$\begin{aligned} \boldsymbol{U}_0 &= \boldsymbol{I}_{N_\mathsf{t}} \otimes \bar{\boldsymbol{F}} \\ \boldsymbol{\lambda}_\ell^{(r,t)} &= [\lambda^{(r,t)}(-\tfrac{N_\mathsf{f}-1}{2}, \ell), \dots, \lambda^{(r,t)}(\tfrac{N_\mathsf{f}-1}{2}, \ell)]^\top \\ \boldsymbol{\lambda}^{(r,t)} &= [\boldsymbol{\lambda}_0^{(r,t)\top} \cdots \boldsymbol{\lambda}_{N_\mathsf{t}-1}^{(r,t)\top}]^\top, \end{aligned}$$

(7) becomes $\boldsymbol{h}^{(r,t)} = \boldsymbol{U}_0 \boldsymbol{\lambda}^{(r,t)}$, which is the Karhunen-Loeve (KL) expansion of $\boldsymbol{h}^{(r,t)}$, since $\boldsymbol{U}_0^\mathsf{H} \boldsymbol{U}_0 = \boldsymbol{I}_{N_\mathsf{f} N_\mathsf{t}}$ and since $\boldsymbol{R}_\lambda^{(r,t)} = E\{\boldsymbol{\lambda}^{(r,t)} \boldsymbol{\lambda}^{(r,t)\mathsf{H}}\} > 0$ is diagonal. Since we assume the same second order statistics for different transmit-receive antenna pairs, we abbreviate $\boldsymbol{R}_\lambda^{(r,t)}$ by $\boldsymbol{R}_\lambda$. Now, with $\bar{\boldsymbol{U}} = \boldsymbol{I}_R \otimes \boldsymbol{U}$, $\boldsymbol{U} = \boldsymbol{I}_T \otimes \boldsymbol{U}_0$, $\bar{\boldsymbol{\lambda}} = [\boldsymbol{\lambda}^{(0)\top}, ..., \boldsymbol{\lambda}^{(R-1)\top}]^\top$, and $\boldsymbol{\lambda}^{(r)} = [\boldsymbol{\lambda}^{(r,0)\top}, ..., \boldsymbol{\lambda}^{(r,T-1)\top}]^\top$, we have $\boldsymbol{h}^{(r)} = \boldsymbol{U} \boldsymbol{\lambda}^{(r)}$ and

$$\bar{\boldsymbol{y}} = (\bar{\boldsymbol{P}} + \bar{\boldsymbol{D}}) \bar{\boldsymbol{U}} \bar{\boldsymbol{\lambda}} + \bar{\boldsymbol{v}}. \tag{8}$$

From the channel independence assumptions between different antenna pairs, we have $\boldsymbol{R}_{\bar{\lambda}} = E\{\bar{\boldsymbol{\lambda}} \bar{\boldsymbol{\lambda}}^\mathsf{H}\} = \boldsymbol{I}_{RT} \otimes \boldsymbol{R}_\lambda$.

III. MIMO-MMSE-PAT DESIGN

In this section, we present the MMSE-PAT design requirements for the MIMO DSC and propose novel MIMO-MMSE-PAT schemes.

A. MSE Lower Bound

The linear-MMSE (LMMSE) estimate of $\bar{\boldsymbol{h}}$ given the knowledge of $\{\bar{\boldsymbol{y}}, \bar{\boldsymbol{P}}\}$ and the knowledge of the second-order statistics of $\{\bar{\boldsymbol{h}}, \bar{\boldsymbol{D}}, \bar{\boldsymbol{v}}\}$ is [7]

$$\hat{\boldsymbol{h}} = \boldsymbol{R}_{\bar{y},\bar{h}}^\mathsf{H} \boldsymbol{R}_{\bar{y}}^{-1} \bar{\boldsymbol{y}}, \tag{9}$$

where $\boldsymbol{R}_{\bar{y},\bar{h}} = E\{\bar{\boldsymbol{y}} \bar{\boldsymbol{h}}^\mathsf{H}\}$ and $\boldsymbol{R}_{\bar{y}} = E\{\bar{\boldsymbol{y}} \bar{\boldsymbol{y}}^\mathsf{H}\}$. Given our assumptions,

$$\begin{aligned} \boldsymbol{R}_{\bar{y},\bar{h}} &= \bar{\boldsymbol{P}} \bar{\boldsymbol{U}} \boldsymbol{R}_{\bar{\lambda}} \bar{\boldsymbol{U}}^\mathsf{H} \\ \boldsymbol{R}_{\bar{y}} &= \bar{\boldsymbol{P}} \bar{\boldsymbol{U}} \boldsymbol{R}_{\bar{\lambda}} \bar{\boldsymbol{U}}^\mathsf{H} \bar{\boldsymbol{P}}^\mathsf{H} + E\{\bar{\boldsymbol{D}} \bar{\boldsymbol{U}} \boldsymbol{R}_{\bar{\lambda}} \bar{\boldsymbol{U}}^\mathsf{H} \bar{\boldsymbol{D}}^\mathsf{H}\} + \sigma_v^2 \boldsymbol{I}_{NR}. \end{aligned}$$

The channel estimation error $\tilde{\boldsymbol{h}} = \bar{\boldsymbol{h}} - \hat{\boldsymbol{h}}$ has total MSE $\sigma_e^2 = E\{\|\tilde{\boldsymbol{h}}\|^2\}$, which is given by [7]

$$\sigma_e^2 = \operatorname{tr}\{\bar{\boldsymbol{U}} \boldsymbol{R}_{\bar{\lambda}} \bar{\boldsymbol{U}}^\mathsf{H} - \boldsymbol{R}_{\bar{y},\bar{h}}^\mathsf{H} \boldsymbol{R}_{\bar{y}}^{-1} \boldsymbol{R}_{\bar{y},\bar{h}}\}. \tag{10}$$

We are interested in finding the energy constrained pilot vectors $\{\boldsymbol{p}^{(t)}\}$ and data modulation matrices $\{\boldsymbol{B}^{(t)}\}$ such that the resulting MSE σ_e^2 is minimal. We refer to these schemes as MIMO-MMSE-PAT and proceed with their design.

Theorem 1 (MSE Lower Bound). *For T-transmit R-receive antenna N-block CP PAT over the CE-BEM DSC, the channel estimate MSE obeys*

$$\sigma_e^2 \geq \operatorname{tr}\left\{ \left(\boldsymbol{R}_{\bar{\lambda}}^{-1} + \frac{E_p}{NT\sigma_v^2} \boldsymbol{I}_{RTN_\mathsf{f}N_\mathsf{t}} \right)^{-1} \right\}, \tag{11}$$

where equality in (11) occurs if and only if the following conditions hold:

1) Pilot-Data Orthogonality:

$$(\boldsymbol{P}\boldsymbol{U})^\mathsf{H} \boldsymbol{D}\boldsymbol{U} = \boldsymbol{0}, \forall \boldsymbol{D} \tag{12}$$

2) Optimal Excitation:

$$(\boldsymbol{P}\boldsymbol{U})^\mathsf{H} \boldsymbol{P}\boldsymbol{U} = \frac{E_p}{NT} \boldsymbol{I} \tag{13}$$

When (12)-(13) are met, $\boldsymbol{R}_{\tilde{h}} = E\{\tilde{\boldsymbol{h}} \tilde{\boldsymbol{h}}^\mathsf{H}\}$ is given by

$$\boldsymbol{R}_{\tilde{h}} = \bar{\boldsymbol{U}} \left(\boldsymbol{R}_{\bar{\lambda}}^{-1} + \frac{E_p}{NT\sigma_v^2} \boldsymbol{I}_{N_\mathsf{f}N_\mathsf{t}RT} \right)^{-1} \bar{\boldsymbol{U}}^\mathsf{H}. \tag{14}$$

Proof. The proof is similar to that of the SISO case [4] and hence is omitted for brevity. □

Denoting the q^{th} column of $\boldsymbol{B}^{(t)}$ as $\boldsymbol{b}_q^{(t)}$, we rephrase the MSE optimality requirements (12)-(13) in terms of the pilot sequence $\{p^{(t)}(i)\}$ and data basis sequence $\{b_q^{(t)}(i)\}$, using the index sets $\mathcal{N}_\mathsf{t} = \{-N_\mathsf{t}+1, ..., N_\mathsf{t}-1\}$, $\mathcal{N}_\mathsf{f} = \{-N_\mathsf{f}+1, ..., N_\mathsf{f}-1\}$, and $\mathcal{T} = \{0, ..., T-1\}$.

Lemma 1 (Time Domain). *For N-block CP PAT over the CE-BEM DSC, the following are necessary and sufficient conditions for equality in (11).* $\forall k \in \mathcal{N}_\mathsf{t}, \forall m \in \mathcal{N}_\mathsf{f}, \forall t_i \in \mathcal{T}$,

$$\sum_{i=0}^{N-1} b_q^{(t_1)}(i) p^{(t_2)*}(i-k) e^{-j\frac{2\pi}{N}mi} \;=\; 0. \tag{15}$$

$$\sum_{i=0}^{N-1} p^{(t_1)}(i) p^{(t_2)*}(i-k) e^{-j\frac{2\pi}{N}mi} \;=\; \frac{E_p}{T}\delta(k)\delta(m)\delta(t_1-t_2). \tag{16}$$

Proof. The proof is similar to that of the SISO case [4] and hence is omitted for brevity. □

Note that the number of receive antennas R does not affect the MIMO-MMSE-PAT design requirements in Lemma 1.

As in the SISO case, there exists a duality between time- and frequency-domain MIMO-MMSE-PAT. Denoting the N-point unitary FFT matrix by $\boldsymbol{F}_N$, we state the duality as follows.

Theorem 2 (Duality). *With $t \in \mathcal{T}$, if $(\{\boldsymbol{p}^{(t)}\}, \{\boldsymbol{B}^{(t)}\})$ parameterizes MIMO-MMSE N-block CP-PAT for T-transmit antenna system over the CE-BEM DSC with time spread N_1 and frequency spread N_2, then $(\{\boldsymbol{F}_N\boldsymbol{p}^{(t)}\}, \{\boldsymbol{F}_N\boldsymbol{B}^{(t)}\})$ parameterizes MIMO-MMSE N-block CP-PAT for T-transmit antenna system over the CE-BEM DSC with time spread N_2 and frequency spread N_1.*

Proof. The proof is similar to that of the SISO case [4] and hence is omitted for brevity. □

B. Data Dimension Analysis

We now present the bounds on the number of (linearly) independent data symbols that can be transmitted from each antenna, i.e., the rank of the data modulation matrix $\boldsymbol{B}^{(t)}$. Given a pilot vector $\boldsymbol{p}^{(t)}$ satisfying (16), a matrix $\boldsymbol{B}^{(t)} \in \mathbb{C}^{N\times N_s}$ which satisfies (15) can be constructed as follows. Defining the $(2N_\mathsf{f}-1)\times N$ matrix $\breve{\boldsymbol{F}}$ element-wise as $[\breve{\boldsymbol{F}}]_{n,m} = \frac{1}{\sqrt{N}}e^{-j\frac{2\pi}{N}(n-N_\mathsf{f}+1)m}$, and then defining

$$\boldsymbol{P}_k^{(t)} \;=\; \mathrm{diag}(p^{(t)}(k), ..., p^{(t)}(k+N-1)) \tag{17}$$

$$\boldsymbol{W}_k^{(t)} \;=\; \breve{\boldsymbol{F}}\boldsymbol{P}_k^{(t)\mathsf{H}} \tag{18}$$

$$\boldsymbol{W}^{(t)} \;=\; [\boldsymbol{W}_{-N_\mathsf{t}+1}^{(t)\top} \cdots \boldsymbol{W}_{N_\mathsf{t}-1}^{(t)\top}]^\top \tag{19}$$

$$\boldsymbol{W} \;=\; [\boldsymbol{W}^{(0)\top} \cdots \boldsymbol{W}^{(T-1)\top}]^\top, \tag{20}$$

condition (15) becomes $\boldsymbol{W}\boldsymbol{b}_q^{(t)} = \boldsymbol{0}$, implying that, for each q, the vector $\boldsymbol{b}_q^{(t)}$ must lie in the null space of $\boldsymbol{W}$. This can be achieved by choosing the columns of $\boldsymbol{B}^{(t)}$ as an orthonormal basis for $\mathrm{null}(\boldsymbol{W})$, yielding "data dimension" $N_s = \dim(\mathrm{null}(\boldsymbol{W}))$. Due to the structure of $\boldsymbol{W}$, the data dimension N_s is the same for each transmit antenna.

The data dimension N_s, i.e., the number of information symbols per N-block CP-PAT per transmit antenna, can be bounded as follows. Note from (20) that the $TN_\mathsf{f}N_\mathsf{t}$ rows of $(\boldsymbol{P}\boldsymbol{U})^\mathsf{H}$ are contained within the $T(2N_\mathsf{f}-1)(2N_\mathsf{t}-1)$ rows of $\boldsymbol{W}$. In order to satisfy (13), those rows must be orthogonal. Thus, $TN_\mathsf{f}N_\mathsf{t} \le \mathrm{rank}(\boldsymbol{W}) \;\le\; T(2N_\mathsf{f}-1)(2N_\mathsf{t}-1)$, which means that (15)-(16) imply

$$N - T(2N_\mathsf{f}-1)(2N_\mathsf{t}-1) \;\le\; N_s \;\le\; N - TN_\mathsf{f}N_\mathsf{t}. \tag{21}$$

From (21), we see that MIMO-MMSE-PAT sacrifices at least $TN_\mathsf{t}N_\mathsf{f}$, but no more than $T(2N_\mathsf{t}-1)(2N_\mathsf{f}-1)$, signaling dimensions per transmit antenna. Recall that $N_\mathsf{f}N_\mathsf{t}$ describes the number of degrees of freedom in the DSC per each transmit-receive antenna pair and $TN_\mathsf{f}N_\mathsf{t}$ denotes the number of (independent) channel coefficients to be estimated at each receive antenna.

C. MIMO-MMSE-PAT Examples

Here we give several examples of N-block CP MIMO-MMSE-PAT schemes for the CE-BEM DSC using the $\{\boldsymbol{p}^{(t)}, \boldsymbol{B}^{(t)}\}_{t\in\mathcal{T}}$ parameterization. It is straightforward to verify that the following examples satisfy the MIMO-MMSE-PAT conditions (15)-(16).

Example 1 (TDKD). *Assuming $\frac{N}{N_\mathsf{f}} \in \mathbb{Z}$, define the pilot index sets $\mathcal{P}_\mathsf{t}^{(t)}$ and the guard index set $\mathcal{G}_\mathsf{t}$:*

$$\mathcal{P}_\mathsf{t}^{(t)} \;=\; \{i+tN_\mathsf{t}, i+tN_\mathsf{t}+\tfrac{N}{N_\mathsf{f}}, ..., i+tN_\mathsf{t}+\tfrac{(N_\mathsf{f}-1)N}{N_\mathsf{f}}\}$$
$$\mathcal{G}_\mathsf{t} \;=\; \bigcup_{t\in\mathcal{T}} \bigcup_{k\in\mathcal{P}_\mathsf{t}^{(t)}} \{-N_\mathsf{t}+1+k, ..., N_\mathsf{t}-1+k\}.$$

An N-block CP MIMO MMSE-PAT scheme for the CE-BEM DSC is given by

$$p^{(t)}(q) \;=\; \begin{cases} \sqrt{\frac{E_p}{TN_\mathsf{f}}}e^{j\theta(q)} & q\in\mathcal{P}_\mathsf{t}^{(t)} \\ 0 & q\notin\mathcal{P}_\mathsf{t}^{(t)} \end{cases} \tag{22}$$

and by $\boldsymbol{B}^{(t)}$ constructed from the columns of $\boldsymbol{I}_N$ with indices not *in the set $\mathcal{G}$. Both $i \in \{0,\ldots,\frac{N}{N_\mathsf{f}}-1\}$ and $\theta(q) \in \mathbb{R}$, are arbitrary. The corresponding data dimension per transmit antenna is $N_s = N-(T+1)N_\mathsf{f}N_\mathsf{t}+N_\mathsf{f}$.*

Example 2 (FDKD). *Assuming $\frac{N}{N_\mathsf{t}} \in \mathbb{Z}$, define the (frequency domain) pilot index sets $\mathcal{P}_\mathsf{f}^{(t)}$ and the guard index set $\mathcal{G}_\mathsf{f}$:*

$$\mathcal{P}_\mathsf{f}^{(t)} \;=\; \{i+tN_\mathsf{f}, i+tN_\mathsf{f}+\tfrac{N}{N_\mathsf{t}}, ..., i+tN_\mathsf{f}+\tfrac{(N_\mathsf{t}-1)N}{N_\mathsf{t}}\}$$
$$\mathcal{G}_\mathsf{f} \;=\; \bigcup_{t\in\mathcal{T}} \bigcup_{k\in\mathcal{P}_\mathsf{f}^{(t)}} \{-N_\mathsf{f}+1+k, ..., N_\mathsf{f}-1+k\}.$$

An N-block CP MIMO MMSE-PAT scheme for the CE-BEM DSC is given by $\boldsymbol{p}^{(t)} = \boldsymbol{F}_N^\mathsf{H}\breve{\boldsymbol{p}}^{(t)}$, with

$$\breve{p}^{(t)}(q) \;=\; \begin{cases} \sqrt{\frac{E_p}{TN_\mathsf{t}}}e^{j\theta(q)} & q\in\mathcal{P}_\mathsf{f}^{(t)} \\ 0 & q\notin\mathcal{P}_\mathsf{f}^{(t)} \end{cases} \tag{23}$$

and by $\boldsymbol{B}^{(t)}$ constructed from the columns of IDFT matrix $\boldsymbol{F}_N^\mathsf{H}$ with indices not *in the set $\mathcal{G}_\mathsf{f}$. Both $i \in \{0,\ldots,\frac{N}{N_\mathsf{t}}-1\}$ and*

$\theta(q) \in \mathbb{R}$, are arbitrary. The corresponding data dimension per transmit antenna is $N_s = N - (T+1)N_\mathsf{f}N_\mathsf{t} + N_\mathsf{t}$.

Example 3 (Superimposed Chirps). *Assuming even N, an N-block CP MMSE-PAT scheme for the CE-BEM DSC is given by*

$$p^{(t)}(q) = \sqrt{\frac{E_p}{NT}} e^{j\frac{2\pi}{N}\left(\frac{N_\mathsf{t}}{2}q^2 + tN_\mathsf{f}N_\mathsf{t}q\right)} \quad (24)$$

$$[\boldsymbol{B}^{(t)}]_{q,m} = \frac{1}{\sqrt{N}} e^{j\frac{2\pi}{N}\left(\frac{N_\mathsf{t}}{2}q^2 + (m+TN_\mathsf{f}N_\mathsf{t})q\right)}, \quad (25)$$

for $q \in \{0,\ldots,N-1\}$ and $m \in \{0,\ldots,N_s-1\}$, where the data dimension per transmit antenna is $N_s = N-(T+1)N_\mathsf{f}N_\mathsf{t}+1$.

The MSE-optimality of TDKD among zero-prefix block transmissions with non-superimposed pilots was established in [5]. For CP transmissions with (time) superimposed pilots, we see that there exist other MMSE-PAT schemes.

IV. Achievable Rates of MIMO-MMSE-PAT

We now calculate ergodic achievable rates of MIMO-MMSE-PAT for the CE-BEM DSC, paying special attention to the high-SNR regime. In the finite SNR case, we obtain the bounds on the achievable rates considering i.i.d. Gaussian input distribution. In the high-SNR asymptotic case, we characterize the maximum achievable rates of the MIMO-MMSE-PAT optimized over all the input probability distributions.

A. System Model

In our achievable-rate analysis, we focus on the CE-BEM [recall (7)] with i.i.d. coefficients, which we refer as uniform CE-BEM (UCE-BEM). Specifically, we assume

$$\boldsymbol{R}_\lambda = \frac{N}{N_\mathsf{f}N_\mathsf{t}} \boldsymbol{I}_{N_\mathsf{f}N_\mathsf{t}}. \quad (26)$$

The UCE-BEM DSC approximates wide-sense stationary uncorrelated scattering (WSSUS) with uniform power delay profile and uniform Doppler power spectrum. Since the UCE-BEM is a special case of the CE-BEM, the MIMO-MMSE-PAT design requirements (15)-(16) and the MIMO-MMSE-PAT examples in Section III-C apply to the UCE-BEM.

Suppose that the MMSE-PAT scheme $\{\boldsymbol{p}^{(t)}, \boldsymbol{B}^{(t)}\}_{t\in\mathcal{T}}$ for the UCE-BEM-DSC has pilot energy E_p [recall (5)] and yields data dimension N_s. Suppose also that the average data energy per block is $E\{\sum_{t=0}^{T-1}\|\boldsymbol{d}^{(t)}\|^2\} = E\{\sum_{t=0}^{T-1}\|\boldsymbol{s}^{(t)}\|^2\} = E_s$. Thus, the total transmit energy per block is $E_\mathsf{tot} = E_p + E_s$. From this, we define the average transmit power $\sigma^2_\mathsf{tot} = \frac{E_\mathsf{tot}}{N}$ and the signal-to-noise ratio $\mathsf{SNR} = \frac{\sigma^2_\mathsf{tot}}{\sigma^2_v}$. In addition, we define the normalized signal and pilot powers $\sigma_s^2 = \frac{E_s}{N_s}$ and $\sigma_p^2 = \frac{E_p}{TN_\mathsf{t}N_\mathsf{f}}$, respectively. In the sequel, we analyze the ergodic capacity of MMSE-PAT schemes $\{\boldsymbol{p}^{(t)}, \boldsymbol{B}^{(t)}\}_{t\in\mathcal{T}}$ over the UCE-BEM-DSC with average transmit power constraint σ^2_tot.

It will be convenient to define $\boldsymbol{H}^{(r,t)} \in \mathbb{C}^{N\times N}$ element-wise as $[\boldsymbol{H}^{(r,t)}]_{n,m} = h^{(r,t)}(n, \langle n-m\rangle_N)$, so that the input-output relation (2) becomes

$$\boldsymbol{y}^{(r)} = \underbrace{\left[\boldsymbol{H}^{(r,0)} \cdots \boldsymbol{H}^{(r,T-1)}\right]}_{\boldsymbol{H}^{(r)}} (\bar{\boldsymbol{p}} + \bar{\boldsymbol{B}}\bar{\boldsymbol{s}}) + \boldsymbol{v}^{(r)}, \quad (27)$$

where $\bar{\boldsymbol{p}} = [\boldsymbol{p}^{(0)\top}, ..., \boldsymbol{p}^{(T-1)\top}]^\top$, $\bar{\boldsymbol{s}} = [\boldsymbol{s}^{(0)\top}, ..., \boldsymbol{s}^{(T-1)\top}]^\top$, and $\bar{\boldsymbol{B}} = \operatorname{diag}(\boldsymbol{B}^{(0)}, ..., \boldsymbol{B}^{(T-1)})$. Then, defining $\bar{\boldsymbol{H}} = [\boldsymbol{H}^{(0)\top}, ..., \boldsymbol{H}^{(R-1)\top}]^\top$, we collect the observations of all receive antennas into $\bar{\boldsymbol{y}} = [\boldsymbol{y}^{(0)\top}, ..., \boldsymbol{y}^{(R-1)\top}]^\top$, such that

$$\bar{\boldsymbol{y}} = \bar{\boldsymbol{H}}\bar{\boldsymbol{p}} + \bar{\boldsymbol{H}}\bar{\boldsymbol{B}}\bar{\boldsymbol{s}} + \bar{\boldsymbol{v}}, \quad (28)$$

with $\bar{\boldsymbol{v}} = [\boldsymbol{v}^{(0)\top}, ..., \boldsymbol{v}^{(R-1)\top}]^\top$.

In our ergodic analysis, we assume a Gaussian channel coefficient vector $\boldsymbol{h}^{(r,t)}$ that varies independently from block to block. Independent fading across blocks can be achieved by time-domain interleaving.

B. Achievable-Rate Bounds - Finite SNR Case

In this section, we obtain achievable-rate bounds under an i.i.d. Gaussian input distribution.

Theorem 3 (Achievable-Rate Bounds). *For the N-block CP MMSE-PAT scheme $\{\boldsymbol{p}^{(t)}, \boldsymbol{B}^{(t)}\}_{t\in\mathcal{T}}$ with i.i.d. Gaussian $\bar{\boldsymbol{s}} \in \mathbb{C}^{N_s}$ over the UCE-BEM DSC, the per-channel-use ergodic achievable-rate $\mathcal{R}_{\textsf{mse-gau}}$ obeys $\mathcal{R}_{\textsf{mse-gau-lb}} \le \mathcal{R}_{\textsf{mse-gau}} \le \mathcal{R}_{\textsf{mse-gau-ub}}$:*

$$\mathcal{R}_{\textsf{mse-gau-lb}} = \frac{1}{N} E\{\log\det[\boldsymbol{I}_{TN_S} + \rho_l \bar{\boldsymbol{B}}^\mathsf{H}\bar{\boldsymbol{H}}^\mathsf{H}\bar{\boldsymbol{H}}\bar{\boldsymbol{B}}]\} \quad (29)$$

$$\mathcal{R}_{\textsf{mse-gau-ub}} = \frac{1}{N} E\{\log\det[\boldsymbol{I}_{TN_s} + \rho_u \bar{\boldsymbol{B}}^\mathsf{H}\bar{\boldsymbol{H}}^\mathsf{H}\bar{\boldsymbol{H}}\bar{\boldsymbol{B}}]\} \quad (30)$$

$$\rho_l = \frac{\sigma_s^2}{T\sigma_v^2}\left(\frac{\sigma_p^2}{\sigma_p^2+\sigma_s^2+\sigma_v^2}\right), \quad \rho_u = \frac{\sigma_s^2}{T\sigma_v^2} \quad (31)$$

Proof. The proof is similar to that of the SISO case [4] and hence is omitted for brevity. □

The lower bound (29) describes the “worst case” scenario of channel estimation error acting as AWGN. This concept was previously used in, e.g., [2] and [8]. The upper bound (30) describes the “best case” scenario of perfect channel estimates.

C. Pilot/Data Power Allocation

Until now, the MMSE-PAT schemes were designed using fixed pilot energy E_p. Now we consider the problem of allocating the total transmit energy E_tot between pilots and data. We approach this problem through $\mathcal{R}_{\textsf{mse-gau-lb}}$ maximization.

Let $\alpha \in [0,1]$ denote the fraction of energy allocated to the data symbols, i.e., $E_s = \alpha E_\mathsf{tot}$ and $E_p = (1-\alpha)E_\mathsf{tot}$. We are interested in finding $\alpha_\star = \arg\max_\alpha \mathcal{R}_{\textsf{mse-gau-lb}}(\alpha)$. Because α affects $\mathcal{R}_{\textsf{mse-gau-lb}}$ only through the term ρ_l, and because $\mathcal{R}_{\textsf{mse-gau-lb}}$ is strictly increasing in ρ_l, it suffices to maximize

ρ_l w.r.t. α. The value of $\alpha_\star$ is readily obtained by finding the value of α which sets $\partial \rho_l / \partial \alpha = 0$. This can be shown to be

$$\alpha_\star = \begin{cases} \beta - \sqrt{\beta^2 - \beta} & \text{if } N_s > TN_\mathsf{f}N_\mathsf{t} \\ \beta + \sqrt{\beta^2 - \beta} & \text{if } N_s < TN_\mathsf{f}N_\mathsf{t} \\ \frac{1}{2} & \text{if } N_s = TN_\mathsf{f}N_\mathsf{t} \end{cases} \tag{32}$$

$$\beta = \frac{1 + \frac{TN_\mathsf{f}N_\mathsf{t}\sigma_v^2}{E_{\text{tot}}}}{1 - \frac{TN_\mathsf{f}N_\mathsf{t}}{N_s}}. \tag{33}$$

While capacity bounds and pilot/data power allocation for the TDKD scheme have been obtained in [5], the results here hold for arbitrary MIMO-MMSE-PAT schemes.

D. High-SNR Asymptotic Achievable Rates

In Section IV-B, we derived achievable rates constraining that the input distribution is i.i.d. Gaussian. Now, focusing on the high-SNR regime, we obtain the maximal achievable- rate $\mathcal{R}_{\mathsf{mse}}$ (up to a constant) for MIMO-MMSE-PAT optimized over all the input distributions.

Theorem 4 (Asymptotic Achievable Rates). *For an N-block CP MIMO-MMSE-PAT scheme operating over the UCE-BEM DSC with T transmit and R receive antennas, and with data dimension N_s, the ergodic maximal achievable-rate $\mathcal{R}_{\mathsf{mse}}$ -optimized over all the input distributions, obeys*

$$\mathcal{R}_{\mathsf{mse}}(\mathit{SNR}) = \frac{\min\{R(N - TN_\mathsf{f}N_\mathsf{t}), TN_s\}}{N} \log \mathit{SNR} + O(1), \tag{34}$$

as $\mathit{SNR} \to \infty$. Also, the Gaussian input distribution achieves the same pre-log factor as that of the maximal rate $\mathcal{R}_{\mathsf{mse}}$ (34).

Proof. See Appendix I. □

The pre-log factor in $\mathcal{R}_{\mathsf{mse}}$ can be interpreted as follows. Since each transmit antenna uses only N_s data dimensions per N-length block, the total number of independent data symbols transmitted is TN_s. Similarly, each receive antenna estimates $TN_\mathsf{f}N_\mathsf{t}$ channel coefficients. In MIMO-MMSE-PAT, the observations corresponding to pilot symbol $\boldsymbol{P}$ use up $TN_\mathsf{f}N_\mathsf{t}$ observations [recall (13)], leaving only $N - TN_\mathsf{f}N_\mathsf{t}$ observations per receive antenna, or $R(N - TN_\mathsf{f}N_\mathsf{t})$ total observations, for data. Since at most $\min\{R(N - TN_\mathsf{f}N_\mathsf{t}), TN_s\}$ data symbols per N-block can be "resolved" at the receiver, this number determines the pre-log factor of $\mathcal{R}_{\mathsf{mse}}(\mathsf{SNR})$.

Theorem 4 implies that, among MIMO-MMSE-PAT schemes, those with lower N_s can not yield higher pre-log factor. Among the examples in Section III-C, TDKD maximizes N_s when $N_\mathsf{f} > N_\mathsf{t}$, and FDKD maximizes N_s when $N_\mathsf{f} < N_\mathsf{t}$. Apart from the trivial case $N_\mathsf{f} = N_\mathsf{t} = 1$, the Chirp PAT is dominated by both TDKD and FDKD.

As is evident from Theorem 4, the asymptotic achievable rates do not necessarily increase with the number of antennas. Hence, it is worthwhile to determine the number of active transmit and receive antennas, $T_\star \in \{0, \ldots, T\}$ and $R_\star \in \{0, \ldots, R\}$, which maximize the pre-log factor in $\mathcal{R}_{\mathsf{mse}}$. Note that $T_\star$ and $R_\star$ depend on the MIMO-MMSE-PAT scheme through the data dimension N_s. For the MIMO-MMSE-PAT Examples 1-3, it can be shown that

$$R_\star = R \tag{35}$$

$$T_\star = \begin{cases} \left\lfloor \min\left(T, \frac{\psi}{2N_\mathsf{f}N_\mathsf{t}} - \frac{R_\star}{2}\right) + \frac{1}{2} \right\rfloor, & \text{if } \psi^2 < 4R_\star NN_\mathsf{t}N_\mathsf{f} \\ \left\lfloor \min\left(T, \frac{\psi - \sqrt{\psi^2 - 4R_\star NN_\mathsf{f}N_\mathsf{t}}}{2N_\mathsf{f}N_\mathsf{t}}\right) + \frac{1}{2} \right\rfloor, & \text{otherwise.} \end{cases} \tag{36}$$

$$\psi = N + (R_\star - 1)N_\mathsf{f}N_\mathsf{t} + \kappa \tag{37}$$

where $\kappa = 1$, N_t, and N_f for Chirp, FDKD, and TDKD, respectively. Note that, in some cases, the transmitter uses strictly less than T antennas. Similar results were obtained for MIMO flat-fading and time-selective channels in [1] and [6], respectively.

V. Numerical Results

In this section, we numerically evaluate the achievable-rate bounds (29) and (30) with $N_\mathsf{t} = 9$, $N_\mathsf{f} = 3$, $R = 2$, and $N = 126$. Figure 1 compares the performance of the TDKD, FDKD and Chirp schemes with $T = 2$. Since this channel has $N_\mathsf{t} > N_\mathsf{f}$, FDKD shows a gain over both TDKD and the Chirp scheme. In Fig. 2, we study the performance of FDKD with $T \in \{1, 2, 3\}$ transmit antennas. We find that the high-SNR achievable-rate is maximum when $T = 2$, coinciding with $T_\star$ from (36).

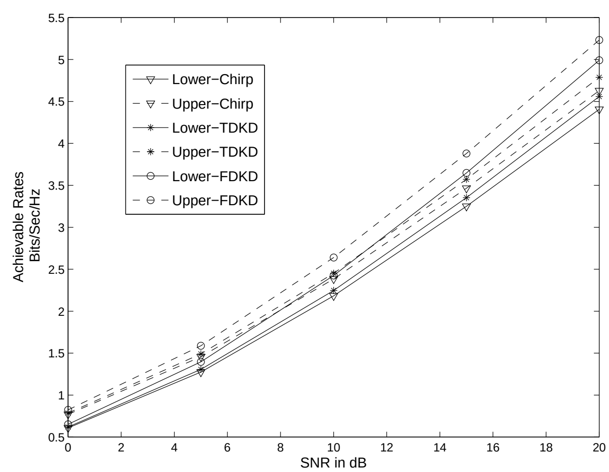

Fig. 1. Performance of different schemes $T = 2$.

VI. Conclusion

We extended our previous work [4] on cyclic prefix block-based MMSE-PAT for the SISO CE-BEM-DSC to the MIMO case. Specifically, we found necessary and sufficient conditions on PAT design, presented novel PAT examples, and derived bounds on the achievable rates. The pilot/data power allocation maximizing the lower bound on achievable-rate was also presented. A high-SNR achievable-rate analysis was then presented which suggested that the channel's spreading parameters should be taken into account when choosing among

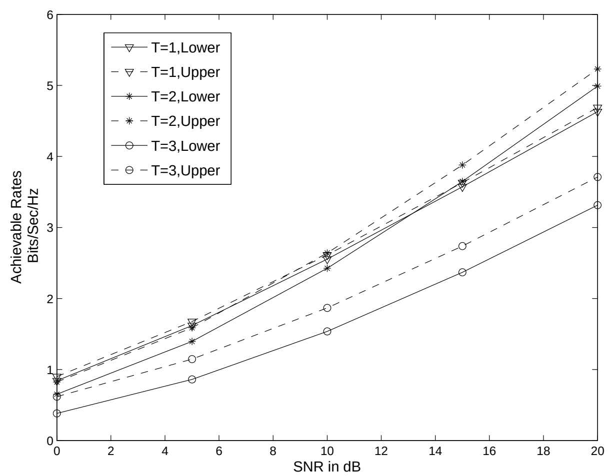

Fig. 2. Performance of FDKD with $T \in \{1, 2, 3\}$ transmit antennas.

MMSE-PAT schemes. Finally, the number of antennas maximizing the high-SNR rates was also derived.

Appendix I
Proof of Theorem 4

We start with the lower bound $\mathcal{R}_{\text{mse-gau-lb}} \leq \mathcal{R}_{\text{mse}}$. First, let us characterize the asymptotic behavior of $\mathcal{R}_{\text{mse-gau-lb}}$. Let the columns of $\boldsymbol{B}_p \in \mathbb{C}^{N \times TN_{\text{f}}N_{\text{t}}}$ form an orthonormal basis for $\text{col}(\boldsymbol{PU})$ and the columns of $\boldsymbol{B}_d \in \mathbb{C}^{N \times (N - TN_{\text{f}}N_{\text{t}})}$ form an orthonormal basis for the left null space of $\boldsymbol{PU}$. The pilot-data orthogonality of MMSE-PAT [recall (12)] implies that $\boldsymbol{B}_p^{\mathsf{H}} \boldsymbol{H}^{(r)} \bar{\boldsymbol{B}} = \boldsymbol{0}$ and $\boldsymbol{B}_d^{\mathsf{H}} \boldsymbol{H}^{(r)} \bar{\boldsymbol{p}} = \boldsymbol{0}$, $\forall\ r$. Projecting the r^{th} antenna observation vector $\boldsymbol{y}^{(r)}$ onto the data subspaces, we obtain

$$\boldsymbol{y}_d^{(r)} \ = \ \boldsymbol{B}_d^{\mathsf{H}} \boldsymbol{H}^{(r)} \bar{\boldsymbol{B}} \bar{\boldsymbol{s}} + \boldsymbol{v}_d^{(r)}, \tag{38}$$

where $\boldsymbol{v}_d^{(r)} = \boldsymbol{B}_d^{\mathsf{H}} \boldsymbol{v}^{(r)}$ and is CWGN with variance σ_v^2. Defining $\bar{\boldsymbol{y}}_d = [\boldsymbol{y}_d^{(0)\top}, ..., \boldsymbol{y}_d^{(R-1)\top}]^\top$, we have

$$\bar{\boldsymbol{y}}_d \ = \ \bar{\boldsymbol{B}}_d^{\mathsf{H}} \bar{\boldsymbol{H}} \bar{\boldsymbol{B}} \bar{\boldsymbol{s}} + \bar{\boldsymbol{v}}_d, \tag{39}$$

where $\bar{\boldsymbol{B}}_d = \boldsymbol{I}_R \otimes \boldsymbol{B}_d$, $\bar{\boldsymbol{v}}_p = [\boldsymbol{v}_p^{(0)\top}, ..., \boldsymbol{v}_p^{(R-1)\top}]^\top$ and $\bar{\boldsymbol{v}}_d = [\boldsymbol{v}_d^{(0)\top}, ..., \boldsymbol{v}_d^{(R-1)\top}]^\top$. Since the projection (39) does not compromise data energy, there is no loss in mutual information, i.e., $I(\bar{\boldsymbol{y}}; \bar{\boldsymbol{s}}) = I(\bar{\boldsymbol{y}}_d; \bar{\boldsymbol{s}})$. Hence, with the "effective channel" $\boldsymbol{H}_e = \bar{\boldsymbol{B}}_d^{\mathsf{H}} \bar{\boldsymbol{H}} \bar{\boldsymbol{B}}$, it easily follows that $\mathcal{R}_{\text{mse-gau-lb}} = \frac{1}{N} E\{\log\det[\boldsymbol{I} + \rho_l \bar{\boldsymbol{H}}_e^{\mathsf{H}} \bar{\boldsymbol{H}}_e]\}$. The structure and statistics of $\bar{\boldsymbol{H}}$, and the structures of $\bar{\boldsymbol{B}}$ and $\boldsymbol{B}_d$ imply that $\bar{\boldsymbol{H}}_e$ is full rank with probability 1. Since $\bar{\boldsymbol{H}}_e$ is an $R(N - TN_{\text{f}}N_{\text{t}}) \times TN_s$ matrix, let $\{\mu_i\}_{i=0}^{Q-1}$ denote the positive eigen values of $\bar{\boldsymbol{H}}_e^{\mathsf{H}} \bar{\boldsymbol{H}}_e$, where $Q = \min\{R(N - TN_{\text{f}}N_{\text{t}}), TN_s\}$. Now, (29) can be written as

$$\mathcal{R}_{\text{mse-gau-lb}} \ = \ \frac{1}{N} E\{\log \prod_{i=0}^{Q-1} (1 + \rho_l \mu_i)\}. \tag{40}$$

With power allocation fraction $\alpha \in (0, 1)$, $E_s = \alpha E_{\text{tot}}$ and $E_p = (1 - \alpha) E_{\text{tot}}$, we have

$$\rho_l \ = \ \frac{N}{T^2 N_s} \alpha (1 - \alpha) \mathsf{SNR} \left(\frac{1}{\frac{N_{\text{f}}N_{\text{t}}}{N_s}\alpha + \frac{1-\alpha}{T} + \frac{N_{\text{f}}N_{\text{t}}}{N\mathsf{SNR}}} \right). \tag{41}$$

Equation (41) implies that there exists a constant k such that $\rho_l \geq k\mathsf{SNR}$ for all $\mathsf{SNR} \geq 1$. We can use this and the fact that $\mu_i > 0$ in (40) to claim

$$\mathcal{R}_{\text{mse}} \ \geq \ \frac{\min\{R(N - TN_{\text{f}}N_{\text{t}}), TN_s\}}{N} \log \mathsf{SNR} + O(1) \tag{42}$$

as $\mathsf{SNR} \to \infty$.

Now, to bound the asymptotic achievable rate from above, we consider the "coherent" case of zero channel estimation error, i.e., $\bar{\boldsymbol{H}}$ is perfectly known at the receiver. Since $I(\bar{\boldsymbol{y}}; \bar{\boldsymbol{s}}) = I(\bar{\boldsymbol{y}}_d; \bar{\boldsymbol{s}})$, using $\mathcal{C}_{\text{coh}}$ to denote the capacity of (39) with power constraint σ_{tot}^2, it is evident that $\mathcal{R}_{\text{mse}} \leq \mathcal{C}_{\text{coh}}$. In fact, for (39), the capacity maximizing input distribution is zero-mean Gaussian [9], so that

$$\mathcal{C}_{\text{coh}} \ = \ \frac{1}{N} \sup_{\text{tr}\{\boldsymbol{R}_{\bar{s}}\} \leq N\sigma_{\text{tot}}^2} E\{\log\det[\boldsymbol{I} + \frac{1}{\sigma_v^2} \bar{\boldsymbol{H}}_e^{\mathsf{H}} \boldsymbol{R}_{\bar{s}} \bar{\boldsymbol{H}}_e]\}, \tag{43}$$

where $\boldsymbol{R}_{\bar{s}}$ denotes the covariance of $\bar{\boldsymbol{s}}$. Note that, for any $\boldsymbol{R}_{\bar{s}}$ satisfying the power constraint in (43), we have

$$\boldsymbol{R}_{\bar{s}} \ \leq \ N\sigma_{\text{tot}}^2 \boldsymbol{I}_{TN_s}, \tag{44}$$

in the positive semi-definite sense. Using (44) in (43), we find

$$\mathcal{R}_{\text{mse}} \ \leq \ \frac{1}{N} E\{\log\det[\boldsymbol{I}_{N_s} + \frac{N\sigma_{\text{tot}}^2}{\sigma_v^2} \bar{\boldsymbol{H}}_e^{\mathsf{H}} \bar{\boldsymbol{H}}_e]\} \tag{45}$$

$$= \ \frac{\min\{R(N - TN_{\text{f}}N_{\text{t}}), TN_s\}}{N} \log \mathsf{SNR} + O(1), \tag{46}$$

as $\mathsf{SNR} \to \infty$. Theorem 4 follows from (42) and (46).

References

[1] L. Zheng and D. Tse, "Communication on the grassmann manifold: a geometric approach to the noncoherent multiple-antenna channel," *IEEE Trans. on Information Theory*, vol. 48, pp. 359–383, Feb. 2002.

[2] B. Hassibi and B. Hochwald, "How much training is needed in multiple-antenna wireless links," *IEEE Trans. on Information Theory*, vol. 49, pp. 951–963, Apr. 2003.

[3] H. Vikalo, B. Hassibi, B. Hochwald, and T. Kailath, "On the capacity of frequency-selective channels in training-based transmission schemes," *IEEE Trans. on Signal Processing*, vol. 52, pp. 2572–2583, Sep. 2004.

[4] A. P. Kannu and P. Schniter, "Minimum mean squared error pilot aided transmission for doubly selective channels," *IEEE Trans. on Signal Processing*. (in submission).

[5] L. Yang, X. Ma, and G. B. Giannakis, "Optimal training for MIMO fading channels with time- and frequency- selectivity," in *Proc. IEEE Internat. Conf. on Acoustics, Speech, and Signal Processing*, 2004.

[6] Y. Liang and V. V. Veeravalli, "Capacity of noncoherent time-selective Rayleigh-fading channels," *IEEE Trans. on Information Theory*, vol. 50, pp. 3095–3110, Dec. 2004.

[7] L. L. Scharf, *Statistical Signal Processing*. Reading, MA: Addison-Wesley, 1991.

[8] M. Medard, "The effect upon channel capacity in wireless communication of perfect and imperfect knowledge of the channel," *IEEE Trans. on Information Theory*, vol. 46, pp. 933–946, May 2000.

[9] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. on Telecommunications*, vol. 10, pp. 585–595, Nov. 1999.