Achievable Rate Analysis of Millimeter Wave Channels Using Random Coding Error Exponents

Shaocheng Huang, Ming Xiao, Senior Member, IEEE, and H. Vincent Poor, Life Fellow, IEEE

Abstract—With emerging applications, e.g., factory automation, autonomous driving and augmented/virtual reality, there have been increasing technical challenges regarding reliability, latency and data rates for existing communication systems. Owing to abundant available bandwidth, millimeter Wave (mmWave) communications can potentially provide reliable communication with an order of magnitude capacity improvement relative to microwave, e.g., sub 6 GHz communications. Though there are many research results showing improved throughputs, the latency and reliability performance of mmWave communications is still not quite clear, especially for finite blocklength regimes. In this paper, we investigate achievable rates of mmWave channels using random coding error exponents. Under the assumption of perfect and imperfect channel state information at the receiver (CSIR), exact and approximate analytical expressions of achievable rates are derived to capture the relationships among rate, latency and reliability. Furthermore, we show that the achievable rate always increases as the bandwidth increases with perfect CSIR. However, there exists a critical bandwidth that maximizes the achievable rate for non-line-of-sight mmWave signals with imperfect CSIR, beyond which the achievable rate will decrease with increasing bandwidth. For imperfect CSIR, the training symbol length and power allocation factor for maximizing the achievable rate at the training phase are investigated and closed-form expressions for special cases are derived.

Index Terms—Millimeter wave, random coding error exponent, achievable rate, reliability, latency.

I. INTRODUCTION

A. Background and Motivation

RECENTLY, emerging applications, e.g., factory automation, autonomous driving, smart grids and augmented/virtual reality, have posed tremendous technical challenges to existing communication systems in terms of reliability, latency and rates [1]–[4]. For instance, reliability and latency requirements for factory automation can be as low as $10^{-9}$ in packet loss rates (PLR) and 250 ms in latency, respectively, while the requirements for smart grids are $10^{-6}$ in PLR and 3 ms–20 ms in latency [5]. Meanwhile, stringent requirements of latency and high data rates are of utmost importance for augmented/virtual reality applications, which require very high data rates (> 1 giga-bit per second) and low latency (1 ms–10 ms) [3], [6]. To meet the stringent requirements, a number of candidate technologies have been investigated: (1) higher frequency spectrum, e.g., millimeter Wave (mmWave) communications [1], [7]; (2) finite block-length coding [8]; and (3) shorter transmission time interval (TTI). For instance, one can reduce latency by shortening orthogonal frequency-division multiplexing (OFDM) symbol duration via wider subcarrier spacing, which can be implemented with larger bandwidth [2].

There has been increasing research interest in investigating the performance of mmWave communications under reliability and latency constraints. From a network layer perspective, several critical challenges and possible strategies for low-latency mmWave networks were reviewed in [9]. By invoking the Lyapunov optimization framework, the problem of ultra-reliability and low-latency in mmWave enabled massive multiple-input multiple-output (MIMO) networks was investigated in [10]. Furthermore, by applying stochastic network calculus theory, the delay performance of buffer-aided mmWave systems was investigated and the bound for probabilistic delay was provided in [11]. Considering dynamic blockage due to mobile blockers, the expected blockage duration for mmWave cellular systems was studied in [3].

However, the results are still not clear on what the performance in terms of latency and reliability will be for mmWave communications with finite blocklength, especially from a fundamental theoretical perspective. Although capacity has been regarded as a typical performance metric, it assumes infinitely long blocklength. Thus, capacity may not be a suitable metric for low-latency communications [2]. Instead, the error exponent [12] and channel dispersion [13] are used to analyze the achievable rate from a finite blocklength information-theoretic perspective. The channel dispersion in [13] is generalized to wireless fading channels with perfect channel state information.
at the receiver (CSIR). The channel dispersion of single-input single-output (SISO) scalar coherent fading channels was investigated in [14], and the result has been generalized to isotropic multiple-input single-output fading channels [15] and MIMO block-fading channels [16]. For situations in which neither the transmitter nor the receiver has a priori channel state information (CSI), the channel dispersion of MIMO Rayleigh fading channels has been studied in [17], [18]. However, it is hard to find the analytical channel dispersion for general fading channels with imperfect CSIR at the transceiver, since the optimal input distribution is in general unknown [19]. The classical lower bound derived by Gallager [20], known as the random coding error exponent (RCEE) or Gallager’s exponent, is available to investigate the achievable rate \( R(N, \epsilon) \) with given blocklength \( N \) and decoding error probability \( \epsilon \). There have been a few results on the RCEE of fading channels with perfect or imperfect CSIR. With perfect CSIR, the exact RCEE for Rayleigh fading channels was presented in [21], [22], while the approximate RCEE for independent and identically distributed (i.i.d.) MIMO Rayleigh fading channels in the low-signal-to-noise ratio (SNR) regime was derived in [23]. In addition, the exact RCEE and the approximate expression in the high-SNR regime for the orthogonal space-time block coded Nakagami-m fading MIMO channels were studied in [24]. For the scenario with imperfect CSIR, the RCEE of SISO Rayleigh fading channels was studied with channel estimation [25]. The approximate RCEE of i.i.d. MIMO Rayleigh block-fading channels in the low-SNR regime was investigated in [26], which is based on the assumption that the coherence length is scaled with the inverse of SNR. Moreover, the approximate RCEE of non-coherent correlated MIMO Rayleigh block-fading channels at the low-SNR regime was derived in [27]. However, the investigations of RCEE above cannot be directly applied to mmWave communications since they do not consider the influence of bandwidth and the propagation features of mmWave signals.

Many measurement results for mmWave communications have been reported because the available bandwidth is orders of magnitude larger than microwave communication systems [7], [28]–[30]. For instance, a total of 12.9 GHz bandwidth is available in E-band (60 – 90 GHz). However, comparing to microwave communication systems, the following propagation features should be taken into account in mmWave communications: (a) mmWave channels can be high line-of-sight (LoS) or high non-line-of-sight (NLoS) as shown in the measurements [28], [29]; (b) mmWave communications are vulnerable to blockage by obstructions due to high penetration loss. For instance, bricks can attenuate signals as much as 40 – 80 dB and the human body itself can result in 20 – 35 dB signal loss [30]; (c) mmWave communications suffer from frequency selective absorption due to their large bandwidth since there are varying oxygen absorption and water vapor absorption properties in mmWave bands [7]. On the other hand, the capacity of mmWave fading channels with imperfect CSIR is studied in [31], [32] and it is shown that, with a power constraint, the capacity reaches a peak rate at some finite critical bandwidth, and might decrease with increasing bandwidth.

Based on above observations, it is valuable to study achievable rates for mmWave communications from a finite block-length information-theoretic perspective, which may serve as a design guideline for future latency, reliability and rates constrained applications with mmWave. With the large available bandwidth in mmWave communications, the packet duration or TTI might be reduced by shortening the symbol duration via increased subcarrier spacing. In addition, it is also interesting to investigate the critical bandwidth of mmWave fading channels with imperfect CSIR in the finite blocklength regime.

B. Main Contributions

In what follows, we will leverage the RCEE to investigate the achievable rate of mmWave channels under reliability and packet duration (finite blocklength) constraints. The main contributions are as follows:

1. With perfect CSIR, the ergodic capacity and the achievable rate under the constraints of reliability and packet duration for mmWave channels are investigated. We first derive analytical expressions for the ergodic capacity and the achievable rate for mmWave SISO channels. Then, we further investigate their approximate expressions in the low-SNR regime. The rate loss caused by the constraints on reliability and packet duration is evaluated and the lower bound is given. We show that, for a given achievable rate, the required packet duration can be decreased with increasing bandwidth. Furthermore, we extend our analysis to mmWave MIMO channels. Specifically, we first derive the achievable rate for an mmWave MIMO channel with an arbitrary-rank LoS channel component. Then, we derive the achievable rate for a special mmWave MIMO channel with an orthogonal LoS channel component.

2. With imperfect CSIR, we first derive analytical expressions for the training based ergodic rate and the training based achievable rate under the constraints of reliability and packet duration. Then, the optimal power allocation factor and the optimal training symbol length at the training phase are investigated. Results show that, for SISO cases, one training symbol is enough in each coherence block. We also provide closed-form expressions for the optimal power allocation factor for special cases. Results show that there might exist a critical bandwidth to maximize the training based achievable rate for the NLoS scenario.

The rest of the paper is organized as follows: In Section II, we give the system model. In Section III, we first provide analytical expressions for the ergodic capacity and the achievable rate for mmWave SISO channels with perfect CSIR. Then, we provide their approximations in low-SNR regimes. In Section IV, we provide the achievable rate with imperfect CSIR. In Section V, achievable rates for mmWave MIMO channels with perfect CSIR are investigated. Numerical results are given in Section VI, while conclusions are given in Section VII.

Notation: For a matrix \( A \), the determinant is denoted by \( \det(A) \) and transpose conjugate is denoted by \( (A)\dagger \). The trace operator is denoted by \( \text{tr}(A) \) and \( \text{etr}(A) = e^{\text{tr}(A)} \). The expectation and variance of a random variable are denoted by \( \mathbb{E}(\cdot) \) and \( \mathbb{V}(\cdot) \), respectively. The complex normal distribution with mean \( \mu \) and variance \( \sigma^2 \) is denoted by \( \mathcal{CN}(\mu, \sigma^2) \).

II. SYSTEM MODEL

We consider a block-fading channel with coherence time \( T_c \) and coherence bandwidth \( W_c \), and assume that the fading
coefficients within each coherence block are identical and change independently.\textsuperscript{1} For instance, coherence time in the scenarios of the walking speed (1 m/s) and the train speed (100 m/s) for 60 GHz radio, are approximately 2.5 ms and 25 μs, respectively, \cite{31}. Measuring at 60 GHz in an indoor environment, the coherence bandwidth of NLoS and LoS channels are around 6 MHz and 20 MHz, respectively \cite{33}. Let \(L_t = T/T_c\) and \(K_w = W/W_c\) denote the number of coherence blocks in time and frequency domain, respectively, and \(N_t = T_c/T_0\) and \(N_w = W_c/W_0\) denote the number of dimensions of each coherence block in time and frequency domain, respectively, where \(T_0\) and \(W_0\) are the effective signaling duration and frequency separation, respectively.\textsuperscript{2} We assume that each codeword (packet) spans \(TW\) time-frequency (T-F) domain with the dimension of \(N = TW/T_0W_0L_tK_wN_c\) consisting of \(L_tK_w\) sub-codewords and each sub-codeword consisting of \(N_c = N_tN_w\) symbols. The transceiver is equipped with a single antenna. The channel input-output relation within the \((i,j)\)-th T-F block, \(i = 1, 2, \ldots, L_t - 1\) and \(j = 1, 2, \ldots, K_w\), can be expressed as

\[
y_{ij} = h_{ij}x_{ij} + n_{ij},
\]

where \(x_{ij}, y_{ij}, n_{ij} \in \mathbb{C}^{N_c}\). We also assume that all entries of \(n_{ij}\) are i.i.d. complex Gaussian with distribution \(\mathcal{CN}(0, N_0)\). This discrete model in (1) can be derived from the discretization of a continuous-time model through transmission and reception over a time interval \([0, T]\) and frequency interval \([0, W]\) of a Weyl-Heisenberg set \cite{37}. The reader is referred to \cite{31, 37} for a detailed exposition of the mapping from the continuous to the discrete representations.

To capture the features of mmWave bands, we use the mmWave channel model in \cite{31} and the fading coefficient of one coherence block is defined as

\[
h_{ij} = a_i v_j g_{ij},
\]

where \(a_i \sim \text{Bern}(1 - P_B)\) denotes the effect of blockage in time block \(i\), and blockage probability is \(P_B\); \(v_j\) denotes the average frequency-selective power absorption at sub-band \(j\) (value can be measured by the gaseous attenuation model in \cite{38}); \(g_{ij} = \sqrt{\frac{K_w}{1 + k}} + \sqrt{\frac{1}{1 + k}} g_w\) where \(k\) is Rician factor expressing the ratio of the energy in the LoS path to the energy in the NLoS path, and \(g_w \sim \mathcal{CN}(0, 1)\) represents the NLoS scattering component. We assume that each codeword is under the average power constraint

\[
\frac{1}{T} \sum_{i=1}^{L_t} \sum_{j=1}^{K_w} E[|x_{ij}|^2] \leq P_r = pW,
\]

where \(P_r\) is the total power per second and \(p\) is the average power per symbol. We assume that the transmitter has no CSI and the receiver has perfect/imperfect CSI.

\textsuperscript{1}In practice, the simplified wideband block-fading channel is not very physically accurate with regard to mmWave channels since the inter-block correlation and intra-block variation are ignored for tractable analysis. However, it is a common assumption, which can be used to reveal information-theoretic relation between bandwidth and achievable rate \cite{31, 32}, and we use it for tractable analysis.

\textsuperscript{2}For OFDM systems without guard interval, \(T_0W_0 = 1\), while \(T_0W_0 > 1\) with guard interval or cyclic prefix between successive symbols \cite{34}. For instance, \(T_0\) and \(W_0\) are approximately 71.4 μs and 15 KHz in LTE standard \cite{35}, and approximately 218 ns and 5.16 MHz in IEEE 802.15.3c \cite{36}.

## III. Achievable Rates of mmWave SISO Channels With Perfect CSIR

In what follows, we will analyze the ergodic capacity and the achievable rate under the reliability and packet duration constraints for the mmWave channel with perfect CSIR. We first determine the exact expressions of achievable rates and ergodic capacity. Then, the approximate expressions of achievable rates and ergodic capacity are derived in the low-SNR regime. Finally, the rate loss caused by the constraints of reliability and packet duration is investigated.

### A. Achievable Rates With Perfect CSIR

Using the formulation defined in \cite{20, 22}, we obtain the random coding bound of error probability of maximum likelihood decoding as

\[
\epsilon \leq \epsilon^U(R) = e^{-\frac{T}{E}\rho R(R, T)},
\]

where \(E_R(R, T)\) is RCEC, which can be expressed as

\[
E_R(R, T) = \max_{0 \leq \rho \leq 1} \{ \max_{p_X(x)} E_0(p_X(x), T, \rho) - \rho R \},
\]

where \(p_X(x)\) is the distribution of \(x\), rate \(R\) is measured by unit nats per second, and \(E_0(p_X(x), T, \rho)\) is given in (6), shown at the bottom of the page.

With Gaussian inputs, the ergodic capacity can be obtained by taking the derivative of \(E_0(p_X(x), T, \rho)\) with respect to \(\rho\) at \(\rho = 0\) \cite{39}, given by

\[
C = \frac{\partial E_0(p_X(x), T, \rho)}{\partial \rho} \bigg|_{\rho=0}.
\]

Let \(\epsilon^U_{th}\) denote the threshold of decoding error probability \(\epsilon^U\). The achievable rate below channel capacity and under the constraints of \(T\) and \(\epsilon^U_{th}\) can be defined as \cite{40}

\[
R^*(T, \epsilon^U_{th}) = \max_{\rho \in RC} \{ R : \epsilon^U(R) \leq \epsilon^U_{th} \}.\]

Then, the exact expression for achievable rates is provided in the following theorem.

Theorem 1: For the mmWave channel in (1) with perfect CSIR and i.i.d. Gaussian inputs under the power constraint in (3), the achievable rate is

\[
R_c(T, \epsilon^U_{th}, k) = \max_{0 \leq \rho \leq 1} \frac{1}{\rho} \left( E_0(v, T, \rho) + \frac{\log(\epsilon^U_{th})}{T} \right),
\]

\[
E_0(v, T, \rho) = -\frac{L_t}{T} \sum_{j=1}^{K_w} \sum_{n=1}^{\infty} \frac{k}{n!} U(N_c; N_c \rho - n; \varphi_j) + P_B,
\]

where \(\varphi_j = \frac{(1+k)(1+\rho)N_0}{v_j P_B} v = [v_1, \ldots, v_{K_w}]\) is the power absorption vector; \(U(\cdot)\) is the Tricomi hypergeometric function.

Proof: See Appendix A.

To assess the convergence of the infinite series in (10), we define

\[
f(N_a) = \sum_{n=0}^{N_a} \frac{k^n}{n!} U(N_c; N_c \rho - n; \varphi_j).
\]
Assuming that $n_0$ terms are used in $f(N_0)$, the truncated error $f(\infty) - f(n_0 - 1)$ can be upper bounded as

$$f(\infty) - f(n_0 - 1) = \sum_{n=n_0}^{\infty} \frac{k^n}{n!} U(N_c\rho; N_c\rho - n; \varphi_j)$$

where (a) follows from the property that $U(b; b - n; c)$ is a monotonically decreasing function of $n$ when Re($b$) > 0 (A detailed proof of this property is provided in Appendix B). Then, the ergodic capacity is given in the following corollary.

**Corollary 1:** For the mmWave channel in (1) with perfect CSIR and i.i.d. Gaussian inputs under the power constraint in (3), the ergodic capacity is given as

$$C = \frac{(1 - P_B) L_{t} N_{e}}{T} \sum_{j=0}^{K_w} \sum_{n=0}^{\infty} \frac{k^n e^{-k}}{(n!)^2} G^{1,3}_{3,2} \left[ \frac{p c_j^2}{N_0 (1 + k)} \left| -n, 1, 1 \right| 1, 0 \right],$$

where $G^{p,q}_{r,s} \left[ x \left| b_1, \ldots, b_q \right| c_1, \ldots, c_p \right]$ is Meijer’s G-function [41].

**Proof:** See Appendix C.

The convergence of the infinite series of ergodic capacity in (13) is provided in Appendix D.

### B. Achievable Rates With Perfect CSIR in the Low-SNR Regime

For the power constrained mmWave communication, SNR per degree of freedom may be low due to both large bandwidth and severe attenuation at mmWave bands. Thus, we will investigate the approximate $R_c(T, \epsilon_{th}^U, k)$ and $C$ in the low-SNR regime. We first derive the approximate expressions of $E_0(v, T, \rho)$ in (10) at low-SNR regimes in the following corollary.

**Corollary 2:** $E_0(v, T, \rho)$ in (10) at low SNRs can be approximated as

$$E_0(v, T, \rho) = -L_t \frac{K_w}{T} \log \left( \frac{1 - P_B}{1 + k + \Lambda_j} e^{-\Lambda_j} + P_B \right),$$

where $\Lambda_j = \frac{N_c \rho v^2}{N_0 (1 + \rho)}$.

**Proof:** To obtain the low-SNR approximation of $E_0(v, T, \rho)$, we need to determine the low-SNR approximation of the term $F_g(v_j, \rho, N_c, k)$ in (40), namely,

$$F_g(v_j, \rho, N_c, k) = \mathbb{E}_g \left[ \left( 1 + \frac{v_j^2 |y_j|^2}{N_0 (1 + \rho)} \right)^{-N_c \rho} \right] \approx \mathbb{E}_g \left[ \exp \left( -\frac{N_c \rho v_j^2 |y_j|^2}{N_0 (1 + \rho)} \right) \right]$$

where (a) follows from the log(1 + $x$) $\approx$ $x$ if $x$ $\approx$ 0, and (b) follows from the infinite series representation of the modified Bessel function of the first kind $I_0(\cdot)$.

With the result of Corollary 2, the low-SNR approximation of the achievable rate is given as follows

$$R_c^U(T, \epsilon_{th}^U, k) = \max_{0 < \rho \leq 1} \frac{1}{\rho} \left( E_0^U(v, T, \rho) + \frac{\log(\epsilon_{th}^U)}{T} \right).$$

Meanwhile, based on the definition in (7), the ergodic capacity at low-SNR regimes can be obtained by taking the derivative of $E_0^U(v, T, \rho)$ with respect to $\rho$ at $\rho = 0$, given as follows

$$C^U = \frac{\partial E_0^U(v, T, \rho)}{\partial \rho} \bigg|_{\rho = 0} = \frac{(1 - P_B) P K_w N_0}{T} \sum_{j=0}^{K_w} \sum_{n=0}^{\infty} u_j^2.$$  

Note that the result in (17) shows that the ergodic capacity at the low-SNR regime is not affected by Rician factor $k$. Although the ergodic capacity has been regarded as a typical performance metric, it assumes infinite long blocklength. It is valuable to investigate rate loss, i.e., $C - R_c(T, \epsilon_{th}^U, k)$, caused by the finite packet duration and reliability constraints.

Combining the results in Corollary 2 and $C^U$ in (17), the rate loss at the low-SNR regime is given in the following corollary.

**Corollary 3:** With packet duration $T$ and decoding error probability $\epsilon_{th}^U$, rate loss at the low-SNR regime is positive and lower bounded as

$$C^U - R_c^U(T, \epsilon_{th}^U, k) \geq \frac{\log(\epsilon_{th}^U)}{T} > 0.$$  

**Proof:** See Appendix E.

From Corollary 3, we can see that there exists positive rate loss under finite packet duration and reliability constraints. Rate loss is greater than $\log(\epsilon_{th}^U)/T$, which depends on factors $T$ and $\epsilon_{th}^U$. These results imply that the rate loss bound is inversely proportional to the signal duration $T$. Thus, it is inversely proportional to blocklength $N$. 

$$E_0(\rho X(x), T, \rho) = -\frac{1}{T} \sum_{i=1}^{L_t} \sum_{j=1}^{K_w} \log \left( \mathbb{E}_{h_{ij}} \left[ \int_{y_{ij}} \left( \int_{x_{ij}} p(x_{ij}) p(y_{ij} | x_{ij}, h_{ij}) y_{ij}^{1+\rho} d_{x_{ij}}^{1+\rho} d_{y_{ij}} \right) \right] \right)$$
for a given bandwidth $W$. In particular, when $T \to \infty$, \[ \lim_{T \to \infty} \log \left( \frac{1}{T} \right) = 0. \] On the other hand, results also show that the rate loss bound is inversely and logarithmically proportional to $\epsilon_{th}$. Thus, when the reliability requirements decrease, the rate loss bound decreases. In particular, when $\epsilon_{th}^U \to 1$, \[ \lim_{\epsilon_{th}^U \to 1} \log \left( \frac{1}{\epsilon_{th}^U} \right) / T = 0. \]

IV. ACHIEVABLE RATES OF mmWAVE SISO CHANNELS WITH IMPERFECT CSIR

In what follows, we will analyze the achievable rate under the reliability and packet duration constraints for the mmWave channel without perfect CSIR. We first determine the analytical expressions of training based achievable rates. Then, the optimal power allocation factor and the optimal training symbol length at training phase are investigated.

We shall investigate the achievable rate under the training based communication scheme, which uses a part of the transmission resources as a training signal to learn the channel. More specifically, we assume that within the $(i,j)$-th coherence block for $i = 1, 2, \ldots, L_t - 1$ and $j = 1, 2, \ldots, K_n$, $N_r^r$ dimensions are used for training, and $D_{ij} = N_c - N_r^r$ dimensions are used for information. We further assume that $\eta_{ij} \in (0, 1)$ is the power allocation factor at the information transmission phase. For simplicity, we use $D$ with $\{D\}_{ij} = D_{ij}$ and $\eta_{ij}$ with $\{\eta\}_{ij} = \eta_{ij}$ to denote the information symbol length matrix and power allocation factor matrix, respectively. Thus, the channel model for the $(i,j)$-th coherence block is denoted by $Y_{ij} = h_{ij}X_{ij} + n_{ij}$, where

\[
Y_{ij} = [X_{ij}^r \ X_{ij}^d]; \quad Y_{ij} = [\gamma_{ij}^r \ \gamma_{ij}^d] \quad \text{and} \quad n_{ij} = [n_{ij}^r \ n_{ij}^d],
\]

are input, output and noise vector, respectively, and superscripts $r$ and $d$ denote training phase and information transmission phase, respectively. Let $E_{ij}^r = (1 - \eta_{ij})E_{N_c}$ denote the allocated power at the training phase, where $E_{N_c} = N_cP$ is the total power in one deterministic coherence block. With an minimum mean-square error (MMSE) estimator [42], we can obtain the estimated channel given by

\[
\hat{h}_{ij} = \frac{\mathbb{E}(h_{ij})}{(1 + \mathbb{V}(h_{ij})E_{ij}^r)}(X_{ij}^r)^\dagger \times ((h_{ij} - \mathbb{E}(h_{ij}))X_{ij}^r + n_{ij}^r) + \mathbb{E}(h_{ij}),
\]

where $\mathbb{E}(h_{ij}) = (1 - P_B)v_j \sqrt{\frac{1}{1 + k}}$ and $\mathbb{V}(h_{ij}) = (1 - P_B)v_j^2(1 - k(1 - P_B))$ are the expectation and variance of the true channel $h_{ij}$ in the $(i,j)$-th coherence block, respectively. Assuming $\hat{h}_{ij} = \hat{h}_{ij} + \hat{n}_{ij}$, the output at the information transmission phase is

\[
\gamma_{ij}^d = \hat{h}_{ij}X_{ij}^d + \hat{n}_{ij}X_{ij}^r + n_{ij}^d.
\]

Since the term $\hat{\gamma}_{ij}^d$ in (20) may not be Gaussian and the distribution of $\hat{\gamma}_{ij}^d$ is unknown, it is in general very hard to obtain closed-form RCE. According to [42]–[44], the capacity of the channel in (20) is minimized if $\hat{\gamma}_{ij}^d$ is assumed to be i.i.d. Gaussian. With the same assumption, the RCE of Rayleigh fading channels is studied in [26]. In practice, as the SNR is high, $\hat{\gamma}_{ij}^d$ can be approximated as being Gaussian. For tractable analysis, we assume that $\hat{\gamma}_{ij}^d$ has i.i.d. Gaussian entries, i.e., $\hat{\gamma}_{ij}^d \sim \mathcal{CN}(0, \hat{N}_0I_{D_{ij}})$ with $\hat{N}_0 = \hat{N}_0 + \mathbb{V}(h_{ij})\mathbb{E}(X_{ij}^d\gamma_{ij}^d)$. Under this provision, we provide the training based achievable rate of the estimated channel (19) in the following theorem.

Theorem 2: With the estimated channel in (19) and $\hat{\gamma}_{ij}^d \sim \mathcal{CN}(0, \hat{N}_0I_{D_{ij}})$, the training based achievable rate under i.i.d. Gaussian inputs with the power constraint in (3), is

\[
R_T(T, \epsilon_{th}^U) = \max_{0 < \epsilon \leq 1} \frac{1}{\rho} \left( \mathbb{E}_0 \{v, T, \rho, \eta, D\} + \frac{\log(\epsilon_{th}^U)}{T} \right),
\]

where $\mathbb{E}_0 \{v, T, \rho, \eta, D\} = \frac{1}{T} \sum_i \sum_j \log \left( (1 - P_B)F_{h_{ij} | a = 1} \right) + P_BF_{h_{ij} | a = 0}.

Theorem 2: \[ R_T(T, \epsilon_{th}^U) = \max_{0 < \epsilon \leq 1} \frac{1}{\rho} \left( \mathbb{E}_0 \{v, T, \rho, \eta, D\} + \frac{\log(\epsilon_{th}^U)}{T} \right), \]

where $\mathbb{E}_0 \{v, T, \rho, \eta, D\} = \frac{1}{T} \sum_i \sum_j \log \left( (1 - P_B)F_{h_{ij} | a = 1} \right) + P_BF_{h_{ij} | a = 0}.

Proof: See Appendix F.

As shown in Theorem 2, it is clear that, within each coherence block, the choice of information symbol length $D_{ij}$ and power allocation factor $\eta_{ij}$ will affect the performance of the training based achievable rate. We first determine the optimal information symbol length $D_{ij}^*$ in the following corollary.

Corollary 4: The optimal information symbol length $D_{ij}^*$ for $R_T(T, \epsilon_{th}^U)$ in Theorem 2 is $D_{ij} = N_c - 1$ for $\eta_{ij} \neq 1$ and $D_{ij} = N_c$ for $\eta_{ij} = 1$.

Proof: See Appendix G.

Note that the result in Corollary 4 is consistent with that in [42], in which the capacity lower bound of Rayleigh fading channels is analyzed in infinite blocklength regimes. Corollary 4 reveals that, with imperfect CSIR, we should choose $D_{ij}$ as large as possible to maximize the training based achievable rate $R_T(T, \epsilon_{th}^U)$. Furthermore, the optimal information symbol length $D_{ij}^*$ in each coherence block only depends on the power allocation factor $\eta_{ij}$. It is difficult to derive the optimal power allocation factor $\eta_{ij}$ in a closed-form for...
\[ R_r(T, c_{th}^I). \] Instead, we provide the closed-form expression of \( \eta_{ij}^* \) for some special cases in the following corollary.

**Corollary 5:** If \( k = 0 \) and \( P_B = 0 \), the optimal power allocation factor \( \eta_{ij}^* \) for \( R_r(T, c_{th}^I) \) in Theorem 2 is given as follows

\[
\eta_{ij}^* = \begin{cases} 
\theta_{ij} - \sqrt{\theta_{ij}^2 - \theta_{ij}}, & D_{ij} > 1, \\
\frac{1}{2}, & D_{ij} = 1,
\end{cases}
\]  

(25)

where \( \theta_{ij} = \frac{D_{ij}V(h_{ij})E_{N_c}N_0}{\sqrt{V(h_{ij})E_{N_c}(D_{ij} - 1)}}. \)

**Proof:** See Appendix H.

Note that the optimal power allocation factor in (25) is consistent with that in [42] for training based ergodic rate of Rayleigh block-fading channels. Then, the training based ergodic rate, and its corresponding optimal \( \eta_{ij} \) and optimal \( D_{ij} \) are given in the following corollary.

**Corollary 6:** With the estimated channel in (19) and \( \hat{W}_{ij} \sim \mathcal{CN}(0, \hat{N}_iL_{D_{ij}}) \), the training based ergodic rate under i.i.d. Gaussian inputs with the power constraint (3), is given as follows

\[
C_r = \sum_{i=1}^{L} \sum_{j=1}^{K} \frac{D_{ij}}{T} \left[ (1 - P_B)C_{h_i[a=1] + P_B C_{h_i[a=0]}} \right],
\]

(26)

\[
C_{h_i[a=\epsilon]} = \exp \left( -\frac{E(\hat{h}_{ij} | a = \epsilon)^2}{\sqrt{V(h_{ij}) | a = \epsilon}} \right) \sum_{n=0}^{\infty} \frac{E(\hat{h}_{ij} | a = \epsilon)2^{n}}{n!} \times G_{ij}^{1,3} \left[ \frac{\eta_{ij}^* E_{N_c} V(h_{ij} | a = \epsilon)}{D_{ij}^* N_0} \right],
\]

(27)

where \( D_{ij}^* \) and \( \eta_{ij}^* \) are given in following three cases:

1. \( D_{ij}^* = N_c, \eta_{ij}^* = 1 \) for \( V(h) = 0 \); (b) \( D_{ij}^* = N_c - 1, \eta_{ij}^* = \frac{1}{2}\phi_{ij} \) for \( V(h) \neq 0 \) and \( N_c = 2 \); (c) \( D_{ij}^* = N_c - 1, \eta_{ij}^* = \theta_{ij} - \sqrt{\theta_{ij}^2 - \theta_{ij}^2} \) for \( V(h) \neq 0 \) and \( N_c > 2 \), where \( \phi_{ij} = 1 + \frac{E(\hat{h}_{ij})^2}{V(h_{ij})} \).

**Proof:** A detailed proof is omitted here since the training based ergodic rate can be obtained with similar approaches to Corollary 1. Meanwhile, \( D_{ij}^* \) and \( \eta_{ij}^* \) can be obtained with similar approaches to Corollary 4 and Corollary 5, respectively.

Note that, if channel variance \( V(h_{ij}) = 0 \), there is no need to allocate power for channel estimation since the channel is already known at the receiver, and \( D_{ij}^* = N_c, \eta_{ij}^* = 1 \). We can also find that the expression of \( D_{ij}^* \) is consistent with that in Corollary 4. Meanwhile, the expression of \( \eta_{ij}^* \) with \( k = 0 \) and \( P_B = 0 \) is consistent with that in Corollary 5. The training based ergodic rate for SISO cases in [32] is a special case (with \( K = 0 \) and \( P_B = 0 \)) of the derived training based ergodic rate in Corollary 6.

V. ACHIEVABLE RATES FOR mmWAVE MIMO CHANNELS WITH PERFECT CSIR

In what follows, we extend our analysis to mmWave MIMO channels with perfect CSIR. Specifically, we first derive the achievable rate for an mmWave MIMO channel with an arbitrary-rank LoS channel component. Then, we derive the achievable rate for a special mmWave MIMO channel with an orthogonal LoS channel component.

With similar channel assumptions to Sec.II, we consider a MIMO block-fading channel with \( M_t \) transmit and \( M_r \) receive antennas. The channel input-output relation within the \((i, j)\)-th T-F block, \( i = 1, 2, \ldots, L \) and \( j = 1, 2, \ldots, K \), can be expressed as

\[
Y_{ij} = H_{ij} X_{ij} + N_{ij},
\]

(28)

where \( X_{ij} \in \mathbb{C}^{M_t \times N_c} \) and \( Y_{ij} \in \mathbb{C}^{M_r \times N_c} \) are the transmitted and received symbol matrices, \( N_{ij} \in \mathbb{C}^{M_r \times N_c} \) is additive noise with i.i.d. zero-mean unit-variance complex Gaussian entries. The channel \( H_{ij} \in \mathbb{C}^{M_r \times M_t} \) can be modeled as

\[
H_{ij} = a_{ij} G_{ij},
\]

(29)

where \( G_{ij} = \sqrt{\frac{1}{N_c}} G + \sqrt{\frac{1}{N_c}} G_w, \) \( G \) represents the (arbitrary) rank-1 deterministic LoS channel component with \( \text{tr}(G G^\dagger) = M_r M_t \), while \( G_w \) accounts for the normalized random NLoS component with i.i.d. zero-mean unit-variance complex Gaussian entries and \( \mathbb{E}[\text{tr}(G_w G_w^\dagger) ] = M_r M_t \). We then define the random matrix \( W \) as

\[
W = \begin{cases} 
(k + 1) G_{ij} G_{ij}^\dagger, & \text{if } M_r \leq M_t, \\
(k + 1) G_{ij}^\dagger G_{ij}, & \text{otherwise}.
\end{cases}
\]

(30)

Here, \( W \in \mathbb{C}^{S \times S} \) follows a complex uncorrelated non-central Wishart distribution with \( M \) degrees of freedom [45], i.e., \( W \sim W_m(M, \Omega, I_s) \), where \( S = \min(M_r, M_t) \), \( M = \max(M_r, M_t) \), and

\[
\Omega = \begin{cases} 
k G G^\dagger, & \text{if } M_r \leq M_t, \\
k G^\dagger G, & \text{otherwise}.
\end{cases}
\]

(31)

is the rank-\( L \) non-centrality matrix. When the non-centrality matrix \( \Omega \) has \( L \) non-zero ordered eigenvalues, i.e., \( 0 < \omega_1 < \ldots < \omega_L \), the joint PDF of the ordered eigenvalues \( \lambda_S \leq \ldots \leq \lambda_1 \) of \( W \) is given by [46]

\[
f(\lambda) = \mathbb{E}[\Phi(\lambda)|\Psi(\lambda)] \prod_{s=1}^{S} \lambda_{s}^{M-S} e^{-\lambda_s},
\]

(32)

where \( \mathbb{E} = \frac{e^{-\mathbb{M}(\lambda)}}{\prod_{s=1}^{S} (\mathbb{M}-s+1)^{-s}} \prod_{s=1}^{S} (\omega_1^{-s+1})|\Psi(\omega)|^2 \) is a Vandermonde matrix, \( \omega = [\omega_1, \ldots, \omega_L], \lambda = [\lambda_1, \ldots, \lambda_S] \), and

\[
\{\Phi(\lambda)\}_{l,u} = \begin{cases} 
0 F_{1} (M - S + 1; \omega_u \lambda_l), & u = 1, \ldots, L, \\
\lambda_{S-u}^{M-S} (M-S)!, & u = L + 1, \ldots, S.
\end{cases}
\]

(33)

The achievable rate of the mmWave channel in (28) under the packet duration and reliability constraints is provided in the following theorem.

**Theorem 3:** For the mmWave channel in (28) with perfect CSIR and i.i.d. Gaussian inputs, if \( \Omega \) has \( L \) non-zero ordered
eigenvalues, i.e., \(0 < \omega_2 < \ldots < \omega_1\), the achievable rate is
\[
R_m(T,c_{th}^l,k) = \max_{0 < c_{th}^l \leq 1} \left( -\frac{L}{T} \sum_{j} \log \left( 1 - P_B \right) \Xi \det(\Delta) + P_B \right) + \frac{\log(c_{th}^l)}{T},
\]
(34)
where \(\{\Delta\}_{l,u}\) for \(l,u = 1, \ldots, S\) is given in (35), shown at the bottom of the page, and \(\Upsilon_j = \frac{p_0\rho^2}{M_o(1+\rho)(1+k)}\).

Proof: See Appendix I.

Note, if \(M_t = M_r = 1\), the result in Theorem 3 is consistent with that in Theorem 1. Meanwhile, Theorem 3 is a general result for MIMO channels with arbitrary-rank LoS channel components. If some eigenvalues of \(\Omega\) are equal or zero, we can obtain the achievable rate by taking the limit of (32) [47]. For instance, when the correlation between LoS responses is eliminated and the LoS channel matrix is spatially orthogonal, we may have \(\Omega = kM_tI_{S}\), i.e., \(\Omega = \ldots = \omega_S = kM_t\). Recent results show that the maximum capacity of LoS-MIMO channels can be obtained by employing specifically designed antenna arrays to preserve the orthogonality of LoS-MIMO channels [48]–[51]. Tractable array design methodologies for maximizing the capacity of mmWave LoS-MIMO channels are provided in [50], [51]. Considering orthogonal LoS channel components, the achievable rate under the packet duration and reliability constraints is provided in the following corollary.

Corollary 7: For the mmWave channel in (28) with perfect CSIR and i.i.d. Gaussian inputs, if \(\Omega = kM_tI_{S}\), the achievable rate is
\[
R_m(T,c_{th}^U,k) = \max_{0 < c_{th}^U \leq 1} \left( -\frac{L}{T} \sum_{j} \log \left( 1 - P_B \right) \Xi_o \det(\Delta_o) + P_B \right) + \frac{\log(c_{th}^U)}{T},
\]
(36)
where \(\{\Delta_o\}_{l,u}\) for \(l,u = 1, \ldots, S\) is given in (35), shown at the bottom of the page, and \(\Upsilon_j = \frac{p_0\rho^2}{M_o(1+\rho)(1+k)}\).

Proof: By taking the limit of (32) with \(\omega_1 = \ldots = \omega_S = kM_t\), we can obtain the joint eigenvalue PDF of \(\Omega\) [52]. Then, following a similar approach to Appendix G, we conclude the result of Corollary 7.

Since there are as yet no analytical distributions of estimated mmWave MIMO channels, it is very hard to derive a closed-form expression for the RCEE or achievable rates of mmWave MIMO channels with imperfect CSIR at this time. This analysis is thus left for future work.

VI. NUMERICAL RESULTS

In this section, numerical results are provided to evaluate the achievable rate of mmWave channels under packet duration and reliability constraints. Specifically, with the experimental results in [29], the achievable rate as a function of bandwidth is evaluated for LoS and NLoS scenarios, in which LoS and NLoS pathloss models in the 5G Channel Model (5GCM) are used to measure the large-scale fading. Furthermore, Monte Carlo simulations are provided to evaluate the effectiveness of analytical results. Subsequently, with given achievable rates, the trade-off between packet duration and bandwidth is investigated. Finally, the impact of coherence time and coherence bandwidth on critical bandwidth is studied. The parameters for two scenarios, i.e., LoS and NLoS, are shown in Table I.

The following numerical results are based on \(f_c = 73\) GHz, and the frequency-selective power absorption is measured by the gaseous attenuation model in [38]. We consider the medium-mobility scenario with a corresponding speed of \(20\) m/s (i.e., Doppler shift about \(10\) KHz) [53], and thus coherence time is \(T_c \approx 0.1\) ms. According to the 3rd Generation Partnership Project (3GPP) standardization [54], a normal delay spread of \(200\) ns in the urban area is reasonable and thus we set \(W_c = 5\) MHz. The received power is normalized by noise power as \(\frac{P}{N_0}\), where \(N_0 = -174\) dBm/Hz.

Fig. 1 shows the ergodic capacity and the achievable rate with different decoding error probability and packet duration constraints, as functions of bandwidth for the NLoS scenario with perfect CSIR. We can see that the ergodic capacity and achievable rate both increase with increasing bandwidth. We can also see that the achievable rate decreases with increasing decoding error probability (i.e., higher reliability requirements). The gap between the ergodic capacity and achievable rate becomes smaller with larger \(c_{th}^l\) or \(T\). This matches the result in Corollary 3. We can observe that, higher achievable rates can be obtained with longer packet duration. The reason is that increasing \(T\) does not change the SNR per degree of freedom but it shall increase codeword length. Thus, increasing \(T\) can reduce the decoding error probability according to the reliability function in (4). In addition, we can observe that to achieve a given rate, one can decrease the required packet duration with the increasing bandwidth. Moreover, we can find that the approximations of the ergodic capacity and the achievable rate become tighter as bandwidth increases since the SNR per degree of freedom is becoming lower with the increasing bandwidth.

\[
\{\Delta\}_{l,u} = \left\{ \begin{array}{ll}
\sum_{m=1}^{\infty} \frac{(M - S)!((M + m - l)!\omega_u^m}{m!((M - S)!(M + m - u - l)!}\frac{1}{\Upsilon_j}U(N_c\rho, N_c\rho - M - m + l, 1, \frac{1}{T_j}), & u = 1, \ldots, L, \\
\sum_{m=1}^{\infty} \frac{(M - S)!((M + m - u)!\omega_u^m}{m!((M - S)!(M + m - u - l)!}\frac{1}{\Upsilon_j}U(N_c\rho, N_c\rho - M - S + u + l, 1, \frac{1}{T_j}), & u = L + 1, \ldots, S,
\end{array} \right.
\]
(35)
Fig. 1. Ergodic capacity and achievable rates vs. bandwidth with different decoding error probability and packet duration constraints for the NLoS scenario with perfect CSIR and $M_r = M_t = 1$.

Fig. 2. Ergodic rates and achievable rates vs. bandwidth with different decoding error probability and packet duration constraints for the NLoS scenario with imperfect CSIR and $M_r = M_t = 1$.

Fig. 3. Achievable rates vs. bandwidth with different decoding error and packet duration constraints for the LoS scenario with/without perfect CSIR and $M_r = M_t = 1$.

Fig. 4 shows the achievable rate as a function of bandwidth for the LoS scenario with multiple antennas and perfect CSIR. Fig. 4 confirms the close agreement between analysis results and simulation results, thereby validating the effectiveness of analytical expressions for achievable rates. Furthermore, it is

Fig. 2 shows the training based ergodic rate and the training based achievable rate with different decoding error probability and packet duration constraints, as functions of bandwidth for the NLoS scenario with imperfect CSIR. It is shown that the training based ergodic rate and the training based achievable rate both increase and then decrease with increasing bandwidth. Thus, there is a critical bandwidth for each of the training based ergodic rate and the training based achievable rate. The critical bandwidth for the training based ergodic rate is smaller than that for the training based achievable rate since we need more bandwidth to achieve the peak rate under the constraints of reliability and packet duration. In addition, we can see that if more stringent reliability (i.e., smaller decoding error probability) is required, then the larger bandwidth is needed to achieve the peak rate. Similar to Fig. 1, we can also see that, a higher rate gain can be obtained with longer packet duration for the scenario with imperfect CSIR. Moreover, it is shown that the shorter packet duration we need, the larger bandwidth is required to achieve the peak rate.

Fig. 3 shows the achievable rate with different decoding error probability and packet duration constraints, as a function of bandwidth for the LoS scenario with/without perfect CSIR. We can see that achievable rates increase as the packet duration and decoding error probability increases. Comparing the result to the NLoS scenario in Fig. 1 and Fig. 2, we can observe that one can obtain a significant rate improvement in the LoS scenario. The reason is that, in mmWave bands, attenuation for NLoS propagation is much greater than that for the LoS propagation. The results for the LoS scenario are based on the unlimited bandwidth, and the achievable rate reaches a peak rate at around $W = 4$ GHz for the case with imperfect CSIR. Thus, it is reasonable to achieve a higher rate with larger bandwidth from a practical perspective for the LoS scenario.

Fig. 4 shows the achievable rate as a function of bandwidth for the LoS scenario with multiple antennas and perfect CSIR. Fig. 4 confirms the close agreement between analysis results and simulation results, thereby validating the effectiveness of analytical expressions for achievable rates. Furthermore, it is
shown that the achievable rate significantly increases when the number of transceiver antennas increase. Finally, we can see that the achievable rate with full-rank LoS channel components outperforms than that with rank-1 LoS channel components. This result is consistent with that in [52].

Fig. 5 shows the ergodic capacity and the achievable rate with different decoding error probability and packet duration constraints for the NLoS scenario with perfect CSIR. We can see that the ergodic capacity and achievable rate both significantly increase with increasing $M_r$. Even with stringent requirements of reliability and packet duration (i.e., $\epsilon_{th} = 10^{-6}$ and $T_d = 0.1$ ms) for the NLoS scenario, we can still achieve communication rates of more than 500 Mbps when $M_r = M_t = 35$.

Fig. 6 shows the required packet duration with a given achievable rate $R_c = R_{th} = 15$ Mbps and different requirements of reliability and blockage probability, i.e., $\epsilon_{th} = \{10^{-6}, 10^{-9}\}$ and $P_B = \{0.1, 0.2\}$, as a function of bandwidth for the NLoS scenario with perfect CSIR. We can see that there is a tradeoff between packet duration and bandwidth, and the required packet duration is inversely proportional to the bandwidth. It is shown that one can decrease the required packet duration with increasing bandwidth, which might be feasible in mmWave bands with abundant available bandwidth. Furthermore, we can see that for a given rate and packet duration, the blockage probability and the decoding error probability can be reduced with increasing bandwidth. Finally, we can observe that the required packet duration significantly decreases when the number of transceiver antennas increases.

Fig. 7 shows the impact of coherence time and coherence bandwidth on the critical bandwidth for the NLoS scenario with packet duration $T_d = 0.4$ ms and imperfect CSIR. The coherence time $T_c = 0.05$ ms and $T_c = 0.4$ ms correspond to the speed of 41 m/s (high speed) and 5 m/s (walking speed), respectively. It is shown that the critical bandwidth increases with increasing coherence time, and the curve is approximately linear. The reason is that the dimension of the coherence block (i.e., $N_c$) increases with increasing $T_c$, and thus the estimated channel approaches the perfect channel. A similar
observation can also be found in [31], [32]. In addition, we can see that if $T_c = 0.4$ ms and $W_c = 10$ MHz for a low-mobility scenario, the critical bandwidth is in the order of several GHz. Thus, even for NLoS scenarios, it is possible to achieve a higher rate gain with a larger bandwidth from a practical perspective. On the other hand, the critical bandwidth is not very large (about 200 MHz) in a high-mobility scenario (e.g., $T_c = 0.05$ ms). This means that if too much spectrum is allocated to the user, it may reduce the achievable rate.

VII. CONCLUSION

In this paper, we have used the RCEE to investigate the achievable rate of mmWave channels under reliability and packet duration constraints. Both exact and approximate expressions for the achievable rate have been derived in a closed-form. The optimal information symbol length and optimal power allocation factor for the training based achievable rate have been investigated. Results show that, at each coherence block, the information symbol length should be as large as possible for the training based achievable rate. In addition, with perfect CSIR, the achievable rate increases via increasing bandwidth. However, with imperfect CSIR, the achievable rate might decrease with increasing bandwidth after some critical bandwidth for NLoS scenarios. Furthermore, results also reveal that, for a given achievable rate, one could reduce the packet duration and increase the reliability with increasing bandwidth. Finally, we have extended our analysis to mmWave MIMO channels with perfect CSIR. Results show that the required packet duration significantly decreases when the number of transceiver antennas increases. The achievable rate with full-rank LoS channel components outperforms than that with rank-1 LoS channel components.

APPENDIX A

PROOF OF THEOREM 1

With perfect CSIR and Gaussian inputs, we have the transition probability density function (PDF) at the $(i, j)$-th block

$$p(y_{ij} | x_{ij}, h_{ij}) = \frac{\text{etr} \left(-N_0(1)(y_{ij} - h_{ij}x_{ij})(y_{ij} - h_{ij}x_{ij})^\dagger\right)}{(\pi N_0)^{N_c}}. \quad (38)$$

Substituting (38) into (6), we have

$$E_0(v, T, \rho) \mathord{\overset{\Delta}{=}} E_0(p_x(x), T, \rho) = \frac{1}{T} \sum_{i=1}^{L_s} \sum_{j=1}^{K_w} \log \left(E_{h_{ij}} \left[\left(1 + \frac{|h_{ij}|^2 \rho}{N_0 (1 + \rho)}\right)^{-N_c \rho}\right]\right),$$

$$= -\frac{L_s K_w}{T} \sum_{j=1}^{K_w} \log \left(1 - P_B\right) \mathbb{E}_g \left[\left(1 + \frac{v_j^2 |g|^2 \rho}{N_0 (1 + \rho)}\right)^{-N_c \rho} + P_B\right]. \quad (39)$$

For simplicity, the term $\mathbb{E}_g \left[\left(1 + \frac{v_j^2 |g|^2 \rho}{N_0 (1 + \rho)}\right)^{-N_c \rho}\right]$ in (39) is denoted as

$$F_g(v_j, \rho, N_c, k) = \mathbb{E}_g \left[\left(1 + \frac{v_j^2 |g|^2 \rho}{N_0 (1 + \rho)}\right)^{-N_c \rho}\right]. \quad (40)$$

Then the PDF of the Rician channel gain $\gamma = |g|^2$ is given as

$$P(\gamma) = (1 + k)e^{-(1+k)\gamma-k}I_0\left(\sqrt{4k}(k+1)\gamma\right). \quad (41)$$

Thereafter, substituting (41) into (40) and with the following identity in (42) [41], we have an analytical expression of $F_g(v_j, \rho, N_c, k)$ given in (43),

$$F_g(v_j, \rho, N_c, k) = \varphi_j^{N_c \rho} e^{-k} \sum_{n=0}^{k^n} \frac{k^n}{n!} U(N_c \rho; N_c \rho - n; \varphi_j), \quad (43)$$

where $\Gamma(\cdot)$ is gamma function and $\varphi_j = \frac{(1+k)(1+\rho)}{v_j^2 \rho}$. Then, substituting (43) into (39), we have desired result in (10).

Since the random coding upper bound of the decoding error probability $e^U$ is an increasing function of rate $R$, the maximum achievable rate can be obtained by solving the equation $e^U(R) = e^U(\hat{R})$ and the proof of Theorem 1 is concluded.

APPENDIX B

Based on [55], we have

$$U(b; b - n; c) = \frac{1}{\Gamma(b)} \int_0^\infty e^{-ct \theta - 1} (1 + t)^{-n-1} dt$$

for Re$(b) > 0$. \quad (44)

In our paper, $b = N_c \rho > 0$. Then, differentiating $U(b; b - n; c)$ with regard to $n$, we have

$$\frac{d}{dn} U(b; b - n; c) \mathord{\overset{(a)}{=}} \frac{1}{\Gamma(b)} \int_0^\infty \frac{\partial}{\partial n} e^{-ct \theta - 1} (1 + t)^{-n-1} dt \mathord{\overset{(b)}{<}} 0,$$
Appendix C

Proof of Corollary 1

According to equation (39) and based on the ergodic capacity definition in (6), we have

\[
C = \frac{\partial E_0(v, T, \rho)}{\partial \rho} \Big|_{\rho = 0} = \frac{N_c L_t K_u}{T} \sum_{i=1}^{L_t} \sum_{j=1}^{K_u} \log \left( 1 + \frac{|h_{ij}|^2 p}{N_0} \right)
\]

Then, substituting (41) into (46), we have

\[
C = \frac{(1 - P_B)L_t N_c}{T} \sum_{j=1}^{K_u} \log \left( 1 + \frac{v_j^2 \rho_j |g|^2 p}{N_0} \right). \quad (46)
\]

With the aid of the following identities of Meijer’s G-function [41]:

\[
\log(1 + x) = \frac{C_{1,2}^{1,2}}{\Gamma(1, 0)} \left[ \begin{aligned}
1, 1 \\
1, 0
\end{aligned} \right], \quad (48)
\]

we can evaluate the integral in (47), and arrive at the desired result in (13).

Appendix D

Proof of the Convergence of the Infinite Series in (13)

To prove the convergence of the infinite series in (13), we first define

\[
f_c(N) = \sum_{n=0}^{K_u} \frac{K_u}{(n!)^2} C_{3,2}^{1,1} \left[ \begin{aligned}
\frac{pu_j^2}{N_0(1 + k)} \\
\frac{1}{1, 0}
\end{aligned} \right] \exp \left( -n \frac{\rho_j}{\rho_j + \rho} \right), \quad (50)
\]

Based on equation (40) in [57] and equation (49) in our paper, we have

\[
G_{3,2}^{1,1} \left[ \begin{aligned}
\frac{pu_j^2}{N_0(1 + k)} \\
\frac{1}{1, 0}
\end{aligned} \right] = \frac{\Gamma(n + 1)}{(1 + k)^n} \exp \left( \frac{(1 + k)N_0}{pu_j^2} \right) \sum_{l=1}^{n+1} E_{n+2-l} \left( \frac{(1 + k)N_0}{pu_j^2} \right), \quad (51)
\]

where \( E_{n}(x) = \int_{1}^\infty e^{-t} \frac{x^{n}}{t^{n+1}} dt \) with \( \text{Re}(x) > 0 \) denotes the exponential integral function of order \( n \). Then, substituting (51) into (50), we have

\[
f_c(N) = \exp \left( \frac{(1 + k)N_0}{pu_j^2} \right) \sum_{n=0}^{K_u} \frac{K_u}{n!} \sum_{l=1}^{n+1} E_{n+2-l} \left( \frac{(1 + k)N_0}{pu_j^2} \right). \quad (52)
\]

Assuming that \( n_0 \) terms are used in \( f_c(N) \), the truncated error

\[
f_c(\infty) - f_c(n_0 - 1) \]

can be upper bounded as

\[
f_c(\infty) - f_c(n_0 - 1) = \exp \left( \frac{(1 + k)N_0}{pu_j^2} \right) \sum_{n=n_0}^{\infty} \frac{k_{n+1}}{n!} \sum_{l=1}^{n+1} E_{n+2-l} \left( \frac{(1 + k)N_0}{pu_j^2} \right) \]

where (a) follows from the fact that \( E_n(x) \) is monotonically decreasing in \( n \) with \( \text{Re}(x) > 0 \).

Appendix E

Proof of Corollary 3

To find the lower bound of rate loss at low-SNR regimes, we can firstly prove that \( C_U - R_c^U(T, \epsilon_{th}, k) \geq \frac{\log(1 + \rho)}{\rho \epsilon_{th}} \). Then the inequality \( C_U - R_c^U(T, \epsilon_{th}, k) \geq \frac{\log(1 + \rho)}{\rho \epsilon_{th}} > 0 \) can be easily obtained based on the previous approach. Let \( \rho^* \) denote the optimal \( \rho \) to maximize the right hand side of (16). Then we have

\[
C_U = C_{\rho^*} + \frac{L_t(1 - P_B)}{\rho^* T} \sum_{j} \log \left( \frac{1 + k}{1 + k + \lambda_j} \right) + P_B
\]

where:

\[
(a) \quad \geq C_{\rho^*} + \frac{L_t(1 - P_B)}{\rho^* T} \sum_{j} \left( \frac{1 + k}{1 + k + \lambda_j} - \frac{\lambda_j k}{1 + k + \lambda_j} \right) + \frac{\log(1 + \rho)}{\rho \epsilon_{th}}
\]

\[
(b) \quad \geq C_{\rho^*} - \frac{L_t(1 - P_B)}{\rho^* T} \sum_{j} \left( \frac{\lambda_j k}{1 + k + \lambda_j} \right) + \frac{\log(1 + \rho)}{\rho \epsilon_{th}}
\]

\[
(c) \quad \geq C_{\rho^*} - \frac{1}{\rho \epsilon_{th}} C_U \frac{\log(1 + \rho)}{\rho^* T}
\]

\[
(d) \quad \geq \frac{\log(1 + \rho)}{\rho \epsilon_{th}} \frac{T}{1}
\]

where:

\[
\text{a) follows from Jensen’s inequality applied to the logarithm function; b) follows from log } \left( \frac{1 + k}{1 + k + \lambda_j} \right) = - \log \left( 1 + \frac{\lambda_j k}{1 + k + \lambda_j} \right) \text{ and utilizing the log inequality}
\]
log(1 + x) ≤ x; (c) follows from the fact that \( \frac{\Lambda_{ij}^k}{1+\alpha + \mu_N} \leq \frac{\Lambda_{ij}^k}{1+\alpha} \) for \( \Lambda_{ij} \geq 0; \) (d) follows from 0 < \( \rho^* \leq 1. \) With the similar approach above, one can obtain that \( C^{U_i}(T, e^{U_i}_{th}, k) \geq \log(\frac{1}{N_0})/T. \) By showing that \( \frac{dE}{d(e^{U_i}_{th})} \) is non-negative for \( k \geq 0, \) we can find that \( R^{U_i}(T, e^{U_i}_{th}, k) \) is a monotonically increasing function of \( k \) for \( k \geq 0. \) Thus, we have \( C^{U_i}(T, e^{U_i}_{th}, k) \geq C^{U_i}(T, e^{U_i}_{th}, k). \) Combining above results, the proof of Corollary 3 is concluded.

**APPENDIX F**

**PROOF OF THEOREM 1**

With the estimated channel in (19), we have \( \hat{h}_{ij} | a = \varepsilon \sim \mathcal{CN}(E(h_{ij} | a = \varepsilon), \mathbb{V}(h_{ij} | a = \varepsilon)) \) for \( \varepsilon \in \{0, 1\}, \) where \( E(h_{ij} | a = \varepsilon) \) and \( \mathbb{V}(h_{ij} | a = \varepsilon) \) are respectively the expectation and variance of estimated channel \( \hat{h}_{ij} \) under the condition with blockage \( (a = 0) \) and without blockage \( (a = 1). \) Details of \( E(h_{ij} | a = \varepsilon) \) and \( \mathbb{V}(h_{ij} | a = \varepsilon) \) are provided in (24). Thus, the mean and variance of \( \hat{h}_{ij} \) can be \( E(h_{ij}) = E(h_{ij}) - E(h_{ij}) = 0 \) and \( \mathbb{V}(h_{ij}) = \mathbb{V}(h_{ij}) - \mathbb{V}(h_{ij}) = \frac{\mathbb{V}(h_{ij})}{1+\rho}, \) respectively. The variance of noise \( \mathbb{V}_{ij} \) is \( \tilde{N}_0 = N_0 + \frac{\mathbb{V}(h_{ij})}{1+\rho}D_{ij}. \) Then, the transition PDF is

\[
p(y_{ij} | x_{ij}, \tilde{h}_{ij}) = \text{erf}\left(-\frac{(N_0 \mathbb{V}_{ij})^{-1/2}(y_{ij} - h_{ij}x_{ij})(y_{ij} - h_{ij}x_{ij})}{(N_0 \mathbb{V}_{ij})^{1/2}}\right). \tag{55}
\]

Thereafter, substituting (55) and the estimated channel in (19) into (6), and then using a similar approach to Theorem 1, we can obtain the training based achievable rate in Theorem 2.

**APPENDIX G**

**PROOF OF COROLLARY 4**

We examine the case \( D_{ij} < N_c \) of \( \eta_{ij} \neq 1, \) since there is obviously no need to assign training data and \( D_{ij} = N_c \) if \( \eta_{ij} = 1. \) Finding the optimal \( D^{*} \) for \( R_{\tau}(T, e^{U_i}_{th}) \) equals to maximize \( E_{\mathbb{V}}(v, T, \rho, \eta, \mathbf{D}) \) over \( \mathbf{D} \) for any given \( \rho \in (0, 1). \) Thus, substituting (55) and the estimated channel in (19) into (6), we have

\[
E_{\mathbb{V}}(v, T, \rho, \eta, \mathbf{D}) = -\frac{1}{T} \sum_{i=1}^{L_t} \sum_{j=1}^{K_w} \log \left( E_{h_{ij}} \left[ \left( 1 + \frac{|\hat{h}_{ij}|^2 \eta_{ij} E_{N_c}}{(1 + \rho) N_0 D_{ij}} \right)^{-D_{ij} \rho} \right] \right). \tag{56}
\]

Following a similar procedure to [42], we can define

\[
E_B = E_{\hat{h}_{ij}} \left[ \left( 1 + \frac{|\hat{h}_{ij}|^2 \eta_{ij} E_{N_c}}{(1 + \rho) N_0 D_{ij}} \right)^{-D_{ij} \rho} \right] = E_B \left[ \left( 1 + \frac{\mathbb{V}(h_{ij})^2 \mathbb{V}(h_{ij})}{N_0} \right)^{-D_{ij} \rho} \right], \tag{57}
\]

where \( \hat{H} \) is the normalized estimated channel coefficient in one coherence block and \( SNR_{eff} = \frac{\mathbb{V}(h_{ij})^2 \mathbb{V}(h_{ij})}{N_0} \) is effective SNR. Then, differentiating \( E_B \) over \( D_{ij} \) leads to

\[
\frac{dE_B}{dD_{ij}} = -\rho E_B \left[ \left( 1 + |\hat{H}|^2 B \right)^{-D_{ij} \rho} \times \left( \log \left( 1 + |\hat{H}|^2 B \right) - \frac{1}{1 + |\hat{H}|^2 B} J \right) \right], \tag{58}
\]

where \( B = \frac{SNR_{eff}}{1+\rho}, \) and \( J = \frac{E_B}{D_{ij} + (1+\rho) E_{N_c}(1-\eta_{ij})}. \) It is obvious that \( J < 1 \) for \( D_{ij} \geq 1. \) With the inequality \( \log(1+x) < \frac{x}{1+x} < 0 \) for \( x \geq 0, \) we have \( \frac{dE_B}{dD_{ij}} < 0. \) Thus, \( E_B \) is a monotonically increasing function over \( D_{ij}. \) Therefore, we should chose \( D_{ij} \) as large as possible to maximize \( E_{\mathbb{V}}(v, T, \rho, \eta, \mathbf{D}). \) Hence, we have \( D_{ij} = N_c - 1 \) if \( \eta \neq 1 \) and \( D_{ij} = N_c \) if \( \eta = 1. \)

**APPENDIX H**

**PROOF OF COROLLARY 5**

When \( P_B = 0, \) finding the optimal \( \eta_{ij} \) to maximize \( R_{\tau}(T, e^{U_i}_{th}) \) in Theorem 2 equals to minimize the term \( F_{h_{ij},\alpha=1} \) in (23) over \( \eta_{ij}. \) For simplicity, we define

\[
F_{\eta_{ij}} = F_{h_{ij},\alpha=1} = (\psi(\eta_{ij}))^{D_{ij} \rho} U(D_{ij} \rho; D_{ij} \rho; \psi(\eta_{ij})), \tag{59}
\]

where

\[
\psi(\eta_{ij}) = \frac{(1 + \rho)}{(\mathbb{V}(h_{ij}) E_{N_c})^2 \times D_{ij} + (1 - \eta_{ij}) D_{ij} \mathbb{V}(h_{ij}) E_{N_c} + \eta_{ij} \mathbb{V}(h_{ij}) E_{N_c} \eta_{ij} (1 - \eta_{ij})} > 0. \]

Differentiating \( F_{\eta_{ij}} \) over \( \psi(\eta_{ij}) \) leads to

\[
\frac{dF_{\eta_{ij}}}{d\psi(\eta_{ij})} = D_{ij} \rho (\psi(\eta_{ij}))^{(D_{ij} \rho - 1)} \left( U(D_{ij} \rho; D_{ij} \rho; \psi(\eta_{ij})) - \psi(\eta_{ij}) \times U(D_{ij} \rho + 1; D_{ij} \rho + 1; \psi(\eta_{ij})) \right) \tag{a}
\]

\[
\leq \frac{D_{ij} \rho (\psi(\eta_{ij}))^{(D_{ij} \rho - 1)}}{\mathbb{E}_{h_{ij}}(\psi(\eta_{ij})) + \mathbb{E}_{h_{ij}}(\psi(\eta_{ij}))} \times (\mathbb{E}_{h_{ij}}(\psi(\eta_{ij})))^{D_{ij} \rho + 1; \psi(\eta_{ij})), \tag{b}
\]

where: (a) follows from the fact that \( U(u; n; z) = \mathcal{E}_n(z) z^{1-n} e^z \) with \( \mathcal{E}_n(z) \) is the exponential integral function and \( U(u; n; z) > 0 \) if \( z > 0; \) (b) follows from the fact that the exponential integral function \( \mathcal{E}_n(z) \) is a monotonically decreasing function of \( n \) as \( z > 0 \) and thus \( \mathcal{E}_{D_{ij} \rho}(\psi(\eta_{ij})) \geq 1. \) The equality in (b) only holds for \( \rho = 0. \) Thereafter, differentiating \( \psi(\eta_{ij}) \) over \( \eta_{ij} \) yields

\[
\frac{d\psi(\eta_{ij})}{d\eta_{ij}} = \frac{1 + \rho}{(\mathbb{V}(h_{ij}) E_{N_c})^2 \times D_{ij} (1 + \mathbb{V}(h_{ij}) E_{N_c}) \eta_{ij} (1 - \eta_{ij})^2 \left( \frac{D_{ij} \rho + 1; \psi(\eta_{ij})}{\mathbb{E}_{h_{ij}}(\psi(\eta_{ij}))} \right) \times (\mathbb{E}_{h_{ij}}(\psi(\eta_{ij})))^{D_{ij} \rho + 1; \psi(\eta_{ij})}}, \tag{61}
\]
Then, the optimal information transmission power factor can be obtained by solving the equation \( \frac{dF}{d\eta} = \frac{dF}{d\eta_0} = \frac{dF}{d\eta_1} = 0 \) with some algebraic operations. Thus, \( \eta^* \) can be obtained in two cases: (1) if \( D_{ij} = 1 \), \( \eta^*_i = \frac{1}{2} \); (2) if \( D_{ij} > 1 \), \( \eta^*_i = \sqrt{\theta_i} - \sqrt{\theta_j} \), where \( \theta_i = \frac{D_{ij} + D_{ji}v(h_{ij})E_N}{\sqrt{v(h_{ij})E_N} + (D_{ij} - 1)} \).

**APPENDIX I**

**PROOF OF THEOREM 3**

Following similar approaches to Appendix A, we can obtain

\[
E_f(m) = \frac{1}{T} \sum_{i=1}^{L_i} \sum_{j=1}^{K_u} \log \left( E_{H_{ij}} \left( \int_{X_i} \eta \left( \int_{X_i} p(x) \right) ^{1+\rho} \right. \right) \left. d\eta_1 \right),
\]

\[
F_G = \Xi \int_{D_0} \left( \eta \left( \int_{X_i} \eta \right) ^{1+\rho} \right. d\eta_1 \right) \left. \right) \left( \int_{D_0} \left( \eta \left( \int_{X_i} \eta \right) ^{1+\rho} \right. \right. d\eta_1 \right),
\]

where \( D_0 = \{0 < \lambda_1 \leq \ldots \leq \lambda_1 < \infty \} \) and (a) follows by invoking the generic approach of Corollary 2 in [58] to simplify the multiple integral into a scalar integral. Then, by invoking the integral equation (42), we obtain the expression of \( \Delta \eta_1 \) given in (35). Thereafter, substituting (63) into (62), we conclude the proof of Theorem 3.

## REFERENCES

1. R. Ford, M. Zhang, M. Mezzavilla, S. Dutta, S. Rangan, and M. Zorzi, “Achieving ultra-low latency in 5G millimeter wave cellular networks,” *IEEE Commun. Mag.*, vol. 55, no. 3, pp. 196–203, Mar. 2017.
2. M. Bennis, M. Debbah, and H. V. Poor, “Ultrareliable and low-latency wireless communication: Tail, risk, and scale,” *Proc. IEEE*, vol. 106, no. 10, pp. 1834–1853, Oct. 2018.
3. I. K. Jain, R. Kumar, and S. S. Panwar, “The impact of mobile blockers on millimeter wave cellular systems,” *IEEE J. Sel. Areas Commun.*, vol. 37, no. 4, pp. 854–868, Apr. 2019.
4. Z. Ma, M. Xiao, Y. Xiao, Z. Pang, H. V. Poor, and B. Vucetic, “Guest editorial special issue on low-latency high-reliability communications for the IoT,” *IEEE Internet Things J.*, vol. 6, no. 5, pp. 7811–7815, Oct. 2019.
5. P. Schultz et al., “Latency critical IoT applications in 5G: Perspective on the design of radio interface and network architecture,” *IEEE Commun. Mag.*, vol. 55, no. 2, pp. 70–78, Feb. 2017.
6. M. S. Elbamby, C. Perfecto, M. Bennis, and K. Doppler, “Toward low-latency and ultra-reliable virtual reality,” *IEEE Netw.*, vol. 32, no. 2, pp. 78–84, Mar. 2018.
7. M. Xiao et al., “Millimeter wave communications for future mobile networks,” *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 1909–1935, Sep. 2017.
8. B. Lee, S. Park, D. J. Love, H. Ji, and B. Shim, “Packet structure and receiver design for low latency wireless communications with ultra-short packets,” *IEEE Trans. Commun.*, vol. 66, no. 2, pp. 796–807, Feb. 2018.
9. G. Yang, M. Xiao, M. Alam, and Y. Huang, “Low-latency heterogeneous networks with millimeter-wave communications,” *IEEE Commun. Mag.*, vol. 56, no. 6, pp. 124–129, Jun. 2018.
10. T. K. Vu, C.-F. Liu, M. Bennis, M. Debbah, M. Latva-Aho, and C. Simon, “Ultra-reliable and low latency communication in mmWave-enabled massive MIMO networks,” *IEEE Commun. Lett.*, vol. 21, no. 9, pp. 2041–2044, Sep. 2017.
11. G. Yang, M. Xiao, and H. V. Poor, “Low-latency millimeter-wave communications: Traffic dispersion or network densification?” *IEEE Trans. Commun.*, vol. 66, no. 8, pp. 3526–3539, Aug. 2018.
12. C. E. Shannon, “Probability of error for optimal codes in a Gaussian channel,” *Bell Syst. Tech. J.*, vol. 38, no. 3, pp. 611–656, May 1959.
13. Y. Polanski, H. V. Poor, and S. Verdú, “Channel coding rate in the finite blocklength regime,” *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, 2010.
14. Y. Polanski, S. Verdú, “Scalar correlated fading channel: Dispersion analysis,” in *Proc. IEEE Int. Symp. Info. Theory*, 2011, pp. 2399–2399.
15. A. Collins and Y. Polanski, “Orthogonal designs optimize achievable dispersion for coherent MISO channels,” in *Proc. IEEE Int. Symp. Info. Theory*, Jun. 2014, pp. 2524–2528.
16. A. Collins and Y. Polanski, “Coherent multiple-antenna block-fading channels at finite blocklength,” *IEEE Trans. Inf. Theory*, vol. 65, no. 1, pp. 380–405, Jan. 2019.
17. J. Ostman, W. Yang, G. Durisi, and T. Koch, “Diversity versus multiplexing at finite blocklength,” in *Proc. 11th Int. Symp. Wireless Commun. Syst. (ISWCS)*, Aug. 2014, pp. 702–706.
18. G. Durisi, T. Koch, J. Ostman, Y. Polanski, and W. Yang, “Short-packet communications over multiple-antenna Rayleigh-fading channels,” *IEEE Trans. Commun.*, vol. 64, no. 2, pp. 618–629, Feb. 2016.
19. A. Lanchio, T. Koch, and G. Durisi, “On single-antenna Rayleigh block-fading channels at finite blocklength,” *IEEE Trans. Inf. Theory*, vol. 66, no. 1, pp. 496–519, Jan. 2020.
20. R. G. Gallager, *Information Theory and Reliable Communication*, vol. 2. New York, NY, USA: Wiley, 1968.
21. X. Wu and R. Srikant, “Asymptotic behavior of error exponents in the wideband regime,” *IEEE Trans. Inf. Theory*, vol. 53, no. 4, pp. 1310–1325, Apr. 2007.
22. E. Malkamaki and H. Leib, “Coded diversity on block-fading channels,” *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 771–781, Mar. 1999.
23. S. Ray, M. Médard, and L. Zheng, “On error probability for wideband MIMO channels,” in *Proc. Conf. Inf. Sci. Syst. (CISS)*, 2005, pp. 1–6.
24. X. Jue, M. Z. I. Sarkar, and T. Ratnarajah, “Error exponents for Nakagami-m fading channels,” *Problems Inf. Transmiss.*, vol. 50, no. 2, pp. 144–170, 2014.
25. W. K. M. Ahmed and P. J. McLean, “Random coding error exponents for flat fading channels with realistic channel estimation,” *IEEE J. Sel. Areas Commun.*, vol. 18, no. 3, pp. 369–379, Mar. 2000.
26. S. Ray, M. Médard, and L. Zheng, “On noncoherent MIMO channels in the wideband regime: Capacity and reliability,” *IEEE Trans. Inf. Theory*, vol. 53, no. 6, pp. 1983–2009, 2007.
27. X. Wu and R. Srikant, “MIMO channels in the low-SNR regime: Communication rate, error exponent, and signal peakness,” *IEEE Trans. Inf. Theory*, vol. 53, no. 4, pp. 1290–1309, Apr. 2007.
28. M. R. Akdeniz et al., “Millimeter wave channel modeling and cellular capacity evaluation,” *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1164–1179, Jun. 2014.
29. X. Wu and R. Srikant, “MIMO channels in the low-SNR regime: Communication rate, error exponent, and signal peakness,” *IEEE Trans. Inf. Theory*, vol. 53, no. 4, pp. 1290–1309, Apr. 2007.
30. M. R. Akdeniz et al., “Millimeter wave channel modeling and cellular capacity evaluation,” *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1164–1179, Jun. 2014.
31. T. S. Rappaport, Y. Xing, G. R. MacCartney, A. F. Molisch, and E. Mellios, “Overview of millimeter wave communications for fifth-generation (5G) wireless networks—With a focus on propagation models,” *IEEE Trans. Antennas Propag.*, vol. 65, no. 12, pp. 6213–6230, Dec. 2017.
32. S. Rangan, T. S. Rappaport, and E. Erkip, “Millimeter-wave cellular wireless networks: Potentials and challenges,” *Proc. IEEE*, vol. 102, no. 3, pp. 366–385, Mar. 2014.
33. C. S. Hong, “Ultra-reliable and low latency communication in mmWave networks with millimeter-wave communications,” *IEEE Commun. Mag.*, vol. 66, no. 8, pp. 3526–3539, Aug. 2018.
\[ \text{B. Hassibi and B. M. Hochwald, “How much spectrum is too much in millimeter wave wireless access,” IEEE J. Sel. Areas Commun., vol. 35, no. 7, pp. 1444–1458, Jul. 2017.} \]

\[ \text{S. Wyne, K. Haneda, S. Ranvier, F. Tufvesson, and A. F. Molisch, “Beaming effects on measured mm-wave channel characteristics,” IEEE Wireless Commun., vol. 10, no. 11, pp. 3553–3559, Nov. 2011.} \]

\[ \text{F.-M. Han and X.-D. Zhang, “Wireless multicarrier digital transmission via Weyl-Heisenberg frames over time-frequency dispersive channels,” IEEE Trans. Commun., vol. 57, no. 6, pp. 1712–1733, Jun. 2009.} \]

\[ \text{L. Yan, X. Fang, Y. Fang, and X. Cheng, “Dual-scheduler design for C/I-plane decoupled railway wireless networks,” IEEE Trans. Mobile Comput., vol. 16, no. 7, pp. 1842–1855, Jul. 2017.} \]

\[ \text{Z. Jin, A. Doufexi, and T. Kocak, “On the performance of IEEE 802.15.3c millimeter-wave WPANs: PHY and MAC,” in Proc. 6th Conf. Wireless Adv. (WiAD), Jun. 2010, pp. 1–6.} \]

\[ \text{G. Durisi, U. G. Schuster, H. Boleskei, and S. Shamai, “Noncoherent capacity of underspread fading channels,” IEEE Trans. Inf. Theory, vol. 56, no. 1, pp. 367–395, Jan. 2010.} \]

\[ \text{X. Zou, W. Li, B. Lu, W. Pan, L. Yan, and L. Shao, “Photonic approach for wide-frequency-range high-resolution microwave/millimeter-wave Doppler frequency shift estimation,” IEEE Trans. Microw. Theory Techn., vol. 63, no. 4, pp. 1421–1430, Apr. 2015.} \]

\[ \text{Study on Channel Model for Frequency Spectrum Above 6 GHz; Document 3GPP TR 38.900 V15.1.0, 2018.} \]

\[ \text{K. P. Oldham, J. Myland, and J. Spanier, An Atlas of Functions: With Equator, the Atlas Function Calculator. New York, NY, USA: Springer, 2010.} \]

\[ \text{H. Vincent Poor (Life Fellow, IEEE) received the Ph.D. degree in EECS from Princeton University in 1977. From 1977 to 1990, he was on the faculty of the University of Illinois at Urbana-Champaign. Since 1990, he has been on the faculty at Princeton, where he is currently the Michael Henry Strater University Professor. From 2006 to 2016, he served as the Dean of the Princeton’s School of Engineering and Applied Science. He has held visiting appointments at several other universities, including most recently at Berkeley and Cambridge. His research interests are in the areas of information theory, machine learning and network science, and their applications in wireless networks, energy systems, and related fields. Among his publications in these areas is the forthcoming book Machine Learning and Wireless Communications (Cambridge University Press, 2021). Dr. Poor is a member of the National Academy of Engineering and the National Academy of Sciences, and he is a foreign member of the Chinese Academy of Sciences, the Royal Society, and other national and international academies. He received the 2017 IEEE Alexander Graham Bell Medal.} \]