Hybrid Training Signal Design for Multiuser OFDM Channel Estimation in mmWave Bands

H. D. Tuan[1], A. A. Nasir[2], A. V. Savkin[3], and L. Hanzo[4]

***Abstract*—Hybrid training signal design is conceived for the estimation of multi-user frequency-selective mmWave channels, where hybrid architectures are adopted due to the limited number of available radio frequency (RF) chains. In this setting, the training signal has a hybrid structure, expressed as the product of a high-dimensional, low-resolution analog beamformer (ABF) and low-dimensional subcarrier-wise digital pilots (SDPs). All pilot subcarriers share the same analog beamformer. Consequently, the joint design of ABF and SDPs to minimize the minimum mean square error (MMSE) of channel estimation leads to a challenging high-dimensional mixed continuous-discrete optimization problem. The main contribution is the development of a new path-following algorithm that leverages closed-form updates of scalable complexity for both ABF and SDPs. We further examine the impact of channel estimation accuracy on the capacity of multi-user wideband mmWave communication systems.**

***Index Terms*—MmWave communication, frequency selectivity, channel estimation, hybrid structure**

I. Introduction

The rapid escalation of data-hungry applications, such as high-definition video streaming, immersive virtual reality, and autonomous systems, has fueled the research of millimeter-wave (mmWave) communication in multi-user multiple-input multiple-output (MU-MIMO) systems [1]–[4]. Orthogonal frequency-division multiplexing (OFDM) is often employed in these systems for counter-acting frequency-selective fading, enabling reliable exploitation of the large bandwidths available in mmWave bands [5]. At these frequencies, radio frequency (RF) chains relying on phase shifters become power-hungry and cannot be assigned to each antenna element. Consequently, mmWave transceivers use far fewer RF chains than antenna elements. This limitation has made hybrid beamforming, in which the transmit signal is generated by a low-dimensional digital beamformer combined with a high-dimensional analog beamformer (ABF), a popular approach for mmWave transmission. As noted in [6]–[9], channel estimation under hybrid architectures differs fundamentally from the fully digital case because digital pilots must pass through a limited number of RF chains. This bottleneck alters identifiability conditions and complicates multi-user CSI acquisition.

Significant research efforts have invested in flat-fading mmWave channel estimation (see, e.g., [10] and references therein). For frequency-selective mmWave channels, [11] employed MUSIC to estimate the angles of arrival, angles of departure, and propagation delays, while [12] and [13] studied least-squares estimation and alternating direction method of multipliers (ADMM) estimation, respectively. Joint estimation of angles and delays was investigated in [14], while [15] analyzed single-user MIMO-OFDM channel estimation based on the Cramér-Rao bound (CRB). Other works, such as [7], [16]–[24], focus on mmWave channel estimation in general, but do not address signal design for OFDM-based mmWave channel estimation.

In the context of pilot signal design conceived for estimating frequency-selective mmWave channels, existing studies [25]–[27] primarily focus their attention on fully digital pilot sequences. As discussed earlier, transmitting fully digital pilots is impractical in mmWave systems due to prohibitive hardware cost and power consumption associated with fully digital RF chains. To alleviate these limitations, the authors of [28] investigated ABFs to enhance the so-called isometry property of compressed sensing. However, the ABF design considered in [28] relies on a fully connected architecture, which requires a prohibitively large number of phase shifters and it is therefore difficult to implement in practice.

To the best of our knowledge, no existing work addresses training signal design conceived for multi-user mmWave OFDM channels under hybrid architectural constraints, i.e., the joint design of an ABF with a limited number of phase shifters and subcarrier-based digital pilots (SDPs). In such hybrid architectures, the ABF is shared across all subcarriers, which inherently couples the spatial and frequency dimensions and fundamentally alters the structure of the transmitted training signal.

Consequently, in contrast to the fully digital case [29], [30], digital pilot signals of hybrid architectures cannot be independently assigned across antennas and subcarriers. As a result, existing fully digital pilot design methods [25], [26], [30] cannot be directly applied. These fundamental challenges motivate the development of a new training signal design that explicitly accounts for hybrid architectures and enables accurate channel estimation with improved power efficiency.

Against the above background, the paper offers the following new contributions:

- **Development of a scalable algorithm for signal design**: We consider the joint design of ABF and SDPs

The financial support of the following Engineering and Physical Sciences Research Council (EPSRC) projects is gratefully acknowledged: Platform for Driving Ultimate Connectivity (TITAN) (EP/X04047X/1; EP/Y037243/1); Robust and Reliable Quantum Computing (RoaRQ, EP/W032635/1); PerCom (EP/X012301/1); India-UK Intelligent Spectrum Innovation ICON UKRI-1859.

[1]School of Electrical and Data Engineering, University of Technology Sydney, Broadway, NSW 2007, Australia (email:Tuan.Hoang@uts.edu.au); [2]Department of Electrical Engineering, King Fahd University of Petroleum and Minerals (KFUPM), Dhahran, Saudi Arabia (email: anasir@kfupm.edu.sa); [3]School of Electrical Engineering and Telecommunications, The University of New South Wales Sydney, NSW 2052, Australia (email: a.savkin@unsw.edu.au); [4]School of Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, UK (email: lh@ecs.soton.ac.uk)

to minimize the minimum mean square error (MMSE) of channel estimation. The ABF adopts an array-of-subarrays (AoSA) structure to use a reduced number of low-resolution phase shifters for controlling power consumption, while the SDPs are constrained in terms of their total power to limit transmit power. The number of pilot subcarriers is also limited to leave more room for information transmission. This design problem constitutes a high-dimensional mixed continuous-discrete optimization problem. We develop a new path-following algorithm that updates both the ABF and SDPs using closed-form expressions of scalable computational complexity.

- **Impact of channel estimation on system performance**: We consider the problem of maximizing system capacity while accounting for the MMSE of channel estimation and propose computational algorithms. Simulations illustrate the sensitivity of system performance to the MMSE.

The rest of the paper is organized as follows. Section II is devoted to the problem of jointly designing ABF and SDPs for estimating multi-user channels. Section III is devoted to the problem of maximizing the system capacity, while incorporating the channel estimation effect. Section IV is devoted to our simulations and Conclusions are given in Section V.

Notation. Only the optimization variables are boldfaced. For a complex number x, $\angle x$ presents its argument. Then $e^{\jmath x} \triangleq (e^{\jmath x_1}, \ldots, e^{\jmath x_n})^T \in \mathbb{C}^N$ for $x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$. The inner product between vectors x and y is defined as $\langle x, y\rangle = x^H y$. Analogously, $\langle X, Y\rangle \triangleq \mathsf{trace}(X^H Y)$ for the matrices X and Y. $\mathsf{diag}[X_1, \ldots, X_N]$ is a matrix with diagonal blocks X_n, $n = 1, \ldots, N$, while $\big[X_{m,n}\big]_{(m,n)\in\mathcal{M}\times\mathcal{N}}$ or $\big[X(m,n)\big]_{(m,n)\in\mathcal{M}\times\mathcal{N}}$ is a matrix with entries $X_{m,n}$ or $X(m,n)$ which are also matrices. $\mathsf{Row}[X_n]_{n=1,\ldots N}$ denotes the matrix formed by stacking X_1 through X_N horizontally. We also use $\langle X\rangle$ for the trace of X. The operator $\mathsf{vec}(X)$ stacks the columns of X from the first to the last into a single column vector. $||X||$ is the Frobenius norm of the matrix X, which is defined by $\sqrt{\langle X^H X\rangle}$. $[X]^2$ denotes the positive semidefinite matrix XX^H. Then $\ln X$ is the natural logarithm of the determinant (log-det) of X. I_n is the $n \times n$ identity matrix. When the size of the identity matrix is clear from the context, we may omit the subscript n in expressions. Lastly, $\mathbb{R}_+^N$ is the set of N-dimensional real vectors with positive entries.

Mathematical ingredient. The following inequalities for all matrices $\boldsymbol{X}$ and $\bar{X}$ of size $n\times m$ and positive definite matrices $\mathbf{Y}$ and $\bar{Y}$ of size $n \times n$ have been established in [30]–[32]:

$$\langle [\boldsymbol{X}]^2 \mathbf{Y}^{-1}\rangle \geq 2\Re\{\langle \bar{X}^H \bar{Y}^{-1} \boldsymbol{X}\rangle\} - \langle [\bar{Y}^{-1}\bar{X}]^2 \mathbf{Y}\rangle, \tag{1}$$

and

$$\begin{aligned}\ln\left(I_n + [\boldsymbol{X}]^2\mathbf{Y}^{-1}\right) \geq \ln\left(I_n + [\bar{X}]^2\bar{Y}^{-1}\right) \\ - \langle [\bar{X}]^2\bar{Y}^{-1}\rangle + 2\Re\{\langle \bar{X}^H\bar{Y}^{-1}\boldsymbol{X}\rangle\} \\ - \left\langle \bar{Y}^{-1} - (\bar{Y} + [\bar{X}]^2)^{-1}, [\boldsymbol{X}]^2 + \mathbf{Y}\right\rangle.\end{aligned} \tag{2}$$

Considering both sides of (1) and (2) as functions of the variables $(\boldsymbol{X}, \mathbf{Y})$, they coincide at $(\bar{\boldsymbol{X}}, \bar{\mathbf{Y}})$. That is, the functions defined by the left-hand sides (LHSs) provide tight minorants of their counterparts on the right-hand sides (RHSs) at $(\bar{\boldsymbol{X}}, \bar{\mathbf{Y}})$ [33, p. 366]. Consequently, maximizing the LHS functions yields an improved point for the RHS functions.

II. Hybrid pilot design under AoSA-ABF for channel estimation

Let us consider the downlink (DL) scenario, where a base station (BS) serves K users, each indexed by $k \in \mathcal{K} \triangleq \{1, 2, \ldots, K\}$. The BS is equipped with a massive N-element uniform linear array (ULA), while each user equipment (UE) k is equipped with an N_u-antenna array.

Let N_c denote the number of RF chains at the BS used for ABF. Following the array-of-subarrays (AoSA) architecture [34], each RF chain is connected to $L = N/N_c$ transmit antennas (TAs). Accordingly, the ABF matrix is given by $\mathbf{V}_A \in \mathbb{C}^{N\times N_c}$ in the form

$$\mathbf{V}_A = \mathcal{D}(\boldsymbol{\theta}), \tag{3}$$

with

$$\mathcal{D}(\boldsymbol{\theta}) \triangleq \mathsf{diag}[e^{\jmath\boldsymbol{\theta}_{n_c}}]_{n_c\in\mathcal{N}_c}, \tag{4}$$

where $\mathcal{N}_c \triangleq \{1, \ldots, N_c\}$, and

$$\boldsymbol{\theta}_{n_c} \triangleq (\boldsymbol{\theta}_{n_c,1}, \ldots, \boldsymbol{\theta}_{n_c,L})^T \in \mathbb{R}^L, n_c \in \mathcal{N}_c. \tag{5}$$

For practical implementations, the phase shifts $\boldsymbol{\theta}_{n_c,\ell_c}$ are quantized at b-bit resolution [35], i.e.

$$\boldsymbol{\theta}_{n_c,\ell_c} \in \mathcal{B} \triangleq \{b'\frac{2\pi}{2^b}, b' = 0, 1, \ldots, 2^b - 1\}. \tag{6}$$

In what follows, the projection of $\alpha \in [0, 2\pi)$ into $\mathcal{B}$ denoted by $\lfloor\alpha\rceil_b$ is referred to as its b-bit rounded version:

$$\lfloor\alpha\rceil_b = \nu_\alpha \frac{2\pi}{2^b} \tag{7}$$

with

$$\nu_\alpha \triangleq \arg\min_{\nu=0,1,\ldots,2^b} \left|\nu\frac{2\pi}{2^b} - \alpha\right|. \tag{8}$$

When $b = \infty$, it is true that

$$\alpha = \lfloor\alpha\rceil_\infty. \tag{9}$$

In the sequence, the 2D index (k, n_u) refers to the n_u-th antenna of UE k. The system bandwidth Δ is divided into $\mu = 2^{\bar{\mu}}$ subbands with frequency grid points $f_{\mu'} \triangleq \mu'\Delta/\mu$, $\mu' = 0, \ldots, \mu - 1$, thereby enabling OFDM transmission with μ sub-carriers. Accordingly, the sampling time interval is $T = 1/\Delta$.

The channel between the n-th DL BS TA and the (k, n_u)-th UE receive antenna (RA) is assumed to be frequency-selective over Δ, and it is characterized by the transfer function:

$$h_{k,n_u}^{(n)}(\zeta) \triangleq \sum_{m=0}^{M-1} h_{k,n_u,m}^{(n)} \zeta^{-m}, \tag{10}$$

where $h_{k,n_u,m}^{(n)} \in \mathbb{C}$ denotes the gain of the m-th path $h_{k,n_u}^{(n)}(t - mT)$ between the n-th DL BS TA and the (k, n_u)-th RA, M is the number of resolvable paths, and $\zeta \in \mathbb{C}$ lies on the unit circle $|\zeta| = 1$.

Accordingly, the DL channel spanning from the BS to the (k, n_u)-th RA is characterized by the transfer co-vector function

$$\mathbb{C}^{1\times N} \ni h_{k,n_u}(\zeta) \triangleq \sum_{m=0}^{M-1} h_{k,n_u,m}\zeta^{-m}, \tag{11}$$

where we have:

$$h_{k,n_u,m} \triangleq \begin{bmatrix} h_{k,n_u,m}^{(1)} & \dots & h_{k,n_u,m}^{(N)} \end{bmatrix} \in \mathbb{C}^{1\times N}. \tag{12}$$

In short, the frequency selective fading MIMO channel spanning from the BS to UE k is described by the transfer matrix

$$H_k(\zeta) = \sum_{m=0}^{M-1} H_{k,m}\zeta^{-m} \tag{13}$$

where we have:

$$\begin{aligned} \mathbb{C}^{N\times N_u} \ni H_{k,m} &\triangleq \begin{bmatrix} h_{k,1,m} \\ \dots \\ h_{k,N_u,m} \end{bmatrix} \\ &= \begin{bmatrix} h_{k,1,m}^{(1)} & \dots & h_{k,1,m}^{(N)} \\ \dots & \dots & \dots \\ h_{k,N_u,m}^{(1)} & \dots & h_{k,N_u,m}^{(N)} \end{bmatrix}, \\ &(k,m) \in \mathcal{K} \times \{0,\dots,M-1\}, \end{aligned} \tag{14}$$

and the stacked channel vector for all receive antennas and resolvable channel taps is given by

$$h_k \triangleq \begin{bmatrix} \mathsf{vec}(H_{k,0}) \\ \dots \\ \mathsf{vec}(H_{k,M-1}) \end{bmatrix} \in \mathbb{C}^{MNN_u} \tag{15}$$

Bearing in mind that the mmWave propagation can be characterized by a scattered channel model, each resolvable path m of the channel $H_{k,m}$ can be modeled as a summation of N_s scattering rays per resolvable path (cluster) [36, Eq. (3)]:

$$H_{k,m} = \frac{\bar{\rho}_k}{N_s} \sum_{s=1}^{N_s} \alpha_{k,m,s}\ a_r\left(\phi_{k,m,s}^r\right) a_t^H\left(\phi_{k,m,s}^t\right), \tag{16}$$

where $\bar{\rho}_k \triangleq \sqrt{10^{-\frac{\rho_k}{10}}}$ is the path-loss experienced by the UE k. Furthermore, ρ_k is the path-loss (in dB), $\alpha_{k,m,s} \sim \mathcal{CN}(0,1)$ is the complex gain of the sth scattering ray in the mth cluster between the BS and the UE k, $\phi_{k,m,s}^t$ is the angle of departure for the sth scattering ray in the mth cluster arriving from the BS to the UE k, $\phi_{k,m,s}^r$ is the angle of arrival for the sth scattering ray in the mth cluster from the BS to the UE k, and $a_t\left(\phi_{k,m,s}^t\right) \in \mathbb{C}^{N\times 1}$ and $a_r\left(\phi_{k,m,s}^r\right) \in \mathbb{C}^{N_u\times 1}$ represent the normalized transmit and receive antenna array response vectors, respectively, where we have $a(\phi) = \frac{1}{\sqrt{\bar{N}}}\left[1, e^{j\frac{2\pi}{\lambda}d\sin(\phi)}, \dots, e^{j\frac{2\pi}{\lambda}(\bar{N}-1)d\sin(\phi)}\right]^T$ for antenna spacing d and $\bar{N} \in \{N, N_u\}$ for the transmit and receive antenna response vectors, respectively.

At the n-th TA, each block of symbol

$$x_n \triangleq \begin{bmatrix} x_n(0) \\ \dots \\ x_n(\mu-1) \end{bmatrix} \in \mathbb{C}^{\mu}, n \in \mathcal{N}, \tag{17}$$

of length μ goes through an OFDM operator to form the OFDM block of length $\mu + M$:

$$\tilde{x}_n \triangleq \begin{bmatrix} x_{n,T} \\ x_{n,H} \\ x_{n,T} \end{bmatrix} \in \mathbb{C}^{\mu+M} \tag{18}$$

along with[1]

$$\begin{bmatrix} x_{n,H} \\ x_{n,T} \end{bmatrix} = F_\mu^H x_n, x_{n,H} \in \mathbb{C}^{\mu-M}, x_{n,T} \in \mathbb{C}^M, \tag{19}$$

where the OFDM cyclic prefix (CP) length is set to M to avoid inter-block interference (IBI).

The OFDM block $\tilde{x}_n$ of the length $\mu + M$ is transmitted from the n-th BS antenna. By discarding the first M entries of the received block and then applying FFT, the signal received at the (k, n_u)-th TA is formulated as:

$$\begin{bmatrix} y_{k,n_u}(0) \\ \dots \\ y_{k,n_u}(\mu-1) \end{bmatrix} = \begin{bmatrix} h_{k,n_u}(e^{\jmath 2\pi f_0})x(0) \\ \dots \\ h_{k,n_u}(e^{\jmath 2\pi f_{\mu-1}})x(\mu-1) \end{bmatrix} + F_\mu \nu_{k,n_u}, \tag{20}$$

where $\nu_{k,n_u} \in \mathbb{C}^\mu$ is the background noise having variance of σI_μ, and

$$x(\mu') \triangleq \begin{bmatrix} x_1(\mu') \\ \dots \\ x_N(\mu') \end{bmatrix} \in \mathbb{C}^N, \mu' = 1,\dots,\mu-1. \tag{21}$$

Let us now write (20) in the sub-carrier-wise form

$$y_{k,n_u}(\mu') = h_{k,n_u}(e^{\jmath 2\pi f_{\mu'}})x(\mu') + n_{k,n_u}(\mu'), \mu' = 0,\dots,\mu-1, \tag{22}$$

where $n_{k,n_u}(\mu')$ is the background noise of power σ. Therefore, the sub-carrier-wise receiver equation at UE k becomes:

$$\begin{aligned} y_k(\mu') &\triangleq \begin{bmatrix} y_{k,1}(\mu') \\ \dots \\ y_{k,N_u}(\mu') \end{bmatrix} && (23) \\ &= H_k^f(\mu')x(\mu') + n_k(\mu'), && (24) \end{aligned}$$

where we have:

$$\mathbb{C}^{N_u\times N} \ni H_k^f(\mu') \triangleq H_k(e^{\jmath 2\pi f_{\mu'}}) = \begin{bmatrix} h_{k,1}(e^{\jmath 2\pi f_{\mu'}}) \\ \dots \\ h_{k,N_u}(e^{\jmath 2\pi f_{\mu'}}) \end{bmatrix} \tag{25}$$

which is also partitioned as

$$\begin{aligned} H_k^f(\mu') &= \begin{bmatrix} H_{k,1}^f(\mu') & \dots & H_{k,N_c}^f(\mu') \end{bmatrix}, \\ H_{k,n_c}^f(\mu') &= \begin{bmatrix} H_{k,n_c}^f(\mu',1) \\ \dots \\ H_{k,n_c}^f(\mu',N_u) \end{bmatrix} \in \mathbb{C}^{N_u\times L}. \end{aligned} \tag{26}$$

Furthermore,

$$n_k(\mu') \triangleq \begin{bmatrix} n_k(\mu',1) \\ \dots \\ n_k(\mu',N_u) \end{bmatrix} \in \mathcal{C}(0,\sigma I_{N_u}). \tag{27}$$

[1] $F_\mu \triangleq \frac{1}{\mu}[e^{-\jmath 2\pi kp/\mu}]_{k,p=1,\dots,\mu-1}$ which is the the fast Fourier transform (FFT) matrix of order μ, and accordingly F_μ^H is the inverse FFT (IFFT) matrix

Let Q training symbols be inserted at the ψ_0-th, ψ_1-th, ..., ψ_{Q-1}-th sub-carriers for channel estimation. Note that we have

$$H_k^f(\psi_q) = \sum_{m=0}^{M-1} H_{k,m} W_\mu^{m\psi_q}, q = 0, \ldots, Q-1, \tag{28}$$

for

$$W_\mu \triangleq e^{-\jmath 2\pi\Delta/\mu}. \tag{29}$$

Therefore, we can write (24) at ψ_q as

$$\begin{aligned} y_k(\psi_q) &= \left(\sum_{m=0}^{M-1} H_{k,m} W_\mu^{m\psi_q}\right) x(\psi_q) + n_k(\psi_q) \\ &= \Lambda_{\psi_q}(x(\psi_q))h_k + n_k(\psi_q), \end{aligned} \tag{30}$$

where Λ_{ψ_q} is a linear operator from $\mathbb{C}^N$ to $\mathbb{C}^{N_u \times (MNN_u)}$ defined by

$$\begin{aligned} \Lambda_{\psi_q}(x(\psi_q)) \triangleq \Big[W_\mu^{\psi_q \cdot 0} x^T(\psi_q) \quad \ldots \quad W_\mu^{\psi_q (M-1)} x^T(\psi_q)\Big] \\ \otimes I_{N_u}. \end{aligned} \tag{31}$$

Upon defining

$$X_{\rm tr} \triangleq \begin{bmatrix} x^T(\psi_0) \\ \ldots \\ x^T(\psi_{Q-1}) \end{bmatrix} \in \mathbb{C}^{Q \times N}, \tag{32}$$

we can write (30) as

$$y_{k,\rm tr} = \widehat{\Lambda}(X_{\rm tr})h_k + n_k, \tag{33}$$

where we have

$$y_{k,\rm tr} \triangleq \begin{bmatrix} y_k(\psi_0) \\ \ldots \\ y_k(\psi_{Q-1}) \end{bmatrix} \in \mathbb{C}^{QN_u}, n_k \triangleq \begin{bmatrix} n_k(\psi_0) \\ \ldots \\ n_k(\psi_{Q-1}) \end{bmatrix}$$

and

$$\begin{aligned} \widehat{\Lambda}(X_{\rm tr}) &\triangleq \begin{bmatrix} \Lambda_{\psi_0}(x(\psi_0)) \\ \ldots \\ \Lambda_{\psi_{Q-1}}(x(\psi_{Q-1})) \end{bmatrix} \\ &= \mathcal{F}(X_{\rm tr}) \otimes I_{N_u} \in \mathbb{C}^{(QN_u)\times(MNN_u)} \end{aligned} \tag{34}$$

along with

$$\mathcal{F}(X_{\rm tr}) \triangleq \begin{bmatrix} F_0 X_{\rm tr} & \ldots & F_{M-1} X_{\rm tr} \end{bmatrix} \in \mathbb{C}^{Q \times (MN)} \tag{35}$$

and

$$F_m \triangleq \mathsf{diag}[W_\mu^{\psi_q m}]_{q=0,\ldots,Q-1} \in \mathbb{C}^{Q\times Q}. \tag{36}$$

To harness Wiener filter for channel estimation, we need to find the channel correlation matrix. In this regard, the channel matrix $H_{k,m}$ can be modeled as

$$H_{k,m} = R_{R,k,m}^{1/2} H_{W,k,m} (R_{T,k,m}^{1/2})^{1/2}, \tag{37}$$

where the deterministic Hermitian $0 \preceq R_{T,k,m}^{1/2} \in \mathbb{C}^{N\times N}$ ($0 \preceq R_{R,k,m}^{1/2} \in \mathbb{C}^{N_u \times N_u}$, resp.) models the correlation among the TAs (RAs, resp.). The matrix $H_{W,k,m}$ in (37) is a stationary process, whose elements are independent and identical distributed circularly symmetric complex Gaussian random variables with unit variance. Thus, covariance of channel h_k is given by

$$R_{h_k} \triangleq \mathbb{E}(h_k (h_k)^H) = \mathsf{diag}[R_{T,k,m} \otimes R_{R,k,m}]_{m=0,\ldots,M-1}. \tag{38}$$

Thus, the Wiener filter harnessed for the estimation of h_k is formulated as:

$$\hat{h}_k = R_{h_k} \widehat{\Lambda}^H(X_{tr}) \left([\widehat{\Lambda}(X_{tr})\sqrt{R_{h_k}}]^2 + \sigma I_{QN_u}\right)^{-1} y_{k,\rm tr}. \tag{39}$$

The computational complexity of the matrix multiplications and inverse operation in (39) is $\mathcal{O}\big((MNN_u)^2 QN_u + (MNN_u)(QN_u)^2 + (MNN_u)QN_u + (QN_u)^3\big)$. Since the product MN is of similar or order as the training size Q, the overall dominant complexity can be expressed as $\mathcal{O}\big((MNN_u)^2 QN_u\big)$.

The resultant MMSE for the estimator in (39) becomes:

$$\langle R_{h_k}\rangle - \langle [\widehat{\Lambda}(X_{tr})R_{h_k}]^2 \left([\widehat{\Lambda}(X_{tr})\sqrt{R_{h_k}}]^2 + \sigma I_{QN_u}\right)^{-1} \rangle. \tag{40}$$

Under the above described AoSA, using N_c RF chains, $x(\psi_q)$ is sought in the form of

$$x(\psi_q) = \mathcal{D}(\boldsymbol{\theta}) s(\psi_q) \tag{41}$$

with $\mathcal{D}(\boldsymbol{\theta}) \in \mathbb{C}^{N\times N_c}$ defined in (4) and $s(\psi_q) \in \mathbb{C}^{N_c}$. So

$$\mathbb{C}^{Q\times N} \ni X_{\rm tr} = \begin{bmatrix} s^T(\psi_0)\mathbf{V}_A^T \\ \ldots \\ s^T(\psi_0)\mathbf{V}_A^T \end{bmatrix} = \mathbf{S}_{\rm tr}\mathbf{V}_A^T$$

for

$$\begin{aligned} \mathbf{S}_{\rm tr} &\triangleq \begin{bmatrix} \mathbf{S}_{\rm tr}(1,1) & \ldots & \mathbf{S}_{\rm tr}(1,N_c) \\ \ldots & \ldots & \ldots \\ \mathbf{S}_{\rm tr}(Q,1) & \ldots & \mathbf{S}_{\rm tr}(Q,N_c) \end{bmatrix} \\ &= \begin{bmatrix} s^T(\psi_0) \\ \ldots \\ s^T(\psi_{Q-1}) \end{bmatrix} \in \mathbb{C}^{Q\times N_c}. \end{aligned} \tag{42}$$

Given the training power budget $P_{\rm tr}$, the DSPs are constrained by the power limit of

$$\sum_{q=0}^{Q-1} ||x(\psi_q)||^2 = L \sum_{q=0}^{Q-1} ||s(\psi_q)||^2 = L||\mathbf{S}_{\rm tr}||^2 \le P_{\rm tr}. \tag{43}$$

Thus, the problem of jointly designing AoSA-based AB $\mathcal{D}(\boldsymbol{\theta})$ and DSPs $\mathbf{S}_{\rm tr}$ to minimize the MMSE in (40) is formulated as:

$$\begin{aligned} \min_{\boldsymbol{\theta}, \mathbf{S}_{\rm tr}} \sum_{k=1}^{K} \Big[\langle R_{h_k}\rangle - \Big\langle [\widehat{\Lambda}(\mathbf{S}_{\rm tr}\mathcal{D}^T(\boldsymbol{\theta}))R_{h_k}]^2 \\ \left([\widehat{\Lambda}(\mathbf{S}_{\rm tr}\mathcal{D}^T(\boldsymbol{\theta}))\sqrt{R_{h_k}}]^2 + \sigma I_{QN_u}\right)^{-1} \Big\rangle\Big] \\ \text{s.t.} \quad (6), (43), \end{aligned} \tag{44}$$

or equivalently asL

$$\max_{\boldsymbol{\theta}, \mathbf{S}_{\rm tr}} F(\boldsymbol{\theta}, \mathbf{S}_{\rm tr}) \quad \text{s.t.} \quad (6), (43), \tag{45}$$

where we have

$$F(\boldsymbol{\theta}, \mathbf{S}_{\rm tr}) \triangleq \sum_{k=1}^{K} f_k(\boldsymbol{\theta}, \mathbf{S}_{\rm tr}) \tag{46}$$

along with

$$f_k(\boldsymbol{\theta}, \mathbf{S}_{\rm tr}) \triangleq \Big\langle [\widehat{\Lambda}(\mathbf{S}_{\rm tr}\mathcal{D}^T(\boldsymbol{\theta}))R_{h_k}]^2 \Big(\big[\widehat{\Lambda}(\mathbf{S}_{\rm tr}\mathcal{D}^T(\boldsymbol{\theta}))\sqrt{R_{h_k}}\big]^2 + \sigma I_{QN_u}\Big)^{-1}\Big\rangle, \quad (47)$$

and according to (34)-(35), we have:

$$\widehat{\Lambda}(\mathbf{S}_{\rm tr}\mathcal{D}^T(\boldsymbol{\theta})) = \big[F_0\mathbf{S}_{\rm tr}\mathcal{D}^T(\boldsymbol{\theta}) \quad \ldots \quad F_{M-1}\mathbf{S}_{\rm tr}\mathcal{D}^T(\boldsymbol{\theta})\big] \otimes I_{N_u} \quad (48)$$

$$= \big[(F_0\mathbf{S}_{\rm tr}\mathcal{D}^T(\boldsymbol{\theta})) \otimes I_{N_u} \quad \ldots \quad (F_{M-1}\mathbf{S}_{\rm tr}\mathcal{D}^T(\boldsymbol{\theta})) \otimes I_{N_u}\big]. \quad (49)$$

It becomes immediately apparent that (45) represents a high-dimensional optimization problem involving both a continuous matrix variable $\mathbf{S}_{\rm tr}$ and a discrete vector variable $\boldsymbol{\theta}$. This combination of continuous and discrete elements across potentially large dimensions makes the problem exceptionally challenging from a computational perspective.

To address (45), we propose a new path-following procedure. Starting from a feasible point $(S_{\rm tr}^{(0)}, \theta^{(0)})$ for (45), let $(S^{(\kappa)}, \theta^{(\kappa)})$ denote a feasible point obtained at the κ-th iteration. The path-following iteration updates $(S^{(\kappa+1)}, \theta^{(\kappa+1)})$ as follows.

A. Digital pilot path-following

With $\boldsymbol{\theta}$ held fixed at $\theta^{(\kappa)}$ we consider the following program of maximizing over $\mathbf{S}_{\rm tr}$:

$$\max_{\mathbf{S}_{\rm tr}} \; F_1(\mathbf{S}_{\rm tr}) \quad \text{s.t.} \quad (43), \quad (49)$$

where we have:

$$F_1(\mathbf{S}_{\rm tr}) \triangleq F(\theta^{(\kappa)}, \mathbf{S}_{\rm tr}) = \sum_{k=1}^{K} f_{1,k}(\mathbf{S}_{\rm tr}) \quad (50)$$

along with

$$f_{1,k}(\mathbf{S}_{\rm tr}) \triangleq \langle [\widehat{\Lambda}_1(\mathbf{S}_{\rm tr})]^2 \Big([\widehat{\Lambda}_2(\mathbf{S}_{\rm tr})]^2 + \sigma I_{QN_u}\Big)^{-1}\rangle \quad (51)$$

and according to (38), we have:

$$\widehat{\Lambda}_1(\mathbf{S}_{\rm tr}) \triangleq \widehat{\Lambda}(\mathbf{S}_{\rm tr}\mathcal{D}^T(\theta^{(\kappa)}))R_{h_k} = \mathsf{Row}[(F_m\mathbf{S}_{\rm tr}D_1^T R_{T,k,m}) \otimes R_{R,k,m}]_{m=0,\ldots,M-1}, \quad (52)$$

for $\widehat{\Lambda}_1(\mathbf{S}_{\rm tr}) \in \mathbb{C}^{(QN_u)\times(MNN_u)}$, and

$$\widehat{\Lambda}_2(\mathbf{S}_{\rm tr}) \triangleq \widehat{\Lambda}(\mathbf{S}_{\rm tr}\mathcal{D}^T(\theta^{(\kappa)}))\sqrt{R_{h_k}} = \mathsf{Row}[(F_m\mathbf{S}_{\rm tr}D_1^T\sqrt{R_{T,k,m}}) \otimes \sqrt{R_{R,k,m}}]_{m=0,\ldots,M-1}. \quad (53)$$

for $\widehat{\Lambda}_2(\mathbf{S}_{\rm tr}) \in \mathbb{C}^{(QN_u)\times(MNN_u)}$ and $D_1 \triangleq \mathcal{D}(\theta^{(\kappa)})$.

Applying the inequality (1) for $(\boldsymbol{X}, \mathbf{Y}) = (\widehat{\Lambda}_1(\mathbf{S}_{\rm tr}), [\widehat{\Lambda}_2(\mathbf{S}_{\rm tr})]^2 + \sigma I_{QN_u}) \in \mathbb{C}^{(QN_u)\times(MNN_u)} \times \mathbb{C}^{(QN_u)\times(QN_u)}$ and

$$(\bar{X}, \bar{Y}) = (X_{1,k}, Y_{1,k}) = (\widehat{\Lambda}_1(S_{\rm tr}^{(\kappa)}), [\widehat{\Lambda}_2(S_{\rm tr}^{(\kappa)})]^2 + \sigma I_{QN_u}) \quad (54)$$

yields the tight minorant of $f_{1,k}(\mathbf{S}_{\rm tr})$ at $S_{\rm tr}^{(\kappa)}$, given by (55) at the top of the page, where

$$a_{1,k} \triangleq -\sigma||(Y_{1,k})^{-1}X_{1,k}||^2, \quad \mathbb{C}^{(MNN_u)\times(QN_u)} \ni A_{1,k} \triangleq (X_{1,k})^H(Y_{1,k})^{-1}, \quad (56)$$

which is partitioned as

$$A_{1,k} = \begin{bmatrix} A_{1,k,0} \\ \cdots \\ A_{1,k,M-1} \end{bmatrix}, \quad A_{1,k,m} \in \mathbb{C}^{(NN_u)\times(QN_u)}, m = 0, \ldots, M-1. \quad (57)$$

For $q = 1, \ldots, Q$ and $n_c = 1, \ldots, N_c$, let $\mathcal{I}_{q,n_c} \in \mathbb{R}^{Q\times N_c}$ be a matrix with zero entries but

$$\mathcal{I}_{q,n_c}(q, n_c) = 1. \quad (58)$$

Then we have:

$$\mathbf{S}_{\rm tr} = \sum_{q=1}^{Q}\sum_{n_c=1}^{N_c} \mathbf{S}_{\rm tr}(q, n_c)\mathcal{I}_{q,n_c}(q, n_c). \quad (59)$$

Let us now define $B_{1,k} \in \mathbb{C}^{N_c\times Q}$ as

$$B_{1,k}(n_c, q) \triangleq \sum_{m=0}^{M-1} \Big\langle A_{1,k,m}\Big[(F_m\mathcal{I}_{q,n_c}\mathcal{D}^T(\theta^{(\kappa)})R_{T,k,m}) \otimes R_{R,k,m}]\Big\rangle, n_c = 1, \ldots, N_c; q = 1, \ldots, Q. \quad (60)$$

Then we can write:

$$\Big\langle \sum_{m=0}^{M-1}\Big(A_{1,k,m}\Big[(F_m\mathbf{S}_{\rm tr}\mathcal{D}^T(\theta^{(\kappa)})R_{T,k,m}) \otimes R_{R,k,m}\Big]\Big)\Big\rangle = \langle B_{1,k}\mathbf{S}_{\rm tr}\rangle = b_{1,k}\mathsf{vec}(\mathbf{S}_{\rm tr}), \quad (61)$$

where we have:

$$b_{1,k} \triangleq (\mathsf{vec}((B_{1,k})^T))^T. \quad (62)$$

Furthermore, upon defining $C_{1,k,q,n_c,m} \in \mathbb{C}^{(MNN_u)\times(NN_u)}$ as

$$C_{1,k,q,n_c,m} \triangleq A_{1,k}\Big[(F_m\mathcal{I}_{q,n_c}\mathcal{D}^T(\theta^{(\kappa)})\sqrt{R_{T,k,m}}) \otimes \sqrt{R_{R,k,m}}\Big], q = 1, \ldots, Q; n_c = 1, \ldots, N_c, \quad (63)$$

we can proceed as follows:

$$\sum_{m=0}^{M-1}||A_{1,k}\Big[(F_m\mathbf{S}_{\rm tr}\mathcal{D}^T(\theta^{(\kappa)})\sqrt{R_{T,k,m}}) \otimes \sqrt{R_{R,k,m}}\Big]||^2 = \sum_{m=0}^{M-1}||\sum_{q=1}^{Q}\sum_{n_c=1}^{N_c}\mathbf{S}_{\rm tr}(q, n_c)C_{1,k,q,n_c,m}||^2 = \sum_{m=0}^{M-1}\sum_{i=1}^{MNN_u}\sum_{j=1}^{NN_u}|\sum_{q=1}^{Q}\sum_{n_c=1}^{N_c}\mathbf{S}_{\rm tr}(q, n_c)C_{1,k,q,n_c,m}(i,j)|^2 = \sum_{m=0}^{M-1}\sum_{i=1}^{MNN_u}\sum_{j=1}^{NN_u}|\langle D_{1,k,i,j,m}\mathbf{S}_{\rm tr}\rangle|^2 =$$

$$\begin{aligned}\tilde{f}_{1,k}(\mathbf{S}_{\rm tr}) &= 2\Re\{\langle (X_{1,k})^H (Y_{1,k})^{-1}\widehat{\Lambda}_1(\mathbf{S}_{\rm tr})\rangle\}\\ &\quad -\Big\langle [(Y_{1,k})^{-1}X_{1,k}]^2\left([\widehat{\Lambda}_2(\mathbf{S}_{\rm tr})]^2+\sigma I_{QN_u}\right)\Big\rangle\\ &= a_{1,k}+2\Re\{\langle A_{1,k}\mathsf{Row}[(F_m\mathbf{S}_{\rm tr}\mathcal{D}^T(\theta^{(\kappa)})R_{T,k,m})\otimes R_{R,k,m}]_{m=0,\ldots,M-1}\rangle\}\\ &\quad -||A_{1,k}\mathsf{Row}[(F_m\mathbf{S}_{\rm tr}\mathcal{D}^T(\theta^{(\kappa)})\sqrt{R_{T,k,m}})\otimes\sqrt{R_{R,k,m}}]_{m=0,\ldots,M-1}||^2\\ &= a_{1,k}+2\sum_{m=0}^{M-1}\Re\{\langle A_{1,k,m}\left[(F_m\mathbf{S}_{\rm tr}\mathcal{D}^T(\theta^{(\kappa)})R_{T,k,m})\otimes R_{R,k,m}\right]\rangle\\ &\quad -\sum_{m=0}^{M-1}||A_{1,k}\left[(F_m\mathbf{S}_{\rm tr}\mathcal{D}^T(\theta^{(\kappa)})\sqrt{R_{T,k,m}})\otimes\sqrt{R_{R,k,m}}\right]||^2, \end{aligned}\tag{55}$$

$$\sum_{m=0}^{M-1}\sum_{i=1}^{MNN_u}\sum_{j=1}^{NN_u}|d_{1,k,i,j,m}\mathsf{vec}(\mathbf{S}_{\rm tr})|^2 = (\mathsf{vec}(\mathbf{S}_{\rm tr})^H D_{1,k}\mathsf{vec}(\mathbf{S}_{\rm tr}),\tag{64}$$

where $D_{1,k,i,j,m}\in\mathbb{C}^{N_c\times Q}$ is defined as

$$D_{1,k,i,j,m}(n_c,q)\triangleq C_{1,k,q,n_c,m}(i,j),\quad n_c=1,\ldots,N_c; q=1,\ldots,Q,\tag{65}$$

and

$$d_{1,k,i,j,m}\triangleq(\mathsf{vec}((D_{1,k,i,j,m})^T))^T,\quad i=1,\ldots,MNN_u; j=1,\ldots,NN_u,\tag{66}$$

along with

$$D_{1,k}\triangleq\sum_{m=0}^{M-1}\sum_{i=1}^{MNN_u}\sum_{j=1}^{NN_u}(d_{1,k,i,j,m})^H d_{1,k,i,j,m}.\tag{67}$$

From (61)-(62) and (64)-(67), we obtain the following tight minorant of the objective function $F_1(\mathbf{S}_{\rm tr})$ in (49):

$$\tilde{F}_1(\mathbf{S}_{\rm tr})\triangleq\sum_{k=1}^{K}[a_{1,k}+2\Re\{b_{1,k}\mathsf{vec}(\mathbf{S}_{\rm tr})\} -(\mathsf{vec}(\mathbf{S}_{\rm tr})^H D_{1,k}\mathsf{vec}(\mathbf{S}_{\rm tr})\big].\tag{68}$$

We then solve the following problem of minorant maximization for (49) to update $S_{\rm tr}^{(\kappa+1)}$:

$$\max_{\mathbf{S}_{\rm tr}}\ \tilde{F}_1(\mathbf{S}_{\rm tr})\quad\text{s.t.}\quad(43),\tag{69}$$

which admits the closed form solution of computational complexity $\mathcal{O}\left(QN_c\log_2(QN_c)\right)$, given by (70) at the top of the page, where $\lambda>0$ is found by bisection such that $||(D_1+\lambda I_{QN_c})^{-1}(b_1)^H||^2=P_{\rm tr}/L$, for

$$b_1\triangleq\sum_{k=1}^{K}b_{1,k}\quad\&\quad D_1\triangleq\sum_{k=1}^{K}D_{1,k}.\tag{71}$$

The monotonic convergence of the digital pilot path-following can be analytically shown as follows:

$$\begin{aligned}F(\theta^{(\kappa)},S_{\rm tr}^{(\kappa+1)}) &= F_1(S_{\rm tr}^{(\kappa+1)})\\ &\geq \tilde{F}_1(S_{\rm tr}^{(\kappa+1)})\\ &> \tilde{F}_1(S_{\rm tr}^{(\kappa)})\\ &= F(\theta^{(\kappa)},S_{\rm tr}^{(\kappa)})\end{aligned}\tag{72}$$

as far as $S_{\rm tr}^{(\kappa+1)}\neq S_{\rm tr}^{(\kappa)}$.

B. Analog beamformer path-following

With $\mathbf{S}_{\rm tr}$ held fixed at $S_{\rm tr}^{(\kappa+1)}$ we consider the following program of maximizing over $\boldsymbol{\theta}$:

$$\max_{\boldsymbol{\theta}}\ F_2(\boldsymbol{\theta})\quad\text{s.t.}\quad(6),\tag{73}$$

where we have

$$F_2(\boldsymbol{\theta})\triangleq F(\boldsymbol{\theta},S_{\rm tr}^{(\kappa+1)})=\sum_{k=1}^{K}f_{2,k}(\boldsymbol{\theta})\tag{74}$$

along with

$$f_{2,k}(\boldsymbol{\theta})\triangleq\langle[\widehat{\Lambda}_3(\boldsymbol{\theta})]^2\left([\widehat{\Lambda}_4(\boldsymbol{\theta})]^2+\sigma I_{QN_u}\right)^{-1}\rangle\tag{75}$$

and according to (38), we have:

$$\widehat{\Lambda}_3(\boldsymbol{\theta})\triangleq\mathsf{Row}[(F_m S_{\rm tr}^{(\kappa+1)}\mathcal{D}^T(\boldsymbol{\theta})R_{T,k,m})\otimes R_{R,k,m}]_{m=0,\ldots,M-1},\tag{76}$$

for $\widehat{\Lambda}_3(\boldsymbol{\theta})\in\mathbb{C}^{(QN_u)\times(MNN_u)}$, and

$$\widehat{\Lambda}_4(\boldsymbol{\theta})\triangleq\mathsf{Row}[(F_m S^{(\kappa+1)}\mathcal{D}^T(\boldsymbol{\theta})\sqrt{R_{T,k,m}})\otimes\sqrt{R_{R,k,m}}]_{m=0,\ldots,M-1}.\tag{77}$$

for $\widehat{\Lambda}_4(\boldsymbol{\theta})\in\mathbb{C}^{(QN_u)\times(MNN_u)}$. Applying the inequality (1) for $(\boldsymbol{X},\mathbf{Y})=(\widehat{\Lambda}_3(\boldsymbol{\theta}),[\widehat{\Lambda}_4(\boldsymbol{\theta})]^2+\sigma I_{QN_u})\in\mathbb{C}^{(QN_u)\times(MNN_u)}\times\mathbb{C}^{(QN_u)\times(QN_u)}$ and

$$\begin{aligned}(\bar{X},\bar{Y}) &= (X_{2,k},Y_{2,k})\\ &= (\widehat{\Lambda}_3(\theta^{(\kappa)}),[\widehat{\Lambda}_4(\theta^{(\kappa)})]^2+\sigma I_{QN_u})\end{aligned}\tag{78}$$

yields the tight minorant of $f_{2,k}(\boldsymbol{\theta})$, which is given by (79) at the top of the page, where

$$\begin{aligned}a_{2,k}&\triangleq-\sigma||(Y_{2,k})^{-1}X_{2,k}||^2,\\ \mathbb{C}^{(MNN_u)\times(QN_u)}\ni A_{2,k}&\triangleq(X_{2,k})^H(Y_{2,k}k)^{-1},\end{aligned}\tag{80}$$

which is partitioned as

$$A_{2,k}=\begin{bmatrix}A_{2,k,0}\\ \cdots\\ A_{2,k,M-1}\end{bmatrix},\quad A_{2,k,m}\in\mathbb{C}^{(NN_u)\times(QN_u)},\ m=0,\ldots,M-1.\tag{81}$$

Let $e_\ell\triangleq(e_\ell(1),\ldots,e_\ell(L))\in\mathbb{R}^{1\times L}$, $\ell=1,\ldots,L$ with all zero entries but

$$e_\ell(\ell)=1.\tag{82}$$

$$\mathsf{vec}(S_{\rm tr}^{(\kappa+1)}) = \begin{cases} (D_1)^{-1}(b_1)^H & \text{if} \quad ||(D_1)^{-1}(b_1)^H|| \leq P_{\rm tr}/L \\ (D_1+\lambda I_{QN_c})^{-1}(b_1)^H & \text{othewise,} \end{cases} \tag{70}$$

$$\begin{aligned} f_{2,k}^{(\kappa)}(\boldsymbol{\theta}) &= 2\Re\{\langle (X_{2,k})^H (Y_{2,k})^{-1} \widehat{\Lambda}_3(\boldsymbol{\theta})\rangle\} \\ &\quad -\Big\langle [(Y_{2,k})^{-1} X_{2,k}]^2 \left([\widehat{\Lambda}_4(\boldsymbol{\theta})]^2 + \sigma I_{QN_u} \right) \Big\rangle \\ &= a_{2,k} + 2\Re\{\langle A_{2,k} \mathsf{Row}[(F_m S_{\rm tr}^{(\kappa+1)} \mathcal{D}^T(\boldsymbol{\theta}) R_{T,k,m}) \otimes R_{R,k,m}]_{m=0,\ldots,M-1},\rangle\} \\ &\quad -||A_{2,k} \mathsf{Row}[(F_m S_{\rm tr}^{(\kappa+1)} \mathcal{D}^T(\boldsymbol{\theta}) R_{T,k,m}) \otimes R_{R,k,m}]_{m=0,\ldots,M-1}||^2 \\ &= a_{2,k} + 2\sum_{m=0}^{M-1} \Re\{\langle A_{2,k,m} \left[(F_m S_{\rm tr}^{(\kappa+1)} \mathcal{D}^T(\boldsymbol{\theta}) R_{T,k,m}) \otimes R_{R,k,m} \right] \rangle \\ &\quad -\sum_{m=0}^{M-1} ||A_{2,k} \left[(F_m S_{\rm tr}^{(\kappa+1)} \mathcal{D}^T(\boldsymbol{\theta}) \sqrt{R_{T,k,m}}) \otimes \sqrt{R_{R,k,m}} \right] ||^2, \end{aligned} \tag{79}$$

Furthermore, let $\mathcal{A}_{n_c,\ell} \in \mathbb{R}^{N_c \times (LN_c)}$ with zero entries but

$$\mathcal{A}_{n_c,\ell}(n_c, (n_c-1)L+1 : n_c L) = e_\ell \mathcal{A}^T. \tag{83}$$

Then we arrive at:

$$\mathcal{D}^T(\boldsymbol{\theta}) = \sum_{n_c=1}^{N_c} \sum_{\ell=1}^{L} e^{\jmath \boldsymbol{\theta}_{n_c,\ell}} \mathcal{A}_{n_c,\ell}. \tag{84}$$

Therefore, we have

$$\begin{aligned} \sum_{m=0}^{M-1} \Big\langle A_{2,k,m} \left[(F_m S_{\rm tr}^{(\kappa+1)} \mathcal{D}^T(\boldsymbol{\theta}) R_{T,k,m}) \otimes R_{R,k,m} \right] \Big\rangle &= \\ \sum_{n_c=1}^{N_c} \sum_{\ell=1}^{L} e^{\jmath \boldsymbol{\theta}_{n_c,\ell}} \sum_{m=0}^{M-1} \Big\langle A_{2,k,m} \Big[(F_m S_{\rm tr}^{(\kappa+1)} \mathcal{A}_{n_c,\ell} & \\ \times R_{T,k,m}) \otimes R_{R,k,m}] \Big\rangle &= \\ \sum_{n_c=1}^{N_c} b_{2,k,n_c} e^{\jmath \boldsymbol{\theta}_{n_c}} &= \\ b_{2,k} e^{\jmath \boldsymbol{\theta}}, & \end{aligned} \tag{85}$$

where b_{2,k,n_c} and $b_{2,k}$ are defined in (86).

Furthermore, we have

$$\begin{aligned} A_{2,k} \left[(F_m S_{\rm tr}^{(\kappa+1)} \mathcal{D}^T(e^{\jmath\boldsymbol{\theta}}) \sqrt{R_{T,k,m}}) \otimes \sqrt{R_{R,k,m}} \right] &= \\ \sum_{n_c=1}^{N_c} \sum_{\ell=1}^{L} e^{\jmath \boldsymbol{\theta}_{n_c,\ell}} A_{2,k} \Big[(F_m S_{\rm tr}^{(\kappa+1)} \mathcal{A}_{n_c,\ell} \sqrt{R_{T,k,m}}) & \\ \otimes \sqrt{R_{R,k,m}} \Big] &= \\ \sum_{n_c=1}^{N_c} \sum_{\ell=1}^{L} e^{\jmath \boldsymbol{\theta}_{n_c,\ell}} C_{2,k,n_c,\ell,m} & \end{aligned} \tag{87}$$

for

$$\begin{aligned} C_{2,k,n_c,\ell,m} &\triangleq A_{2,k} \Big[(F_m S_{\rm tr}^{(\kappa+1)} \mathcal{A}_{n_c,\ell} \sqrt{R_{T,k,m}}) \\ &\quad \otimes \sqrt{R_{R,k,m}} \Big] \in \mathbb{C}^{(MNN_u) \times (NN_u)}, \\ &\quad n_c = 1, \ldots, N_c; \ell = 1, \ldots, L. \end{aligned} \tag{88}$$

Thus,

$$\begin{aligned} \sum_{m=0}^{M-1} \Big\| A_{2,k} \Big[(F_m S_{\rm tr}^{(\kappa+1)} \mathcal{D}^T(e^{\jmath\boldsymbol{\theta}}) \sqrt{R_{T,k,m}}) & \\ \otimes \sqrt{R_{R,k,m}} \Big] \Big\|^2 &= \\ \sum_{m=0}^{M-1} \sum_{i=1}^{MNN_u} \sum_{j=1}^{NN_u} | \sum_{n_c=1}^{N_c} \sum_{\ell=1}^{L} e^{\jmath \boldsymbol{\theta}_{n_c,\ell}} C_{2,k,n_c,\ell,m}(i,j)|^2 &= \\ \sum_{m=0}^{M-1} \sum_{i=1}^{MNN_u} \sum_{j=1}^{NN_u} | \sum_{n_c=1}^{N_c} d_{2,k,i,j,n_c,m} e^{\jmath \boldsymbol{\theta}_{n_c}}|^2 &= \\ \sum_{m=0}^{M-1} \sum_{i=1}^{MNN_u} \sum_{j=1}^{NN_u} |d_{2,k,i,j,m} e^{\jmath\boldsymbol{\theta}}|^2 &= \\ (e^{\jmath\boldsymbol{\theta}})^H D_{2,k} e^{\jmath\boldsymbol{\theta}}, & \end{aligned} \tag{89}$$

where $d_{2,k,i,j,n_c,m}$, $d_{2,k,i,j,m}$, and $D_{2,k}$ are defined in (90) at the top of the page. By (85) and (89), the function $f_{2,k}^{(\kappa)}(\boldsymbol{\theta})$ from (78)-(79) is represented in a compact form as

$$\begin{aligned} f_{2,k}^{(\kappa)}(\boldsymbol{\theta}) &= a_{2,k} + 2\Re\{b_{2,k} e^{\jmath\boldsymbol{\theta}}\} - (e^{\jmath\boldsymbol{\theta}})^H D_{2,k} e^{\jmath\boldsymbol{\theta}} \\ &= a_{2,k} + 2\Re\{b_{2,k} e^{\jmath\boldsymbol{\theta}}\} - (e^{\jmath\boldsymbol{\theta}})^H (D_{2,k} \\ &\quad -\lambda_{\max}(D_{2,k}) I_{N_c L}) e^{\jmath\boldsymbol{\theta}} - \lambda_{\max}(D_{2,k}) \|e^{\jmath\boldsymbol{\theta}}\|^2 \\ &\geq a_{2,k} - N_c L \lambda_{\max}(D_{2,k}) + 2\Re\{(b_{2,k} \\ &\quad -(e^{\jmath\theta^{(\kappa)}})^H (D_{2,k} - \lambda_{\max}(D_{2,k}) I_{N_c L}) \Big) e^{\jmath\boldsymbol{\theta}} \Big\} \\ &\quad +(e^{\jmath\theta^{(\kappa)}})^H (D_{2,k} - \lambda_{\max}(D_{2,k}) I_{N_c L}) e^{\jmath\theta^{(\kappa)}} \\ &= \tilde{a}_{2,k} + 2\Re\{\widetilde{b}_{2,k}^{(\kappa)} e^{\jmath\boldsymbol{\theta}}\}, \end{aligned} \tag{91}$$

where $\lambda_{\max}(D_{2,k})$ is the maximal eigenvalue of $D_{2,k}$ and

$$\begin{aligned} \tilde{a}_{2,k} &\triangleq a_{2,k} - N_c L \lambda_{\max}(D_{2,k}) + (e^{\jmath\theta^{(\kappa)}})^H \\ &\quad \times (D_{2,k} - \lambda_{\max}(D_{2,k}) I_{N_c L}) e^{\jmath\theta^{(\kappa)}}, \\ \widetilde{b}_{2,k}^{(\kappa)} &\triangleq b_{2,k} - (e^{\jmath\theta^{(\kappa)}})^H (D_{2,k} - \lambda_{\max}(D_{2,k}) I_{N_c L}). \end{aligned} \tag{92}$$

$$\begin{array}{rcl} b_{2,k,n_c} & \triangleq & \Big[\sum_{m=0}^{M-1}\langle A_{2,k,m}\Big[(F_m S_{\rm tr}^{(\kappa+1)}\mathcal{A}_{n_c,1}R_{T,k,m})\otimes R_{R,k,m}\Big]\rangle \ \ldots \\ & & \sum_{m=0}^{M-1}\langle A_{2,k,m}\Big[(F_m S_{\rm tr}^{(\kappa+1)}\mathcal{A}_{n_c,L}R_{T,k,m})\otimes R_{R,k,m}\Big]\rangle\Big] \in \mathbb{C}^{1\times L}, \\ b_{2,k} & \triangleq & \big[b_{2,k,1} \quad \ldots \quad b_{2,k,N_c}\big] \in \mathbb{C}^{1\times (LN_c)}. \end{array} \tag{86}$$

$$\begin{array}{c} d_{2,k,i,j,n_c,m} \triangleq \big[C_{2,k,n_c,1,m}(i,j) \quad \ldots \quad C_{2,k,n_c,L,m}(i,j)\big] \in \mathbb{C}^{1\times L}, n_c=1,\ldots,N_c, \\ d_{2,k,i,j,m} \triangleq \big[d_{2,k,i,j,1,m} \quad \ldots \quad d_{2,k,i,j,N_c,m}\big] \in \mathbb{C}^{1\times (N_cL)}, \\ D_{2,k} \triangleq \sum_{m=0}^{M-1}\sum_{i=1}^{MNN_u}\sum_{j=1}^{NN_u}(d_{2,k,i,j,m})^H d_{2,k,i,j,m}, \end{array} \tag{90}$$

Therefore, the following function provides a tight minorant of the objective $F_2(\boldsymbol{\theta})$ in (73):

$$\tilde{F}_2(\boldsymbol{\theta}) \triangleq \sum_{k=1}^{K}[\tilde{a}_{2,k} + 2\Re\{\widetilde{b}_{2,k}^{(\kappa)} e^{\jmath\boldsymbol{\theta}}\}]. \tag{93}$$

We now solve the following problem of minorant maximization for (73) to update $\theta^{(\kappa+1)}$;

$$\max_{\boldsymbol{\theta}} \ \tilde{F}_2(\boldsymbol{\theta}) \quad \text{s.t.} \quad (6), \tag{94}$$

which admits the closed form solution

$$\theta^{(\kappa+1)}(i) = \lfloor -\angle \widetilde{b}_2^{(\kappa)}(i)\rceil_b, i=1,\ldots,N_cL \tag{95}$$

of computational complexity $\mathcal{O}(N)$, for

$$\widetilde{b}_2^{(\kappa)} \triangleq \sum_{k=1}^{K}\widetilde{b}_{2,k}^{(\kappa)}. \tag{96}$$

We have

$$\begin{aligned} F(\theta^{(\kappa+1)}, S_{\rm tr}^{(\kappa+1)}) &= F_2(\theta^{(\kappa+1)}) \\ &\geq \tilde{F}_2(\theta^{(\kappa+1)}) \\ &> \tilde{F}_2(\theta^{(\kappa)}) \\ &= F(\theta^{(\kappa)}, S_{\rm tr}^{(\kappa+1)}) \end{aligned} \tag{97}$$

as far as $\theta^{(\kappa+1)} \neq \theta^{(\kappa)}$.

C. Path-following channel estimation algorithm

The pseudo-code outlining the implementation of the path-following method is presented in Algorithm 1. It follows from (72) and (97) that

$$F(\theta^{(\kappa+1)}, S_{\rm tr}^{(\kappa+1)}) > F(\theta^{(\kappa)}, S_{\rm tr}^{(\kappa)}) \tag{98}$$

holds whenever $(\theta^{(\kappa+1)}, S_{\rm tr}^{(\kappa+1)}) \neq (\theta^{(\kappa)}, S_{\rm tr}^{(\kappa)})$. Consequently, the sequence $(\theta^{(\kappa)}, S_{\rm tr}^{(\kappa)})$ of progressively improved points is path-following toward a solution of (45).

The scalability of the proposed algorithm is reflected by the computational complexities of the pilot and ABF optimizations presented in Sections II-B and II-C, respectively, a benefit of the closed-form solutions derived in (70) and (95). In particular, the complexity of the digital pilot optimization is $\mathcal{O}\big(QN_c\log_2(QN_c)\big)$, which scales log-linearly with the pilot size, while that of the ABF optimization is $\mathcal{O}(N)$, which scales linearly with the number of phase shifters. Both complexities are significantly lower than the cubic or higher-order polynomial complexity associated with iterative convex quadratic solvers.

Finally, the complexity of obtaining the channel estimate $\hat{h}_k$ using (39) is $\mathcal{O}\big((MNN_u)^2QN_u\big)$, which also remains computationally manageable. Note that the channel estimate is evaluated only once at the end (Step 3 of Alg. 1); therefore, its computational complexity is less critical compared to those of the pilot and ABF optimizations, which are updated iteratively in Step 2 of the path-following Alg. 1 over multiple iterations until convergence. As discussed above, the pilot and ABF optimizations exhibit low and scalable computational complexity, a benefit of the closed-form solutions derived.

Algorithm 1 Path-following channel estimation algorithm

1: **Initialization.** Randomly generate a feasible point $\Big(\theta^{(0)}, S_{\rm tr}^{(0)}\Big)$ for (45). Set $\kappa=0$.
2: **Path-following:** Update $S_{\rm tr}^{(\kappa+1)}$ by the closed-form expression (70)-(71), and $\theta^{(\kappa+1)}$ by (95)-(96). Increment $\kappa := \kappa+1$ until the objective in (45) converges.
3: **Output** $\Big(S_{\rm tr}^{(opt)}, \theta^{(opt)}\Big) = \Big(S_{\rm tr}^{(\kappa)}, \theta^{(\kappa)}\Big)$, the channel estimate $\hat{h}_k$ using (39), and the achieved MSE error $\sum_{k=1}^{K}[\langle R_{h_k}\rangle - f_k(\theta^{(opt)}, S_{\rm tr}^{(opt)}]$.

III. Impact of channel estimation on system performance

For channel estimation, we use Q subcarriers ψ_q, $q=0,\ldots,Q-1$, leaving the remaining $\mu-Q$ subcarriers ψ_q, $q=\in \mathcal{N}_I \triangleq \{Q,\ldots,\mu-1\}$ for information transmission to the users. Observe that the problem (45) is solved at the BS prior to signal transmission so its ABF solution $\bar{\mathcal{D}} \triangleq \mathcal{D}(\theta^{(opt)})$ is employed for both training and information transmission. In this section, we evaluate the system capacity of delivering information using the channel estimates, in order to assess the impact of channel estimation.

A. Single-user case

In this case, the information symbol $\bar{s}_q \in \mathcal{C}(0, I_{N_u})$, $q \in \mathcal{N}_I$ is beamformed by $\mathbf{P}_q \in \mathbb{C}^{N_c\times N_u}$, so by omitting the subscript k in (24) the receiver equation over sub-carrier ψ_q is formulated as

$$\begin{aligned} y_q &= H^f(\psi_q)\bar{\mathcal{D}}\mathbf{P}_q\bar{s}_q + n(\psi_q) \qquad (99) \\ &= \hat{H}^f(\psi_q)\bar{\mathcal{D}}\mathbf{P}_q\bar{s}_q + (H^f(\psi_q) - \hat{H}^f(\psi_q))\bar{\mathcal{D}}\mathbf{P}_q\bar{s}_q \\ &\quad + n(\psi_q), \qquad (100) \end{aligned}$$

where $H^f(\psi_q)$ is the estimate of $H^f(\psi_q)$. Note that $(H^f(\psi_q) - \hat{H}^f(\psi_q))\bar{\mathcal{D}}\mathbf{P}_q\bar{s}_q$ is uncorrelated with

$$P_q^{(\kappa+1)} = \begin{cases} \left(\Psi_q^H \Phi_q \Psi_q + \epsilon_q \langle \Phi_q \rangle I_{N_c}\right)^{-1} \left(\bar{X}_q^H \bar{Y}_q^{-1} \Psi_q\right)^H & \text{if} \\ \sum_{q=1}^{\mu} || \left(\Psi_q^H \Phi_q \Psi_q + \epsilon_q \langle \Phi_q \rangle I_{N_c}\right)^{-1} \left(\bar{X}_q^H \bar{Y}_q^{-1} \Psi_q\right)^H ||^2 \leq P_d, \\ \left(\Psi_q^H \Phi_q \Psi_q + \epsilon_q \langle \Phi_q \rangle I_{N_c} + \gamma I_{N_c}\right)^{-1} \left(\bar{X}_q^H \bar{Y}_q^{-1} \Psi_q\right)^H \\ \text{otherwise,} \end{cases} \tag{106}$$

$\hat{H}^f(\psi_q)\bar{\mathcal{D}}\mathbf{P}_q\bar{s}_q$ and considered as a noise with the zero mean and covariance of

$$\epsilon_q ||\mathbf{P}_q||^2 I_{N_u},$$

where $\epsilon_q \triangleq \mathbb{E}(||H^f(\psi_q) - \hat{H}^f(\psi_q)||^2) = \langle R_h \rangle - F(\theta^{opt}, S_{\text{tr}}^{opt})$ with $F(\theta^{opt}, S_{\text{tr}}^{opt})$ defined from (47). As such, we define the throughput of $\bar{s}_q$ as

$$r_q(\mathbf{P}_q) = \ln\left(I_{N_u} + [\Psi_q \mathbf{P}_q]^2 (\epsilon_q ||\mathbf{P}_q||^2 I_{N_u} + \sigma I_{N_u})^{-1}\right), \tag{101}$$

where $\Psi_q \triangleq \hat{H}^f(\psi_q)\bar{\mathcal{D}}$. We thus consider the following problem

$$\max_{\mathbf{P} \triangleq (\mathbf{P}_Q, \mathbf{P}_{\mu-1})} \quad g(\mathbf{P}) \triangleq \sum_{q=Q}^{\mu} r_q(\mathbf{P}_q) \quad \text{s.t.} \tag{102a}$$

$$\sum_{q=Q}^{\mu-1} ||\mathbf{P}_q||^2 \leq P_d, \tag{102b}$$

where P_d is the power budget of information transmission. Initialized by $P^{(0)}$ feasible for the power constraint (102b), let $P^{(\kappa)}$ be a feasible point obtained from the $\kappa-1$-th iteration. Applying the inequality (2) for $(\boldsymbol{X}, \mathbf{Y}) = (\Psi_q \mathbf{P}_q, \epsilon_q ||\mathbf{P}_q||^2 I_{N_u} + \sigma I_{N_u})$ and $(\bar{X}, \bar{Y}) = (\bar{X}_q, \bar{Y}_q) = (\Psi_q P_q^{(\kappa)}, \epsilon_q ||P_q^{(\kappa)}||^2 I_{N_u} + \sigma I_{N_u})$ yields the following tight minorant of $r_q(\mathbf{P}_q)$ at $P_q^{(\kappa)}$

$$\begin{aligned} \tilde{r}_q(\mathbf{P}_q) \quad \triangleq \quad & a_q + 2\Re\{\langle \bar{X}_q^H \bar{Y}_q^{-1} \Psi_q \mathbf{P}_q \rangle\} \\ & - \left\langle \Phi_q, [\Psi_q \mathbf{P}_q]^2 + \epsilon_q ||\mathbf{P}_q||^2 I_{N_u} \right\rangle, \end{aligned} \tag{103}$$

where we have:

$$\begin{aligned} \Phi_q &\triangleq \bar{Y}_q^{-1} - (\bar{Y}_q + [\bar{X}_q]^2)^{-1}, \\ a_q &\triangleq r_q(P_q^{(\kappa)}) - \langle [\bar{X}_q]^2 \bar{Y}_q^{-1} \rangle - \sigma \langle \Phi_q \rangle. \end{aligned} \tag{104}$$

We thus solve the following convex problem of tight minorant maximization for (102) to update $P^{(\kappa+1)}$:

$$\max_{\mathbf{P}} \sum_{q=Q}^{\mu-1} \tilde{r}_q(\mathbf{P}_q) \quad \text{s.t.} \quad (102b), \tag{105}$$

which admits the closed-form solution given by (106), where $\gamma > 0$ is found from bisection such that

$$\sum_{q=Q}^{\mu} \left\| \left(\Psi_q^H \Phi_q \Psi_q + \epsilon_q \langle \Phi_q \rangle I_{N_c} + \gamma I_{N_c}\right)^{-1} \left(\bar{X}_q^H \bar{Y}_q^{-1} \Psi_q\right)^H \right\|^2 = P_d. \tag{107}$$

The pseudo-code outlining the implementation of the path-following method for solving (102) is presented in Algorithm 2.

Algorithm 2 Path-following algorithm for solving (102)

1: **Initialization.** Randomly generate a feasible point $P(0)$ for (102). Set $\kappa = 0$.
2: **Path-following:** Update $P^{(\kappa+1)}$ by the closed-form expression (106). Increment $\kappa := \kappa + 1$ until the objective in (102) converges.
3: **Output** $P^{(opt)} \triangleq P^{(\kappa)}$ and the achieved rate $g(P^{(opt)})$.

B. Multi-user case:

Let $\mathbf{D}_{k,q} \in \mathbb{C}^{N_u \times N_u}$ be the covariance of the signal carrying information for user k over the q-th sub-carrier, and $\mathbf{D} \triangleq [\mathbf{D}_{k,q}]_{k=1,\ldots,K;q=Q,\ldots,\mu-1}$. By using [37], the capacity of the multi-user OFDM channels is bounded below by

$$\begin{aligned} f(\mathbf{D}) \triangleq \sum_{q=Q}^{\mu-1} \ln \Bigg(& I_{N_c} + \Big[\sum_{k=1}^{K} \Psi_{k,q} \mathbf{D}_{k,q} \Psi_{k,q}\Big] \\ & \times \Big(\sigma I_{N_c} + \sum_{k=1}^{K} \epsilon_{k,q} \langle \mathbf{D}_{k,q} \rangle I_{N_c}\Big)^{-1} \Bigg), \end{aligned} \tag{108}$$

where we have $\Psi_{k,q} \triangleq \hat{H}_k^f(\psi_q)\bar{\mathcal{D}}$ with $\hat{H}_k^f(\psi_q)$ denoting the estimate of $H_k^f(\psi_q)$, and $\epsilon_{k,q} \triangleq \mathbb{E}(||H_k^f(\psi_q) - \hat{H}_k^f(\psi_q)||^2) = \langle R_{h_k} \rangle - f_k(\theta^{opt}, S_{\text{tr}}^{opt})$ with $f_k(\theta^{opt}, S_{\text{tr}}^{opt})$ defined from (47).

By setting $\mathbf{D}_{k,q} = [\mathbf{P}_{k,q}]^2$ with $\mathbf{P}_{k,q} \in \mathbb{C}^{N_u \times N_u}$ we consider the following problem

$$\max_{\mathbf{P}} \sum_{q=Q}^{\mu-1} g_q(\mathbf{P}_q) \tag{109a}$$

$$\text{s.t.} \quad \sum_{q=Q}^{\mu} \sum_{k=1}^{K} ||\mathbf{P}_{k,q}||^2 \leq P_d. \tag{109b}$$

where $g_q(\mathbf{P}_q) \triangleq \ln\left(I_{N_c} + [\sum_{k=1}^{K} \Psi_{k,q}^H [\mathbf{P}_{k,q}]^2 \Psi_{k,q}](\sigma I_{N_c} + \sum_{k=1}^{K} \epsilon_{k,q} ||\mathbf{P}_{k,q}||^2 I_{N_c})^{-1}\right)$ for $\mathbf{P}_q \triangleq (\mathbf{P}_{1,q}, \ldots, \mathbf{P}_{K,q})$ and $\mathbf{P} \triangleq (\mathbf{P}_Q, \ldots, \mathbf{P}_{\mu-1})$.

Let $P_{k,q}^{(0)}$ be the initial feasible point for (109b) and $P_{k,q}^{(\kappa)}$ be obtained from the $(\kappa-1)$-th iteration. Applying the inequality (2) for

$$\begin{aligned} (\boldsymbol{X}, \mathbf{Y}) = \Big(& [\Psi_{1,q}^H \mathbf{P}_{1,q} \ldots \Psi_{K,q}^H \mathbf{P}_{K,q}], \\ & \sigma I_{N_c} + \sum_{k=1}^{K} \epsilon_{k,q} ||\mathbf{P}_{k,q}||^2 I_{N_c} \Big) \end{aligned}$$

and

$$\begin{aligned} (\bar{X}, \bar{Y}) &= (\bar{X}_q, \bar{Y}_q) \\ &\triangleq \Big([\bar{X}_{1,q} \ldots \bar{X}_{K,q}], \sigma I_{N_c} + \sum_{k=1}^{K} \epsilon_{k,q} ||P_{k,q}^{(\kappa)}||^2 I_{N_c} \Big), \end{aligned}$$

$$P_{k,q}^{(\kappa+1)} = \begin{cases} \left(\Psi_{k,q}\Phi_q\Psi_{k,q}^H + \epsilon_{k,q}\langle\Phi_q\rangle I_{N_u}\right)^{-1} B_{k,q}^H \\ \text{if} \quad \sum_{q=Q}^{\mu}\sum_{k=1}^{K} || \left(\Psi_{k,q}\Phi_q\Psi_{k,q}^H + \epsilon_{k,q}\langle\Phi_q\rangle I_{N_u}\right)^{-1} B_{k,q}^H||^2 \leq P_d, \\ \left(\Psi_{k,q}\Phi_q\Psi_{k,q}^H + \epsilon_{k,q}\langle\Phi_q\rangle I_{N_u} + \gamma I_{N_u}\right)^{-1} B_{k,q}^H \quad \text{otherwise}, \end{cases} \tag{113}$$

$$\bar{X}_{k,q} \triangleq \Psi_{k,q}^H P_{k,q}^{(\kappa)}$$

yields the following tight minorant of $g_q(\mathbf{P}_q)$:

$$\tilde{g}_q(\mathbf{P}) \triangleq a_q + 2\sum_{k=1}^{K} \Re\{\langle B_{k,q}\mathbf{P}_{k,q}\rangle\} - \left\langle \Phi_q, \sum_{k=1}^{K}[\Psi_{k,q}^H\mathbf{P}_{k,q}]^2 + \sum_{k=1}^{K}\epsilon_{k,q}||\mathbf{P}_{k,q}||^2 I_{N_c}\right\rangle, \tag{110}$$

where we have:

$$\begin{gathered} \Phi_q \triangleq \bar{Y}_q^{-1} - (\bar{Y}_q + \textstyle\sum_{k=1}^{K}[\bar{X}_{k,q}]^2)^{-1}, \\ B_{k,q} \triangleq \bar{X}_{k,q}^H \bar{Y}_q^{-1}\Psi_{k,q}^H, \\ a_q \triangleq f_q(P^{(\kappa)}) - \textstyle\sum_{k=1}^{K}\langle[\bar{X}_{k,q}]^2\bar{Y}_q^{-1}\rangle - \sigma\langle\Phi_q\rangle. \end{gathered} \tag{111}$$

We thus solve the following program of tight minorant maximization for (109) to update $P^{(\kappa+1)}$:

$$\max_{\mathbf{P}} \sum_{q=Q}^{\mu} \tilde{f}_q(\mathbf{P}) \quad \text{s.t.} \quad (109b), \tag{112}$$

which admits the closed-form solution given by (113) at the top of the page, where $\gamma > 0$ is found from bisection such that

$$\sum_{q=Q}^{\mu}\sum_{k=1}^{K} || \left(\Psi_{k,q}\Phi_q\Psi_{k,q}^H + \epsilon_{k,q}\langle\Phi_q\rangle I_{N_u} + \gamma I_{N_u}\right)^{-1} B_{k,q}^H||^2 = P_d.$$

The pseudo-code outlining the implementation of the path-following method for solving (109) is presented in Algorithm 3.

Algorithm 3 Path-following algorithm for solving (109)

1: **Initialization.** Randomly generate a feasible point $P(0)$ for (109). Set $\kappa = 0$.
2: **Path-following:** Update $P^{(\kappa+1)}$ by the closed-form expression (113). Increment $\kappa := \kappa + 1$ until the objective in (109) converges.
3: **Output** $P^{(opt)} \triangleq P^{(\kappa)}$ and the achieved rate $g(P^{(opt)})$.

IV. Numerical Results

This section evaluates the performance of our proposed training signal design proposed in Sections II and III. We assume a uniform linear array (ULA) at the BS, where K UEs are randomly distributed within a cell having 60-meter radius. We characterize the mmWave propagation by adopting the geometric channel model (16) having $N_s = 10$ scattering rays per cluster [38], whose angles of departure and arrival are generated according to the Laplacian distribution in conjunction with random mean cluster angles $\phi_{k,m,s} \in [0, 2\pi)$ and angular spreads of 10 degrees within each cluster. Assuming a carrier frequency of 30 GHz and 16.5 dB gain for multiple-antenna beamforming-aided mmWave transmission [38], the path-loss for the BS-to-UE k link at distance d_k is set to $\rho_k = 36.72 + 35.3\log_{10}(d_k)$ dB [39, Table IV], [40, Table 7.4.1-1]. For the calculation of transmit and receive correlation matrices, $R_{T,k,m}$ and $R_{R,k,m}$, in (37), we harness the spatial correlation model of clustered MIMO channel, where the angle-of-departure or arrivals follow Laplacian distribution for different path-rays in a cluster [36, Eq. (10)].

Unless stated otherwise, we set the total transmit power budget to $P = 50$ dBm, the number of TAs at the BS and RAs of the UE is set to $N = 16$ and $N_u = 2$, respectively. The number of pilot subcarriers is set to $Q = 64$, out of a total of $\mu = 128$ subcarriers. A portion of the total power budget, $P_{\text{tr}} = Q/\mu P$, is allocated for training during the transmission of Q pilot subcarriers, while the remaining portion, $P_d = (\mu - Q)/\mu P$, is used for data transmission. The number of RF chains is assumed to be half the number of TAs, i.e., $N_c = N/2$, and the number of multipath components is assumed to be $M = 4$.

In order to asses the performance of the proposed training signal design algorithm in Section II, we have used the normalized mean-square error (MSE) metric of:

$$\text{Normalized MSE} = \sum_{k=1}^{K} \mathbb{E}\left\{ \left| h_k - \hat{h}_k \right|^2 / |h_k|^2 \right\}. \tag{114}$$

To observe the impact of channel estimation on the achievable data rate, we simulate the algorithms 2 and 3 in Section III. The average achievable rate metric is adopted, which is given by

$$\text{Average Rate} = \frac{1}{\ln 2}\frac{f(\mathbf{P})}{\mu} \quad \text{bits/sec/Hz} \tag{115}$$

where $f(\mathbf{P})$ is defined in (102) and (109) for the single-user and multi-user cases, respectively.

The performance cannot be directly compared to existing benchmarks because no prior work has proposed training design and channel estimation for multi-user mmWave OFDM channels under the considered AoSA architecture. Consequently, we compare the proposed training algorithm (Alg. 1) against random training, while the proposed precoding algorithms (Alg. 2 and Alg. 3) are compared against random precoding. In addition, to provide a lower benchmark for training performance, we will compare the proposed AoSA architecture-based training algorithm with a counterpart based on a fully connected architecture. Moreover, to provide an upper benchmark on the achievable rate of the proposed precoding schemes, we will later compare the proposed training

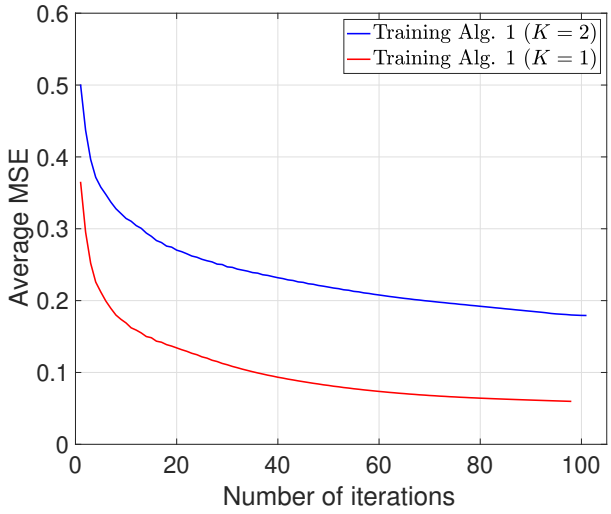


Fig. 1: Convergence of the proposed training Alg. 1.

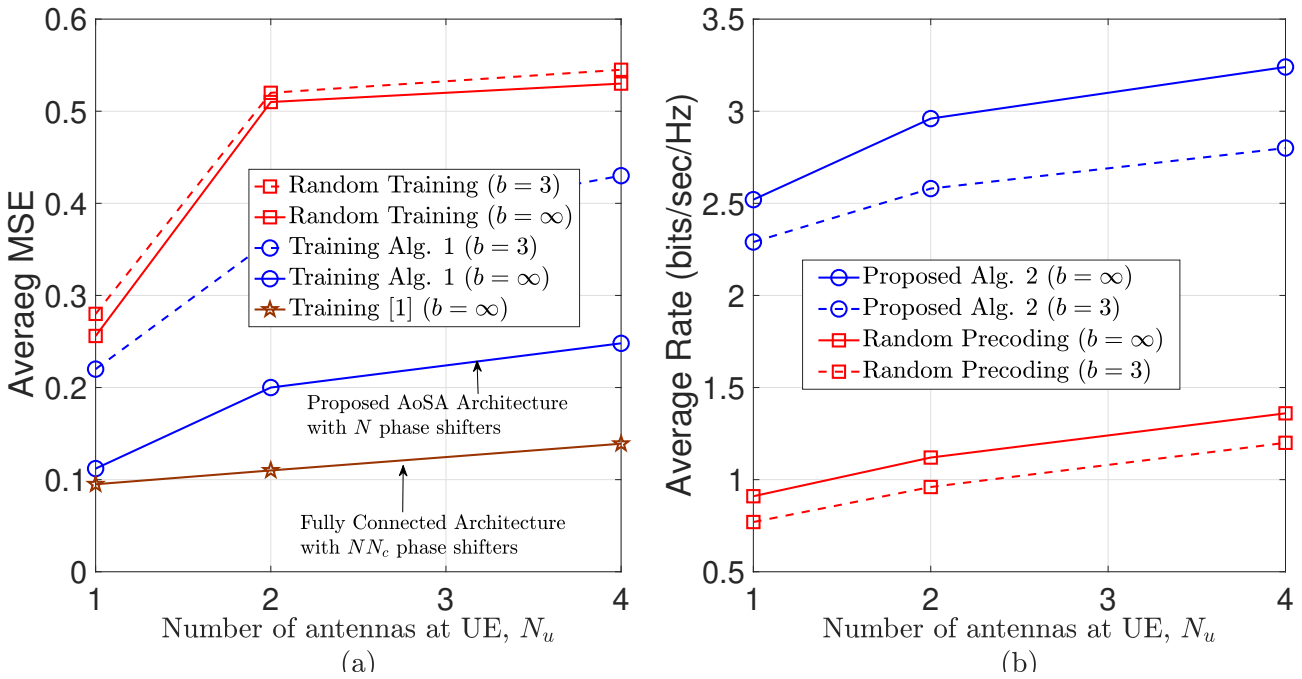


Fig. 2: Impact of the number of antennas at the UE N_u and low-resolution quantization at the AP on (a) MSE and (b) achievable rate performance for the single-user case.

and precoding algorithms to a benchmark that assumes perfect CSI.

Fig. 1 shows the convergence of the proposed training Alg. 1 for a particular simulation, for both single-user and two-users cases. We can observe that the average MSE of the proposed training Alg. 1 monotonically converges after around 100 iterations. When extending the system to the two-user case, the MSE approximately doubles, as the metric in (114) aggregates the estimation errors across multiple users.

Fig. 2 illustrates the impact of the number of RAs N_u at the UE on the MSE and on the achievable rate performance for the single-user case. Both infinite-resolution ($b = \infty$) and low-resolution ($b = 3$) quantization schemes are considered for the design of analog precoding $\mathbf{V}_A$. Fig. 2(a) shows that the MSE increases with N_u, due to the larger number of channel parameters that must be estimated, governed by the size of the channel vector MNN_u. By contrast, Fig. 2(b) indicates that this increase in MSE does not impede the improvement in achievable rate, which continues to grow with N_u as a result of the proposed precoding design and the multi-antenna diversity at the receiver. In general, Fig. 2 highlights the benefits of the proposed training and precoding algorithms over random training and precoding in terms of MSE and

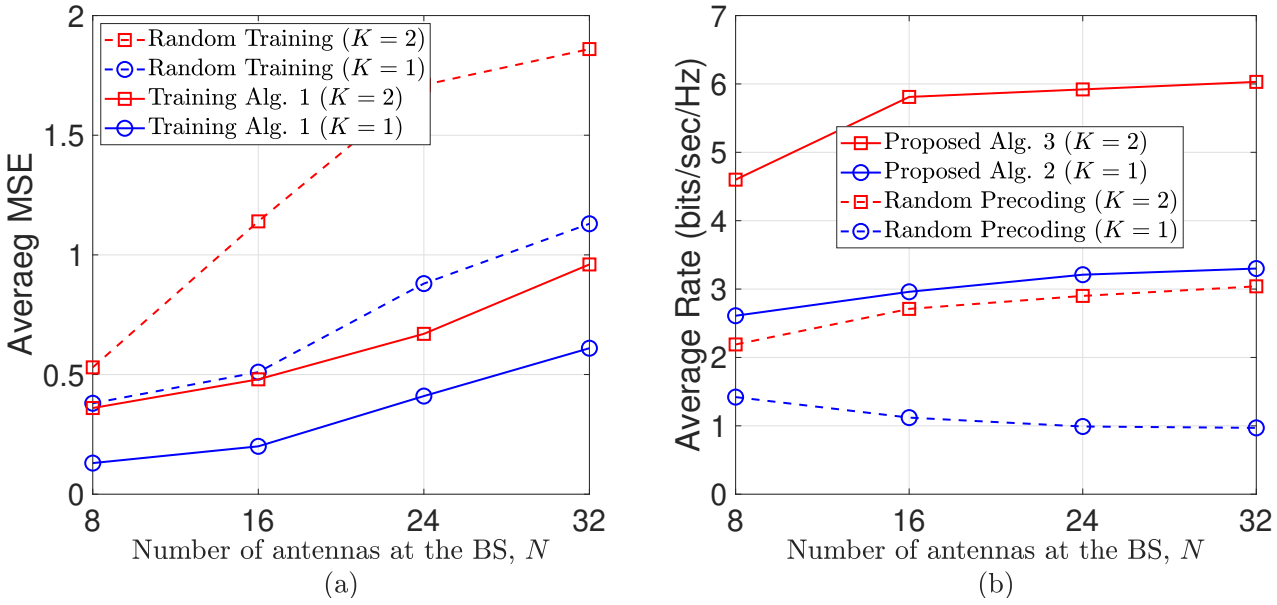


Fig. 3: Impact of the number of antennas at the BS N on (a) MSE and (b) achievable rate performance.

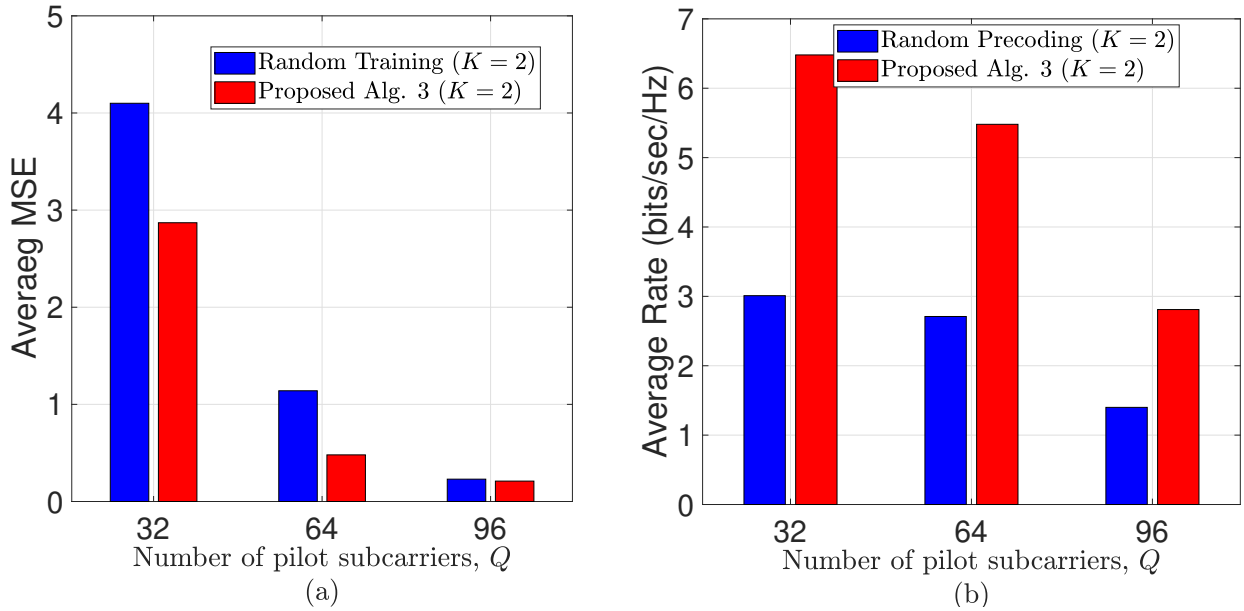


Fig. 4: Impact of varying the number of pilot subcarriers Q, under fixed total number of subcarriers $\mu = 128$, on (a) MSE and (b) achievable rate performance.

rate performance, respectively. Focusing on the impact of quantizing the analog precoder design, Fig. 2(a) shows that the MSE of the proposed algorithm nearly doubles when using 3-bit resolution compared to ∞-resolution analog precoding. By contrast, Fig. 2(b) shows that the achievable rate decreases by only about 10% when transitioning from ∞-resolution to 3-bit resolution in the design of analog precoding.

In Fig. 2(a), we compare the performance of the proposed training algorithm to that of the training algorithm in [13]. It can be observed that the MSE achieved by the proposed Algorithm 1 is almost twice that of the algorithm in [13] for $N_u \geq 2$ UE antennas. This performance difference arises because Algorithm 1 is based on an AoSA architecture, which requires only N phase shifters, whereas the algorithm in [13] relies on a fully connected (FC) architecture employing $N_c N$ phase shifters. Consequently, the power consumption associated with the phase shifters in the FC architecture, as adopted in [13], is N_c times higher than that of the AoSA-based architecture considered in this work. Moreover, the algorithm in [13] is limited to single-UE scenarios.

The effect of varying the number of TAs N at the BS on the MSE and on the achievable rate is depicted in Fig. 3. The single-user and $K = 2$ user cases are compared. The MSE and achievable rate approximately double when

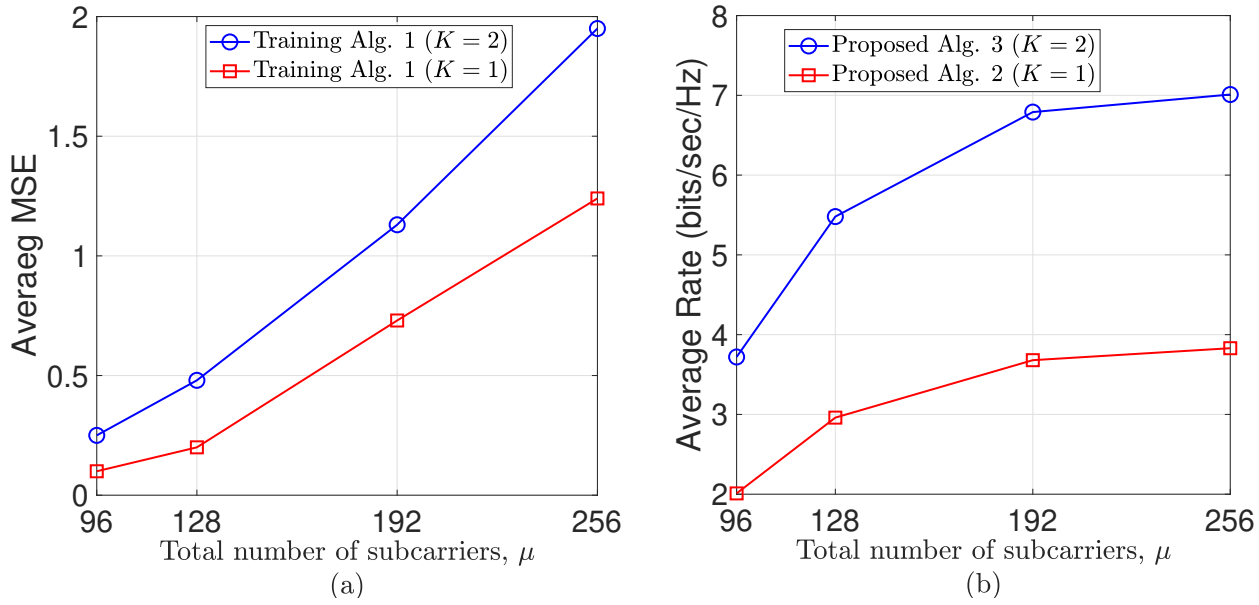


Fig. 5: Impact of varying the total number of subcarriers μ, under fixed number of pilot subcarriers $Q = 64$, on (a) MSE and (b) achievable rate performance.

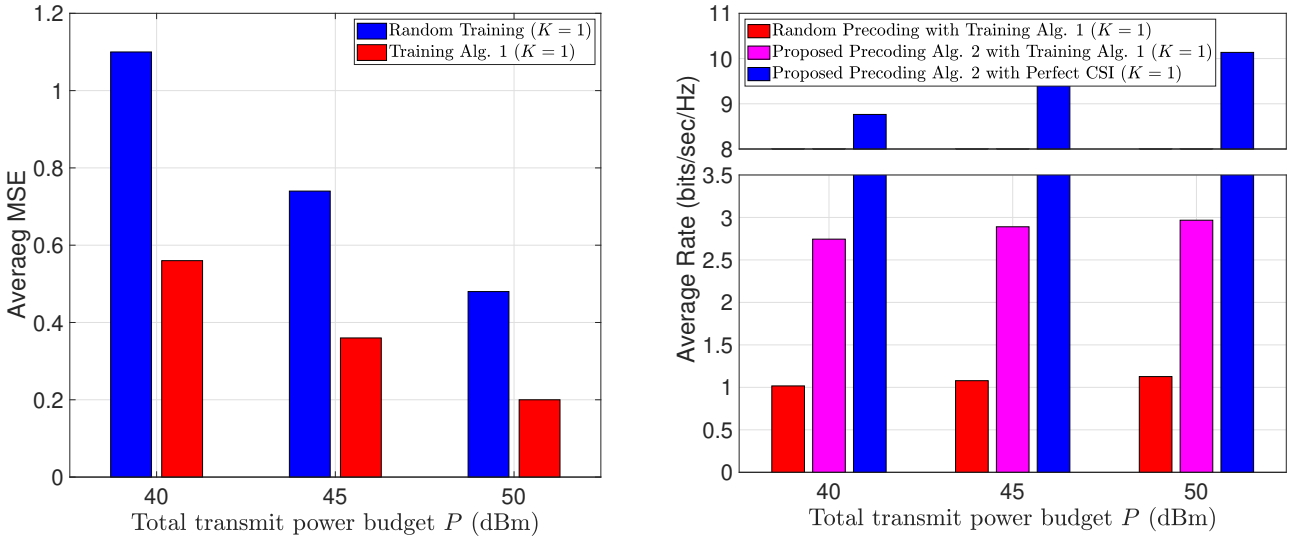
moving from the single-user to the two-user case. This is because the metrics defined in (114) and (115) sum the MSE and achievable rate across all K users. Overall, Fig. 3 demonstrates the effectiveness of the proposed training and precoding algorithms compared to their random counterparts, with noticeable gains in both MSE and achievable rate. As seen in Fig. 3(a), the MSE increases with N, owing to the growing number of channel parameters that must be estimated proportional to the size of the channel vector MNN_u. Despite this, Fig. 3(b) indicates that the achievable rate continues to improve as N increases, showing resilience to the rise in estimation error.

Fig. 4 shows the impact of varying the number of pilot subcarriers Q on both the MSE and achievable rate, under a fixed total number of subcarriers $\mu = 128$. As shown in Fig. 4(a), the average MSE decreases upon increasing Q, due to the availability of increased training resources. By contrast, Fig. 4(b) shows that the average achievable rate also decreases as Q increases. One might expect the rate to improve, given the reduction in MSE. However, the decrease in rate is due to the reduced number of data subcarriers, $\mu - Q$, as more subcarriers are allocated for training. Fig. 4 also provides clear evidence of the superiority of the proposed training and precoding algorithms over their random counterparts, yielding significant improvements in both MSE and achievable rate.

Fig. 5 shows the impact of varying the total number of subcarriers μ on both the MSE and on the achievable rate, under a fixed number of pilot subcarriers $Q = 64$. The analysis includes both single-user and two-user scenarios. Transitioning to the two-user case results in roughly double the MSE and achievable rate, since the metrics in (114) and (115) aggregate these quantities for two users. As seen in Fig. 5(a), the MSE increases for larger values of μ, owing to the relative reduction in the number of pilot subcarriers, i.e., if the ratio of pilot subcarriers to total subcarriers, $\frac{Q}{\mu}$, decreases. One might expect the achievable rate to decrease as a result of the increased MSE. However, Fig. 5(b) shows that the rate actually increases, due to the growing number of data subcarriers, $\mu - Q$, which allows for more data transmission. It is worth noting that the improvement in achievable rate becomes marginal as μ continues to increase.

Fig. 6 illustrates the impact of the maximum power budget $P_{\max}$ on both the MSE and achievable rate. As shown in Fig. 6a, the average MSE decreases, while the average rate increases upon increasing P, due to the greater availability of resources. Fig. 6a clearly demonstrates the advantage of the training algorithm proposed in Sec. II in terms of achievable MSE, compared to random training. A similar benefit is observed in Fig. 6b, where the precoding algorithm proposed in Sec. III outperforms random precoding, yielding more than a two-fold improvement in the achievable rate. Fig. 6b also depicts the performance of the proposed precoding algorithm under perfect CSI, when the achievable rate becomes about $3 - 4$ times higher than in the case, where the proposed precoding relies on channel estimation obtained through Sec. II. Considering the case of $K = 2$ users, results for varying power budgets are shown in Fig. 7, where a trend in average MSE and achievable rate similar to that in Fig. 6 is observed for the single-user case. As anticipated, moving to the two-user scenario leads to approximately twice the MSE and achievable rate, as the metrics in (114) and (115) represent aggregated values across both users.

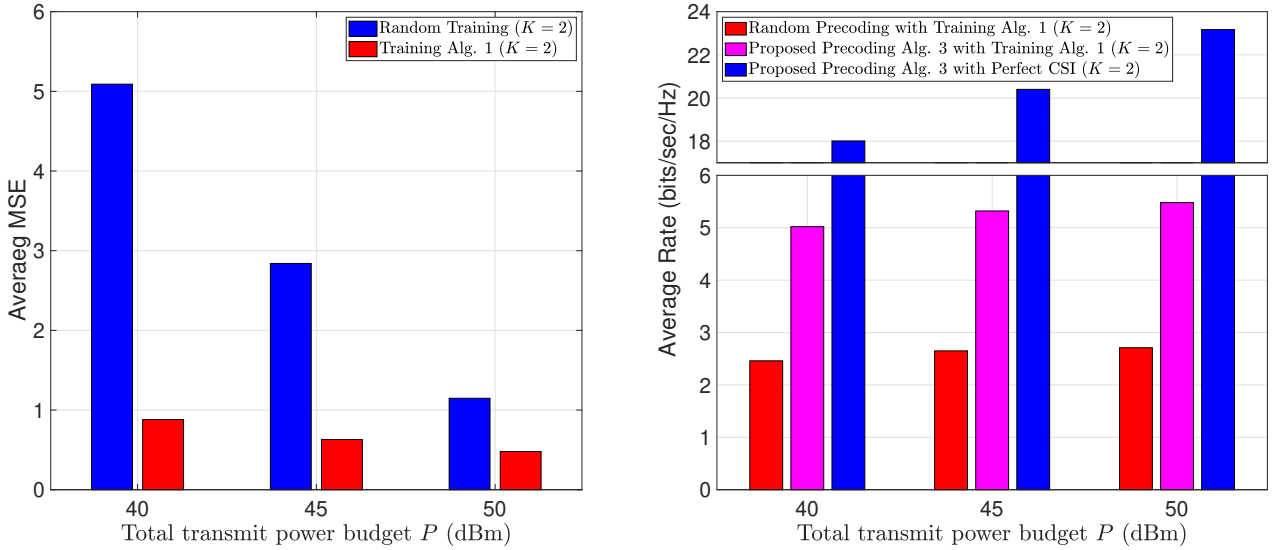
Fig. 8 illustrates the impact of the number of RF chains N_c for $K = 2$ users and $N = 32$ TAs on both the MSE and the achievable rate. It can be observed that the average MSE decreases, and consequently the achievable average rate increases, as the number of RF chains grows, this is owing to the availability of more resources. Thus, having fewer RF



(a) MSE performance versus power budget

(b) Achievable rate versus power budget

Fig. 6: Impact of maximum power budget $P_{\max}$ (dBm) for single-user case $K = 1$.



(a) MSE performance versus power budget

(b) Achievable rate versus power budget

Fig. 7: Impact of maximum power budget $P_{\max}$ (dBm) for multi-user case $K = 2$

chains (i.e., $N_c < N/2$) results in poor channel estimation accuracy. Therefore, we have assumed $N_c = N/2$ RF chains in all other simulation results.

Fig. 9 illustrates the impact of the number of users K on both the MSE and the achievable rate. It can be observed that both the average MSE and the achievable rate increase with K. Naturally, this is expected because the metrics defined in (114) and (115) sum the MSE and achievable rate across all K users. Overall, Fig. 9 demonstrates the effectiveness of the proposed training and precoding algorithms: the former handles multiple users efficiently, limiting the MSE growth to approximately linear with K, while the latter manages multiuser interference effectively, allowing the overall achievable rate to continue improving as K increases.

V. Conclusions

We have conceived training signal designs through the joint optimization of ABF and SDPs for estimating multi-user wideband mmWave channels based on their statistics. A path-following algorithm was developed to update ABF and SDPs via closed-form expressions. Furthermore, we proposed path-following algorithms to investigate how the channel estimation error affects the system performance, particularly in terms of capacity for decoding user information. An extension to a new class of ABF that saves phase shifts to reduce power consumption [41] is currently under investigation.

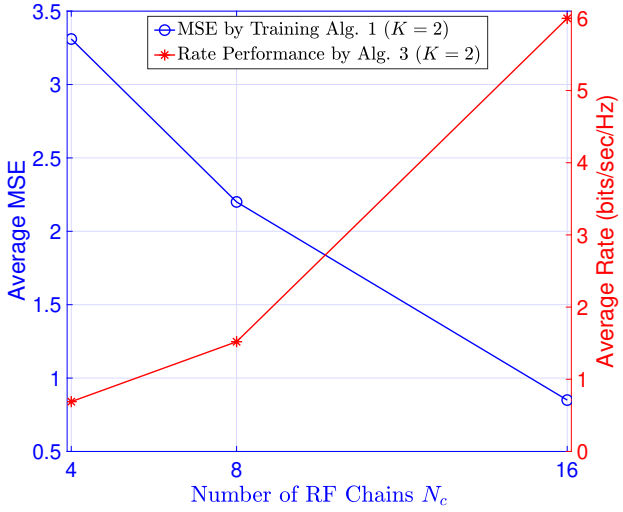


Fig. 8: MSE and achievable rate versus the number of RF chains N_c for $K = 2$ users and $N = 32$ TAs.

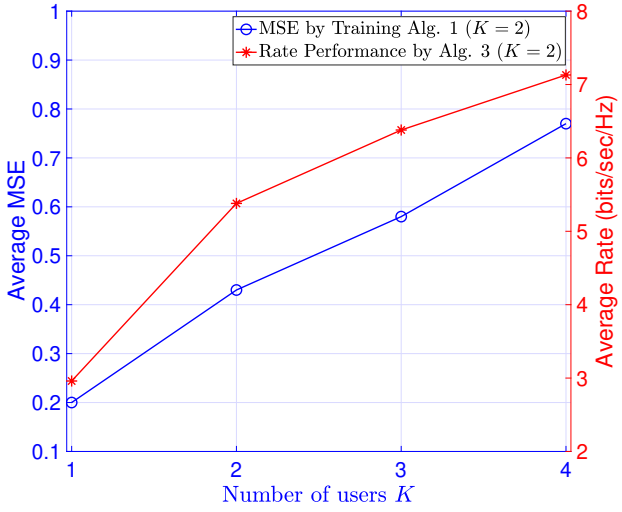


Fig. 9: MSE and achievable rate versus the number of users K.

REFERENCES

[1] K. Zheng, H. Yang, Z. Ying, P. Wang, and L. Hanzo, "Vision-assisted millimeter-wave beam management for next-generation wireless systems: Concepts, solutions, and open challenges," *IEEE Veh. Techn. Mag.*, vol. 18, no. 3, pp. 58–68, Mar. 2023.

[2] R. W. Heath, N. Gonzalez-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Select. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, April 2016.

[3] M. Xiao et. al., "Millimeter wave communications for future mobile networks," *IEEE J. Select. Areas Commun.*, vol. 35, no. 9, pp. 1909–1935, Sept. 2017.

[4] J. Li et. al, "Mobility support for millimeter wave communications: Opportunities and challenges," *IEEE Commun. Surv. Tut.*, vol. 24, no. 3, pp. 1816–1842, 3rd Q. 2022.

[5] T. S. Rappaport, R. W. Heath, R. C. Daniels, and J. N. M. 2015., *Millimeter Wave Wireless Communications*. Upper Saddle River, NJ, USA: Prentice Hall, 2015.

[6] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Tran. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.

[7] A. Alkhateeb, O. E. Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Select. Topics Signal Process.*, vol. 8, no. 5, pp. 831–846, May 2014.

[8] X. Gao, L. Dai, S. Han, C.-L. I, and R. W. Heath, "Energy-efficient hybrid analog and digital precoding for MmWave MIMO systems with large antenna arrays," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 998–1009, Apr. 2016.

[9] J. Lee, G.-T. Gil, and Y. H. Lee, "Channel estimation via orthogonal matching pursuit for hybrid MIMO systems in millimeter wave communications," *IEEE Trans. Commun.*, vol. 64, no. 6, pp. 2370–2386, June 2016.

[10] Y.-T. Chiew and Y.-P. Lin, "Channel estimation for hybrid mmwave systems using generalized Kronecker compressive sensing (G-KCS) with successive decision-aided recovery," *IEEE Trans. Signal Process.*, vol. 72, pp. 2970–2981, 2024.

[11] P. Zhu, H. Lin, J. Li, D. Wang, and X. You, "High-performance channel estimation for mmWave wideband systems with hybrid structures," *IEEE Trans. Commun.*, vol. 71, no. 4, pp. 2503–2516, Apr. 2023.

[12] K. Zheng, Z. Gu, X. Xia, Z. Zhang, D. Wang, and P. Zhu, "Joint phase-noise and channel estimation in mmWave massive MIMO systems with hybrid structures using nested tensor decomposition," *IEEE Trans. Veh. Techn.*, vol. 73, no. 12, pp. 19 890–19 895, Dec. 2024.

[13] E. Vlachos, G. C. Alexandropoulos, and J. Thompson, "Wideband mimo channel estimation for hybrid beamforming millimeter wave systems via random spatial sampling," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 5, pp. 1136–1150, 2019.

[14] Z. Lin, T. Lv, W. Ni, J. A. Zhang, J. Zeng, and R. P. Liu, "Joint estimation of multipath angles and delays for millimeter-wave cylindrical arrays with hybrid front-ends," *IEEE Trans. Wireless Commun.*, vol. 20, no. 7, pp. 4631–4645, Jul. 2021.

[15] Y. Li and A. Hu, "Channel estimation via subcarrier grouping for wideband mmwave hybrid massive MIMO-OFDM systems," *IEEE Trans. Green Commun. Netw.*,, vol. 8, no. 1, pp. 64–78, Jan. 2024.

[16] K. Venugopal, A. Alkhateeb, N. Gonzalez-Prelcic, and R. W. Heath, "Channel estimation for hybrid architecture-based wideband millimeter wave systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 1996–2009, Sep. 2017.

[17] J. Rodriguez-Fernandez, N. Gonzalez-Prelcic, K. Venugopal, and R. W. Heath, "Frequency-domain compressive channel estimation for frequency-selective hybrid millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 2946–2960, May 2018.

[18] H. Kim, G.-T. Gil, and Y. H. Lee, "Two-step approach to time-domain channel estimation for wideband millimeter wave systems with hybrid architecture," *IEEE Trans. Commun.*, vol. 67, no. 7, pp. 5139–5152, July 2019.

[19] B. Wang, M. Jian, F. Gao, G. Y. Li, and H. Lin, "Beam squint and channel estimation for wideband mmWave massive MIMO-OFDM systems," *IEEE Trans. Signal Process.*, vol. 67, no. 23, pp. 5893–5908, Dec. 2019.

[20] Y. Lin, S. Jin, M. Matthaiou, and X. You, "Tensor-based channel estimation for millimeter wave MIMO-OFDM with dual-wideband effects," *IEEE Trans. Commun.*, vol. 68, no. 7, pp. 4218–4232, July 2020.

[21] K. F. Masood, J. Tong, J. Xi, J. Yuan, Q. Guo, and Y. Yu, "Low-rank matrix sensing-based channel estimation for mmWave and THz hybrid MIMO systems," *IEEE J. Select. Topics Signal Process.*, vol. 17, no. 4, pp. 777–793, July 2023.

[22] X. Gong, W. Chen, L. Sun, J. Chen, and B. Ai, "An ESPRIT-based supervised channel estimation method using tensor train decomposition for mmWave 3-D MIMO-OFDM systems," *IEEE Trans. Signal Process.*, vol. 71, pp. 555–570, 2023.

[23] Y. Zhang, M. El-Hajjar, and L.-L. Yang, "Multi-layer sparse Bayesian learning for mmWave channel estimation," *IEEE Trans. Veh. Techn.*, vol. 73, no. 3, pp. 3485–3498, Mar. 2024.

[24] J. Du, Y. Chen, P. Zhang, S. Mumtaz, X. Li, and D. B. da Costa, "An effective simultaneous channel estimation and sensing algorithm for mmWave MIMO-OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 23, no. 11, pp. 17 054–17 069, Nov. 2024.

[25] X. Ma, F. Yang, S. Liu, J. Song, and Z. Han, "Design and optimization on training sequence for mmWave communications: A new approach for sparse channel estimation in massive MIMO," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1486–1497, July 2017.

[26] X. Li, J. Fang, H. Li, and P. Wang, "Millimeter wave channel estimation via exploiting joint sparse and low-rank structures," *IEEE Trans. Wireless Commun.*, vol. 17, no. 2, pp. 1123–1133, Feb. 2018.

[27] C. Lin, J. Gao, R. Jin, and C. Zhong, "Self-adaptive measurement matrix

design and channel estimation in time-varying hybrid MmWave massive MIMO-OFDM systems," *IEEE Trans. Commun.*, vol. 72, no. 1, pp. 618–629, Jan. 2024.

[28] H. Wang, J. Fang, P. Wang, G. Yue, and H. Li, "Efficient beamforming training and channel estimation for millimeter wave OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 20, no. 5, pp. 2805–2819, May 2021.

[29] H. D. Tuan, H. H. Kha, H. H. Nguyen, and V.-J. Luong, "Optimized training sequences for spatially correlated MIMO-OFDM," *IEEE Trans. Wireless Commun.*, vol. 9, no. 9, pp. 2768–2778, Sept. 2010.

[30] Z. Sheng, H. D. Tuan, H. H. Nguyen, and M. Debbah, "Optimal training sequences for large-scale MIMO-OFDM systems," *IEEE Trans. Signal Process.*, vol. 65, no. 13, pp. 3329–3342, July 2017.

[31] H. H. M. Tam, H. D. Tuan, and D. T. Ngo, "Successive convex quadratic programming for quality-of-service management in full-duplex MU-MIMO multicell networks," *IEEE Trans. Commun.*, vol. 64, no. 6, pp. 2340–2353, June 2016.

[32] H. D. Tuan, H. H. M. Tam, H. H. Nguyen, T. Q. Duong, and H. V. Poor, "Superposition signaling in broadcast interference networks," *IEEE Trans. Commun.*, vol. 65, no. 11, pp. 4646–4656, Nov. 2017.

[33] H. Tuy, *Convex Analysis and Global Optimization (second edition)*. Springer International, 2017.

[34] O. E. Ayach, R. W. Heath, S. Rajagopal, and Z. Pi, "Multimode precoding in millimeter wave MIMO transmitters with multiple antenna sub-arrays," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2013, pp. 3476–3480.

[35] J. D. Krieger, C.-P. Yeang, and G. W. Wornell, "Dense delta-sigma phased arrays," *IEEE Trans. Antenn. Propag.*, vol. 61, no. 4, pp. 1825–1837, Apr. 2013.

[36] A. Forenza, D. J. Love, and R. W. Heath, "Simplified spatial correlation models for clustered MIMO channels with different array configurations," *IEEE Trans. Veh. Technol.*, vol. 56, no. 4, pp. 1924–1934, 2007.

[37] N. Jindal, S. Vishwanath, and A. Goldsmith, "On the duality of Gaussian multiple-access and broadcast channels," *IEEE Trans. Info. Theory*, vol. 50, no. 4, pp. 768–783–1580, 2004.

[38] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1164–1179, June 2014.

[39] T. S. Rappaport, Y. Xing, G. R. MacCartney, A. F. Molisch, E. Mellios, and J. Zhang, "Overview of millimeter wave communications for fifth-generation (5G) wireless networks–With a focus on propagation models," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 12, pp. 6213–6230, Dec 2017.

[40] 3GPP, "Study on channel model for frequencies from 05 to 100 GHz," 3rd Generation Partnership Project (3GPP), Tech. Rep. TR 38.901 V14.1.1 Release 14, Aug. 2017, [Online]. Available: http://www.3gpp.org/DynaReport/38901.htm.

[41] W. Zhu, H. D. Tuan, E. Dutkiewicz, H. V. Poor, and L. Hanzo, "A new class of analog precoding for multi-antenna multi-user communications over high-frequency bands," *IEEE Trans. Wireless Commun.*, vol. 23, no. 9, pp. 11 493–11 507, Sept. 2024.