Frequency-Selective Multiuser Downlink Channels Under Mismatched Coherence Conditions

Mohamed Fadel, *Student Member, IEEE*, and Aria Nosratinia, *Fellow, IEEE*

***Abstract*— Downlink transmission into frequency-selective links experiencing different coherence bandwidths and coherence times has obvious practical relevance, but is often addressed via assuming for all users the coherence conditions corresponding to the fastest, most dispersive user. This paper demonstrates an alternative approach that directly exploits the differences between coherence bandwidth and coherence time of users under frequency-selective conditions, showing that it can lead to significant new efficiencies and gains. More specifically, we study a two-user downlink frequency-selective channel under three broad conditions of disparity between the link qualities: when the disparity is in coherence time, in coherence bandwidth, and in both coherence time and coherence bandwidth. Each of the coherence scenarios calls for a distinct treatment; for each, we provide an analysis demonstrating the gains in achievable rates due to exploiting coherence disparity. Numerical simulations demonstrate the advantages of the proposed schemes.**

***Index Terms*— Channel estimation, coherence diversity, downlink channel, fading channel, multiuser, OFDM, product superposition.**

I. Introduction

WIDEBAND wireless channels, which are important for high data rate applications, are often frequency-selective. Downlink frequency-selective channels in particular are the subject of significant interest [2]–[6], and are often analyzed via OFDM by dividing the time-frequency plane into blocks. This facilitates resource allocation and channel training, and is the underlying motivation of organizing transmission strategies around *resource blocks* [7], [8].

The analysis of multiuser networks in the literature has mostly concentrated on fading links with equal coherence time and equal coherence bandwidth. In this paper, we analyze a multi-user frequency-selective downlink channel where the coherence conditions of the different links are not identical. We show that by leveraging the differences between the coherence conditions in frequency-selective channels, a gain can be obtained in the transmission rate. This gain is denoted *coherence diversity*, since its source is the disparity in the coherence conditions. To demonstrate the coherence diversity gains, we propose to use a product superposition that allows the signal of the less-selective receiver to disappear into the equivalent channel seen by the more-selective receiver. This product superposition utilizes all the available degrees of freedom in the more-selective link and in addition transmits simultaneously into the less-selective link without causing interference to the more-selective one. We show that in all the conditions where the two channels have non-identical dynamical characteristics, coherence diversity applies and leads to gains.

Mismatch in coherence conditions in multi-receiver downlink scenarios have been considered so far only in a very narrow set of conditions. Disparity in coherence time under flat-fading scenario was analyzed in [9]–[16]. In particular, in [9] and [10], a downlink channel for two users, one time-static and time-dynamic, was investigated, and a product superposition was used to achieve gains in the transmission rate. In [11] and [12], the degrees of freedom region for an arbitrary number of users with different coherence times were investigated. Product superposition was combined with interference alignment and beamforming in [13]. The disparity of the multi-user fading conditions in frequency-selective channels was referred to in [14]–[16], but these works use orthogonal transmission and utilize disparities only for classifying low-mobility and high-mobility receivers for the purposes of scheduling.

Our results are obtained in the framework of OFDM transmission with pilots [17]–[19]. We address the case of pilots on all sub-carriers as well as pilot interpolation. We also consider different channel estimation techniques including frequency-domain and time-domain estimation [19]–[21]. The effects of the disparity of coherence conditions are studied in several distinct cases: when frequency-selective users have different coherence times but the same coherence bandwidth, when they have different coherence bandwidth but the same coherence time, and finally when they have different coherence bandwidth and different coherence times. For each case, we outline and analyze an appropriate transmission strategy that matches the coherence conditions. In each case we show that proper exploitation of the differences in coherence conditions of the links gives rise to gains. These gains are further demonstrated via numerical simulations.

The organization of this paper is as follows (please also see Table I). Section II describes the system model, Section III studies the disparity in coherence time, Section IV analyzes

Manuscript received February 5, 2018; revised August 9, 2018; accepted September 28, 2018. Date of publication October 30, 2018; date of current version March 15, 2019. This work was supported in part by the grants CIF1527598 and CIF1718551 from the National Science Foundation. This paper was presented in part at the IEEE Global Telecommunication Conference (GLOBECOM), Washington, DC, USA, December 2016 [1]. The associate editor coordinating the review of this paper and approving it for publication was N. B. Mehta. *(Corresponding author: Aria Nosratinia.)*

The authors are with the Department of Electrical Engineering, The University of Texas at Dallas, Richardson, TX 75083-0688 USA (e-mail: mohamed.fadel@utdallas.edu; aria@utdallas.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCOMM.2018.2878692

TABLE I
DISPARITY SCENARIOS

Selectivity	None	Time	Frequency	Both
None		[9]	Section IV	Section III
Time	[9]		Section V	Section IV
Frequency	Section IV	Section V		Section III
Both	Section III	Section IV	Section III	

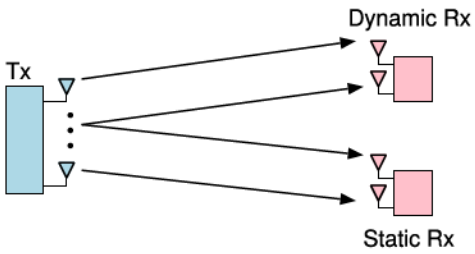

Fig. 1. MIMO downlink channel with coherence diversity, with one link being more dynamic in time or frequency (or both).

disparity in coherence bandwidth, Section V considers disparity in both coherence time and coherence bandwidth, Section VI presents numerical results and discussion, and Section VII concludes the paper.

The following notations are used in this paper: $\mathrm{diag}\{\mathbf{a}\}$ denotes a diagonal matrix whose entries consist of the elements of the vector $\mathbf{a}$. $\mathbf{I}_M$ is the identity matrix, $\mathbf{0}_{N\times M}$ denotes the all-zero matrix, and $\mathbf{1}_N$ denotes the all-one vector, with subscripts indicating dimensions. Furthermore, $\otimes$ denotes the Kronecker product. $\mathbb{C}^{N\times M}$ is the set of complex matrices, $\mathbb{E}\{\cdot\}$ represents expectation, and log is taken to the base 2.

II. System Model

Consider an OFDM downlink transmission with an M-antenna transmitter and two receivers equipped with N_1, N_2 antennas and having coherence times T_1 and T_2, respectively, as depicted in Fig. 1. Define $\mathbf{X}_k(t) \in \mathbb{C}^{M\times T_o}$ to be the transmitted space-time block code at subcarrier k, where $t = 1, 2, \ldots$ is the time index of the space-time code and T_o is the number of channel uses per block. The corresponding received signals at User 1 and User 2 are

$$\mathbf{Y}_{1,k}(t) = \sqrt{\frac{\rho}{M}}\mathbf{H}_{1,k}(t)\mathbf{X}_k(t) + \mathbf{W}_{1,k}(t),$$
$$\mathbf{Y}_{2,k}(t) = \sqrt{\frac{\rho}{M}}\mathbf{H}_{2,k}(t)\mathbf{X}_k(t) + \mathbf{W}_{2,k}(t), \quad k = 1, \ldots, K, \quad (1)$$

where $\mathbf{W}_{1,k}(t) \in \mathbb{C}^{N_1\times T_o}$ and $\mathbf{W}_{2,k}(t) \in \mathbb{C}^{N_2\times T_o}$ are the additive Gaussian noise matrices whose elements are i.i.d. circularly symmetric with zero mean and unit variance, ρ represents the received SNR, and K is the number of subcarriers.

The matrices $\mathbf{H}_{1,k}(t) \in \mathbb{C}^{N_1\times M}$ and $\mathbf{H}_{2,k}(t) \in \mathbb{C}^{N_2\times M}$ represent the MIMO channel frequency responses at subcarrier k of the two receivers. Assume $T_o = \min\{T_1, T_2\}$; hence the elements of $\mathbf{H}_{1,k}(t)$ and $\mathbf{H}_{2,k}(t)$ remain invariant during the period T_o. In the remainder of the paper, the index t is omitted for economy of notation. Define $\mathbf{h}_{1,ij} \in \mathbb{C}^{L_1}$ and $\mathbf{h}_{2,ij} \in \mathbb{C}^{L_2}$ to be the impulse response vectors between transmit antenna j and receive antenna i of User 1 and User 2, having L_1 and L_2 taps respectively. The elements of $\mathbf{h}_{1,ij}$ and $\mathbf{h}_{2,ij}$ are assumed to be i.i.d. zero-mean unit-variance complex-Gaussian, modeling a rich scattering environment. The MIMO channel impulse response matrices $\bar{\mathbf{H}}_1 \in \mathbb{C}^{N_1L_1\times M}$ and $\bar{\mathbf{H}}_2 \in \mathbb{C}^{N_2L_2\times M}$ are defined as

$$\bar{\mathbf{H}}_1 = \begin{bmatrix} \mathbf{h}_{1,11} & \cdots & \mathbf{h}_{1,1M} \\ \vdots & \ddots & \vdots \\ \mathbf{h}_{1,N_11} & \cdots & \mathbf{h}_{1,N_1M} \end{bmatrix},$$
$$\bar{\mathbf{H}}_2 = \begin{bmatrix} \mathbf{h}_{2,11} & \cdots & \mathbf{h}_{2,1M} \\ \vdots & \ddots & \vdots \\ \mathbf{h}_{2,N_21} & \cdots & \mathbf{h}_{2,N_2M} \end{bmatrix}. \quad (2)$$

Define $\bar{\mathbf{F}} \in \mathbb{C}^{K\times K}$ to be the standard DFT matrix, with $\frac{1}{\sqrt{K}}\exp^{-i\frac{2\pi kn}{K}}$ being the entry at row k and column n. The standard DFT matrix, $\bar{\mathbf{F}}$, relates the channel frequency response and the channel impulse response for one antenna for all subcarriers. To facilitate analysis, we further define $\mathbf{F}_{1,k} \in \mathbb{C}^{N_1\times N_1L_1}$ and $\mathbf{F}_{2,k} \in \mathbb{C}^{N_2\times N_2L_2}$ as matrices that relate the channel frequency and impulse response for all the antennas at one subcarrier, i.e.,

$$\mathbf{F}_{1,k} \triangleq \bar{\mathbf{F}}_{k,L_1} \otimes \mathbf{I}_{N_1},$$
$$\mathbf{F}_{2,k} \triangleq \bar{\mathbf{F}}_{k,L_2} \otimes \mathbf{I}_{N_2}, \quad (3)$$

where $\bar{\mathbf{F}}_{k,L_1} \in \mathbb{C}^{1\times L_1}$ denotes the vector consisting of the first L_1 elements of row k of $\bar{\mathbf{F}}$. Similarly, $\bar{\mathbf{F}}_{k,L_2} \in \mathbb{C}^{1\times L_2}$ denotes the vector consisting of the first L_2 elements of row k of $\bar{\mathbf{F}}$. Therefore, the MIMO channel frequency response matrices, $\mathbf{H}_{1,k}$ and $\mathbf{H}_{2,k}$ in (1) are related as follows to the MIMO channel impulse response matrices $\bar{\mathbf{H}}_1$ and $\bar{\mathbf{H}}_2$, defined in (2).

$$\mathbf{H}_{1,k} = \mathbf{F}_{1,k}\bar{\mathbf{H}}_1,$$
$$\mathbf{H}_{2,k} = \mathbf{F}_{2,k}\bar{\mathbf{H}}_2. \quad (4)$$

The elements of $\mathbf{h}_{1,ij}$ and $\mathbf{h}_{2,ij}$ are i.i.d. due to rich scattering. Thus,

$$\mathbb{E}\left\{\mathbf{H}_{1,k}\mathbf{H}_{1,k}^H\right\} = M\mathbf{F}_{1,k}\mathbf{F}_{1,k}^H = M\|\mathbf{f}_1\|^2\mathbf{I}_{N_1},$$
$$\mathbb{E}\left\{\mathbf{H}_{2,k}\mathbf{H}_{2,k}^H\right\} = M\mathbf{F}_{2,k}\mathbf{F}_{2,k}^H = M\|\mathbf{f}_2\|^2\mathbf{I}_{N_2}, \quad (5)$$

where $\|\mathbf{f}_1\|^2 = \|\bar{\mathbf{F}}_{k,L_1}\|^2$ and $\|\mathbf{f}_2\|^2 = \|\bar{\mathbf{F}}_{k,L_2}\|^2$. This paper employs a system model without CSIT (Channel State Information at Transmitter), i.e., the channel matrices $\mathbf{H}_{1,k}$, $\mathbf{H}_{2,k}$, $\bar{\mathbf{H}}_{1,k}$, and $\bar{\mathbf{H}}_{2,k}$ are unknown at the transmitter. The transmitter only knows the channel statistics including the number of taps L_1, L_2 and the coherence times T_1, T_2 (see [22] and the references therein).

III. Disparity in Coherence Time

In this section, we investigate downlink channels that have disparity in coherence time, where $T_1 \ll T_2$, meaning that User 1 is *time-dynamic* and User 2 is *time-static*. We model this scenario by assuming $\mathbf{H}_{2,k}$ to be known at User 2 whereas

$\mathbf{H}_{1,k}$ is assumed to be unknown at both users. The links have the same coherence bandwidth since they have the same delay spread ($L_1 = L_2 = L$). The conditions analyzed in this section can occur in practical multiuser networks as shown in the following example.

Example 1: Consider an urban environment where a macro-cell may serve two users: one time-dynamic user that is in a moving car and another time-static user that is sitting at a desk. Based on the WINNER radio channel models [23], the propagation of the time-dynamic user is *C2: typical urban macro-cell* with 120 Km/h mobility and 234 ns r.m.s. delay spread. The propagation of the time-static user is *C4: outdoor to indoor macro-cell* with 5 Km/h mobility and 240 ns r.m.s. delay spread. Hence, the coherence time of the time-dynamic and the time-static user are 2.25 ms and 54 ms, respectively. The coherence bandwidth of both links is approximately 85 KHz. We use the definition of coherence time $T = \frac{1}{2\ f_d}$, where f_d is the Doppler spread at frequency 2 GHz and that of the coherence bandwidth $B_{c,90\%} = \frac{1}{50\sigma_\tau}$, where σ_τ is the delay spread [8].

In the following, Section III-A revisits the conventional OFDM pilot-based transmission. Section III-B investigates frequency-domain channel estimation when pilots are available for all subcarriers. Section III-C studies frequency domain channel estimation via pilot interpolation. Section III-D investigates time-domain channel estimation, i.e., when the channel impulse response $\tilde{\mathbf{H}}_1$ defined in (2) is estimated.

A. Conventional OFDM Downlink Transmission

In conventional OFDM downlink transmission, TDMA is used to serve multiple users in the absence of CSIT [24]. In the following, we calculate the channel estimate in the frequency domain, estimation error, and achievable rate for User 1. The transmitted space-time code of length T_1 at subcarrier k is

$$\mathbf{X}_k = \left[\sqrt{M}\mathbf{I}_M,\ \mathbf{U}_{1,k}\right], \quad k = 1, \cdots, K, \tag{6}$$

where pilot signals are sent during the first M time instances to estimate $\mathbf{H}_{1,k}$, and $\mathbf{U}_{1,k} \in \mathbb{C}^{M\times(T_1-M)}$ contains the symbols intended for User 1, and the symbols are i.i.d. Gaussian. The corresponding received signal is

$$\mathbf{Y}_{1,k} = \left[\sqrt{\rho_\tau}\mathbf{H}_{1,k}, \sqrt{\frac{\rho_\delta}{M}}\mathbf{H}_{1,k}\mathbf{U}_{1,k}\right] + \mathbf{W}_{1,k}, \tag{7}$$

where ρ_τ and ρ_δ are the received SNR during pilot and symbol transmission, respectively. The MMSE estimate of $\mathbf{H}_{1,k}$ is

$$\hat{\mathbf{H}}_{1,k} = \mathbf{\Sigma}_{HY}\mathbf{\Sigma}_Y^{-1}\left(\sqrt{\rho_\tau}\mathbf{H}_{1,k} + \mathbf{W}_{1,\tau,k}\right), \tag{8}$$

where $\mathbf{W}_{\tau,k} \in \mathbb{C}^{N_1\times M}$ is the additive noise during the first M time instances and

$$\begin{aligned}\mathbf{\Sigma}_{HY} &= M\sqrt{\rho_\tau}\|\mathbf{f}_1\|^2\mathbf{I}_{N_1},\\ \mathbf{\Sigma}_Y &= \left(M\rho_\tau\|\mathbf{f}_1\|^2 + M\right)\mathbf{I}_{N_1}.\end{aligned} \tag{9}$$

Hence,

$$\hat{\mathbf{H}}_{1,k} = \gamma\mathbf{H}_{1,k} + \frac{\gamma}{\sqrt{\rho_\tau}}\mathbf{W}_{1,\tau,k}, \tag{10}$$

where

$$\gamma = \frac{\rho_\tau\|\mathbf{f}_1\|^2}{\rho_\tau\|\mathbf{f}_1\|^2 + 1}. \tag{11}$$

Define $\hat{\mathbf{\Sigma}}_k$ to be,

$$\hat{\mathbf{\Sigma}}_k = \mathbb{E}\left\{\hat{\mathbf{H}}_k\hat{\mathbf{H}}_k^H\right\} = M\left(\gamma^2\|\mathbf{f}_1\|^2 + \frac{\gamma^2}{\rho_\tau}\right)\mathbf{I}_{N_1}, \tag{12}$$

and

$$\hat{\sigma}_k^2 = \frac{1}{N_1 M}\mathrm{Tr}\left\{\hat{\mathbf{\Sigma}}_k\right\} = \gamma^2\|\mathbf{f}_1\|^2 + \frac{\gamma^2}{\rho_\tau} = \gamma\|\mathbf{f}_1\|^2. \tag{13}$$

The estimation error is

$$\begin{aligned}\tilde{\mathbf{H}}_{1,k} &= \mathbf{H}_{1,k} - \hat{\mathbf{H}}_{1,k}\\ &= (1-\gamma)\,\mathbf{H}_{1,k} - \frac{\gamma}{\sqrt{\rho_\tau}}\mathbf{W}_{1,\tau,k},\end{aligned} \tag{14}$$

and hence the estimation error covariance matrix is

$$\begin{aligned}\tilde{\mathbf{\Sigma}}_k &= \mathbb{E}\left\{\tilde{\mathbf{H}}_{1,k}\tilde{\mathbf{H}}_{1,k}^H\right\}\\ &= M\left(\|\mathbf{f}_1\|^2(1-\gamma)^2 + \frac{\gamma^2}{\rho_\tau}\right)\mathbf{I}_{N_1}\\ &= \frac{M\gamma}{\rho_\tau}\mathbf{I}_{N_1}.\end{aligned} \tag{15}$$

The normalized MMSE is

$$\tilde{\sigma}_k^2 = \frac{1}{MN_1}\mathrm{Tr}\left\{\tilde{\mathbf{\Sigma}}_k\right\} = \frac{\gamma}{\rho_\tau}. \tag{16}$$

After estimating $\mathbf{H}_{1,k}$, User 1 decodes $\mathbf{U}_{1,k}$ coherently during $(T_1 - M)$ time instances. The received signal during this time is

$$\mathbf{Y}_{1,\delta,k} = \sqrt{\frac{\rho_\delta}{M}}\mathbf{H}_{1,k}\mathbf{U}_{1,k} + \mathbf{W}_{1,\delta,k}, \tag{17}$$

where $\mathbf{W}_{1,\delta,k} \in \mathbb{C}^{N_1\times(T_1-M)}$ is the corresponding noise. From (14),

$$\mathbf{Y}_{1,\delta,k} = \sqrt{\frac{\rho_\delta}{M}}\hat{\mathbf{H}}_{1,k}\mathbf{U}_{1,k} + \sqrt{\frac{\rho_\delta}{M}}\tilde{\mathbf{H}}_{1,k}\mathbf{U}_{1,k} + \mathbf{W}_{1,\delta,k}. \tag{18}$$

Based on [25], the achievable rate at subcarrier k is

$$\begin{aligned}R_{1,k} &= I\left(\mathbf{U}_{1,k}; \mathbf{Y}_{1,\delta,k}|\hat{\mathbf{H}}_{1,k}\right)\\ &\geq \left(1-\frac{M}{T_1}\right)\mathbb{E}\left\{\log\det\left\{\mathbf{I}_{N_1} + \frac{\rho_\delta}{1+\rho_\delta\tilde{\sigma}_k^2}\frac{\hat{\mathbf{H}}_{1,k}\hat{\mathbf{H}}_{1,k}^H}{M}\right\}\right\},\end{aligned} \tag{19}$$

where $\tilde{\sigma}_k^2$ is given in (16). Define the normalized estimated channel

$$\check{\mathbf{H}}_{1,k} = \frac{\hat{\mathbf{H}}_{1,k}}{\hat{\sigma}_k}, \tag{20}$$

where $\hat{\sigma}_k$ is given in (13). Thus, the achievable rate in T_1 time slots and K subcarriers is

$$R_1 \geq \frac{1}{K}\left(1-\frac{M}{T_1}\right)\sum_{k=1}^{K}\mathbb{E}\left\{\log\det\left\{\mathbf{I}_{N_1} + \rho_k\frac{\check{\mathbf{H}}_{1,k}\check{\mathbf{H}}_{1,k}^H}{M}\right\}\right\}, \tag{21}$$

where

$$\rho_k = \frac{\rho_\delta \hat{\sigma}_k^2}{1 + \rho_\delta \tilde{\sigma_k}^2}. \tag{22}$$

According to Appendix A, the above rate is maximized when

$$\rho_\delta = (1-\alpha)\rho \frac{T_1}{T_{1,\delta}},$$
$$\rho_\tau = \alpha \rho \frac{T_1}{T_{1,\tau}}, \tag{23}$$

where $T_{1,\delta} = T_1 - M$, $T_{1,\tau} = M$, and

$$\alpha = -\beta + \sqrt{\beta(\beta+1)},$$
$$\beta = \frac{T_{1,\delta} T_{1,\tau} + \rho T_1 T_{1,\tau} \|\mathbf{f}_1\|^2}{T_{1,\delta} \rho T_1 \|\mathbf{f}_1\|^2 \left(1 - \frac{T_{1,\tau}}{T_{1,\delta}}\right)}. \tag{24}$$

B. Product Superposition OFDM Transmission

Product superposition was first introduced in the context of flat-fading [9]. In the following, a product superposition is introduced for frequency-selective fading channels. We investigate the case of frequency-domain channel estimation. From (6), the transmitted space-time code is

$$\mathbf{X}_k = \left[\sqrt{M}\mathbf{U}_{2,k}, \mathbf{U}_{21,k}\right], \tag{25}$$

where $\mathbf{U}_{21,k} = \mathbf{U}_{2,k}\mathbf{U}_{1,k}$, and $\mathbf{U}_{2,k} \in \mathbb{C}^{M\times M}$ contains symbol intended for User 2, where the symbols are i.i.d. Gaussian and $\mathbb{E}\left\{\mathbf{U}_{2,k}\mathbf{U}_{2,k}^H\right\} = \mathbf{I}_M$.[1] Hence, the received signal at User 1 at subcarrier k during the first M time instances is

$$\mathbf{Y}_{1,\tau,k} = \sqrt{\rho_\tau}\mathbf{G}_{1,k} + \mathbf{W}_{1,\tau,k}, \tag{26}$$

where $\mathbf{G}_{1,k} = \mathbf{H}_{1,k}\mathbf{U}_{2,k}$ is the equivalent channel seen by User 1 that can be estimated during the first M time instances. Following the analysis of Section III-A, The MMSE estimate of the equivalent channel is

$$\hat{\mathbf{G}}_{1,k} = \gamma \mathbf{G}_{1,k} + \frac{\gamma}{\sqrt{\rho_\tau}}\mathbf{W}_{1,\tau,k}, \tag{27}$$

where γ, $\hat{\mathbf{\Sigma}}_k, \hat{\sigma}_k^2, \tilde{\mathbf{\Sigma}}_k, \tilde{\sigma}_k^2$ are given in (11), (12), (13), (15), and (16), respectively.

During the remaining $(T_1 - M)$ time instances, $\mathbf{U}_{21,k}$ is sent, and hence the received signal is

$$\mathbf{Y}_{1,\delta,k} = \sqrt{\frac{\rho_\delta}{M}}\mathbf{G}_{1,k}\mathbf{U}_{1,k} + \mathbf{W}_{1,\delta,k}$$
$$= \sqrt{\frac{\rho_\delta}{M}}\hat{\mathbf{G}}_{1,k}\mathbf{U}_{1,k} + \sqrt{\frac{\rho_\delta}{M}}\tilde{\mathbf{G}}_{1,k}\mathbf{U}_{1,k} + \mathbf{W}_{1,\delta,k}. \tag{28}$$

Therefore, for K subcarriers and T_1 time instances the achievable rate is R_1 as given in (21). User 2, knowing its channel, achieves the rate [9]

$$R_2 \geq \frac{M}{T_1 K}\sum_{k=1}^{K}\mathbb{E}\left\{\log\det\left\{\mathbf{I}_{N_2} + \frac{1}{\mathbb{E}\left\{\lambda_i^{-2}\right\}}\mathbf{H}_{2,k}\mathbf{H}_{2,k}^H\right\}\right\}, \tag{29}$$

[1]Recall that in Section III, no pilots are transmitted within the block for User 2 since $T_1 \ll T_2$.

where λ_i^{-2} is an eigenvalue of $\tilde{\mathbf{U}}_{1,k}\tilde{\mathbf{U}}_{1,k}^H$ and $\tilde{\mathbf{U}}_{1,k} = [\rho_\tau \mathbf{I},\ \rho_\delta \mathbf{U}_{1,k}]$. For both conventional transmission and product superposition, User 1 achieves the same rate in (21). The rate of User 2 in (29) demonstrates the coherence diversity gain that is achieved by using product superposition and the utilization of the differences in the fading variations in the two links.

C. Channel Estimation With Interpolation

In some systems, pilots are sent at a subset of subcarriers and other subcarrier channels are estimated via interpolation, utilizing the correlation of the channel across frequency [8], [18], [26]–[33]. Channel estimation with interpolation has been investigated in the literature via nearest neighbor, linear, and low-pass interpolation [27], [28]. In particular, [27] considered OFDM systems and [28] considered space-time block coding with spatial modulation. It was shown that linear interpolation has low complexity and low-pass interpolation has the best performance.

Two-dimensional interpolation over frequency and time has been investigated in [18] and [29]–[32]. A two-dimensional Wiener filter interpolation was shown to be the optimal in terms of mean-squared error [29], [34]. An adaptive channel estimator based on raised cosine frequency-domain interpolation along with time-domain windowing was proposed in [33].

In this section, we introduce pilot-interpolated channel estimation for the proposed scheme and analyze its performance. The location of the pilots with respect to time and frequency has been studied in [35] and [36], where it was shown that an equidistant arrangement of pilots achieves MMSE estimation of the channel. Assume $\{i, j\}$ are subcarriers containing pilots and $\mathbf{A}_k \in \mathbb{C}^{N_1\times 2N_1}$ to be the interpolation matrix of $\mathbf{H}_{1,k}$ estimation at subcarriers k. Also, define $\mathbf{U}_{2,i} \in \mathbb{C}^{M\times M}$ and $\mathbf{U}_{2,j} \in \mathbb{C}^{M\times M}$ to be two matrices that contain symbols intended for User 2 and sent at subcarriers i and j, respectively. The transmitted space-time code is

$$\mathbf{X}_k = \begin{cases}\left[\sqrt{M}\mathbf{U}_{2,k},\ \mathbf{U}_{2,k}\mathbf{U}_{1,k}\right], & k = i, j\\ \mathbf{U}_{2,k}\mathbf{U}_{1,k}, & k \neq i, j,\end{cases} \tag{30}$$

where

$$\mathbf{U}_{1,k} \in \begin{cases}\mathbb{C}^{M\times(T_1-M)}, & k = i, j\\ \mathbb{C}^{M\times T_1}, & k \neq i, j,\end{cases}$$
$$\mathbf{U}_{2,k} = \mathbf{B}_k\begin{bmatrix}\mathbf{U}_{2,i}\\ \mathbf{U}_{2,j}\end{bmatrix}, \tag{31}$$

$\mathbf{B}_k \in \mathbb{C}^{M\times 2M}$ is the interpolation matrix for the symbols intended for User 2. The difference between the interpolation matrices $\mathbf{B}_k$ and $\mathbf{A}_k$ is that the former is used by the transmitter to design the transmitted space-time code, whereas the latter is used at User 1 to estimate $\mathbf{H}_{1,k}$. User 1 estimates the equivalent channel $\mathbf{G}_{1,k}$,

$$\hat{\mathbf{G}}_{1,k} = \mathbf{A}_k\begin{bmatrix}\hat{\mathbf{H}}_{1,i}\mathbf{U}_{2,i}\\ \hat{\mathbf{H}}_{1,j}\mathbf{U}_{2,j}\end{bmatrix} = \mathbf{A}_k\begin{bmatrix}\hat{\mathbf{G}}_{1,i}\\ \hat{\mathbf{G}}_{1,j}\end{bmatrix}$$
$$= \gamma\mathbf{A}_k\begin{bmatrix}\mathbf{G}_{1,i}\\ \mathbf{G}_{1,j}\end{bmatrix} + \frac{\gamma}{\sqrt{\rho_\tau}}\mathbf{A}_k\begin{bmatrix}\mathbf{W}_{1,\tau,i}\\ \mathbf{W}_{1,\tau,j}\end{bmatrix}, \tag{32}$$

and hence,

$$\begin{aligned}\hat{\boldsymbol{\Sigma}}_k &= \mathbb{E}\left\{\hat{\mathbf{G}}_{1,k}\hat{\mathbf{G}}_{1,k}^H\right\} \\ &= \gamma^2\mathbf{A}_k\left(\begin{bmatrix}\mathbb{E}\{\mathbf{G}_{1,i}\mathbf{G}_{1,i}^H\} & \mathbf{0} \\ \mathbf{0} & \mathbb{E}\{\mathbf{G}_{1,j}\mathbf{G}_{1,j}^H\}\end{bmatrix} + \frac{\gamma^2}{\rho_\tau}\mathbf{I}\right)\mathbf{A}_k^H \\ &= M\gamma^2\left(\|\mathbf{f}_1\|^2 + \frac{\gamma^2}{\rho_\tau}\right)\mathbf{A}_k\mathbf{A}_k^H, \end{aligned} \tag{33}$$

and furthermore

$$\hat{\sigma}_k^2 = \frac{1}{N_1 M}\mathrm{Tr}\left\{\hat{\boldsymbol{\Sigma}}_k\right\}. \tag{34}$$

The estimation error is

$$\begin{aligned}\tilde{\mathbf{G}}_k &= \mathbf{G}_{1,k} - \hat{\mathbf{G}}_{1,k} \\ &= \mathbf{H}_{1,k}\mathbf{B}_k\begin{bmatrix}\mathbf{U}_{2,i} \\ \mathbf{U}_{2,j}\end{bmatrix} - \gamma\mathbf{A}_k\begin{bmatrix}\mathbf{H}_{1,i}\mathbf{U}_{2,i} \\ \mathbf{H}_{1,j}\mathbf{U}_{2,j}\end{bmatrix} - \frac{\gamma}{\sqrt{\rho_\tau}}\mathbf{A}_k\begin{bmatrix}\mathbf{W}_{1,\tau,i} \\ \mathbf{W}_{1,\tau,j}\end{bmatrix} \\ &= \left(\mathbf{H}_{1,k}\mathbf{B}_k - \gamma\mathbf{A}_k\begin{bmatrix}\mathbf{H}_{1,i} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_{1,j}\end{bmatrix}\right)\begin{bmatrix}\mathbf{U}_{2,i} \\ \mathbf{U}_{2,j}\end{bmatrix} \\ &\quad - \frac{\gamma}{\sqrt{\rho_\tau}}\mathbf{A}_k\begin{bmatrix}\mathbf{W}_{1,\tau,i} \\ \mathbf{W}_{1,\tau,j}\end{bmatrix}. \end{aligned} \tag{35}$$

The error covariance matrix is

$$\begin{aligned}\tilde{\boldsymbol{\Sigma}}_k = \mathbb{E}\left\{\left(\mathbf{H}_{1,k}\mathbf{B}_k - \gamma\boldsymbol{\Gamma}_k\right)\left(\mathbf{H}_{1,k}\mathbf{B}_k - \gamma\boldsymbol{\Gamma}_k\right)^H\right\} \\ + \frac{\gamma^2}{\rho_\tau}M\mathbf{A}_k\mathbf{A}_k^H, \end{aligned} \tag{36}$$

where

$$\boldsymbol{\Gamma}_k = \mathbf{A}_k\begin{bmatrix}\mathbf{H}_{1,i} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_{1,j}\end{bmatrix}, \tag{37}$$

and furthermore

$$\tilde{\sigma}_k^2 = \frac{1}{N_1 M}\mathrm{Tr}\left\{\tilde{\boldsymbol{\Sigma}}_k\right\}. \tag{38}$$

Therefore, the achievable rate is

$$\begin{aligned}R_1 \geq \frac{1}{K}\left(1-\frac{M}{T_1}\right)\sum_{k\in\{i,j\}}\mathbb{E}\left\{\log\det\left\{\mathbf{I}_{N_1} + \rho_k\frac{\check{\mathbf{G}}_{1,k}\check{\mathbf{G}}_{1,k}^H}{M}\right\}\right\} \\ + \frac{1}{K}\sum_{k=1,k\notin\{i,j\}}^{K}\mathbb{E}\left\{\log\det\left\{\mathbf{I}_{N_1} + \rho_k\frac{\check{\mathbf{G}}_{1,k}\check{\mathbf{G}}_{1,k}^H}{M}\right\}\right\}, \end{aligned} \tag{39}$$

where

$$\rho_k = \frac{\rho_\delta\hat{\sigma}_k^2}{1+\rho_\delta\tilde{\sigma}_k^2}, \tag{40}$$

$\check{\mathbf{G}}_{1,k} = \frac{\hat{\mathbf{G}}_{1,k}}{\hat{\sigma}_k}$, $\hat{\mathbf{G}}_{1,k}$ is defined in (32), and $\hat{\sigma}_k^2, \tilde{\sigma}_k^2$ are given in (34) and (38), respectively. On the other hand, User 2 decodes $\mathbf{U}_{2,i}$ and $\mathbf{U}_{2,j}$ sent on subcarriers i and j achieving the rate

$$R_2 \geq \frac{M}{T_1 K}\sum_{k\in\{i,j\}}\mathbb{E}\left\{\log\det\left\{\mathbf{I}_{N_2} + \frac{1}{\mathbb{E}\left\{\lambda_i^{-2}\right\}}\mathbf{H}_{2,k}\mathbf{H}_{2,k}^H\right\}\right\}. \tag{41}$$

Now, we investigate the design of the matrix $\mathbf{B}_k$. The estimation error, $\tilde{\sigma}_k^2$, in (35) depends on $\mathbf{B}_k$. Hence, we can maximize the achievable rate by finding the optimal $\mathbf{B}_k$,

$$\begin{aligned}\mathbf{B}_k = \arg\min\ &\tilde{\sigma}_k^2 \\ \text{subject to } &\mathrm{Tr}\left\{\mathbf{B}_k\mathbf{B}_k^H\right\} \leq M, \end{aligned} \tag{42}$$

where the above constraint is used to maintain the transmission power constraint. Using the identity $\mathrm{Tr}\{\mathbf{XY}\} = \mathrm{Tr}\{\mathbf{YX}\}$,

$$\begin{aligned}\tilde{\sigma}_k^2 = \frac{1}{N_1 M}\mathrm{Tr}\Bigg\{&\mathbf{B}_k^H\boldsymbol{\Psi}_k\mathbf{B}_k - \mathbf{B}_k\boldsymbol{\Phi}_k - \boldsymbol{\Phi}_k^H\mathbf{B}_k^H \\ &+ M\gamma^2\left(\|\mathbf{f}_1\|^2 + \frac{\gamma^2}{\rho_\tau}\right)\mathbf{A}_k\mathbf{A}_k^H\Bigg\}, \end{aligned} \tag{43}$$

where $\boldsymbol{\Phi}_k = \mathbb{E}\left\{\gamma\boldsymbol{\Gamma}_k^H\mathbf{H}_{1,k}\right\}$ and $\boldsymbol{\Psi}_k = \mathbb{E}\left\{\mathbf{H}_{1,k}^H\mathbf{H}_{1,k}\right\}$. Hence,

$$\begin{aligned}\mathbf{B}_k = \arg\min\ &\mathrm{Tr}\left\{\mathbf{B}_k^H\boldsymbol{\Psi}_k\mathbf{B}_k\right\} - \mathrm{Tr}\left\{\mathbf{B}_k\boldsymbol{\Phi}_k\right\} - \mathrm{Tr}\left\{\boldsymbol{\Phi}_k^H\mathbf{B}_k^H\right\} \\ \text{subject to } &\mathrm{Tr}\left\{\mathbf{B}_k\mathbf{B}_k^H\right\} \leq M. \end{aligned} \tag{44}$$

The above optimization problem is convex and therefore is computationally feasible.

D. Time-Domain Channel Estimation

For comparison, we first give the achievable rate in conventional transmission when channel impulse response is estimated.

1) Conventional Transmission With Time-Domain Channel Estimation: Define $\tilde{\mathbf{Y}}_{1,i} \in \mathbb{C}^K$ to be the received signal of User 1 at receive antenna $i = 1, \cdots, N_1$. Hence,

$$\tilde{\mathbf{Y}}_{1,i} = \sqrt{\frac{\rho}{M}}\sum_{j=1}^{M}\mathrm{diag}\left\{\bar{\mathbf{F}}[\mathbf{h}_{1,ij}^H, \mathbf{0}_{1\times(K-L_1)}^H]^H\right\}\tilde{\mathbf{X}}_j + \tilde{\mathbf{W}}_{1,i}, \tag{45}$$

where $\tilde{\mathbf{X}}_j \in \mathbb{C}^K$ is the transmitted signal at antenna j and $\tilde{\mathbf{W}}_{1,i} \in \mathbb{C}^K$ is the additive Gaussian noise at receive antenna i. Therefore,

$$\tilde{\mathbf{Y}}_{1,i} = \sqrt{\frac{\rho}{M}}\sum_{j=1}^{M}\mathrm{diag}\left\{\tilde{\mathbf{X}}_j\right\}\bar{\mathbf{F}}[\mathbf{h}_{1,ij}^H, \mathbf{0}_{1\times(K-L_1)}^H]^H + \tilde{\mathbf{W}}_{1,i}. \tag{46}$$

For estimating $\mathbf{h}_{1,ij}$, at least L pilots are needed for each transmit antenna j. Assume that LM pilots are sent at the first L subcarriers where M pilots are sent during the first M time instances of each subcarrier. Furthermore,

$$\tilde{\mathbf{X}}_{\tau,r} = \begin{cases}\sqrt{M}\mathbf{1}_L, & r = j, \\ \mathbf{0}_{L\times 1}, & r \neq j, \end{cases} \tag{47}$$

is the transmitted pilot signal to estimate $\mathbf{h}_{1,ij}$, and hence the received signal at antenna i during pilot transmission is

$$\tilde{\mathbf{Y}}_{1,\tau,i} = \sqrt{\rho_\tau}\bar{\mathbf{F}}\mathbf{h}_{1,ij} + \tilde{\mathbf{W}}_{1,\tau,i}, \tag{48}$$

where $\mathbf{F}_L \in \mathbb{C}^{L\times L}$ is the submatrix consists of the first L columns and the first L rows of $\bar{\mathbf{F}}$, and $\tilde{\mathbf{W}}_{1,\tau,i} \in \mathbb{C}^{L\times 1}$ is the corresponding additive noise.

The MMSE channel estimate is

$$\hat{\mathbf{h}}_{1,ij} = \sqrt{\rho_\tau}\mathbf{F}_L^H \left(\rho_\tau \mathbf{F}_L \mathbf{F}_L^H + \mathbf{I}\right)^{-1} \tilde{\mathbf{Y}}_{1,\tau,i}, \tag{49}$$

and hence

$$\begin{aligned}\hat{\mathbf{\Sigma}}_h &= \mathbb{E}\left\{\hat{\mathbf{h}}_{1,ij}\hat{\mathbf{h}}_{1,ij}^H\right\} \\ &= \rho_\tau^2 \mathbf{F}_L^H \left(\rho_\tau \mathbf{F}_L \mathbf{F}_L^H + \mathbf{I}\right)^{-1} \mathbf{F}_L \mathbf{F}_L^H \left(\rho_\tau \mathbf{F}_L \mathbf{F}_L^H + \mathbf{I}\right)^{-1} \mathbf{F}_L \\ &\quad + \rho_\tau \mathbf{F}_L^H \left(\rho_\tau \mathbf{F}_L \mathbf{F}_L^H + \mathbf{I}\right)^{-2} \mathbf{F}_L. \end{aligned} \tag{50}$$

The estimation error is $\tilde{\mathbf{h}}_{1,ij} = \mathbf{h}_{1,ij} - \hat{\mathbf{h}}_{1,ij}$, and hence

$$\begin{aligned}\tilde{\mathbf{\Sigma}}_h &= \mathbb{E}\left\{\tilde{\mathbf{h}}_{1,ij}\tilde{\mathbf{h}}_{1,ij}^H\right\} \\ &= \hat{\mathbf{\Sigma}}_h + \mathbf{I} - 2\rho_\tau \mathbf{F}_L^H \left(\rho_\tau \mathbf{F}_L \mathbf{F}_L^H + \mathbf{I}\right)^{-1} \mathbf{F}_L. \end{aligned} \tag{51}$$

Based on (4), the estimated channel, and the estimation error are

$$\hat{\mathbf{H}}_{1,k} = \mathbf{F}_{1,k}\hat{\bar{\mathbf{H}}}_{1,k}, \quad \tilde{\mathbf{H}}_{1,k} = \mathbf{F}_{1,k}\tilde{\bar{\mathbf{H}}}_{1,k}, \tag{52}$$

where

$$\begin{aligned}\hat{\bar{\mathbf{H}}}_{1,k} &= \begin{bmatrix} \hat{\mathbf{h}}_{1,11} & \cdots & \hat{\mathbf{h}}_{1,1M} \\ \vdots & \ddots & \vdots \\ \hat{\mathbf{h}}_{1,N_1 1} & \cdots & \hat{\mathbf{h}}_{1,N_1 M} \end{bmatrix}, \\ \tilde{\bar{\mathbf{H}}}_{1,k} &= \begin{bmatrix} \tilde{\mathbf{h}}_{1,11} & \cdots & \tilde{\mathbf{h}}_{1,1M} \\ \vdots & \ddots & \vdots \\ \tilde{\mathbf{h}}_{1,N_1 1} & \cdots & \tilde{\mathbf{h}}_{1,N_1 M} \end{bmatrix}. \end{aligned} \tag{53}$$

Furthermore,

$$\begin{aligned}\hat{\sigma}_k^2 &= \frac{1}{N_1 M}\mathrm{Tr}\left\{\mathbb{E}\left\{\mathbf{F}_{1,k}\hat{\bar{\mathbf{H}}}_{1,k}\hat{\bar{\mathbf{H}}}_{1,k}^H\mathbf{F}_{1,k}^H\right\}\right\} \\ &= \frac{1}{N_1}\mathrm{Tr}\left\{\mathbf{F}_{1,k}\left(\mathbf{I}_{N_1}\otimes\hat{\mathbf{\Sigma}}_{\mathcal{H}}\right)\mathbf{F}_{1,k}^H\right\}, \\ \tilde{\sigma}_k^2 &= \frac{1}{N_1 M}\mathrm{Tr}\left\{\mathbb{E}\left\{\mathbf{F}_{1,k}\tilde{\bar{\mathbf{H}}}_{1,k}\tilde{\bar{\mathbf{H}}}_{1,k}^H\mathbf{F}_{1,k}^H\right\}\right\} \\ &= \frac{1}{N_1}\mathrm{Tr}\left\{\mathbf{F}_{1,k}\left(\mathbf{I}_{N_1}\otimes\tilde{\mathbf{\Sigma}}_{\mathcal{H}}\right)\mathbf{F}_{1,k}^H\right\}. \end{aligned} \tag{54}$$

Thus, the achievable rate of the dynamic receiver is

$$\begin{aligned}R_1 \geq \frac{1}{K}\left(1-\frac{M}{T_1}\right)\sum_{k=1}^{L}\mathbb{E}\left\{\log\det\left\{\mathbf{I}_{N_1}+\rho_k\frac{\check{\mathbf{H}}_{1,k}\check{\mathbf{H}}_{1,k}^H}{M}\right\}\right\} \\ + \frac{1}{K}\sum_{k=L+1}^{K}\mathbb{E}\left\{\log\det\left\{\mathbf{I}_{N_1}+\rho_k\frac{\check{\mathbf{H}}_{1,k}\check{\mathbf{H}}_{1,k}^H}{M}\right\}\right\}, \end{aligned} \tag{55}$$

where $\check{\mathbf{H}}_{1,k} = \frac{\hat{\mathbf{H}}_{1,k}}{\hat{\sigma}_k}$,

$$\rho_k = \frac{\rho_\delta \hat{\sigma}_k^2}{1+\rho_\delta \tilde{\sigma_k}^2}, \tag{56}$$

and $\hat{\sigma}_k^2, \tilde{\sigma_k}^2$ are given in (54).

2) Product Superposition With Time-Domain Channel Estimation: In product superposition, the transmitted pilot signal is

$$\tilde{\mathbf{X}}_{\tau,r} = \begin{cases} \sqrt{M}u_{2,j}\mathbf{1}_L, & r = j, \\ \mathbf{0}_{1\times L}, & r \neq j, \end{cases} \tag{57}$$

where $u_{2,j} \in \mathbb{C}$ is a symbol intended for User 2 that is sent by transmit antenna j. Hence, the received signal at antenna i of User 1 during pilot transmission is given by

$$\begin{aligned}\tilde{\mathbf{Y}}_{1,\tau,i} &= \sqrt{\rho_\tau}\mathbf{F}_L\mathbf{h}_{1,ij}u_{2,j} + \tilde{\mathbf{W}}_{1,\tau,i} \\ &= \sqrt{\rho_\tau}\mathbf{F}_L\mathbf{g}_{1,ij} + \tilde{\mathbf{W}}_{1,\tau,i} \end{aligned} \tag{58}$$

where $\mathbf{g}_{1,ij} = \mathbf{h}_{1,ij}u_{2,j}$ is the equivalent channel impulse response seen by User 1. The MMSE estimation of the equivalent channel follows the same steps as in Section III-D.1. In particular, the estimated channel is

$$\hat{\mathbf{g}}_{1,ij} = \sqrt{\rho_\tau}\mathbf{F}_L^H \left(\rho_\tau \mathbf{F}_L \mathbf{F}_L^H + \mathbf{I}\right)^{-1} \tilde{\mathbf{Y}}_{1,\tau,i}, \tag{59}$$

and the estimation error is $\tilde{\mathbf{g}}_{1,ij} = \mathbf{g}_{1,ij} - \hat{\mathbf{g}}_{1,ij}$. Thus, the achievable rate of User 1 is

$$\begin{aligned}R_1 \geq \frac{1}{K}\left(1-\frac{M}{T_1}\right)\sum_{k=1}^{L}\mathbb{E}\left\{\log\det\left\{\mathbf{I}_{N_1}+\rho_k\frac{\check{\mathbf{G}}_{1,k}\check{\mathbf{G}}_{1,k}^H}{M}\right\}\right\} \\ + \frac{1}{K}\sum_{k=L+1}^{K}\mathbb{E}\left\{\log\det\left\{\mathbf{I}_{N_1}+\rho_k\frac{\check{\mathbf{G}}_{1,k}\check{\mathbf{G}}_{1,k}^H}{M}\right\}\right\}. \end{aligned} \tag{60}$$

The received signal of User 2 at the pilot subcarriers, i.e., $k = 1, \cdots, L$, is

$$\mathbf{Y}_{2,k} = \sqrt{\rho_\tau}\mathbf{H}_{2,k}\mathbf{u}_2 + \mathbf{W}_{2,k}, \tag{61}$$

where $\mathbf{u}_2 = [u_{2,1}, \cdots, u_{2,M}]^H$ is the vector containing the symbols intended for User 2. Define the received signal at the L subcarriers containing pilots to be

$$\bar{\bar{\mathbf{Y}}}_2 = \sqrt{\rho_\tau}\bar{\bar{\mathbf{H}}}_2\mathbf{u}_2 + \bar{\bar{\mathbf{W}}}_2, \tag{62}$$

where $\bar{\bar{\mathbf{Y}}}_2 = \left[\mathbf{Y}_{2,1}^H, \cdots, \mathbf{Y}_{2,L}^H\right]^H$, $\bar{\bar{\mathbf{H}}}_2 = \left[\mathbf{H}_{2,1}^H, \cdots, \mathbf{H}_{2,L}^H\right]^H$, and $\bar{\bar{\mathbf{W}}}_2 = \left[\mathbf{W}_{2,1}^H, \cdots, \mathbf{W}_{2,L}^H\right]^H$. Thus the achievable rate of User 2 is

$$R_2 \geq \frac{1}{T_1 K}\mathbb{E}\left\{\log\det\left\{\mathbf{I}_{N_2}+\mathbb{E}_{\left\{\lambda_i^{-2}\right\}}\bar{\bar{\mathbf{H}}}_2\bar{\bar{\mathbf{H}}}_2^H\right\}\right\}. \tag{63}$$

The sum-rate of the conventional transmission is R_1 given in (55), whereas the sum-rate of product superposition is $R_1 + R_2$. Thus, (63) demonstrates the gain of the proposed product superposition.

The analysis in Section III-C and Section III-D can be connected as follows. Since $\mathbf{G}_{1,k} = \mathbf{H}_{1,k}\mathbf{U}_{2,k}$, (32) can be written as,

$$\hat{\mathbf{G}}_{1,k} = \gamma\mathbf{A}_k\begin{bmatrix}\hat{\mathbf{H}}_{1,i}\mathbf{U}_{2,i} \\ \hat{\mathbf{H}}_{1,j}\mathbf{U}_{2,j}\end{bmatrix} + \frac{\gamma}{\sqrt{\rho_\tau}}\mathbf{A}_k\begin{bmatrix}\mathbf{W}_{1,\tau,i} \\ \mathbf{W}_{1,\tau,j}\end{bmatrix}. \tag{64}$$

From (4),

$$\hat{\mathbf{G}}_{1,k} = \gamma\mathbf{A}_k\begin{bmatrix}\mathbf{F}_{1,i}\bar{\mathbf{H}}_1\mathbf{U}_{2,i} \\ \mathbf{F}_{1,j}\bar{\mathbf{H}}_1\mathbf{U}_{2,j}\end{bmatrix} + \frac{\gamma}{\sqrt{\rho_\tau}}\mathbf{A}_k\begin{bmatrix}\mathbf{W}_{1,\tau,i} \\ \mathbf{W}_{1,\tau,j}\end{bmatrix}. \tag{65}$$

By setting $\mathbf{U}_{2,i} = \mathbf{U}_{2,j} = \mathbf{U}_2$, and $\mathbf{B}_k$, defined in (31),

$$\mathbf{B}_k = [\mathbf{I}, \mathbf{0}], \tag{66}$$

and (65) becomes

$$\hat{\mathbf{G}}_{1,k} = \gamma \mathbf{A}_k \begin{bmatrix} \mathbf{F}_{1,i} \\ \mathbf{F}_{1,j} \end{bmatrix} \bar{\mathbf{H}}_1 \mathbf{U}_2 + \frac{\gamma}{\sqrt{\rho_\tau}} \mathbf{A}_k \begin{bmatrix} \mathbf{W}_{1,\tau,i} \\ \mathbf{W}_{1,\tau,j} \end{bmatrix}. \tag{67}$$

If $\mathbf{A}_k$ is selected to be the IFFT interpolation matrix, we can obtain (59) by using (2), $\mathbf{g}_{1,ij} = h_{1,ij} u_{2,j}$, and

$$\hat{\mathbf{G}}_{1,k} = \mathbf{F}_{1,k} \begin{bmatrix} \hat{\mathbf{g}}_{1,11} & \cdots & \hat{\mathbf{g}}_{1,1M} \\ \vdots & \ddots & \vdots \\ \hat{\mathbf{g}}_{1,N_1 1} & \cdots & \hat{\mathbf{g}}_{1,N_1 M} \end{bmatrix}. \tag{68}$$

IV. Disparity in Coherence Bandwidth

In this section, we study downlink transmission under disparity in coherence bandwidth, where User 1 is frequency-static and User 2 is frequency-dynamic, i.e., $L_2 \geq L_1$. The coherence time of both users is the same, i.e., $T_1 = T_2 = T$. The coherence time T could be either short, where both channels are time-dynamic, or long, where both channels are time-static. In this section, we consider channel estimation in the frequency domain without interpolation. Disparity in coherence bandwidth can occur in practical systems as shown in the following example.

Example 2: Consider an urban environment, where a micro-cell is serving two users: a frequency-static user with *B1: typical urban micro-cell* with 76 ns r.m.s. delay spread, and a frequency-dynamic user with *B2: bad urban micro-cell* with 480 ns r.m.s. delay spread [8], [23]. Therefore, the coherence bandwidth of the two users are 263 KHz and 41 KHz. Both users have 70 Km/h mobility leading to a coherence time of 3.85 ms.

For simplicity, we consider $\mathbf{H}_{1,k} = \mathbf{H}_1, \forall k$ meaning that the channel is flat-fading, whereas $\mathbf{H}_{2,k}$ is frequency-selecive. For product superposition, every T time instances, $\mathbf{H}_1$ can be estimated at subcarrier i whereas $\mathbf{H}_{2,k}$ needs channel estimation at all subcarriers. Therefore, at subcarrier i, pilot signals are sent for estimating $\mathbf{H}_{2,i}$ and $\mathbf{H}_1$. At the other subcarriers $k \neq i$, product superposition can be used to provide gain. Hence, at subcarriers $k \neq i$, the transmitted space-time code is

$$\mathbf{X}_k = \left[\sqrt{M}\mathbf{U}_{1,k}, \mathbf{U}_{12,k}\right], \tag{69}$$

where $\mathbf{U}_{12,k} = \mathbf{U}_{1,k}\mathbf{U}_{2,k}$. Hence, User 2 received signal at the first M time instances is

$$\begin{aligned} \mathbf{Y}_{2,\tau,k} &= \sqrt{\rho_\tau}\mathbf{H}_{2,k}\mathbf{U}_{1,k} + \mathbf{W}_{2,\tau,k}, \\ &= \sqrt{\rho_\tau}\mathbf{G}_{2,k} + \mathbf{W}_{2,\tau,k} \end{aligned} \tag{70}$$

where $\mathbf{G}_{2,k} = \mathbf{H}_{2,k}\mathbf{U}_{1,k}$ is the equivalent channel seen by User 2. Following the same analysis as in Section III, the achievable rate of User 2 over K subcarriers and during T time slots is

$$R_2 \geq \frac{1}{K}\left(1 - \frac{M}{T}\right) \sum_{k=1}^{K} \mathbb{E}\left\{ \log\det\left\{ \mathbf{I}_{N_2} + \frac{\rho_\delta \hat{\sigma}_G^2}{1 + \rho_\delta \tilde{\sigma}_G^2} \frac{\bar{\mathbf{G}}_{2,k}\bar{\mathbf{G}}_{2,k}^H}{M} \right\} \right\}, \tag{71}$$

where $\bar{\mathbf{G}}_{2,k} = \frac{\hat{\mathbf{G}}_{2,k}}{\hat{\sigma}_G}$. User 1 achievable rate is

$$R_1 \geq \frac{M}{TK} \sum_{k=1, k\neq i}^{K} \mathbb{E}\left\{ \log\det\left\{ \mathbf{I}_{N_1} + \rho_k \frac{\bar{\mathbf{H}}_1 \bar{\mathbf{H}}_1^H}{M} \right\} \right\}, \tag{72}$$

where

$$\rho_k = \frac{\hat{\sigma}_k^2}{\mathbb{E}\left\{\epsilon_i^{-2}\right\} + \tilde{\sigma}_k^2}, \tag{73}$$

and ϵ_i^{-2} is an eigenvalue of $\tilde{\mathbf{U}}_{2,k}\tilde{\mathbf{U}}_{2,k}^H$, $\tilde{\mathbf{U}}_{2,k} = [\rho_\tau \mathbf{I},\ \rho_\delta \mathbf{U}_{2,k}]$. The achievable rate of User 1, i.e., R_1 given in (72), demonstrates the gain in the sum-rate due to product superposition.

V. Disparity in Both Coherence Time and Coherence Bandwidth

Disparity in both coherence time and coherence bandwidth occurs when $T_1 \neq T_2$ and $L_1 \neq L_2$. An example of this scenario is as follows.

Example 3: Consider an urban environment and a micro-cell serving two users: User 1 is time-dynamic and frequency-dynamic where it has a propagation of *B2 bad urban micro-cell*, where User 2 is time-static and frequency-static where it has a propagation of *B4 outdoor to indoor*. User 1 has 70 Km/h mobility and 480 ns r.m.s. delay spread, whereas User 2 has 5 Km/h mobility and 49 ns r.m.s. delay spread.

If $T_1 \leq T_2$ and $L_1 \geq L_2$, i.e., User 2 is time-static and frequency-static, the proposed product superposition in Section III and Section IV can be used, resulting in the same achievable rates. If $T_1 \leq T_2$ and $L_1 \leq L_2$, User 1 is time-dynamic and frequency-static. In this case, since none of the users is static in both time and frequency, two transmission strategies are proposed: product superposition over coherence time and product superposition over coherence bandwidth. In the sequel, we assume $T_2 = \ell T_1$, where $\ell \in \mathbb{N}$, and furthermore $L_1 = 1, L_2 \geq 1$.

A. Product Superposition Over Coherence Time

The interval T_2 can be divided into ℓ subintervals, each of length T_1. During the first subinterval, the transmitter sends pilots during the first M time instances at subcarrier i so that $\mathbf{H}_1$ and $\mathbf{H}_{2,i}$ can be estimated. Hence,

$$\begin{aligned} \hat{\mathbf{H}}_1 &= \frac{\rho_\tau}{\rho_\tau + 1}\mathbf{H}_1 + \frac{\sqrt{\rho_\tau}}{\rho_\tau + 1}\mathbf{W}_{1,\tau,i}, \\ \hat{\mathbf{H}}_{2,i} &= \gamma \mathbf{H}_{2,i} + \frac{\gamma}{\sqrt{\rho_\tau}}\mathbf{W}_{2,\tau,i}. \end{aligned} \tag{74}$$

During the following $(\ell - 1)$ subintervals, $\mathbf{H}_1$ needs pilots to be estimated, whereas $\mathbf{H}_{2,i}$ does not. Therefore, the transmitted space-time code at subcarrier i during the first M time instances is

$$\mathbf{X}_{\tau,i} = \sqrt{M}\mathbf{U}_2, \tag{75}$$

where $\mathbf{U}_2 \in \mathbb{C}^{M \times T_1}$ contains symbols intended for User 2. Hence, the received signal of User 1 during the first M time instances is

$$\mathbf{Y}_{1,\tau,i} = \sqrt{\rho_\tau}\mathbf{G}_1 + \mathbf{W}_{1,\tau,i}, \tag{76}$$

where $\mathbf{G}_1 = \mathbf{H}_1\mathbf{U}_2$ is the equivalent channel seen by User 1. The MMSE channel estimate is

$$\hat{\mathbf{G}}_1 = \frac{\rho_\tau}{\rho_\tau + 1}\mathbf{G}_1 + \frac{\sqrt{\rho_\tau}}{\rho_\tau + 1}\mathbf{W}_{1,\tau,i}. \tag{77}$$

and furthermore

$$\hat{\sigma}^2 = \frac{\rho_\tau}{1+\rho_\tau}, \tag{78}$$

The estimation error is

$$\tilde{\mathbf{G}}_1 = \mathbf{G}_1 - \hat{\mathbf{G}}_1 = \frac{1}{1+\rho_\tau}\mathbf{G}_1 - \frac{\sqrt{\rho_\tau}}{1+\rho_\tau}\mathbf{W}_{1,\tau,i}. \tag{79}$$

and furthermore

$$\tilde{\sigma}^2 = \frac{1}{1+\rho_\tau}. \tag{80}$$

The achievable rate is

$$\begin{aligned} R_1 \geq\ & \frac{1}{\ell K}\left(1-\frac{M}{T_1}\right)\mathbb{E}\left\{\log\det\left\{\mathbf{I}_{N_1} + \rho_i\frac{\bar{\mathbf{G}}_1\bar{\mathbf{G}}_1^H}{M}\right\}\right\} \\ & + \frac{1}{\ell K}\sum_{k=1,k\neq i}^{K}\mathbb{E}\left\{\log\det\left\{\mathbf{I}_{N_1} + \rho_k\frac{\bar{\mathbf{G}}_1\bar{\mathbf{G}}_1^H}{M}\right\}\right\} \\ & + \left(1-\frac{1}{\ell}\right)\left(\frac{T_1-M}{T_1K}\mathbb{E}\left\{\log\det\left\{\mathbf{I}_{N_1} + \check{\rho}_i\frac{\bar{\mathbf{G}}_1\bar{\mathbf{G}}_1^H}{M}\right\}\right\}\right. \\ & \left. + \frac{1}{K}\sum_{k=1,k\neq i}^{K}\mathbb{E}\left\{\log\det\left\{\mathbf{I}_{N_1} + \check{\rho}_k\frac{\bar{\mathbf{G}}_1\bar{\mathbf{G}}_1^H}{M}\right\}\right\}\right), \end{aligned} \tag{81}$$

where

$$\check{\rho}_k = \frac{\rho_\delta\hat{\sigma}_k^2}{1+\rho_\delta\tilde{\sigma}_k^2}. \tag{82}$$

$\hat{\sigma}^2$, $\tilde{\sigma}^2$ are given in (78) and (80), respectively, and ρ_k is defined in (22). The achievable rate for User 2 is

$$R_2 \geq \left(1-\frac{1}{\ell}\right)\frac{M}{T_1K}\mathbb{E}\left\{\log\det\left\{\mathbf{I}_{N_2} + \rho_i\frac{\bar{\mathbf{H}}_{2,i}\bar{\mathbf{H}}_{2,i}^H}{M}\right\}\right\}. \tag{83}$$

B. Product Superposition Over Coherence Bandwidth

This case is similar to that discussed in Section IV. The achievable rates R_1 and R_2 are given in (72) and (71), respectively.

VI. Numerical Results and Discussion

A. Numerical Results

The numerical results are produced via Monte Carlo simulations. Each data point is produced by averaging over 10,000 randomly generated samples. For each sample, the channel coefficients and the noise are generated independently according to the relevant probability distributions.

For disparity in coherence time investigated in Section III-B, the performance of the proposed product superposition is shown in Fig. 2 and Fig. 3. In particular, Fig. 2 shows the sum-rate of the conventional transmission and the proposed product superposition versus SNR for $M = N_1 = N_2 = 4$, $L = 2$, $K = 12$, and $T_1 = 20$. The gain in the sum-rate is 10 b/sec/Hz at 20dB SNR, and the improvement is at least 50% over the range of 0-20 dB SNR. The curves with no power optimization use $\rho_\tau = \rho_\delta = \rho$ and with power optimization use ρ_τ, ρ_δ from (23). Fig. 3 demonstrates the sum-rate gain versus the number of transmit and receive antennas for $\rho = 20$ dB, $K = 8$, $N_1 = N_2 = N$, $T_1 = 11$. This figure shows that our proposed method offers gains even in a SISO configuration, therefore it is distinct from beamforming gains.

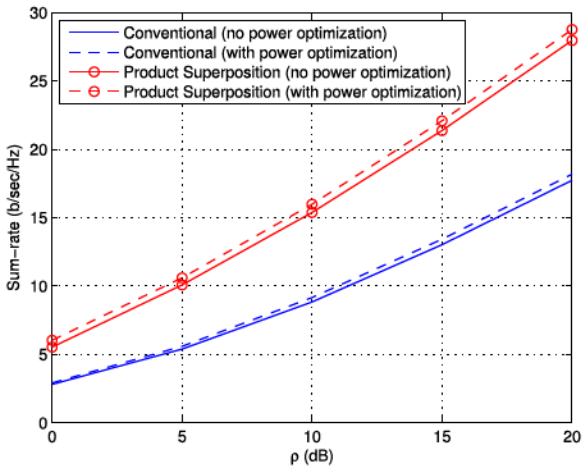

Fig. 2. Sum-rate versus SNR under disparity in coherence time for frequency-domain channel estimation without interpolation.

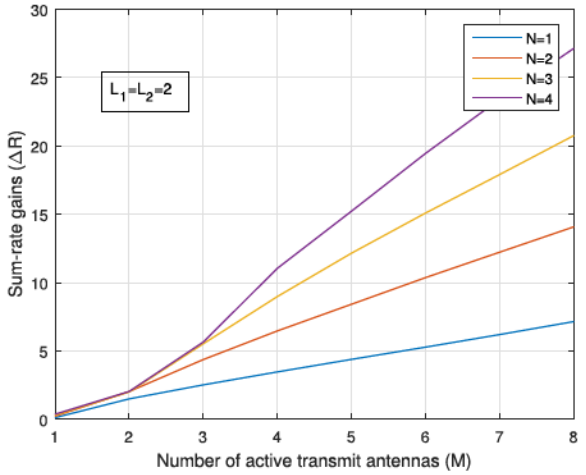

Fig. 3. Sum-rate gains in a variety of system configurations, under frequency-domain channel estimation.

With $M = N_1 = N_2 = 4, L = 2, K = 12, T_1 = 20, i = 1$, and $j = 7$, Fig. 4 demonstrates the gain in the sum-rate for disparity in coherence time when channel estimation is done using interpolation. The interpolation matrix $\mathbf{A}_k$ is piece-wise linear, and the matrix $\mathbf{B}_k$ is the optimizer of (44). This figure shows at least 35% improvement in sum-rate over the range of 0-20 dB SNR. The effect of the interpolation matrix $\mathbf{A}_k$ is shown in Fig. 5 for $L = 3$, where the gain is shown for piece-wise linear, nearest-neighbor, and Wiener interpolation. Fig. 6 and Fig. 7 demonstrate the effect of optimal, linear, and random $\mathbf{B}_k$.

Fig. 8 depicts performance under time domain estimation and $M = 4$, $N_1 = 2$, $N_2 = 4$, $L = 5$, $K = 10$, and $T_1 = 8$, showing sum-rate improvement of at least 10% over the range

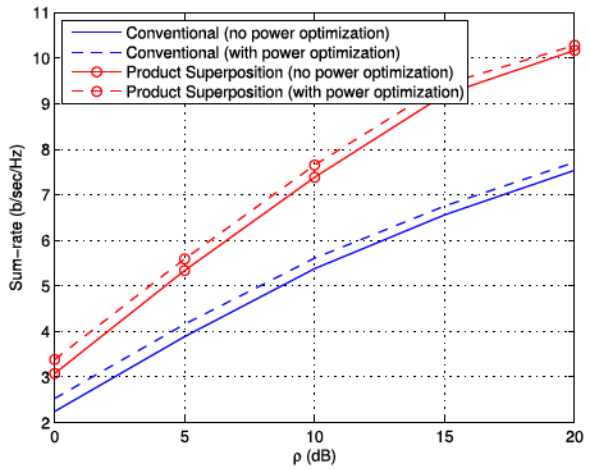

Fig. 4. Sum-rate versus SNR under disparity in coherence time for frequency-domain channel estimation with interpolation.

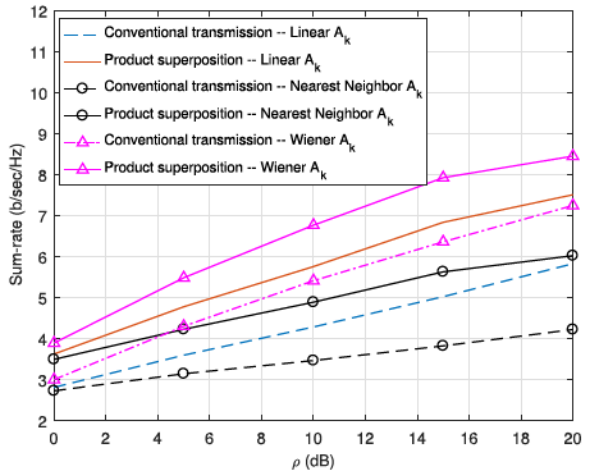

Fig. 5. Sum-rate versus SNR under disparity in coherence time for different interpolation techniques of $\mathbf{A}_k$.

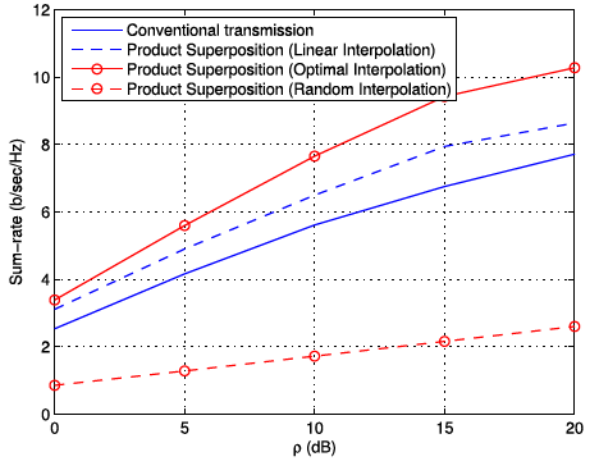

Fig. 6. The effect of the interpolation matrix $\mathbf{B}_k$ on the sum-rate.

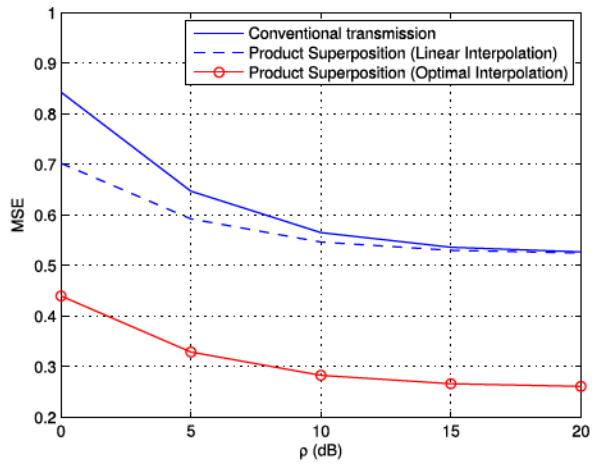

Fig. 7. The effect of the interpolation matrix $\mathbf{B}_k$ on the MMSE.

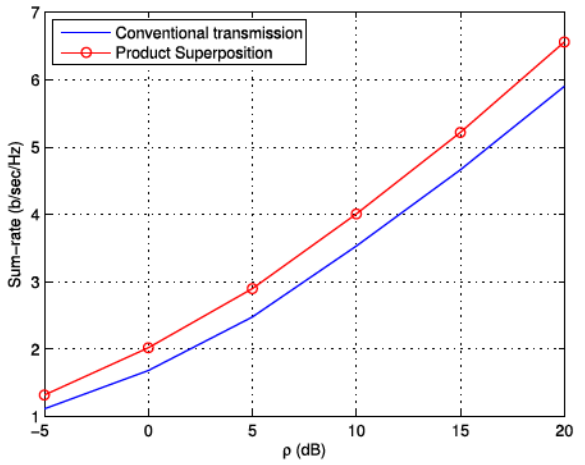

Fig. 8. Sum-rate versus SNR under disparity in coherence time with time-domain channel estimation.

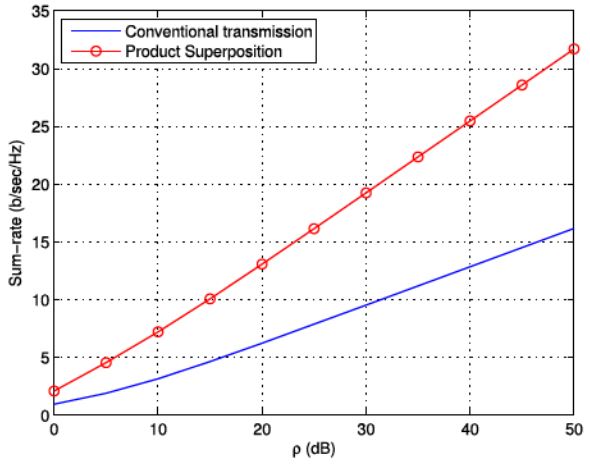

Fig. 9. Sum-rate versus SNR under disparity in coherence bandwidth.

of 0-20 dB SNR. Fig. 9 shows the product superposition gain when the disparity is in coherence bandwidth (Section IV) and when $M = 4$, $N_1 = 2$, $N_2 = 2$, $T_1 = T_2 = T = 8$, $L_1 = 1$, $L_2 = 2$, and $K = 8$.

Fig. 10 and Fig. 11 show the gain of product superposition for disparity in both coherence time and coherence bandwidth (Section V). In these two figures, User 1 is time-dynamic

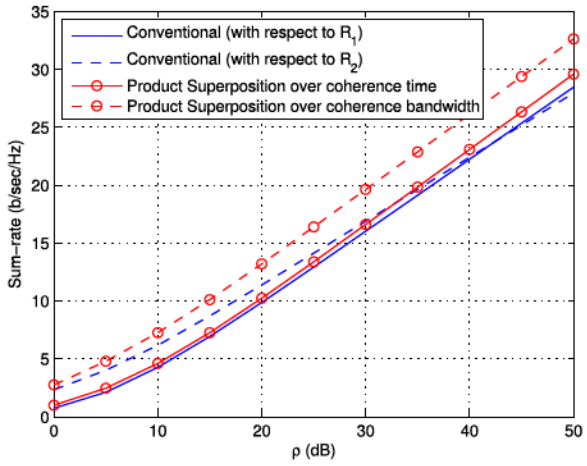

Fig. 10. Sum-rate versus SNR under disparity in both coherence time and coherence bandwidth.

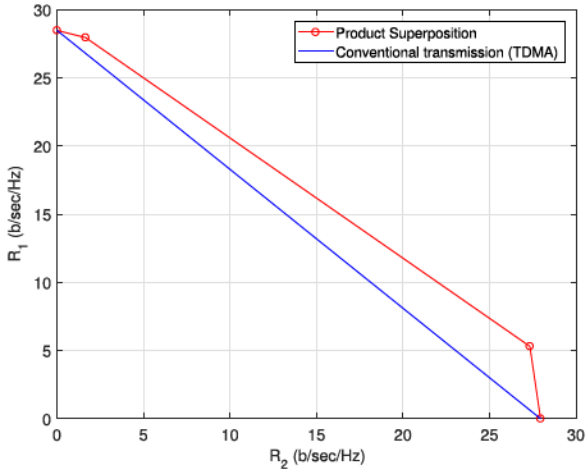

Fig. 11. Rate region under disparity in both coherence time and coherence bandwidth.

and frequency-static and User 2 is time-static and frequency-dynamic. The simulation parameters are $M = 4, N_1 = N_2 = 2, L_1 = 1, L_2 = 3, T_1 = 8$, and $T_2 = 24$. Fig. 10 shows the performance of product superposition over coherence time (Section V-A) and product superposition over coherence bandwidth (Section V-B), and for reference also shows single-user transmission rates to User 1 and User 2. Fig. 11 shows the achievable rate region of orthogonal transmission and the proposed product superposition. The achievable rate region of product superposition is the polygon that represents time-sharing between strategies achieving rates at its vertices. The achieved rate pairs (R_2, R_1) given in Section V-A and Section V-B lead to the two corner points that increase the product superposition rate region compared with the TDMA region.

B. Channel Models and Relation to Contemporary Standards

In some practical communication systems, e.g. 3GPP LTE, the size of the resource blocks of 12 subcarriers and 0.5 ms is taken to be the same for all receivers. Some reference signals, denoted Cell-specific Reference Signals (CRSs), are sent to be used by the receiver in channel estimation. In each resource block, 4 CRSs are sent regardless of the coherence conditions of the link. This design does not provide different number of reference signals to different links to match their individual coherence conditions.

In practice, wireless receivers in the same cell may have widely varying coherence conditions, based on their individual delay and Doppler spread [37], [38]. For instance, in an urban environment, a macro-cell may serve two users: one high-mobility user with 120 Km/h mobility and one low-mobility user with 5 Km/h mobility; hence, the coherence times of the two users are 2.25 ms and 54 ms, respectively [8], [23]. For simplicity, the analysis of multiuser networks in a large part of the literature assumes that all fading links have equal coherence time and equal coherence bandwidth. Moreover, in practical communication systems, some parameters are often designed based on the shortest coherence conditions among users. For instance, in 3GPP LTE system, the number of Cell-specific Reference Signals is designed to support 500 Km/h mobility [7], [8]. These have all been motivating factors in the present study.

C. Complexity

In general, most superposition schemes have a complexity overhead compared with orthogonal transmission, since they allow interference that must at some point be accounted for and/or peeled away. This is to some extent true for product superposition, although the unique aspects of this technique produce some distinct points of interest.

In product superposition, the added complexity is imposed only on the less-selective user, for the following reason: product superposition is designed so that from the viewpoint of the inner code (more selective user), the interfering signal merges multiplicatively with channel coefficients and thus disappears into the channel [9]. Thus the more selective user does not need to do anything differently compared with conventional transmission.

The added complexity for the less selective user (called "static user" in [9]) is due to decoding and cancelling the signal of the more selective user. However, in the framework presented in this paper, the added complexity only occurs in resource blocks that contain a pilot. Overall, the complexity of the less selective user is bounded above by twice the complexity of the conventional transmission, although this might be a pessimistic bound.

Finally, we note that there exist other variations of product superposition that do not require a peeling decoder [9]. These versions will trade off reduced complexity with slightly lower performance. The discussion of these variations has been omitted in this paper for the sake of brevity.

VII. Conclusion

This paper investigates multiuser wireless channels when the receivers have disparity in coherence time and/or bandwidth. The disparity is a source of gains that is denoted coherence diversity. To explore this phenomenon, we investigated OFDM downlink transmission for two users when the nodes

experience differences either in coherence time, in coherence bandwidth, or in both. In each scenario a version of product superposition was proposed and its gains were demonstrated. Numerical simulations were provided to demonstrate the gains compared with the conventional transmission that uses TDMA.

APPENDIX A
POWER OPTIMIZATION OF ρ_τ AND ρ_δ IN (23)

In the following we find the optimal values of ρ_τ and ρ_δ so that the value of ρ_k is maximized, which in turn maximizes the User 1 achievable rate in (21). From the energy constraint,

$$\rho_\delta T_{1,\delta} + \rho_\tau T_{1,\tau} = \rho T_1, \tag{84}$$

and hence

$$\begin{aligned} \rho_\delta T_{1,\delta} &= (1-\alpha)\rho T_1, \\ \rho_\tau T_{1,\tau} &= \alpha \rho T_1, \end{aligned} \tag{85}$$

where $\alpha \in [0,1]$ is the power allocation factor. Therefore,

$$\begin{aligned} \rho_k &= \frac{\rho_\delta \alpha \|\mathbf{f}_1\|^2}{1+\rho_\delta \frac{\gamma}{\rho_\tau}} \\ &= \frac{\alpha(1-\alpha)\frac{(\rho T)^2}{T_{1,\tau}T_{1,\delta}}\|\mathbf{f}_1\|^4}{1+\alpha\frac{\rho T_1}{T_{1,\tau}}\|\mathbf{f}_1\|^2 + (1-\alpha)\frac{\rho T_1}{T_{1,\delta}}\|\mathbf{f}_1\|^2} \\ &= \frac{\rho T_1 \|\mathbf{f}_1\|^2}{(T_{1,\delta}-T_{1,\tau})}\frac{\alpha(1-\alpha)}{\alpha+\beta}, \end{aligned} \tag{86}$$

where

$$\beta = \frac{T_{1,\delta}T_{1,\tau} + \rho T_1 T_{1,\tau}\|\mathbf{f}_1\|^2}{T_{1,\delta}\rho T_1 \|\mathbf{f}_1\|^2 \left(1-\frac{T_{1,\tau}}{T_{1,\delta}}\right)}. \tag{87}$$

By taking the first and the second derivatives of ρ_k with respect to α, we obtain

$$\frac{\partial \rho_k}{\partial \alpha} = \frac{\rho T_1}{T_{1,\delta}-T_{1,\tau}}\frac{(\alpha+\beta)(1-2\alpha)-\alpha(1-\alpha)}{(\alpha+\beta)^2}, \tag{88}$$

and hence

$$\begin{aligned} \frac{\partial^2 \rho_k}{\partial \alpha^2} &= \frac{\rho T_1}{(T_{1,\delta}-T_{1,\tau})(\alpha+\beta)^4}\Big(-2(\alpha+\beta)^3 \\ &\quad -2(\alpha+\beta)\left((\alpha+\beta)(1-2\alpha)-\alpha(1-\alpha)\right)\Big) \\ &= \frac{-2\rho T_1}{(T_{1,\delta}-T_{1,\tau})(\alpha+\beta)^3}\Big((\alpha+\beta)^2-\alpha^2-2\alpha\beta+\beta\Big) \\ &= \frac{-2\rho T_1(\beta+\beta^2)}{(T_\delta - T_{1,\tau})(\alpha+\beta)^3}, \end{aligned} \tag{89}$$

which shows the concavity of ρ_k. Therefore, from the KKT conditions [39], i.e. $\frac{\partial \rho_k}{\partial \alpha} = 0$, and hence the optimal α is

$$\alpha = -\beta + \sqrt{\beta(\beta+1)}. \tag{90}$$

REFERENCES

[1] M. Fadel and A. Nosratinia, "Coherence diversity in time and frequency," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2016, pp. 1–6.

[2] C. Y. Wong, R. S. Cheng, K. B. Lataief, and R. D. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 10, pp. 1747–1758, Oct. 1999.

[3] J. Huang, V. G. Subramanian, R. Agrawal, and R. A. Berry, "Downlink scheduling and resource allocation for OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 8, no. 1, pp. 288–296, Jan. 2009.

[4] S. J. Yi, C. C. Tsimenidis, and B. S. Sharif, "Transmitter precoding in downlink MC-CDMA systems over frequency-selective Rayleigh fading channels," *IEE Proc.-Commun.* , vol. 152, no. 6, pp. 952–958, Dec. 2005.

[5] L.-U. Choi and R. D. Murch, "A pre-blast-DFE technique for the downlink of frequency-selective fading MIMO channels," *IEEE Trans. Commun.*, vol. 52, no. 5, pp. 737–743, May 2004.

[6] M. Fadel, A. S. Ibrahim, and H. Elgebaly, "QoS-aware multi-RAT resource allocation with minimum transmit power in multiuser OFDM system," in *Proc. IEEE Globecom Workshops*, Dec. 2012, pp. 670–675.

[7] E. Dahlman, S. Parkvall, and J. Sköld, *4G: LTE/LTE-Advanced for Mobile Broadband*. New York, NY, USA: Academic, 2013.

[8] S. Sesia, I. Toufik, and M. Baker, *LTE—The UMTS Long Term Evolution*. Hoboken, NJ, USA: Wiley, 2009.

[9] Y. Li and A. Nosratinia, "Coherent product superposition for downlink multiuser MIMO," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1746–1754, Mar. 2015.

[10] M. Fadel and A. Nosratinia, "Coherent, non-coherent, and mixed-CSIR broadcast channels: Multiuser degrees of freedom," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2014, pp. 2574–2578.

[11] M. Fadel and A. Nosratinia, "Coherence disparity in broadcast and multiple access channels," *IEEE Trans. Inf. Theory*, vol. 62, no. 12, pp. 7383–7401, Dec. 2016.

[12] M. Fadel and A. Nosratinia, "Broadcast channel under unequal coherence intervals," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2016, pp. 275–279.

[13] M. Fadel and A. Nosratinia, "Block-fading broadcast channel with hybrid CSIT and CSIR," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2017, pp. 1873–1877.

[14] J. Niu, D. Lee, X. Ren, G. Y. Li, and T. Su, "Scheduling exploiting frequency and multi-user diversity in LTE downlink systems," in *Proc. IEEE 23rd Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, vol. 12, no. 4, pp. 1412–1417, Sep. 2012.

[15] J. Niu, D. Lee, T. Su, G. Y. Li, and X. Ren, "User classification and scheduling in LTE downlink systems with heterogeneous user mobilities," *IEEE Trans. Wireless Commun.*, vol. 12, no. 12, pp. 6205–6213, Dec. 2013.

[16] J. Niu, T. Su, G. Y. Li, D. Lee, and Y. Fu, "Joint transmission mode selection and scheduling in LTE downlink MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 3, no. 2, pp. 173–176, Apr. 2014.

[17] T. Hwang, C. Yang, G. Wu, S. Li, and G. Y. Li, "OFDM and its wireless applications: A survey," *IEEE Trans. Veh. Technol.*, vol. 58, no. 4, pp. 1673–1694, May 2009.

[18] Y. Li, "Pilot-symbol-aided channel estimation for OFDM in wireless systems," *IEEE Trans. Veh. Technol.*, vol. 49, no. 4, pp. 1207–1215, Jul. 2000.

[19] M. Morelli and U. Mengali, "A comparison of pilot-aided channel estimation methods for OFDM systems," *IEEE Trans. Signal Process.*, vol. 49, no. 12, pp. 3065–3073, Dec. 2001.

[20] M. K. Ozdemir and H. Arslan, "Channel estimation for wireless OFDM systems," *IEEE Commun. Surveys Tuts.*, vol. 9, no. 2, pp. 18–48, 2nd Quart. 2007.

[21] A. Vosoughi and A. Scaglione, "Everything you always wanted to know about training: Guidelines derived using the affine precoding framework and the CRB," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 940–954, Mar. 2006.

[22] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 684–702, Jun. 2003.

[23] J. Meinilä, P. Kyösti, T. Jämsä, and L. Hentila, "WINNER II channel models," *Radio Technol. Concepts IMT-Adv.*, vol. 54, no. 3, pp. 39–92, Oct. 2009.

[24] C. Huang, S. A. Jafar, S. Shamai (Shitz), and S. Vishwanath, "On degrees of freedom region of MIMO networks without channel state information at transmitters," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 849–857, Feb. 2012.

[25] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.

[26] S. Omar, A. Ancora, and D. T. M. Slock, "Performance analysis of general pilot-aided linear channel estimation in LTE OFDMA systems with application to simplified MMSE schemes," in *Proc. IEEE 19th Int. Symp. Pers. Indoor Mobile Radio Commun. (PIMRC)*, Sep. 2008, pp. 1–6.

[27] S. Coleri, M. Ergen, A. Puri, and A. Bahai, "Channel estimation techniques based on pilot arrangement in OFDM systems," *IEEE Trans. Broadcast.*, vol. 48, no. 3, pp. 223–229, Sep. 2002.
[28] Y. Acar, H. Dovgan, E. Başar, and E. Panayirci, "Interpolation based pilot-aided channel estimation for STBC spatial modulation and performance analysis under imperfect CSI," *IET Commun.*, vol. 10, no. 14, pp. 1820–1828, Sep. 2016.
[29] P. Hoeher, S. Kaiser, and P. Robertson, "Two-dimensional pilot-symbol-aided channel estimation by Wiener filtering," in *Proc. IEEE Int. Conf. Acoust. Speech, Signal Process.*, Munich, Germany, Apr. 1997, pp. 1845–1848.
[30] G. Liu, L. Zeng, H. Li, L. Xu, and Z. Wang, "Adaptive interpolation for pilot-aided channel estimator in OFDM system," *IEEE Trans. Broadcast.*, vol. 60, no. 3, pp. 486–498, Sep. 2014.
[31] Y. Li, L. J. Cimini, and N. R. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading channels," *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 902–915, Jul. 1998.
[32] F. Pena-Campos, R. Carrasco-Alvarez, O. Longoria-Gandara, and R. Parra-Michel, "Estimation of fast time-varying channels in OFDM systems using two-dimensional prolate," *IEEE Trans. Wireless Commun.*, vol. 12, no. 2, pp. 898–907, Feb. 2013.
[33] P.-Y. Tsai and T.-D. Chiueh, "Adaptive raised-cosine channel interpolation for pilot-aided OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 8, no. 2, pp. 1028–1037, Feb. 2009.
[34] N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*. Cambridge, MA, USA: MIT Press, 1945.
[35] R. Negi and J. Cioffi, "Pilot tone selection for channel estimation in a mobile OFDM system," *IEEE Trans. Consum. Electron.*, vol. 44, no. 3, pp. 1122–1128, Aug. 1998.
[36] I. Barhumi, G. Leus, and M. Moonen, "Optimal training design for MIMO OFDM systems in mobile wireless channels," *IEEE Trans. Signal Process.*, vol. 51, no. 6, pp. 1615–1624, Jun. 2003.
[37] A. Goldsmith, *Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
[38] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
[39] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

Mohamed Fadel (S'14) received the B.Sc. degree in electrical engineering from Alexandria University, Egypt, in 2009, the M.Sc. degree in communication and information technology from Nile University, Egypt, in 2011, and the Ph.D. degree in electrical engineering from The University of Texas at Dallas in 2017, where he received the Industrial Advisory Board Fellowship in 2015 and the Louis Beecherl, Jr. Graduate Fellowship in 2016. In 2011, he was with the Middle East Mobile Innovation Center, part of Intel Labs, where he was involved in multiple radio access technologies. He was also a Systems Engineer Intern at PoLTE Corporation, where he was involved in user positioning over LTE. He is currently with Qualcomm, where he is involved in the design of WiFi chips. His research interests include wireless communication, signal processing, and multiuser information theory.

Aria Nosratinia (S'87–M'97–SM'04–F'10) received the Ph.D. degree in electrical and computer engineering from the University of Illinois at Urbana–Champaign in 1996. He has held visiting appointments at Princeton University, Rice University, and UCLA. He is currently an Erik Jonsson Distinguished Professor and the Associate Head of the Electrical Engineering Department, The University of Texas at Dallas. His interests lie in the broad area of information theory and signal processing, with applications in wireless communications. He was a recipient of the National Science Foundation Career Award. He is the General Co-Chair of ITW 2018, and he was the Secretary for the IEEE Information Theory Society from 2010 to 2011 and the Treasurer for ISIT 2010 in Austin, TX, USA. He has served as an Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY, IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE SIGNAL PROCESSING LETTERS, IEEE TRANSACTIONS ON IMAGE PROCESSING, and IEEE WIRELESS COMMUNICATIONS. In 2016, he was named a Thomson Reuters Highly Cited Researcher.