

Optimizing the mmWave Channel Estimation Duration by Rate Prediction

Jeongwan Kang¹, Graduate Student Member, IEEE, Nil Garcia², Member, IEEE,
Henk Wymeersch³, Senior Member, IEEE, Carlo Fischione⁴, Senior Member, IEEE,
Gonzalo Seco-Granados⁵, Senior Member, IEEE, and Sunwoo Kim⁶, Senior Member, IEEE,

Abstract—In millimeter-wave (mmWave) wireless communications, the duration of the channel estimation plays a major role to establish the links before data transmission. However, fixed or long channel estimation can substantially hinder the achievable transmit data rates. In this letter, we propose a new scheme that optimizes the channel estimation duration to establish the link between the base station (BS) and a mobile station (MS) in mmWave communications. Before pilot transmissions, the BS predicts the downlink effective rate that would be achieved after channel estimation with the pilot precoder and compares this predicted rate with the current rate, based on the current channel estimates. The proposed scheme optimizes the mmWave channel estimation duration by ending pilot transmissions when the predicted rate is lower than the current rate.

Index Terms—Channel estimation duration, rate prediction, CRB, mmWave, MIMO.

I. INTRODUCTION

IN most wireless communication systems, the channel estimation is performed by sending pilots to establish the link [1]. Recently, millimeter-wave (mmWave) communications have become widely accepted in 5G to boost data rates, but require a large number of pilot transmissions to estimate the channel [2], which limit their use to low-mobility applications. This leads to a trade-off between the channel estimation overhead and the downlink effective rate [3].

Various efforts have been made to improve the downlink effective rate by reducing the channel estimation duration in mmWave communications. For example, exhaustive and hierarchical designs are introduced in [4]. System-level (frame

length, training time, and bandwidth) optimization [5], pilot overhead optimization for short-packet transmission [6] and compressed sensing training [7] have been introduced to reduce the channel training time. The fast channel estimation based on overlapped beam pattern design [8] and localization-assisted beam selection [9] have been proposed to reduce the beam training overhead. These approaches can reduce the channel estimation duration, but they still use a decided duration that is fixed (and therefore highly sub-optimal) before pilot transmissions.

In this letter, we seek to find the optimal channel estimation duration for a multipath channel. We consider a geometrical parameterization of the mmWave channel, based on the angle-of-departure (AoD), angle-of-arrival (AoA), and channel gain of each individual paths [1]. The Cramér-Rao bound (CRB) of mmWave channel parameters can be derived as in [10] and the pilot precoder that minimizes the CRB of the angle estimates without beam sweeping is adopted to transmit pilots [11]. The channel error covariance that would result from the transmission by the BS of pilots through the precoder can be predicted without actually realizing the transmission. Therefore, *the downlink effective rate can be predicted before sending pilots*. We develop a scheme where the current rate and predicted rate are compared to each other and based on this, a decision on whether the BS needs to continue transmitting pilots is taken.

To the best of our knowledge, while beam alignment protocols or channel estimation schemes have been developed to optimize the channel estimation duration, this duration is always set a priori, before the channel estimation commences. In contrast, we propose a new adaptive method that decides the channel estimation duration during execution. The proposed scheme is capable of balancing the channel estimation duration and the estimation accuracy to maximize the downlink effective rate.

II. SYSTEM MODEL

We consider a mmWave system where the BS and mobile station (MS) are equipped with uniform linear arrays (ULA) of N_T and N_R antennas, respectively, as shown in Fig. 1. We assume that there is a feedback channel between the BS and MS. The narrow-band flat fading downlink channel $\mathbf{H} \in \mathbb{C}^{N_R \times N_T}$ in which the bandwidth is much smaller than the carrier frequency is [12]

$$\mathbf{H} = \sum_{l=0}^{L-1} \alpha_l \mathbf{a}_{R_x}(\phi_l) \mathbf{a}_{T_x}^H(\theta_l), \quad (1)$$

Manuscript received September 9, 2020; accepted September 26, 2020. Date of publication October 6, 2020; date of current version February 11, 2021. This research was supported in part by the MSIP (Ministry of Science, ICT and Future Planning) under the ITRC (Information Technology Research Center) support program (IITP-2020-2017-0-01637) supervised by the IITP (Institute for Information & communications Technology Promotion), in part by IITP grant funded by the MSIT (Ministry of Science and ICT) (2016-0-00208), and in part by Swedish Research Council, under grant 2018-03701. The associate editor coordinating the review of this letter and approving it for publication was J. Zhang. (Corresponding author: Sunwoo Kim.)

Jeongwan Kang and Sunwoo Kim are with the Department of Electronics and Computer Engineering, Hanyang University, Seoul 04763, South Korea (e-mail: rkdwjddhks77@hanyang.ac.kr; remero@hanyang.ac.kr).

Nil Garcia and Henk Wymeersch are with the Department of Electrical Engineering, Chalmers University of Technology, 41296 Gutenberg, Sweden (e-mail: nilg@chalmers.se; henkw@chalmers.se).

Carlo Fischione is with the Division of Network and Systems Engineering, KTH Royal Institute of Technology, 10044 Stockholm, Sweden (e-mail: carlofi@kth.se).

Gonzalo Seco-Granados is with the Department of Telecommunications and Systems Engineering, Universitat Autònoma de Barcelona, 08193 Barcelona, Spain (e-mail: gonzalo.seco@uab.es).

Digital Object Identifier 10.1109/LCOMM.2020.3029107

1558-2558 © 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.

See <https://www.ieee.org/publications/rights/index.html> for more information.

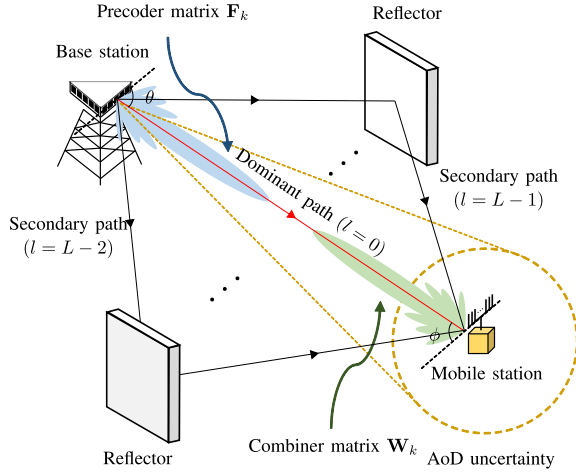


Fig. 1. An illustration of the mmWave downlink channel using precoders and combiners.

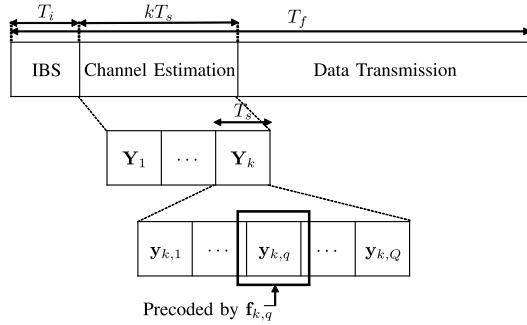


Fig. 2. Frame structure consisting of three phases [3] (initial beam sweeping (IBS), channel estimation and data transmission).

where $\alpha_l \sim \mathcal{CN}(0, 1)$ is the complex channel gain, θ_l and ϕ_l are the AoD and AoA of path l . The array vectors at BS and MS are denoted as $\mathbf{a}_{\text{Tx}}(\theta_l) \in \mathbb{C}^{N_{\text{T}} \times 1}$ and $\mathbf{a}_{\text{Rx}}(\phi_l) \in \mathbb{C}^{N_{\text{R}} \times 1}$ and are given by

$$\mathbf{a}_{\text{Tx}}(\theta_l) = [1, e^{j\pi \cos \theta_l}, \dots, e^{j\pi(N_{\text{T}}-1) \cos \theta_l}]^T, \quad (2)$$

$$\mathbf{a}_{\text{Rx}}(\phi_l) = [1, e^{j\pi \cos \phi_l}, \dots, e^{j\pi(N_{\text{R}}-1) \cos \phi_l}]^T. \quad (3)$$

Note that even though we consider the flat-fading channel model for simplicity, the proposed scheme will be applicable to the frequency-selective channel model.

The channel is assumed to remain constant during a transmission frame duration of T_f . As shown in Fig. 2, we consider the initial access where pilots have to be at the beginning of the frame [3]. At k -th time slot of the channel estimation interval, the BS transmits Q consecutive pilots, and each pilot s_q with $\mathbb{E}[|s_q|^2] = 1$ is precoded by the vector $\mathbf{f}_{k,q} \in \mathbb{C}^{N_{\text{T}} \times 1}$. To avoid the dependence of the transmitted energy on the number of pilots Q , all precoders are normalized to $\|\mathbf{f}_{k,q}\|_2^2 = 1/Q$. The received baseband signal $\mathbf{y}_{k,q} \in \mathbb{C}^{N_{\text{R}} \times 1}$, when s_q is transmitted at k -th time slot, is given by

$$\mathbf{y}_{k,q} = \sqrt{P} \mathbf{W}_k^H \mathbf{H} \mathbf{f}_{k,q} s_q + \mathbf{n}_{k,q}, \quad (4)$$

where $\mathbf{W}_k = [\mathbf{w}_{k,1}, \dots, \mathbf{w}_{k,N_{\text{RF}}}] \in \mathbb{C}^{N_{\text{R}} \times N_{\text{RF}}}$ is a combiner matrix associated with N_{RF} RF chains ($N_{\text{RF}} < N_{\text{R}}$), P is the transmitted power, and $\mathbf{n}_{k,q}$ is circularly symmetric and denoted by $\mathcal{CN}(0, \sigma_n^2 \mathbf{I}_{N_{\text{RF}}})$.

We assume that L paths are well separated in the angular domain [13]. In addition, since the received power of the secondary path is known to be 10 to 20 dB lower than that of the dominant path [14], the secondary paths can be neglected. By multiplying s_q^* (complex conjugate of s_q), we can simplify (4) to

$$\mathbf{y}_{k,q} \approx \sqrt{P} \alpha_0 \mathbf{W}_k^H \mathbf{a}_{\text{Rx}}(\phi_0) \mathbf{a}_{\text{Tx}}^H(\theta_0) \mathbf{f}_{k,q} + \mathbf{n}_{k,q}, \quad (5)$$

provided $\mathbf{a}_{\text{Tx}}^H(\theta_l) \mathbf{f}_{k,q} \approx 0$ for $l \neq 0$ and all q . So, the path index l will be omitted from now on. The observation matrix $\mathbf{Y}_k = [\mathbf{y}_{k,1}, \dots, \mathbf{y}_{k,Q}] \in \mathbb{C}^{N_{\text{RF}} \times Q}$ at k -th time slot is given by

$$\mathbf{Y}_k = \sqrt{P} \alpha \mathbf{W}_k^H \mathbf{a}_{\text{Rx}}(\phi) \mathbf{a}_{\text{Tx}}^H(\theta) \mathbf{F}_k + \mathbf{N}_k, \quad (6)$$

where $\mathbf{F}_k = [\mathbf{f}_{k,1}, \dots, \mathbf{f}_{k,Q}] \in \mathbb{C}^{N_{\text{T}} \times Q}$ is the pilot precoder and $\mathbf{N}_k = [\mathbf{n}_{k,1}, \dots, \mathbf{n}_{k,Q}] \in \mathbb{C}^{N_{\text{RF}} \times Q}$.

III. OPTIMIZING THE MMWAVE CHANNEL ESTIMATION DURATION BY RATE PREDICTION

The proposed scheme consists of the following steps: the acquisition of initial angle ranges, the choice of precoder, the choice of combiner, pilot transmissions, refinement of the channel estimate, and rate computation. In addition, we perform a rate prediction before transmitting the next pilot, in order to decide whether or not to stop sending additional pilots.

A. Initial Beam Sweeping

At each time slot the BS sends out pilots consisting of Q symbols to do the beam sweeping. In the course of the beam sweeping, each beamformer at the BS and MS steers towards M_{T} and M_{R} equispaced directions within $[0, \pi]$, respectively. The beamformers $\bar{\mathbf{a}}_{\text{Tx}}(\bar{\theta}, M_{\text{T}}) \in \mathbb{C}^{M_{\text{T}} \times 1}$ and $\bar{\mathbf{a}}_{\text{Rx}}(\bar{\phi}, M_{\text{R}}) \in \mathbb{C}^{M_{\text{R}} \times 1}$ are given by [15]

$$\bar{\mathbf{a}}_{\text{Tx}}(\bar{\theta}, M_{\text{T}}) = [1, e^{j\pi \cos \bar{\theta}}, \dots, e^{j\pi(M_{\text{T}}-1) \cos \bar{\theta}}]^T, \quad (7)$$

$$\bar{\mathbf{a}}_{\text{Rx}}(\bar{\phi}, M_{\text{R}}) = [1, e^{j\pi \cos \bar{\phi}}, \dots, e^{j\pi(M_{\text{R}}-1) \cos \bar{\phi}}]^T, \quad (8)$$

where $\bar{\theta}$ and $\bar{\phi}$ are the steering directions of the BS and MS, respectively. The MS determines the best beam pair maximizing the received power and feeds it back to the BS. Initial angle ranges ($\mathcal{R}_{\text{Tx},1}$ and $\mathcal{R}_{\text{Rx},1}$) are the angle coverage of the best beams at the BS and MS, respectively.

B. Pilot Precoder Design With the AoD Uncertainty

While any precoder can be used at k -th time slot, we propose to use the precoder matrix \mathbf{F}_k that minimizes the CRB of the AoD estimate introduced in [11] yields the direction error bound (DEB) as follows:

$$\text{var}(\hat{\theta}) \geq \text{DEB}(\theta, \mathbf{F}) = \left[2 \rho \|\mathbf{b}_{\text{Rx},k}\|_2^2 \left(\|\mathbf{F}^H \dot{\mathbf{a}}_{\text{Tx}}(\theta)\|_2^2 - \frac{|\mathbf{a}_{\text{Tx}}^H(\theta) \mathbf{F} \mathbf{F}^H \dot{\mathbf{a}}_{\text{Tx}}(\theta)|^2}{\|\mathbf{F}^H \mathbf{a}_{\text{Tx}}(\theta)\|_2^2} \right) \right]^{-1}, \quad (9)$$

where $\mathbf{b}_{\text{Rx},k} = \mathbf{W}_k^H \mathbf{a}_{\text{Rx}}(\phi)$, $\dot{\mathbf{a}}_{\text{Tx}}(\theta) \triangleq d\mathbf{a}_{\text{Tx}}(\theta)/d\theta$ and the signal to noise ratio (SNR) is defined as $\rho \triangleq |\alpha|^2 P/\sigma_n^2$. Given the CRB of the AoD estimate and the prior AoD range $\mathcal{R}_{\text{Tx},k}$, the precoder matrix \mathbf{F}_k at k -th time slot is obtained as follows:

$$\mathbf{F}_k = \arg \min_{\mathbf{F}} \max_{\theta \in \mathcal{R}_{\text{Tx},k}} \text{DEB}(\theta, \mathbf{F}), \quad (10)$$

where $\mathcal{R}_{\text{Tx},k} = [\hat{\theta}_{k-1} - 3\sigma_{\theta,k-1}, \hat{\theta}_{k-1} + 3\sigma_{\theta,k-1}]$ with the standard deviation $\sigma_{\theta,k-1}$ of the prior AoD estimates, obtained from the channel estimation (see Section III-D).

C. Combiner Design With the AoA Uncertainty

For a combiner matrix, each combiner steers towards equispaced directions within the prior range of the AoA [16], i.e.,

$$\mathbf{W}_k = \sqrt{\frac{N_{\text{RF}}}{N_{\text{R}}}} \begin{pmatrix} \tilde{\mathbf{a}}_{\text{Rx}}(\tilde{\phi}_{k,1}) & & 0 \\ & \ddots & \\ 0 & & \tilde{\mathbf{a}}_{\text{Rx}}(\tilde{\phi}_{k,N_{\text{RF}}}) \end{pmatrix}, \quad (11)$$

where $\tilde{\phi}_{k,1}, \dots, \tilde{\phi}_{k,N_{\text{RF}}}$ are the equispaced directions within the prior AoA range $\mathcal{R}_{\text{Rx},k} = [\hat{\phi}_{k-1} - 3\sigma_{\phi,k-1}, \hat{\phi}_{k-1} + 3\sigma_{\phi,k-1}]$ with the standard deviation $\sigma_{\phi,k-1}$ of the prior AoA estimates, obtained from the channel estimation (see Section III-D); $\tilde{\mathbf{a}}_{\text{Rx}}(\tilde{\phi}_{k,n_{\text{RF}}}) \in \mathbb{C}^{(N_{\text{R}}/N_{\text{RF}}) \times 1}$ because each RF chain is routed to $N_{\text{R}}/N_{\text{RF}}$ antennas only.

D. Channel Estimation

While any channel estimator can be utilized, we propose to use the maximum likelihood (ML) estimator, the error covariance of which we approximate with the CRB (this approximation is valid for a large number of measurements). At the end of $(k-1)$ -th time slot, the user has access to the consecutive observations $\mathbf{Y}_{1:k-1} = \{\mathbf{Y}_1, \dots, \mathbf{Y}_{k-1}\}$, and the unbiased estimate $\hat{\boldsymbol{\eta}}_{k-1}$ of the $\boldsymbol{\eta}$ can be computed by

$$\hat{\boldsymbol{\eta}}_{k-1} = \arg \max_{\boldsymbol{\eta}} f(\mathbf{Y}_{1:k-1} | \boldsymbol{\eta}). \quad (12)$$

1) *Error Covariance in the Channel Parameter Domain:* The channel parameter vector $\boldsymbol{\eta} \in \mathbb{R}^{4 \times 1}$ can be defined as $\boldsymbol{\eta} = [\theta, \phi, \Re\alpha, \Im\alpha]^T$, where $\Re\alpha$ is real part and $\Im\alpha$ is imaginary part of the channel gain α . The error covariance is bounded as

$$\mathbb{E}_{\boldsymbol{\eta}} \left[(\hat{\boldsymbol{\eta}}_{k-1} - \boldsymbol{\eta}) (\hat{\boldsymbol{\eta}}_{k-1} - \boldsymbol{\eta})^T \right] \succeq \mathbf{J}_{\boldsymbol{\eta}}^{-1}, \quad (13)$$

in which $\mathbf{J}_{\boldsymbol{\eta}} \in \mathbb{R}^{4 \times 4}$ is the Fisher information matrix (FIM). The FIM $\mathbf{J}_{\boldsymbol{\eta}}$ is defined as

$$\mathbf{J}_{\boldsymbol{\eta}} \triangleq -\mathbb{E} \left[\frac{\partial^2 \ln f(\mathbf{Y}_{1:k-1} | \boldsymbol{\eta})}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}^T} \right], \quad (14)$$

where $\ln f(\mathbf{Y}_{1:k-1} | \boldsymbol{\eta})$ is log-likelihood function of the consecutive observations $\mathbf{Y}_{1:k-1}$ up to the $(k-1)$ -th time slot conditioned on $\boldsymbol{\eta}$. Since these observations are conditionally independent, $\ln f(\mathbf{Y}_{1:k-1} | \boldsymbol{\eta}) = \sum_{k'=1}^{k-1} \ln f(\mathbf{Y}_{k'} | \boldsymbol{\eta})$ in which

$$\ln f(\mathbf{Y}_{k'} | \boldsymbol{\eta}) \propto \frac{1}{\sigma_n^2} \left\| \mathbf{Y}_{k'} - \sqrt{P} \alpha \mathbf{W}_{k'}^H \mathbf{a}_{\text{Rx}}(\phi) \mathbf{a}_{\text{Tx}}^H(\theta) \mathbf{F}_{k'} \right\|_{\text{F}}^2, \quad (15)$$

from which $\mathbf{J}_{\boldsymbol{\eta}}$ can be easily computed, following the approach in [17].

Since the true channel is not known at $(k-1)$ -th time slot, $\hat{\boldsymbol{\eta}}_{k-1}$ is used to evaluate the FIM, denoted by $\mathbf{J}_{\boldsymbol{\eta}}(\hat{\boldsymbol{\eta}}_{k-1}, \mathbf{F}_{1:k-1})$. The $\mathbf{J}_{\boldsymbol{\eta}}(\hat{\boldsymbol{\eta}}_{k-1}, \mathbf{F}_{1:k-1})$ is the FIM accumulated from \mathbf{Y}_1 (transmitted by \mathbf{F}_1) to \mathbf{Y}_{k-1} (transmitted by \mathbf{F}_{k-1}) as follows:

$$\mathbf{J}_{\boldsymbol{\eta}}(\hat{\boldsymbol{\eta}}_{k-1}, \mathbf{F}_{1:k-1}) = \sum_{k'=1}^{k-1} \mathbf{J}_{\boldsymbol{\eta}}(\hat{\boldsymbol{\eta}}_{k-1}, \mathbf{F}_{k'}). \quad (16)$$

2) *Error Covariance in the Channel Domain:* The error covariance matrix $\mathbf{C}_{k-1}(\hat{\boldsymbol{\eta}}_{k-1}, \mathbf{F}_{1:k-1})$ of $\mathbf{h} = \text{vec}(\mathbf{H}) \in \mathbb{C}^{N_{\text{R}} N_{\text{T}} \times 1}$ can be calculated by a Jacobian transformation matrix $\mathbf{T} \in \mathbb{C}^{N_{\text{R}} N_{\text{T}} \times 4}$:

$$\mathbf{C}_{k-1}(\hat{\boldsymbol{\eta}}_{k-1}, \mathbf{F}_{1:k-1}) = \mathbf{T}(\hat{\boldsymbol{\eta}}_{k-1}) \mathbf{J}_{\boldsymbol{\eta}}^{-1}(\hat{\boldsymbol{\eta}}_{k-1}, \mathbf{F}_{1:k-1}) \mathbf{T}^H(\hat{\boldsymbol{\eta}}_{k-1}), \quad (17)$$

where the transformation matrix $\mathbf{T} = \partial \mathbf{h} / \partial \boldsymbol{\eta}$. Consequently, we obtain the entries of $\mathbf{T}(\hat{\boldsymbol{\eta}}_{k-1})$ replacing $\boldsymbol{\eta}$ with $\hat{\boldsymbol{\eta}}_{k-1}$ as

$$\mathbf{T} = \left[\frac{\partial \mathbf{h}}{\partial \theta}, \frac{\partial \mathbf{h}}{\partial \phi}, \frac{\partial \mathbf{h}}{\partial \Re\alpha}, \frac{\partial \mathbf{h}}{\partial \Im\alpha} \right], \quad (18)$$

where

$$\frac{\partial \mathbf{h}}{\partial \theta} = -\alpha j \pi \sin \theta (\dot{\mathbf{a}}_{\text{Tx}}(\theta) \otimes \mathbf{a}_{\text{Rx}}(\phi)), \quad (19)$$

$$\frac{\partial \mathbf{h}}{\partial \phi} = \alpha j \pi \sin \phi (\mathbf{a}_{\text{Tx}}^*(\theta) \otimes \dot{\mathbf{a}}_{\text{Rx}}(\phi)), \quad (20)$$

$$\frac{\partial \mathbf{h}}{\partial \Re\alpha} = \mathbf{a}_{\text{Tx}}^*(\theta) \otimes \mathbf{a}_{\text{Rx}}(\phi), \quad (21)$$

$$\frac{\partial \mathbf{h}}{\partial \Im\alpha} = j (\mathbf{a}_{\text{Tx}}^*(\theta) \otimes \mathbf{a}_{\text{Rx}}(\phi)), \quad (22)$$

in which \otimes is the Kronecker product.

E. The Decision to Stop Pilot Transmissions

At the beginning of k -th time slot, the rate can be evaluated. The combiner $\tilde{\mathbf{u}}_k$ and precoder $\tilde{\mathbf{v}}_k$ for data transmission are designed resorting to the maximum gain criteria:

$$\tilde{\mathbf{u}}_k, \tilde{\mathbf{v}}_k = \arg \max_{\substack{\mathbf{u}_k, \mathbf{v}_k \\ \|\mathbf{u}_k\|=1, \|\mathbf{v}_k\|=1}} |\mathbf{u}_k^H \hat{\mathbf{H}}_{k-1} \mathbf{v}_k|^2. \quad (23)$$

The rate is then

$$R_{k-1} = (1 - \frac{T_i + (k-1)T_s}{T_f}) \log_2 \left(1 + \frac{P |\tilde{\mathbf{u}}_k^H \hat{\mathbf{H}}_{k-1} \tilde{\mathbf{v}}_k|^2}{\sigma_n^2 + P \sigma_{e,k-1}^2(\tilde{\mathbf{u}}_k, \tilde{\mathbf{v}}_k)} \right), \quad (24)$$

where T_i is an initial channel acquisition duration, T_s is a time slot period, and

$$\begin{aligned} \sigma_{e,k-1}^2(\tilde{\mathbf{u}}_k, \tilde{\mathbf{v}}_k) \\ = (\tilde{\mathbf{v}}_k^* \otimes \tilde{\mathbf{u}}_k)^H \mathbf{C}_{k-1}(\hat{\boldsymbol{\eta}}_{k-1}, \mathbf{F}_{1:k-1}) (\tilde{\mathbf{v}}_k^* \otimes \tilde{\mathbf{u}}_k), \end{aligned} \quad (25)$$

accounts for the channel estimation error on the down-link effective rate. Note that the channel estimation error

Algorithm 1 Optimizing the Channel Estimation Duration Scheme by Rate Prediction

Initialization: Obtain initial angle ranges ($\mathcal{R}_{\text{Tx},1}$ and $\mathcal{R}_{\text{Rx},1}$) by performing the initial beam sweeping

- 1: **for** $k = 1, 2, \dots$ **do**
- 2: Design \mathbf{F}_k to minimize $\text{DEB}(\theta, \mathbf{F})$ from (10)
- 3: **if** $k = 1$ **then**
- 4: Estimate $\hat{\mathbf{H}}_1$ by transmitting pilots with \mathbf{F}_1
- 5: Compute $\mathbf{C}_1(\hat{\boldsymbol{\eta}}_1, \mathbf{F}_1)$ according to (17)
- 6: Design $\tilde{\mathbf{v}}_2, \tilde{\mathbf{u}}_2$ from (23) given $\hat{\mathbf{H}}_1$
- 7: Compute the effective rate R_1 from (24)
- 8: **else**
- 9: Predict $\hat{\mathbf{C}}_k(\hat{\boldsymbol{\eta}}_{k-1}, \mathbf{F}_{1:k})$ according to (26)
- 10: Compute the predicted rate \hat{R}_k from (27)
- 11: **if** $\hat{R}_k > R_{k-1}$ **then**
- 12: Estimate $\hat{\mathbf{H}}_k$ by transmitting pilots with \mathbf{F}_k
- 13: Compute $\mathbf{C}_k(\hat{\boldsymbol{\eta}}_k, \mathbf{F}_{1:k})$ according to (17)
- 14: Design $\tilde{\mathbf{v}}_{k+1}, \tilde{\mathbf{u}}_{k+1}$ from (23) given $\hat{\mathbf{H}}_k$
- 15: Compute the effective rate R_k from (24)
- 16: **else**
- 17: The BS sends the data by $\tilde{\mathbf{v}}_k$ and $\tilde{\mathbf{u}}_k$
- 18: **end if**
- 19: **end if**
- 20: **end for**

$\sigma_{e,k-1}^2(\tilde{\mathbf{u}}_k, \tilde{\mathbf{v}}_k)$ decreases over the time slot as the FIM accumulates as in (16).

Now a decision should be taken whether or not to transmit the pilot \mathbf{F}_k . To make this decision, we consider two hypothetical scenarios:

- If we don't send further pilots, the rate is R_{k-1} .
- If we were to send at least one more pilot, we can predict the channel error covariance that would be realized after pilot transmissions as

$$\hat{\mathbf{C}}_k(\hat{\boldsymbol{\eta}}_{k-1}, \mathbf{F}_{1:k}) = \mathbf{T}(\hat{\boldsymbol{\eta}}_{k-1}) \mathbf{J}_{\boldsymbol{\eta}}^{-1}(\hat{\boldsymbol{\eta}}_{k-1}, \mathbf{F}_{1:k}) \mathbf{T}(\hat{\boldsymbol{\eta}}_{k-1})^T. \quad (26)$$

From this predicted error covariance, we predict the rate

$$\hat{R}_k = \left(1 - \frac{T_i + kT_s}{T_f}\right) \log_2 \left(1 + \frac{P|\tilde{\mathbf{u}}_k^H \hat{\mathbf{H}}_{k-1} \tilde{\mathbf{v}}_k|^2}{\sigma_n^2 + P\hat{\sigma}_{e,k}^2(\tilde{\mathbf{u}}_k, \tilde{\mathbf{v}}_k)}\right), \quad (27)$$

where

$$\hat{\sigma}_{e,k}^2(\tilde{\mathbf{u}}_k, \tilde{\mathbf{v}}_k) = (\tilde{\mathbf{v}}_k^* \otimes \tilde{\mathbf{u}}_k)^H \hat{\mathbf{C}}_k(\hat{\boldsymbol{\eta}}_{k-1}, \mathbf{F}_{1:k}) (\tilde{\mathbf{v}}_k^* \otimes \tilde{\mathbf{u}}_k), \quad (28)$$

accounts for the predicted channel estimation error on the downlink effective rate. Note that both the prediction of the error covariance and the rate are performed prior to transmission of \mathbf{F}_k .

Finally, our decision rule is as follows: If $\hat{R}_k < R_{k-1}$, the pilot \mathbf{F}_k is not sent, the BS keeps the current combiner $\tilde{\mathbf{u}}_k$ and precoder $\tilde{\mathbf{v}}_k$, and switches to data transmission mode with rate R_{k-1} . If $\hat{R}_k \geq R_{k-1}$, the BS transmits the pilot \mathbf{F}_k to generate \mathbf{Y}_k , based on which the channel is estimated and

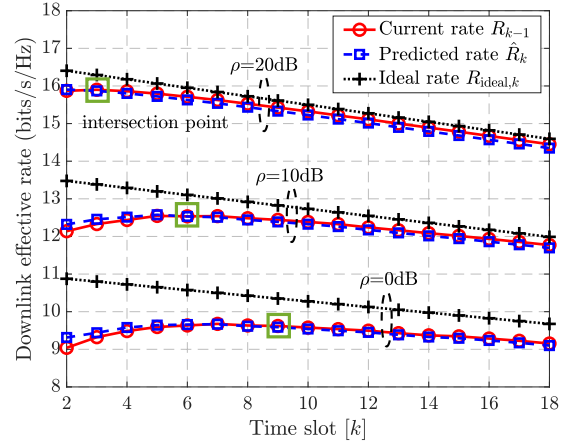


Fig. 3. The comparison between the current rate R_{k-1} , the predicted rate \hat{R}_k and the ideal rate $R_{ideal,k}$ with the time slot.

the CRB $\mathbf{C}_k(\boldsymbol{\eta}, \mathbf{F}_k)$ can be computed. The estimated channel $\hat{\mathbf{H}}_k$ is distributed as $\hat{\mathbf{h}}_k \sim \mathcal{CN}(\mathbf{h}, \mathbf{C}_k(\boldsymbol{\eta}, \mathbf{F}_k))$. The decision process is then repeated at $(k+1)$ -th time slot. A summary of the proposed scheme is presented in Algorithm 1.

Regarding the complexity of the proposed scheme, most of the procedures in the proposed scheme are carried out at the BS where computational complexity is not an issue in general. Also, the proposed scheme includes operations, that are usually considered in a wide range of communication algorithms, such as beamformer design and the CRB calculations. Therefore, the proposed scheme is considered to be practical.

IV. NUMERICAL RESULTS

The channel gain, AoD and AoA of multipath signals are determined based on the 28 GHz mmWave outdoor channel model [14]. According to [14], we set $L = 2$ where one dominant path and one secondary path exist. The power of the secondary path is 10 dB lower than the power of the dominant path. The transmitted power P is fixed to 1, and the noise power σ_n^2 is $1/\rho$. The θ_l and ϕ_l are distributed uniformly on $[0, \pi]$. We set $M_T = M_R = 4$, so four beams ($\pi/4$ first null beamwidth) which are equispaced within $[0, \pi]$ are used for initial beam sweeping [15]. In Fig. 2, we set $T_f = 160 T_s$, $T_i = 16 T_s$ (4 Tx beams \times 4 Rx beams = 16 time slot) and $Q = 3$. Both BS and MS are equipped with $N_T = N_R = 64$ and $N_{RF} = 4$.

Fig. 3 shows the comparison between the current rate R_{k-1} , the predicted rate \hat{R}_k and the ideal rate $R_{ideal,k}$ with the time slot. The ideal rate $R_{ideal,k}$ under the perfect channel state information (CSI) in the initial beam sweeping phase can be defined as

$$R_{ideal,k} = \left(1 - \frac{T_i + kT_s}{T_f}\right) \log_2 \left(1 + \frac{P|\tilde{\mathbf{u}}^H \mathbf{H} \tilde{\mathbf{v}}|^2}{\sigma_n^2}\right), \quad (29)$$

where $\tilde{\mathbf{u}}$ and $\tilde{\mathbf{v}}$ are the combiner and precoder maximizing the gain. As shown in Fig. 3, the ideal rate $R_{ideal,k}$ is an upper bound of the downlink effective rate and decreases over the time slot as pilot transmission time increases. The \hat{R}_k and R_{k-1} are plotted concave due to deficient or excessive pilot transmissions. If the \hat{R}_k is lower than the R_{k-1} , the BS decides to stop transmitting pilots and start sending the data.

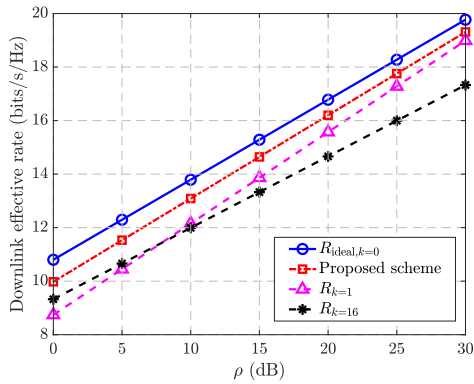


Fig. 4. Downlink effective rate achieved by the proposed scheme vs. the $R_{ideal,k=0}$ and the $R_{k=1,16}$ with the SNR (ρ).

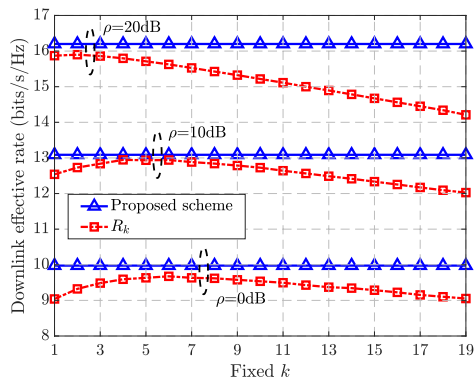


Fig. 5. Downlink effective rate achieved by the proposed scheme vs. R_k with the fixed k .

As ρ increases, a good channel estimation can be achieved with a relatively smaller number of pilot transmissions, which results in the early termination of pilot transmissions (see the intersection points).

In Fig. 4, we compare the proposed scheme to the ideal rate $R_{ideal,k=0}$ and the $R_{k=1,16}$. The proposed scheme approaches to the $R_{ideal,k=0}$ as the ρ increases, and always higher than the $R_{k=1}$ and $R_{k=16}$, as shown in Fig. 4. For the high SNR regime, the intersection point gets smaller, and the downlink effective rate achieved after pilot transmissions stop becomes closer to $R_{k=1}$. Contrarily, for the low SNR regime, the downlink effective rate achieved becomes closer to $R_{k=16}$. In this simulation, 200 Monte Carlo runs were carried out to compute the downlink effective rate achieved by the proposed scheme, and the intersection points were mostly located between the time slot 1 and 16.

Fig. 5 shows the comparison between the downlink effective rate achieved by the proposed scheme and the R_k with the fixed k . Note that R_k with the fixed k is the downlink effective rate achieved by sending pilots k times only, without the rate comparison. It confirms that proposed scheme achieves the highest downlink effective rate for different SNR scenarios. This result was also obtained with 200 Monte Carlo runs.

V. CONCLUSION

We have proposed a scheme that optimizes the channel estimation duration to maximize the downlink effective rate

by rate prediction. Prior to actual pilot transmissions, the BS computes the predicted rate with the prior channel estimates and pilot precoder. Comparing the predicted rate and current rate, the BS determines whether to continue transmitting pilots. The proposed scheme provides an elegant way of balancing the channel estimation duration and the estimation accuracy to maximize the downlink effective rate, and its performance has been analyzed by numerical simulations under various settings. Future work will be its extension to a multi-user MIMO (MU-MIMO) system by taking into account the multi-user interference.

REFERENCES

- [1] T. Nitsche, C. Cordeiro, A. Flores, E. Knightly, E. Perahia, and J. Widmer, "IEEE 802.11ad: Directional 60 GHz communication for multi-Gigabit-per-second Wi-Fi [invited paper]," *IEEE Commun. Mag.*, vol. 52, no. 12, pp. 132–141, Dec. 2014.
- [2] G. Wunder, S. Stefanatos, A. Flinthe, I. Roth, and G. Caire, "Low-overhead hierarchically-sparse channel estimation for multiuser wide-band massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 18, no. 4, pp. 2186–2199, Apr. 2019.
- [3] H. Shokri-Ghadikolaei, L. Gkatzikis, and C. Fischione, "Beam-searching and transmission scheduling in millimeter wave communications," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2015, pp. 1292–1297.
- [4] A. Alkhateeb, Y.-H. Nam, M. S. Rahman, J. Zhang, and R. W. Heath, "Initial beam association in millimeter wave cellular systems: Analysis and design insights," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2807–2821, May 2017.
- [5] H. Zhou, D. Guo, and M. L. Honig, "Beam acquisition and training in millimeter wave networks with narrowband pilots," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 12, pp. 2759–2771, Dec. 2019.
- [6] M. Mousaei and B. Smida, "Optimizing pilot overhead for ultra-reliable short-packet transmission," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–5.
- [7] 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.
- [8] M. Kokshoorn, H. Chen, P. Wang, Y. Li, and B. Vucetic, "Millimeter wave MIMO channel estimation using overlapped beam patterns and rate adaptation," *IEEE Trans. Signal Process.*, vol. 65, no. 3, pp. 601–616, Feb. 2017.
- [9] O. Igbafe, J. Kang, H. Wymeersch, and S. Kim, "Location-aware beam alignment for mmWave communications," 2019, *arXiv:1907.02197*. [Online]. Available: <http://arxiv.org/abs/1907.02197>
- [10] H. L. Van Trees and K. L. Bell, *Detection, Estimation, and Modulation Theory*. Hoboken, NJ, USA: Wiley, 2013.
- [11] N. Garcia, H. Wymeersch, and D. T. M. Slock, "Optimal precoders for tracking the AoD and AoA of a mmWave path," *IEEE Trans. Signal Process.*, vol. 66, no. 21, pp. 5718–5729, Nov. 2018.
- [12] 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. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, Apr. 2016.
- [13] Z. Abu-Shaban, X. Zhou, T. Abhayapala, G. Seco-Granados, and H. Wymeersch, "Error bounds for uplink and downlink 3D localization in 5G millimeter wave systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 8, pp. 4939–4954, Aug. 2018.
- [14] M. K. Samimi and T. S. Rappaport, "3-D millimeter-wave statistical channel model for 5G wireless system design," *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 7, pp. 2207–2225, Jul. 2016.
- [15] V. Desai, L. Krzymien, P. Sartori, W. Xiao, A. Soong, and A. Alkhateeb, "Initial beamforming for mmWave communications," in *Proc. 48th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2014, pp. 1926–1930.
- [16] R. Mendez-Rial, C. Rusu, N. Gonzalez-Prelcic, A. Alkhateeb, and R. W. Heath, "Hybrid MIMO architectures for millimeter wave communications: Phase shifters or switches?" *IEEE Access*, vol. 4, pp. 247–267, Jan. 2016.
- [17] A. Shahmansoori, G. E. Garcia, G. Destino, G. Seco-Granados, and H. Wymeersch, "Position and orientation estimation through millimeter-wave MIMO in 5G systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1822–1835, Mar. 2018.