

Gaussian Mixture based Context-aware Short-Term Characterization of Wireless Channels

Priyadarshi Mukherjee, Deepak Mishra, and Swades De

Abstract—5G wireless communication technologies aim at simultaneously achieving energy efficiency and spectral efficiency. 5G also demands high communication reliability. In this context, fine-grained temporal characterization of wireless channel can be used to enhance both. To this end, we propose a novel context-aware characterization of the temporally-varying wireless channel. Our characterization of temporal variation of the channel is based on the method of finite mixture of Gaussian distributions. However, unlike the classical Gaussian mixture model, the proposed characterization does not use an iterative algorithm for its parameter estimation; it depends on the current channel state and its statistics. Based on this characterization we estimate the quantity of data that can be transferred over the channel in a time interval without knowing the actual channel state in that duration. We propose an application context dependent upper bound on the time interval over which this estimation can be made. Our numerical results demonstrate that the present channel state plays a crucial role. When the proposed characterization is used in the context of channel adaptive communication, energy efficiency obtained is as high as 3.15 times over its nearest approach. A nontrivial trade-off between energy efficiency and precision of the proposed characterization is also investigated.

Index Terms—Fading channel, temporal characterization, ultra-reliable low-latency communications, Gaussian mixture model, inter-feedback duration, energy efficiency

I. INTRODUCTION AND BACKGROUND

The Internet traffic has been growing at an explosive rate in recent years. According to Ericsson, global Internet traffic is expected to increase more than five times between 2018 and 2024 [1]. This increased traffic will also lead to an increase in energy consumption of the wireless devices. However, high energy consumption is unaffordable in energy-constrained scenarios, such as in wireless sensor network (WSN) and Internet of Things (IoT) applications. In this context, energy-efficient green communication [2], [3] have attracted a lot of recent attention in industry as well as academia.

A. Prior Art and Motivation

WSN and IoT applications require energy to be judiciously used. Some of these applications also demand ultra-reliable low-latency communication (URLLC) [4] over wireless channel. Hence the current state of the fading wireless channel plays a crucial role in these reliability as well as battery-constrained scenarios, and therefore it is important to focus

on temporal characteristics of the *unreliable* wireless channel. One of the major objectives of the desired channel characterization is prediction of the fading state of channel in next few slots. To the best of our knowledge, none of the existing wireless channel characterizations serve the purpose.

However, in the context of future states, wireless channel state prediction has been a well-investigated topic [5]–[11]. The authors in [6], [7] proposed the use of autoregressive models that use minimum mean square error criterion in predicting the future state of channel. A method based on sum-of-sinusoids was proposed in [8] for the same purpose, but specifically for a Rayleigh fading scenario. It was shown that when the channel is modeled as a sum of complex sinusoids, then it can be effectively estimated by frequency estimation techniques, e.g., estimation of signal parameters via rotational invariance techniques and multiple signal classification. An adaptive Kalman algorithm was proposed in [9] for modeling and prediction of mobile radio channels. Channel state prediction in the context of OFDM was investigated in [10], where a low complexity estimation method using polynomial approximation was proposed. A predictive resource allocation technique was proposed in [11] in context of energy-efficient video streaming. The work in [12] revisited wireless channel prediction and proposed a relay selection scheme in context of URLLC. Various channel-aware link layer retransmissions strategies have also been proposed in literature [13]–[16] that aim at reducing chances of blind data retransmissions and also maximize throughput.

We observe that, although there have been prior studies on channel state prediction, they were mainly channel characteristics independent in nature. To improve upon the prior art, in this work we propose a generalized Gaussian Mixture (GM) distribution based temporal characterization of channel that is also able to estimate the future states of the channel depending on its current state.

The proposed characterization will be helpful in achieving short-term goals, such as estimation of the possible amount of data transferred in a certain time interval without requiring the actual channel state information (CSI) in that interval. Hence this characterization will meet the demands of URLLC while the cost incurred due to CSI feedback will be saved, resulting in significant energy saving [17]. As instantaneous CSI feedback in each slot incurs high cost, the proposed scheme effectively bridges the gap between fully-blind [18] and fully-instantaneous [19] schemes. Further, the proposed characterization approach can be efficiently used in a system that applies channel adaptive communication to increase the energy efficiency in an environment with varying channel.

P. Mukherjee and S. De are with the Department of Electrical Engineering, Indian Institute of Technology Delhi, New Delhi, India (e-mail: {priyadarshi.mukherjee, swadesd}@ee.iitd.ac.in). D. Mishra is with the School of Electrical Engineering and Telecommunications, University of New South Wales Sydney, Australia (e-mail: d.mishra@unsw.edu.au)

There are a few existing works that aimed to fill the gap between fully-blind and fully-instantaneous schemes. The study in [20] proposed a Huffman coding based feedback compression algorithm. The authors in [21] investigated various feedback subsampling based schemes for exploiting time diversity in temporally-correlated channels. In [22], a new limited feedback framework was proposed using dictionary-based sparse channel estimation algorithms. However, the studies [20]–[22] assumed the channel to be invariant over a time interval equal to coherence time. On the contrary, the proposed characterization allows the user to have a control on the inter-feedback interval, resulting in a better utilization of the channel irrespective of the time scale of its dynamics.

Though Rayleigh model is a reasonable assumption in context of fading experienced in most wireless scenarios, it does not consider the effect of line-of-sight (LOS) signal component. A more generalized fading model that accounts for the LOS component is the Rician model, which is considered in this work. *It is important to observe that, though the Rician fading model is taken in the analysis, the observations made in this study are applicable in all fading scenarios.*

Note that, GM modeling of wireless channel is not new. The studies in [23]–[25] modeled the channel in terms of a GM to meet different objectives. Specifically, high-complexity algorithms were employed in these works to mathematically approximate the channel model by a convex combination of Gaussian random variables. In contrast, *this is the first work that models the short-term temporal variation of a wireless fading channel using GM distribution without the use of any such algorithm.* In particular, we prove that temporal variations of the channel actually follow a GM, and unlike [23]–[25], it need not be approximated by a GM distribution.

B. Contributions

The key contributions in this work are as follows:

- Temporal variation of a wireless channel is modeled as a finite sum of Gaussian distributions, i.e., in the form of a GM distribution that depends on current channel state. The proposed characterization is applicable for all the existing fading channel models.
- Unlike the conventional definitions of outage probability and ergodic capacity, our proposed definitions of *short term outage probability* and *short term channel capacity* accurately capture short-term temporal variations of the channel. Convergence of the short-term outage and capacity definitions to the outage probability and ergodic capacity are also proven.
- An application-dependent upper bound is obtained on the time interval over which the short-term characterization is done. An efficient algorithm is proposed to obtain the upper bound. The impact of temporally varying channel on the proposed upper bound and energy efficiency of the system is also discussed. A non-trivial trade-off that exists between the quality of prediction and the upper bound is investigated.
- Detailed investigation is carried out to validate the analysis and provide insights on the proposed characterization. Our results demonstrate that, when the proposed

characterization is used in context of channel-adaptive communication, energy efficiency obtained is as high as 3.15 times over its nearest competitive approach.

C. Paper Organization

The paper is organized as follows. Section II introduces the system model, the temporal variation of wireless channel, and the Gaussian modeling of both the received signal envelope as well as the signal-to-noise ratio (SNR) at the receiver (Rx). Section III presents the proposed characterization of temporal variation of channel. In Section IV, definitions of *short term outage probability* and *short term channel capacity* are introduced and their respective closed-form expressions are derived. Section V proposes an upper bound on time interval over which the characterization is done, Section VI presents the effect of the proposed characterization on adaptive communication system, followed by discussion and concluding remarks in Section VII.

II. SYSTEM MODEL AND TEMPORAL VARIATION OF WIRELESS CHANNEL

A. System Model

We consider communication between a node pair in a mobile environment. The system is assumed slotted, with slot duration T_p seconds. Without any loss of generality, we consider a quasi-static block fading channel model [26], [27], i.e., the channel remains constant for a block duration (which in this case is taken as T_p), but vary from one block to another.

Before the transmitter (Tx) sends data to the Rx, it sends a pilot signal to Rx¹. Rx responds with useful CSI feedback to Tx along with maximum Doppler frequency f_D [29].

f_D corresponding to node velocity v of Rx is $f_D \cong \frac{vf_c}{c}$, where f_c is the carrier frequency and c is the velocity of light in vacuum ($f_c = 900$ MHz is considered in this work). The product $f_D T_p$ signifies the temporal variation rate of the wireless channel. From [30] we consider that $f_D T_p < 0.1$ indicates correlated slow fading channel, whereas $f_D T_p > 0.2$ implies that the two samples of the channel are almost independent, i.e., the channel is fast fading.

The CSI requirement at Tx in every time slot results in large demand of energy, which is sometimes unaffordable in energy-constrained scenarios. At the same time if Tx estimates the channel state only from the available channel statistics, it cannot take the advantage of current channel state. Some existing works [20]–[22] consider coherence time based slotting, resulting in under-utilization of channel. Thus there exists a gap, which we aim to address in this work.

B. Wireless Channel and Its Temporal Variation

We consider, data transmission is taking place in a wide-sense stationary (W.S.S.) Rician fading scenario with signals of transmission power P . The received signal y at Rx is:

$$y = \sqrt{P} h x + n, \quad (1)$$

¹Based on the channel reciprocity of a TDD-based wireless system [28], it is assumed that the channel gain of the uplink is the same as that of the downlink estimated by Rx.

where x is the signal transmitted, h is the time varying channel gain, and n is independent Additive White Gaussian Noise (AWGN) at Rx with zero mean and variance σ_o^2 . Channel gain h at time t is generally modeled in the form of a complex random process as $h(t) = h_I(t) + jh_Q(t)$, where $h_I(t)$ and $h_Q(t)$ represent the in-phase and quadrature components, respectively. Based on Clarke's two dimensional isotropic scattering model, $h_I(t)$ and $h_Q(t)$ corresponding to Rician fading scenario are defined as [31]:

$$h_I(t) = \frac{1}{\sqrt{1+K}} \left[\frac{1}{\sqrt{N_P}} \sum_{n=1}^{N_P} \cos(\omega_D t \cos \alpha_n + \phi_n) + \sqrt{K} \cos(\omega_D t \cos \alpha_0 + \phi_0) \right] \quad \text{and} \quad (2)$$

$$h_Q(t) = \frac{1}{\sqrt{1+K}} \left[\frac{1}{\sqrt{N_P}} \sum_{n=1}^{N_P} \sin(\omega_D t \cos \alpha_n + \phi_n) + \sqrt{K} \sin(\omega_D t \cos \alpha_0 + \phi_0) \right].$$

Here N_P is the number of propagation paths, $\omega_D = 2\pi f_D$ is the maximum radian Doppler frequency, α_n and ϕ_n are respectively the azimuth angle of arrival and initial phase of the n th propagation path, K is the ratio of the specular power to the scattered power, and α_0 and ϕ_0 are respectively the azimuth angle of arrival and the initial phase of the specular component.

h (index t is removed for brevity) can be expressed as $h = |h|e^{j\kappa}$ with κ taking values between $-\pi$ and π . In a rich scattering environment, we have $N_P \rightarrow \infty$ and according to the central limit theorem, both $\frac{1}{\sqrt{N_P(1+K)}} \sum_{n=1}^{N_P} \cos(\omega_D t \cos \alpha_n + \phi_n)$ and $\frac{1}{\sqrt{N_P(1+K)}} \sum_{n=1}^{N_P} \sin(\omega_D t \cos \alpha_n + \phi_n)$ tend to be Gaussian. With ϕ_0 being uniformly distributed over $[-\pi, \pi]$ [31], $|h|$ tends to the Rician distribution; its probability distribution function (PDF) is stated below in (3).

Remark 1. *Practical channels do not always have a very rich scattering environment and hence it is impractical to always take the $N_P \rightarrow \infty$ assumption. However, Xiao et al. [31] have shown that the PDF of $|h|$ with N_P as low as $N_P = 8$ is in very good agreement with the theoretical values obtained when considering $N_P \rightarrow \infty$. Hence, the considered model stands valid for practical channels irrespective of number of scattering components present.*

We further assume that $\mathbb{E}[x] = 0$, $\mathbb{E}[x^2] = 1$, and $\mathbb{E}[|h|^2] = 1$, where $\mathbb{E}[\cdot]$ denotes the expectation operator. The PDF of received signal envelope $\Theta = \sqrt{P}|h|$ is [32]:

$$f_{\Theta}(\alpha, \mu_{\Theta}, K) = \frac{2(1+K)\alpha e^{-K}}{\mu_{\Theta}} e^{-\frac{(1+K)\alpha^2}{\mu_{\Theta}}} I_0 \left[2\alpha \sqrt{\frac{K(1+K)}{\mu_{\Theta}}} \right], \quad \alpha \geq 0. \quad (3)$$

Here $\mu_{\Theta} = \mathbb{E}[\Theta]$, K is the Rice factor, and $I_0(\cdot)$ denotes the

zero-order modified Bessel function of the first kind. Estimate of Rice factor K is known from [33]. From [34] we know that the time-derivative of Θ , i.e., $\dot{\Theta} = \lim_{\Delta t \rightarrow 0} \frac{\Theta(t+\Delta t) - \Theta(t)}{\Delta t}$ is a zero mean Gaussian random variable (RV) with variance $\dot{\sigma}_{\Theta}$, i.e., $\dot{\Theta} \sim \mathcal{N}(0, \dot{\sigma}_{\Theta})$, where

$$\dot{\sigma}_{\Theta} = \frac{\pi f_D}{(K+1)} \sqrt{P(2K+1)}. \quad (4)$$

To the best of our knowledge, till date $\dot{\Theta}$ has been used to obtain two key second-order statistics of the channel [35], namely, level crossing rate (LCR) and average fade duration (AFD). Mathematically, $\text{LCR} = \int_0^{\infty} \dot{\Theta} f_{\Theta\dot{\Theta}}(\Theta_{th}, \dot{\Theta}) d\dot{\Theta}$ and $\text{AFD} = \frac{\Pr\{\Theta < \Theta_{th}\}}{\text{LCR}}$, where Θ_{th} is the target threshold and $f_{\Theta\dot{\Theta}}(\Theta, \dot{\Theta})$ is the joint PDF of $\Theta(t)$ and $\dot{\Theta}(t)$ at the same time instant. Θ being Rician, received SNR $Z = \frac{P|h|^2}{\sigma_o^2}$ is a non-central χ^2 RV with two degrees of freedom. The PDF of Z is [32]:

$$f_Z(z, \mu_z, K) = \frac{(1+K)e^{-K}}{\mu_z} e^{-\frac{(1+K)z}{\mu_z}} I_0 \left[2\sqrt{\frac{K(1+K)z}{\mu_z}} \right], \quad z \geq 0. \quad (5)$$

Here $\mu_Z = \mathbb{E}[Z]$. Just like $\dot{\Theta}$, $\dot{Z} \triangleq \frac{dZ}{dt}$ is also a zero mean Gaussian RV, i.e., $\dot{Z} \sim \mathcal{N}(0, \dot{\sigma}_Z)$. The standard deviation of \dot{Z} , i.e., $\dot{\sigma}_Z$ can be obtained as follows.

Proposition 1. *The standard deviation of \dot{Z} in a Rician fading scenario is*

$$\dot{\sigma}_Z = \frac{2P}{\sigma_o^2} \frac{\pi f_D}{(K+1)} \sqrt{2K+1}. \quad (6)$$

Proof. Variance of \dot{Z} , $\text{Var}(\dot{Z})$ is given by:

$$\begin{aligned} \text{Var}(\dot{Z}) &= \text{Var} \left(\frac{d}{dt} \left\{ \frac{P|h|^2}{\sigma_o^2} \right\} \right) = \text{Var} \left(\frac{d}{dt} \left\{ \frac{\Theta^2}{\sigma_o^2} \right\} \right) \\ &= \text{Var} \left(\frac{2}{\sigma_o^2} \Theta \dot{\Theta} \right) \\ &= \frac{4}{\sigma_o^4} \text{Var}(\Theta \dot{\Theta}). \end{aligned} \quad (7)$$

From [36] we know that $\dot{\Theta}$ and Θ are independent. Using the definition of variance for product of two independent RV, we get $\text{Var}(\dot{Z}) = \frac{4}{\sigma_o^4} \left\{ \mathbb{E}[\Theta^2] \mathbb{E}[\dot{\Theta}^2] - \mathbb{E}[\Theta]^2 \mathbb{E}[\dot{\Theta}]^2 \right\}$. From the definition of $\dot{\Theta}$, we have $\mathbb{E}[\dot{\Theta}] = 0$. Accordingly the above equation gets reduced as $\text{Var}(\dot{Z}) = \frac{4}{\sigma_o^4} \left\{ \mathbb{E}[\Theta^2] \mathbb{E}[\dot{\Theta}^2] \right\} = \frac{4P^2}{\sigma_o^4} \left(\frac{\pi f_D}{K+1} \right)^2 (2K+1)$. Hence we get, $\dot{\sigma}_Z = \sqrt{\text{Var}(\dot{Z})} = \frac{2P}{\sigma_o^2} \frac{\pi f_D}{(K+1)} \sqrt{2K+1}$. \square

Note that though Proposition 1 calculates $\text{Var}(\dot{Z})$, \dot{Z} is not analytically proved to be a Gaussian RV. However as observed in Fig. 1(b), considering \dot{Z} to be zero-mean Gaussian results in an root mean square error (RMSE) on the order of 10^{-4} when we compare the probability distribution of the actual signal power and the one obtained by our proposed characterization

(discussed in Section III). This justifies \dot{Z} as a zero-mean Gaussian RV with the calculated variance.

Remark 2. *The approach of isotropic scattering is not popular, especially in modeling 5G systems. The works in [37], [38] based on various real environments show that the scattering nature at Rx is anisotropic. The authors in [39]–[41] investigated the channel at Rx side and proposed various channel models. However, our current work is based on the temporal variation of wireless channel, which is Gaussian in nature [34] irrespective of whether the scattering considered is isotropic or not. Hence the proposed characterization stands valid even in anisotropic scattering scenarios.*

C. Gaussian Modeling of Received Signal Envelope and SNR

If the received signal envelope Θ at time t is $\Theta(t) = \theta_0$, then Θ in the next slot, i.e., after time T_p , is expressed as

$$\Theta(t + T_p) \approx \theta_0 + \dot{\Theta} \cdot T_p, \quad \text{where} \quad (8)$$

$\dot{\Theta} \sim \mathcal{N}(0, \dot{\sigma}_\Theta)$. Since practically $T_p \ll 1$ s, we obtain (8) using first order approximation of Taylor series expansion of $\Theta(t + T_p)$ along the following lines:

$$\begin{aligned} \Theta(t + T_p) &= \Theta(t) + \dot{\Theta} \cdot T_p + \ddot{\Theta} \cdot \frac{T_p^2}{2!} + \dots \\ &= \Theta(t) + \dot{\Theta} \cdot T_p + O(T_p^2) \\ &\approx \theta_0 + \dot{\Theta} \cdot T_p. \end{aligned} \quad (9)$$

Accordingly we define a RV $\Theta_1 (= \dot{\Theta} \cdot T_p)$ that denotes temporal variation of the signal envelope in next one slot.

Theorem 1. Θ_1 is not an unconstrained Gaussian RV, but a truncated Gaussian RV in $[-\theta_0, +\infty)$.

Proof. We prove this theorem by contradiction. As $\Theta_1 = \dot{\Theta} \cdot T_p$, we have $\Theta_1 \sim \mathcal{N}(0, \dot{\sigma}_{\Theta_1})$, where $\dot{\sigma}_{\Theta_1} = \dot{\sigma}_\Theta \cdot T_p = \frac{\pi f_D T_p}{(K+1)} \sqrt{P(2K+1)}$.

$\Theta_1(\theta_1) \sim \mathcal{N}(0, \dot{\sigma}_{\Theta_1}) \forall \theta_1$ implies that $\theta_1 \in (-\infty, +\infty)$. Let $\theta_1 \in (-\infty, -\theta_0)$, which is a subset of $(-\infty, +\infty)$. Then from (9) we have $\Theta(t + T_p) \approx \theta_0 + \theta_1 < 0$. But $\Theta(t + T_p)$ being signal envelope, it is a non-negative quantity. We have thus reached a contradiction, that $\theta_1 \notin (-\infty, -\theta_0)$ but $\theta_1 \in [-\theta_0, +\infty)$. In other words, Θ_1 cannot be an unconstrained Gaussian RV, but it is a truncated Gaussian RV in $[-\theta_0, +\infty)$ with its PDF as:

$$f_{\Theta_1}(\theta_1) = \begin{cases} \frac{1}{\sqrt{2\pi}\dot{\sigma}_{\Theta_1} \left[1 - \Phi\left(-\frac{\theta_0}{\dot{\sigma}_{\Theta_1}}\right)\right]} e^{-\frac{\theta_1^2}{2\dot{\sigma}_{\Theta_1}^2}}, & \theta_1 \geq -\theta_0 \\ 0, & \text{elsewhere.} \end{cases} \quad (10)$$

Here $\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$ is the cumulative distribution function of standard univariate normal distribution. Value of $\dot{\sigma}_{\Theta_1}$ depends upon the fading distribution model. \square

Similarly, Θ after 2 time slots is $\Theta(t + 2T_p) \approx \theta_0 + \Theta_2$, where Θ_2 is also a zero mean Gaussian RV like Θ_1 with $\dot{\sigma}_{\Theta_2}^2 = 2\dot{\sigma}_{\Theta_1}^2$; just the difference being that Θ_2 resembles the

temporal variation of Θ over the next two slots, whereas Θ_1 denotes temporal variation over next one slot only.

Remark 3. *Generalizing this concept over next N slots, we can say that if $\Theta(t) = \theta_0$, then*

$$f_{\Theta_N}(\theta_N) = \begin{cases} \frac{1}{\sqrt{2\pi}\dot{\sigma}_{\Theta_N} \left[1 - \Phi\left(-\frac{\theta_0}{\dot{\sigma}_{\Theta_N}}\right)\right]} e^{-\frac{\theta_N^2}{2\dot{\sigma}_{\Theta_N}^2}}, & \theta_N \geq -\theta_0 \\ 0, & \text{elsewhere.} \end{cases} \quad (11)$$

Here $\dot{\sigma}_{\Theta_N}^2 = N\dot{\sigma}_{\Theta_1}^2$. Extending this concept to the received SNR, we can make an analogous statement.

Remark 4. *If the received SNR at time t is $Z(t) = Z_0$, then*

$$f_{Z_N}(z_N) = \begin{cases} \frac{1}{\sqrt{2\pi}\dot{\sigma}_{Z_N} \left[1 - \Phi\left(-\frac{Z_0}{\dot{\sigma}_{Z_N}}\right)\right]} e^{-\frac{z_N^2}{2\dot{\sigma}_{Z_N}^2}}, & z_N \geq -Z_0 \\ 0, & \text{elsewhere.} \end{cases} \quad (12)$$

Here $\dot{\sigma}_{Z_N}^2 = N\dot{\sigma}_{Z_1}^2$, where $\dot{\sigma}_{Z_1} = \dot{\sigma}_Z \cdot T_p$.

It may be noted that Remarks 1 and 2 hold for $N \cdot T_p \rightarrow 0$. Taking this into consideration, we propose a GM characterization of the temporal variation of Θ and Z . This characterization has ζ component distributions. Details regarding the choice of ζ (and hence N) is investigated later in Section V.

D. Case of Imperfect CSI at the Transmitter

Previously we characterized $\Theta(t + T_p)$ in terms of $\Theta(t) = \theta_0$ and temporal variation of Θ over the immediate slot, i.e., Θ_1 . It was proved in Theorem 1 that Θ_1 is a truncated Gaussian RV in $[-\theta_0, +\infty)$. Hence, if Tx has imperfect knowledge of $\Theta(t)$, it also affects Θ_1 . In this regard, let $\Theta(t) = \hat{\theta}_0$, where $\hat{\theta}_0 = \sqrt{P}|\hat{h}|$ is the imperfect CSI available at the Tx and \hat{h} is modeled by the Gauss-Markov formulation form as [42]:

$$\hat{h} = \sqrt{1 - \varphi^2} h + \varphi e. \quad (13)$$

CSI imperfection parameter $\varphi \in [0, 1]$ indicates the quality of instantaneous CSI, i.e., $\varphi = 0$ corresponds to perfect CSI and $\varphi = 1$ corresponds to having only statistical channel knowledge. The quantity e represents the varying component of the channel and is assumed to be a circularly symmetric complex Gaussian random variable with zero mean and unit variance, i.e., $e \sim \mathcal{CN}(0, 1)$. Note that the introduced CSI imperfection accounts for channel estimation errors and/or channel variations during the feedback delay.

Thus we observe from (13) that in the imperfect CSI scenario, Θ_1 is a truncated Gaussian RV in $[-\hat{\theta}_0, +\infty)$ and not $[-\theta_0, +\infty)$. However from Theorem 1 we note that the mean and/or variance of Θ_1 remains unaffected.

Generalizing the effect of imperfect CSI at TX, Remark 3 and Remark 4 (as $Z = \frac{P|h|^2}{\sigma_o^2}$) are revised as:

Remark 5. If $\Theta(t) = \widehat{\theta}_0$,

$$f_{\Theta_N}(\theta_N) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma_{\Theta_N} \left[1 - \Phi\left(-\frac{\widehat{\theta}_0}{\sigma_{\Theta_N}}\right)\right]} e^{-\frac{\theta_N^2}{2\sigma_{\Theta_N}^2}}, & \theta_N \geq -\widehat{\theta}_0 \\ 0, & \text{elsewhere.} \end{cases} \quad (14)$$

Remark 6. If $Z(t) = \widehat{Z}_0 = \frac{P|\widehat{h}|^2}{\sigma_z^2}$, then

$$f_{Z_N}(z_N) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma_{Z_N} \left[1 - \Phi\left(-\frac{\widehat{Z}_0}{\sigma_{Z_N}}\right)\right]} e^{-\frac{z_N^2}{2\sigma_{Z_N}^2}}, & z_N \geq -\widehat{Z}_0 \\ 0, & \text{elsewhere.} \end{cases} \quad (15)$$

Since the CSI imperfection affects the truncation interval but not the corresponding moments of Θ_N or Z_N , the proposed model in next section stands valid even with imperfect CSI.

III. PROPOSED GAUSSIAN MIXTURE MODEL

From the last section we see that, if $\Theta(t) = \theta_0$, then the probable received signal envelope after N slots $\Theta(t+NT_p) \approx \theta_0 + \Theta_N$ can be characterized as a truncated Gaussian RV with its mean being θ_0 and variance $\sigma_{\Theta_N}^2$. Thus it can be observed that the mean of $\Theta(t+NT_p)$ is a *deterministic* quantity that can be obtained from the channel estimate sent by Rx to Tx, and the variance is obtained from maximum Doppler frequency f_D and Rice factor K .

It is important to note here that in general, mean, variance, or any moment of a RV are time-invariant quantities that define the RV. But in this case it is not so; here the CSI received at Tx acts as the mean while variance is obtained from f_D and K . In other words, $\Theta(t)$ does not denote the same RV at $t = t_1$ and $t = t_2$. Hence the proposed characterization of the channel is *hybrid* in nature, i.e., it depends on the channel characteristics as well as its instantaneous value. We use this very interesting property of Θ to model its variation over next ζ slots, where the present state of the channel, i.e., $\Theta(t)$ plays a very important role.

Since for a given $\Theta(t) = \theta_0$ the variation of Θ in each of the next ζ slots follows a scaled Gaussian distribution, we propose to model the entire variation of Θ over the next ζ slots in terms of a GM distribution.

According to the theory of classical GM model [43], any arbitrary probability distribution $f_A(a)$ can be expressed as a convex combination of finite number of Gaussian distributions $f_{B_i}(b_i)$ in the following form:

$$f_A(a) = \sum_{i=1}^{\zeta} \pi_i f_{B_i}(b_i), \quad (16)$$

where $f_{B_i}(b_i)$ is the i^{th} component Gaussian distribution, i.e., $f_{B_i}(b_i) = \frac{1}{\sqrt{2\pi}\sigma_{b_i}} e^{-\frac{(b_i - \mu_{b_i})^2}{2\sigma_{b_i}^2}}$ and π_i is its corresponding weight with constraints $0 \leq \pi_i \leq 1$ and $\sum_{i=1}^{\zeta} \pi_i = 1$. The optimal values of the unknown parameters, i.e., π_i , μ_{b_i} , and

$\sigma_{b_i} \forall 1 \leq i \leq \zeta$ are generally obtained by iterative Expectation-Maximization (EM) algorithm [44]. The run-time complexity of the EM algorithm is $O(K^2L)$, where K is the number of Gaussian components in the GM distribution, and L is the size of data set on which GM distribution is considered to obtain the optimal set of parameters. We introduce Theorems 2 and 3 that characterize Θ and Z respectively in terms of GM distribution as follows.

Theorem 2. If Ψ denotes the temporal variation of Θ over the next ζ slots from time t , it can be characterized in terms of $\Theta(t) = \theta_0$ by a GM distribution whose PDF is:

$$f_{\Psi}(\beta | \Theta(t) = \theta_0, \zeta) = \begin{cases} \frac{1}{\zeta} \sum_{i=1}^{\zeta} \frac{1}{\sqrt{2\pi}\sigma_{\Theta_i}} \left[1 - \Phi\left(-\frac{\theta_0}{\sigma_{\Theta_i}}\right)\right]} e^{-\frac{\beta^2}{2\sigma_{\Theta_i}^2}}, & \beta \geq -\theta_0 \\ 0, & \text{elsewhere.} \end{cases} \quad (17)$$

Proof. From [34] we know that irrespective of the underlying fading distribution, the rate of temporal variation of Θ is a zero mean Gaussian RV. Moreover we also know from Remark 3 that based on the current channel state $\Theta(t)$, the temporal variation of the next i^{th} slot can be modeled as a truncated Gaussian distribution. Hence in observing the effective Θ over each of next ζ slots, it will be a mixture combination of ζ different Gaussian distributions. In other words, the temporal variation of Θ over next ζ slots will be a GM distribution as given by (17).

It should be noted that we use the concept of *equal weightage* in Theorem 2. That is, $\frac{1}{\zeta}$ in the theorem implies that equal weightage is given to all of the next ζ slots. This is because we are interested in characterizing the variation of Θ over next ζ slots as a whole, without giving any additional priority or being biased to any of these ζ slots. \square

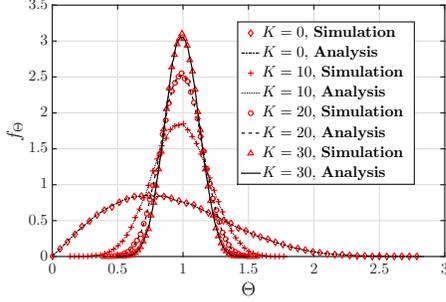
On similar lines, we also characterize the variation of received SNR in next ζ slots as follows.

Theorem 3. If Ω denotes the temporal variation of Z over the next ζ slots from time t , then it can be represented in terms of $Z(t) = Z_0$ by a GM distribution, whose PDF is:

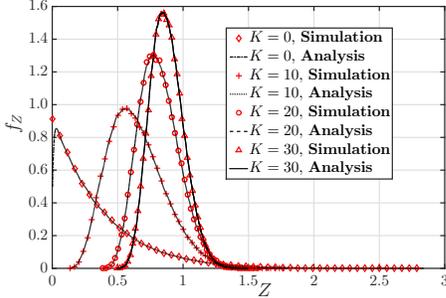
$$f_{\Omega}(\gamma | Z(t) = Z_0, \zeta) = \begin{cases} \frac{1}{\zeta} \sum_{j=1}^{\zeta} \frac{1}{\sqrt{2\pi}\sigma_{Z_j}} \left[1 - \Phi\left(-\frac{Z_0}{\sigma_{Z_j}}\right)\right]} e^{-\frac{\gamma^2}{2\sigma_{Z_j}^2}}, & \gamma \geq -Z_0 \\ 0, & \text{elsewhere.} \end{cases} \quad (18)$$

Proof. The proof is in line with the proof of Theorem 3. \square

Remark 7. It is important to note from Theorem 1 and 2 that both Ψ and Ω actually follow GM distribution. Unlike the classical GM definition, they are not any arbitrary PDFs approximated by a convex combination of Gaussian RVs. Hence we do not require the high complexity iterative EM algorithm to estimate their parameters. Instead, we simply require the channel state at time instant t and the channel



(a) Theorem 2 verification.



(b) Theorem 3 verification.

Fig. 1. Proof of concept for various Rice factor K . System parameters: node velocity $v = 6$ kmph, slot duration $T_p = 500 \mu\text{s}$, $\varphi = 0$, and $\zeta = 10$.

statistics, f_D and Rice factor K . Hence, the proposed model is of minimal complexity and always executes in constant time, i.e., its run time complexity is $O(1)$.

Fig. 1 demonstrates that the probability distributions of the actual signal envelope (and power) match very closely with the probability distributions obtained through the proposed characterization. Extensive Monte-Carlo simulation-based results presented in Fig. 1(a) and (b) validate the Theorems 2 and 3. Monte-Carlo simulation being an exhaustive technique, we have obtained the corresponding PDFs from 10^6 samples of the channel generated by the Clarke's model according to (3) and (5), with the system parameters as stated in Fig. 1. While Fig. 1(a) also shows the special case of Rayleigh distribution as Rician distribution with $K = 0$, Fig. 1(b) demonstrates the transformation of a non-central χ^2 distribution with two degrees of freedom to the well known exponential distribution for $K = 0$.

Remark 8. Though the GM distribution has been proposed for Rician fading channel, the proposed characterization is in general valid for all fading distributions. This generalization is based on the fact that irrespective of the underlying fading distribution of channel, the time derivative of received signal amplitude at Rx is always a zero mean Gaussian RV [34]. Only the variance of the RV changes depending on the slot duration T_p and the underlying fading distribution.

Remark 9. Based on Section II-D if we consider the imperfect CSI scenario here, Theorems 2 and 3 remain the same except for a subtle difference; θ_0 and Z_0 get replaced by $\hat{\theta}_0$ and \hat{Z}_0 respectively, as expressed earlier in Remarks 5 and 6.

Unlike the existing channel characterizations, the proposed

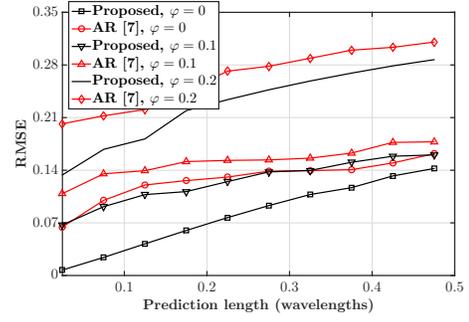


Fig. 2. Variation of RMSE of channel state prediction versus prediction window ζ . System parameters: node velocity = 6 kmph, slot duration $T_p = 500 \mu\text{s}$, and Rice factor $K = 40$.

model also serves the purpose of channel state prediction over the next few slots. In this context, it may be noted that there exist a number of channel prediction techniques that are used for channel prediction. While several tests have been conducted to evaluate and compare the performance of the existing techniques, Duel-Hallen [5] demonstrated that the auto-regressive (AR) model based prediction outperforms the others on real measurement data.

To normalize the prediction range, it is often expressed in spatial units, i.e., wavelengths. When the maximum Doppler frequency is f_D , τ seconds ahead prediction corresponds to a prediction range of $f_D\tau$. Therefore, with slot duration $T_p = 500 \mu\text{sec}$, ζ slot ahead prediction implies a prediction range of $\zeta \cdot f_D T_p$ wavelengths. Accordingly in Fig. 2 we compare the AR-based method with our scheme, where root mean square error (RMSE) is plotted against the prediction range. Note that the figure demonstrates both the cases of perfect and imperfect CSI scenarios, i.e., $\varphi = 0, 0.1$, and 0.2 .

For the AR-based technique, following the parameter values considered in [7] we take an observation window of past 200 samples, a predictor filter of order 20, and Burg method is employed. The plots demonstrate that the proposed characterization significantly outperforms the AR-based approach; the average improvement in performance (ΔImp) being $\Delta\text{Imp} = 36.94\%$, 17.51% , and 13.53% corresponding to $\varphi = 0, 0.1$, and 0.2 , respectively. Hence we observe that ΔImp decreases with increasing φ . We also observe that in both the cases, performance in the perfect CSI feedback scenario i.e., when $\varphi = 0$, acts as an upper bound to the scenarios when $\varphi \neq 0$.

Remark 10. Methods in [5], [7] are for long-term prediction whereas the proposed method here is essentially for short-term characterization. This can also be observed from Fig. 2, where we note that for any particular φ , ΔImp decreases with increasing ζ . For example when $\varphi = 0.1$, $\Delta\text{Imp} = 38.12\%$ and 9.79% corresponding to the wavelengths $\zeta = 0.025$ and 0.4750 , respectively. However, the proposed characterization helps model the temporal channel variations more effectively, which will be discussed subsequently.

Note that the proposed method characterizes the signal envelope using a RV, and thus the outcome is a random quantity that changes in different trials. Hence we have considered the averaged-out performance of the proposed method over 10^5 trials and accordingly obtained the plots in Fig. 2. When

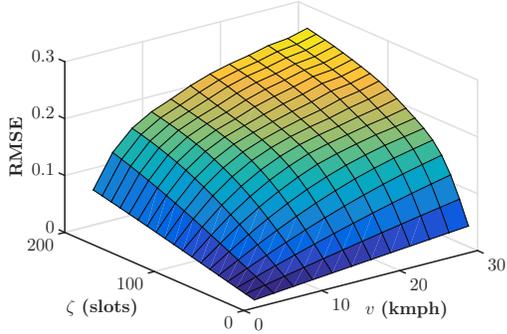


Fig. 3. Joint effect of prediction window ζ and node velocity v on RMSE of channel state estimation. System parameters: $\varphi = 0$, $T_p = 500 \mu\text{s}$, and Rice factor $K = 40$.

$\varphi = 0$, mean and variance of the error in the figure is 0.0080 and 5.1289×10^{-5} , respectively. Similarly the moments can also be calculated from the plot corresponding to $\varphi = 0.1$.

In this case, we are doing the prediction by taking both the current channel state as well as the probability distribution of channel variation into account. On the other hand, the AR-based approach does the same by using a set of update equations on the training data set. Thus, the AR-based approach relies solely on the training data set and not on the statistical properties of the channel or its variation. This is the reason behind significantly lower RMSE of the proposed characterization. It may be noted further that Fig. 2. exhibits an overall increasing trend of RMSE with ζ irrespective of the approach, which is intuitive.

Lastly, we highlight that a W.S.S. scenario is considered in this work which is considered in Fig. 2 plots as well, whereas the study in [7] originally dealt with a non-stationary one. The proposed characterization may or may not outperform the characterization approach in [7] in a non-stationary setup. Therefore, we do not make any claim on the performance of the proposed characterization in a non-stationary channel setup; we keep it aside as a topic of future work.

Fig. 3 shows joint effect of ζ and v on the estimation error, i.e., RMSE. It shows that, for a given ζ , RMSE increases with v ; for $\zeta = 40$ slots, RMSE = 0.0148 and 0.1054, respectively for $v = 2.5$ and 20 kmph. The reason behind this is attributed to the decrease of correlation with increasing v , which is captured by the $f_D T_p$ product [30]. The figure also reaffirms the observation made in Fig. 2 that RMSE increases with ζ when v is constant; for $v = 12.5$ kmph, RMSE = 0.0152 and 0.1592, respectively for $\zeta = 10$ and 110 slots. Note that the error increases with increase in either of ζ or v , or both. We exploit this property in Section V and propose an upper bound $\zeta^{(\max)}$ as a function of the fading scenario.

Remark 11. ‘Node velocity’ v in the proposed characterization may give an impression that it holds only in mobile scenarios. However, v does not necessarily imply mobile Tx and/or Rx. It also applies to scenarios, where the Tx and Rx nodes are static but the scatterers present along the path are in motion [45]. Hence, in applications like IoT, where nodes may be statically placed, scatterers could be the primary reasons for Doppler effect. Thus the characterization applies

to scenarios even with static Tx and/or Rx in presence of slow moving scatterers, like walking pedestrians or foliage movement, which equate to highly correlated scenarios. In fact, the system performance significantly improves in such scenarios, as they represent $v \rightarrow 0$ but $v \neq 0$, which will be discussed in Section V.

IV. PERFORMANCE ANALYSIS OF PROPOSED GM CHARACTERIZATION

The existing performance metrics, such as, outage probability and ergodic capacity are solely based on channel statistics, which do not capture the effects of short-term channel variation. In the other extreme, the conventional adaptive communication strategies (ACS) rely on instantaneous channel state in every transmission moment, i.e., time slot. Thus, the ACS approach accounts for the short-term phenomena of channel variation, but at the cost of high bandwidth and energy consumption overhead.

In contrast to the above approaches we estimate channel state in short-term basis without requiring feedback in every slot. Accordingly, we define *short-term outage probability* and *short-term channel capacity*. The essence of this ‘short-term’ channel characterization lies in exploiting the temporal correlation present in the channel. Thus, the proposed approach intelligently collects the dynamic CSI for an improved trade-off between short-term capacity gain and the overheads.

A. Short-term Outage Probability

Outage probability (\bar{P}_{out}) is defined as the probability that the received SNR Z at Rx falls below a certain specified threshold Z_{th} [35], i.e.,

$$\bar{P}_{\text{out}} = \Pr \{ Z < Z_{th} \} = \int_0^{Z_{th}} f_Z(z) dz. \quad (19)$$

We are interested in estimating the outage probability in next ζ slots based on current $Z(t)$, when $Z(t) = Z_0 (< Z_{th})$. This outage probability is different from the \bar{P}_{out} defined in [35], as it incorporates the knowledge of $Z(t)$ and it holds only for the next ζ slots from the time instant t when $Z(t)$ was obtained. Hence we define *short term outage probability* $P_{\text{out}}^{(\text{ST})}(Z(t), \zeta)$ as follows.

From Theorem 3, we know that $f_{\Omega}(\gamma | Z(t) = Z_0, \zeta)$ is the PDF of variation of Z for next ζ time slots given that $Z(t) = Z_0$ and accordingly we define $P_{\text{out}}^{(\text{ST})}(Z(t) = Z_0, \zeta)$. Hence from (18) and (19) we get

$$\begin{aligned} P_{\text{out}}^{(\text{ST})}(Z(t) = Z_0, \zeta) &= \frac{1}{\zeta} \sum_{j=1}^{\zeta} \frac{1}{1 - \Phi\left(-\frac{Z_0}{\sigma_{Z_j}}\right)} \\ &\times \int_0^{Z_{th}} \frac{1}{\sqrt{2\pi}\sigma_{Z_j}} e^{-\frac{(\gamma - Z_0)^2}{2\sigma_{Z_j}^2}} d\gamma \\ &= \frac{1}{\zeta} \sum_{j=1}^{\zeta} \frac{Q\left(-\frac{Z_0}{\sigma_{Z_j}}\right) - Q\left(\frac{Z_{th} - Z_0}{\sigma_{Z_j}}\right)}{1 - \Phi\left(-\frac{Z_0}{\sigma_{Z_j}}\right)}, \end{aligned} \quad (20)$$

where $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-\frac{t^2}{2}} dt$. The importance of $P_{\text{out}}^{(\text{ST})}(Z(t), \zeta)$ is that, unlike \bar{P}_{out} , it gives us an estimate of

outage probability in next ζ slots when $Z(t) = Z_0 (< Z_{th})$. For instance if $P_{\text{out}}^{(\text{ST})}(Z(t) = Z_0, \zeta) = \epsilon$, this implies that when $Z(t) = Z_0$, on average Z will remain below Z_{th} in ϵ fraction of the next ζ slots.

B. Short-term Channel Capacity

Analogous to *short term outage probability*, we introduce the definition of *short term channel capacity*. Short term channel capacity $C^{(\text{ST})}(Z(t), \zeta)$ is defined as the channel capacity in the next ζ slots, which depends on the current channel state $Z(t)$. $C^{(\text{ST})}(Z(t), \zeta)$ is defined as follows.

Ergodic capacity [35] of a fading channel is defined as

$$\bar{C} = \mathbb{E}_Z [\log_2(1 + z)] = \int_0^\infty \log_2(1 + z) f_Z(z) dz, \quad (21)$$

where $Z = \frac{P|h|^2}{\sigma_o^2}$. Using Theorem 3, we define $C^{(\text{ST})}(Z(t), \zeta)$ on the basis of current sensed SNR $Z(t) = Z_0$ and its variation over next ζ slots. Accordingly from (18) and (21) we get

$$\begin{aligned} C^{(\text{ST})}(Z(t) = Z_0, \zeta) &= \frac{1}{\zeta} \sum_{j=1}^{\zeta} \frac{1}{1 - \Phi\left(-\frac{Z_0}{\sigma_{Z_j}}\right)} \\ &\quad \times \int_0^\infty \log_2(1 + \gamma) \\ &\quad \times \frac{1}{\sqrt{2\pi}\sigma_{Z_j}} e^{-\frac{(\gamma - Z_0)^2}{2\sigma_{Z_j}^2}} d\gamma. \end{aligned} \quad (22)$$

It is important to note that $\bar{C} \neq C^{(\text{ST})}$, as \bar{C} is a statistical average and it does not depend on Z_0 unlike $C^{(\text{ST})}$. $C^{(\text{ST})}$ is especially useful in scenarios where we aim to estimate the future behavior of the channel in the next few time slots depending on the current channel state. As $C^{(\text{ST})}$ is a function of Z_0 , it is intuitive that $C^{(\text{ST})} > \bar{C}$ when $Z_0 \gg 1$ and vice-versa otherwise.

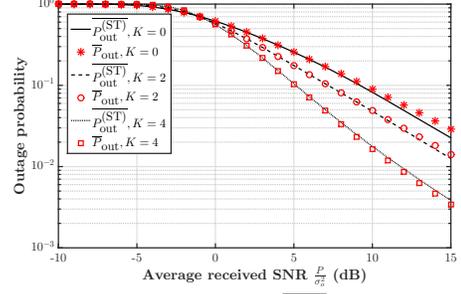
To the best of our knowledge, an exact solution of the integral in (22) cannot be obtained in closed-form. In order to obtain an approximate solution, we take help of Taylor series expansion of $\log_2(1 + \gamma)$ about the mean value of γ , i.e., Z_0 to obtain an approximation [46]:

$$\begin{aligned} \log_2(1 + \gamma) &\approx \frac{1}{\ln 2} \left[\ln(1 + Z_0) + \frac{\gamma - Z_0}{1 + Z_0} - \right. \\ &\quad \left. \frac{1}{2} \left(\frac{\gamma - Z_0}{1 + Z_0} \right)^2 + O[(\gamma - Z_0)^2] \right]. \end{aligned} \quad (23)$$

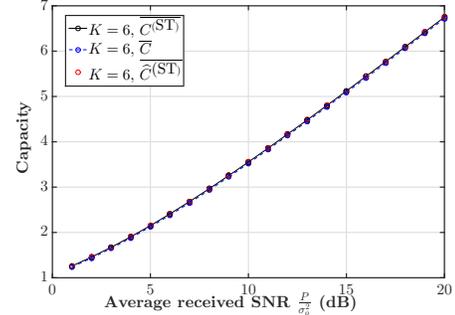
Using (23) in (22), we get a tight approximation of $C^{(\text{ST})}(Z(t) = Z_0, \zeta)$ as shown in (24). RMSE is 0.0076 (very close to 0) in Fig. 4(b) signifies excellent precision of the proposed approximation.

Remark 12. We observe from the definition of $P_{\text{out}}^{(\text{ST})}(Z(t), \zeta)$ and $C^{(\text{ST})}(Z(t), \zeta)$ that both of them are functions of time. Hence their temporal averages for a given value of ζ are:

$$\overline{P_{\text{out}}^{(\text{ST})}}(\zeta) = \lim_{\lambda \rightarrow \infty} \frac{1}{\lambda} \sum_{i=1}^{\lambda} P_{\text{out}}^{(\text{ST})}(Z(t_i), \zeta) \quad \text{and} \quad (25)$$



(a) Convergence of $\overline{P_{\text{out}}^{(\text{ST})}}$ to \bar{P}_{out} .



(b) Convergence of $\overline{C^{(\text{ST})}}$ to \bar{C} .

Fig. 4. Validation of Remark 12. System parameters: node velocity = 12 kmph, slot duration $T_p = 500 \mu\text{s}$, $\varphi = 0$, $Z_{th} = 0$ dB, and $\zeta = 10$.

$$\overline{C^{(\text{ST})}}(\zeta) = \lim_{\lambda \rightarrow \infty} \frac{1}{\lambda} \sum_{i=1}^{\lambda} C^{(\text{ST})}(Z(t_i), \zeta).$$

As the channel considered is wide-sense stationary, $\overline{P_{\text{out}}^{(\text{ST})}}$ and $\overline{C^{(\text{ST})}}$ converge to their respective well-known statistical averages, \bar{P}_{out} (expressed in (19)) and \bar{C} (expressed in (21)).

Numerical Verification: In Fig. 4(a) it is observed that for a given threshold Z_{th} , $\overline{P_{\text{out}}^{(\text{ST})}}$ matches very closely with the \bar{P}_{out} irrespective of the value of Rice factor K . Similarly, $\overline{C^{(\text{ST})}}$ is compared with \bar{C} in Fig. 4(b) to show the validation of Remark 12.

With the help of these definitions we quantify the performance of channel taking into consideration the aspect of its temporal variation, which is discussed next.

Remark 13. $\zeta \cdot T_p$ is a measure of inter-feedback duration as it captures the time duration between two consecutive CSI feedbacks from Rx in a particular fading scenario.

Unlike \bar{C} , $C^{(\text{ST})}(Z(t), \zeta)$ takes into account the present channel condition. Using $C^{(\text{ST})}(Z(t), \zeta)$ we can estimate the amount of data that can be transferred from Tx to Rx in next ζ slots with a fixed transmission power P , without requiring to actually know Z in those ζ slots. Thus, the proposed GM characterization allows to estimate the amount of data that can be transferred in a certain time interval without any investment in obtaining CSI during the entire interval. This in turn also results in enhanced energy efficiency of the system.

C. Effect of $Z(t)$ on Energy Efficiency

In order to estimate the channel, Tx sends a pilot signal. The channel estimate is obtained as feedback from Rx. We

$$\begin{aligned}
\widehat{C}^{(\text{ST})}(Z(t) = Z_0, \zeta) &\approx \frac{1}{\zeta \ln 2} \sum_{j=1}^{\zeta} \frac{1}{1 - \Phi\left(-\frac{Z_0}{\dot{\sigma}_{Z_j}}\right)} \\
&\times \int_0^{\infty} \left[\ln(1 + Z_0) - \frac{1}{2} \left(\frac{\gamma - Z_0}{1 + Z_0} \right)^2 + O[(\gamma - Z_0)^2] \right] \times \frac{1}{\sqrt{2\pi}\dot{\sigma}_{Z_j}} e^{-\frac{(\gamma - Z_0)^2}{2\dot{\sigma}_{Z_j}^2}} d\gamma \\
&\approx \frac{1}{\zeta \ln 2} \sum_{j=1}^{\zeta} \frac{1}{1 - \Phi\left(-\frac{Z_0}{\dot{\sigma}_{Z_j}}\right)} \left[\ln(1 + Z_0) Q\left(-\frac{Z_0}{\dot{\sigma}_{Z_j}}\right) - \frac{\dot{\sigma}_{Z_j}^2}{2\sqrt{\pi}(1 + Z_0)^2} \Gamma\left(1.5, \frac{Z_0^2}{2\dot{\sigma}_{Z_j}^2}\right) \right]. \quad (24)
\end{aligned}$$

Here $\Gamma(s, x) = \int_x^{\infty} t^{s-1} e^{-t} dt$ is the upper incomplete Gamma function.

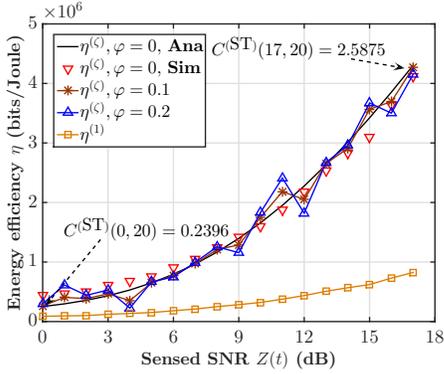


Fig. 5. Variation of η in next ζ slots versus currently sensed SNR $Z(t)$. System parameters: $v = 20$ kmph, $T_p = 500 \mu\text{s}$, $\delta = 0.2$, Rice factor $K = 2$, $P = 0$ dBm, $P_{\text{pilot}} = 10$ dBm, $\frac{P}{\sigma_z^2} = 10$ dB, and $\zeta = 20$.

assume that the pilot signal duration T_f is very small compared to T_p , i.e., $T_f = \delta T_p$ where $\delta \ll 1$ [47]. Accordingly, the achievable throughput for next ζ slots based on the proposed characterization is:

$$\begin{aligned}
\Lambda_{\text{prop}}(Z(t), \zeta) &= \frac{\zeta T_p - T_f}{\zeta T_p} C^{(\text{ST})}(Z(t), \zeta) \\
&= \frac{\zeta - \delta}{\zeta} C^{(\text{ST})}(Z(t), \zeta). \quad (26)
\end{aligned}$$

If P is the transmission power and P_{pilot} is the power of the pilot signal, then energy-efficiency η of the system (in bits/Joule) with respect to the proposed characterization is:

$$\begin{aligned}
\eta^{(\zeta)}(Z(t), \zeta) &= \frac{\text{Estimated data transfer capability } D^{(\text{ST})}}{\text{Energy consumed to transfer the data}} \\
&= \frac{\Lambda_{\text{prop}}(Z(t), \zeta) \zeta}{P[\zeta - \delta] T_p + 2P_{\text{pilot}} T_f}. \quad (27)
\end{aligned}$$

Similarly we consider the case when CSI is required in every slot. Accordingly the achievable throughput is:

$$\begin{aligned}
\Lambda_{\text{ext}}(Z(t)) &= \frac{T_p - T_f}{T_p} \log_2(1 + Z(t)) \\
&= (1 - \delta) \log_2(1 + Z(t)) \quad (28)
\end{aligned}$$

and the corresponding energy-efficiency is $\eta^{(1)}(Z(t)) = \frac{\Lambda_{\text{ext}}(Z(t))}{P[1 - \delta] T_p + 2P_{\text{pilot}} T_f}$ bits/Joule. It is to be noted that, as $T_f = \delta T_p$, energy efficiency η also varies with changing δ , when the remaining parameters are constant.

Fig. 5 shows the variation of η against $Z(t)$. As both $\eta^{(\zeta)}$ and $\eta^{(1)}$ take $Z(t)$ into consideration, they are useful in various applications where $Z(t)$ plays a significant role. We observe that though both of them increase proportionately with $Z(t)$, $\eta^{(\zeta)}$ always exceeds $\eta^{(1)}$ by a monotonically increasing margin. As it can be noted from the figure, the gap consistently increases with increasing $Z(t)$ and it reaches as high as approximately 4 times over the per-slot feedback based communication system. It is important to note from Fig. 5 that the proposed characterization offers not only higher energy efficiency, but also $C^{(\text{ST})}(17, 20)$ is much higher compared to $C^{(\text{ST})}(0, 20)$. Lastly note that, though η corresponding to $Z(t) = \hat{Z}_0$ with $\varphi = 0.1, 0.2$ follows a similar upward trend, unlike the other plots, they are not smooth in nature. This lack of smoothness in the plot can be attributed to the random nature of error term e in (13), i.e., $e \sim \mathcal{CN}(0, 1)$, which increases with increasing φ . This observation brings out the advantage of the proposed GM characterization of temporally-varying wireless channel, and hence we conclude that *the temporal channel variations can be exploited to reduce channel sensing frequency, leading to a considerable gain in energy efficiency of the communication system.*

V. CHOICE OF OPTIMAL ζ

We defined $C^{(\text{ST})}(Z(t), \zeta)$ in Section IV-B, but we have not yet defined its range of ζ for a given $Z(t)$. Though it appears that ζ can take any arbitrary positive integer value, it is actually not the case.

A. The Need for an Optimal ζ

The proposed characterization exploits the temporal correlation present in channel. Hence, ζ is dependent on the nature of fading scenario as well as the application for which it is being estimated, i.e., the error tolerance level e_t . This can be understood from Fig. 6. In this context, we define error as: $\text{Error } \Delta_\zeta = \frac{|C_{\text{ana}}^{(\text{ST})}(Z(t), \zeta) - C_{\text{sim}}^{(\text{ST})}(Z(t), \zeta)|}{C_{\text{sim}}^{(\text{ST})}(Z(t), \zeta)} \cdot 100\%$.

Fig. 6 shows the variation of Δ_ζ with ζ , where $e_t = 5\%$ is considered as an example. The value of this error tolerance is application specific. With respect to e_t we observe from Fig. 6 that allowable $\zeta^{(\text{max})}$ decreases with increasing node velocity v , such that $\Delta_\zeta \leq e_t$ for a given set of system parameters $Z(t), v, T_p, \frac{P}{\sigma_z^2}$, and K . From Fig. 6 we observe that, at $v = 5$

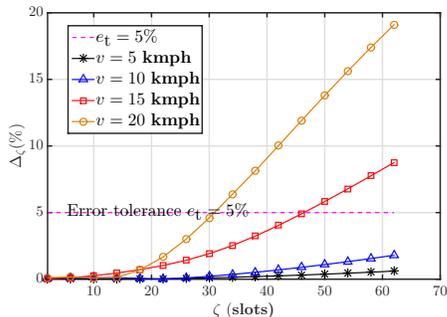


Fig. 6. Variation of Δ_ζ versus ζ . System parameters: current SNR $Z(t) = 10$ dB, Rice factor $K = 4$, $T_p = 500 \mu\text{s}$, $\varphi = 0$, and $\frac{P}{\sigma^2} = 10$ dB.

kmph, Δ_ζ for $\zeta = 60$ slots is far below the chosen $e_t = 5\%$, whereas at $v = 20$ kmph, ζ is restricted to $\zeta^{(\max)} = 32$ slots. This physically means that $\zeta^{(\max)}$ is inversely proportional to node velocity for a given e_t . Hence we conclude that the choice of ζ cannot be arbitrary. Accordingly we propose an upper bound $\zeta^{(\max)}$ as a function of the fading scenario.

We estimate $\zeta^{(\max)}$ based on an acceptable limit of temporal correlation in between the samples. In other words, the application context plays a critical role in determining $\zeta^{(\max)}$. Hence the proposed characterization is termed as *context-aware characterization* of the temporally varying channel.

In the following, we propose two approaches for estimation of $\zeta^{(\max)}$, which are based on temporal correlation.

1) *Correlation based Choice of $\zeta^{(\max)}$* : In this sub-section, we estimate $\zeta^{(\max)}$ based on the temporal correlation present in the channel. From [30] we have, correlation coefficient ρ for a given f_D and T_p is given by:

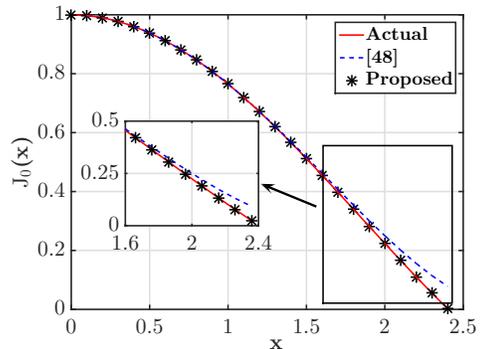
$$\rho(\zeta) = J_0(2\pi f_D \zeta T_p), \quad (29)$$

where $J_0(\cdot)$ is Bessel function of first kind and of zeroth order and ζT_p is time in between samples under consideration.

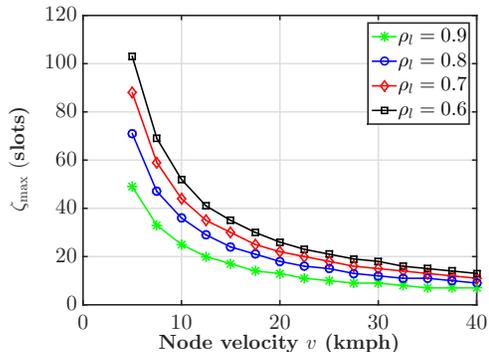
We are interested in saving energy by optimally avoiding unnecessary pilot signal transmission based on CSI feedback. To this end, the number of slots between two pilot signal transmissions is maximized. Hence $\zeta^{(\max)}$ is estimated in a way that the correlation between two successive channel samples separated by a time interval $\zeta^{(\max)} T_p$ seconds does not fall below a user-defined *acceptable reliability* limit $\rho^{(l)}$, i.e., $\zeta^{(\max)} = \max_{\rho(\zeta) \geq \rho^{(l)}} \{\zeta\}$, for a given set of f_D and T_p .

$\rho^{(l)} \in [0, 1]$ is a system parameter that depends on the application at hand. It is an abstraction of the desired accuracy of prediction. $\rho^{(l)} = 1$ implies that absolutely no error is allowed, and $\rho^{(l)} = 0$ represents the other extreme. If ρ between channel samples separated by $\zeta^{(\max)}$ time slots goes below $\rho^{(l)}$, it implies that the previously sensed $Z(t)$ is no longer good for providing the minimum acceptable reliability.

However due to the oscillatory nature of $J_0(\cdot)$, $\rho(\zeta)$ is not a monotonically decreasing function of ζ . It may happen that we obtain multiple values of $\zeta^{(\max)}$ for a particular $\rho^{(l)}$ in a given fading scenario, which is not a practically feasible solution. To obtain a reasonable solution, we consider $J_0(\cdot)$ only upto the point where its argument reaches the first zero. As $J_0(x)$ reaches 0 for the first time at $x = 2.4048$, $[0, 2.4048]$ now



(a) Validation of proposed analytical approximation.



(b) Variation of maximum channel state prediction window $\zeta^{(\max)}$ with node velocity v .

Fig. 7. Illustration of ρ based choice of $\zeta^{(\max)}$.

becomes our domain of interest, where $J_0(\cdot)$ is monotonically decreasing. Accordingly we replace $J_0(\cdot)$ by a tight cosine approximation as:

$$J_0(x) \triangleq 0.32 + 0.68 \cos(0.8565x) \quad x \in [0, 2.4048]. \quad (30)$$

It may be noted from Fig. 7(a) that the proposed analytical approximation very closely follows $J_0(\cdot)$ in $[0, 2.4048]$. The RMSE value 0.0002 (very close to 0) and R-square statistics value 0.9999 (very close to 1) signify goodness of the proposed approximation. The authors in [48] had proposed the following approximation in the same domain of interest:

$$J_0(x) \triangleq 1 - \frac{x^2}{4} + \frac{x^4}{64} + o(z^4) \quad x \in [0, 2.4048]. \quad (31)$$

It can also be observed from Fig. 7(a) that (31) is a good approximation of $J_0(\cdot)$ in $[0, 1.5]$, but not beyond. On the other hand, the proposed approximation in (30) closely follows $J_0(\cdot)$ in the entire $[0, 2.4048]$ range.

Using (29) and (30), we estimate $\zeta^{(\max)}$ for a given set of f_D and T_p as:

$$\zeta^{(\max)} = \left\lceil \frac{0.1858}{f_D T_p} \cos^{-1} \left(\frac{\rho^{(l)} - 0.32}{0.68} \right) \right\rceil. \quad (32)$$

Note that $\zeta^{(\max)}$ is not an independent parameter. It is a function of f_D , T_p , and $\rho^{(l)}$. As $\rho^{(l)} \in [0, 1]$, it is possible to have very high values of $\zeta^{(\max)}$ only when we have extremely small values of $f_D T_p$ and as $T_p > 0$, asymptotically $\zeta^{(\max)} \rightarrow \infty$ as $f_D \rightarrow 0$. This observation is intuitive also; very small value of $f_D T_p$ implies that the channel is varying

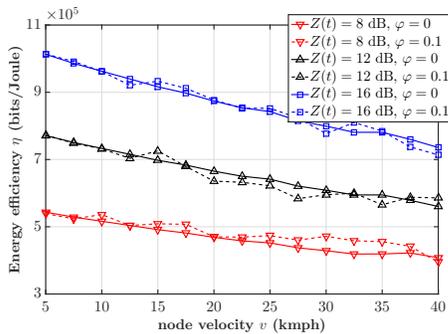


Fig. 8. Variation of energy efficiency with node velocity. System parameters: $\rho^{(l)} = 0.8$, $T_p = 500 \mu\text{s}$, $\delta = 0.2$, Rice factor $K = 2$, $P = 0$ dBm, $P_{\text{pilot}} = 10$ dBm, and $\frac{P}{\sigma_o^2} = 10$ dB.

at a very slow rate, i.e., it is a highly correlated scenario and hence, we obtain a large value of $\zeta^{(\text{max})}$.

This has been explained in Fig. 7(b), which shows the variation of $\zeta^{(\text{max})}$ with node velocity v for various levels of desired accuracy, i.e., $\rho^{(l)}$. It is observed from the figure that $\zeta^{(\text{max})}$ follows a decreasing trend with increase in v , i.e., as the channel changes from slowly-fading to fast fading. This is because increasing v results in high $f_D T_p$ product, i.e., the channel is now less correlated than it was with lower v . Our claim regarding the $\zeta^{(\text{max})} \rightarrow \infty$ scenario is also verified in this figure, where we see that irrespective of $\rho^{(l)}$, $\zeta^{(\text{max})}$ shoots up as v decreases, i.e., $f_D \rightarrow 0$.

It can also be observed from the figure that in case of lower values of v , there is a considerable gap among the $\zeta^{(\text{max})}$ for various $\rho^{(l)}$. For example, at $v = 5$ kmph, $\zeta^{(\text{max})} = 103$ slots for $\rho^{(l)} = 0.6$ whereas $\zeta^{(\text{max})} = 49$ slots for $\rho^{(l)} = 0.9$. But this gap gradually decreases as v increases. At $v = 60$ kmph, $\zeta^{(\text{max})} = 9$ slots for $\rho^{(l)} = 0.6$, whereas $\zeta^{(\text{max})} = 5$ slots for $\rho^{(l)} = 0.9$. It can be further observed from Fig. 7(b) that for a given v , $\zeta^{(\text{max})}$ is higher for lower values of $\rho^{(l)}$. This observation highlights the trade-off that exists in between the accuracy of estimation denoted by higher value of $\rho^{(l)}$ and number of time slots $\zeta^{(\text{max})}$ that can be effectively estimated; the more accurate we are, lesser is the value of $\zeta^{(\text{max})}$.

2) *Error of Estimate based Choice of $\zeta^{(\text{max})}$* : Sometimes we are not provided with $\rho^{(l)}$ but an acceptable *standard error of estimate* threshold $\sigma_e^{(l)}$ is given. In such applications, we estimate $\zeta^{(\text{max})}$ as discussed in this sub-section. The corresponding analysis is presented in Appendix A.

B. Effect of Mobility on System Performance

We now discuss the effect of node velocity v on η . From Fig. 8 it can be noted that η exhibits an overall decreasing trend with v . It is observed that increase in v results in high $f_D T_p$, which implies that the channel is relatively less correlated at a higher $f_D T_p$. This decrease in correlation with increasing v results in smaller $\zeta^{(\text{max})}$ for a given $\rho^{(l)}$. As a result, η decreases with increasing v for a given $\rho^{(l)}$ or $\sigma_e^{(l)}$.

It is also noted from Fig. 8 that higher $Z(t)$ results in a higher η at any particular v . For instance at $v = 20$ kmph with $\varphi = 0$, we have $\eta = 4.701 \times 10^5$, 6.654×10^5 , and 8.772×10^5 bits/Joule for $Z(t) = 8, 12,$ and 16 dB, respectively. Thus, Fig. 8 in a way reinstates the point that was made in Fig. 5

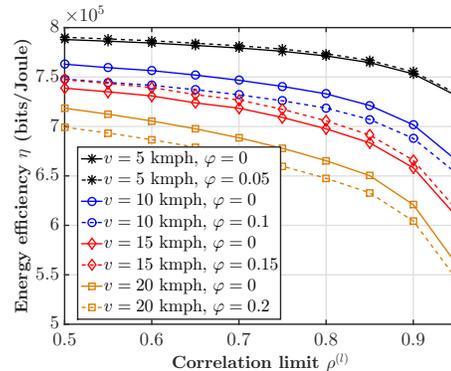


Fig. 9. Variation of energy efficiency with $\rho^{(l)}$. System parameters: sensed SNR $Z(t) = 12$ dB, slot duration $T_p = 500 \mu\text{s}$, $\delta = 0.2$, Rice factor $K = 2$, $P = 0$ dBm, $P_{\text{pilot}} = 10$ dBm, and $\frac{P}{\sigma_o^2} = 10$ dB.

regarding the fact that the channel being in deep fade is not equivalent to when it is in good state. Therefore, it is possible to transfer a large amount of data with higher energy efficiency when the channel is in good state than when it is not. Lastly, the lack of smoothness corresponding to the $\varphi \neq 0$ scenario has been explained earlier in Fig 5.

C. Trade-off between Optimal ζ and Energy Efficiency

It is important to note here that unlike ρ , $\rho^{(l)}$ is not an independent parameter. While ρ denotes the correlation present in the channel, $\rho^{(l)}$ is an application-dependent parameter as defined by the user. $\rho^{(l)}$ is an abstraction of the minimum required degree of accuracy in estimation. As noted in sections V-A1 and V-A2 that a trade-off exists between quality of estimation and $\zeta^{(\text{max})}$, there also exists a trade-off between energy efficiency η and $\rho^{(l)}$.

From (27) we have $\eta^{(\text{max})} = \frac{\Lambda_{\text{prop}}(Z(t), \zeta^{(\text{max})}) \zeta^{(\text{max})}}{P[\zeta^{(\text{max})} - \delta] T_p + 2P_{\text{pilot}} T_f}$ where $\zeta^{(\text{max})}$ is calculated from (32) or (A-2) depending on the respective availability of $\rho^{(l)}$ or $\sigma_e^{(l)}$. Fig. 9 shows that for a given set of system parameters, energy efficiency monotonically decreases with increasing $\rho^{(l)}$.

It can be further observed from Fig. 9 that the rate of decrease of $\eta^{(\text{max})}$ with $\rho^{(l)}$ increases with increasing node velocity. For $\varphi = 0$, we have $\frac{\partial \eta^{(\text{max})}}{\partial \rho^{(l)}} = -1.3178 \times 10^5$ and -3.5156×10^5 corresponding to $v = 5$ and 20 , kmph respectively. The above observation implies that the correlation in channel decreases with increasing v . As a result the inter-feedback duration also decreases with increasing v for a given $\rho^{(l)}$, and hence the trade-off. A similar variation of η against $\sigma_e^{(l)}$ can also be obtained. It is further interesting to note that although the nature of plots remain same, the relative gap in performance between the $\varphi = 0$ and $\varphi \neq 0$ scenarios show an increasing trend with increasing φ , which is intuitive also.

Finally, as ζ denotes the number of components in the proposed GM characterization, unlike the classical GM, it is the application at hand and not any iterative algorithm that determines the number of components of the GM distribution.

VI. EFFECT OF PROPOSED CHARACTERIZATION ON ADAPTIVE COMMUNICATION SYSTEM

In this section we discuss the effect of the proposed characterization on channel adaptive communication.

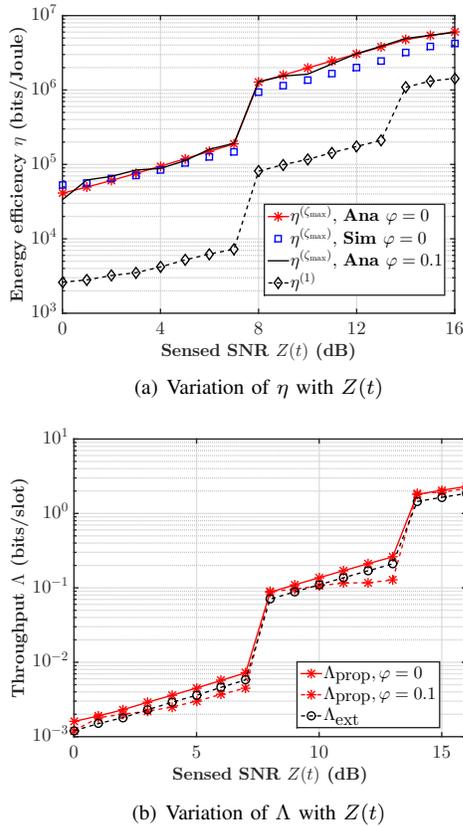


Fig. 10. Comparison of conventional and proposed multi-level power allocation strategy. System parameters: $v = 20$ kmph, $T_p = 500 \mu\text{s}$, Rice factor $K = 4$, $\delta = 0.1$, $P_{\text{pilot}} = 10$ dBm, $\mathbb{P} = \{-20, -10, 0\}$ dBm, thresholds $= \{8, 14\}$ dB, and $\rho^{(l)} = 0.8$.

Limited battery energy being one of the tightest constraints, transmission power level control is important for maximizing data throughput and node lifetime simultaneously. Hence we invoke a power adaptive transmission strategy [49] based on the proposed characterization that increases the overall energy efficiency of the system.

Discrete set of transmission power levels $\mathbb{P} = \{P_1, \dots, P_{|\mathbb{P}|}\}$ is considered where $P_1 < P_2 < \dots < P_{|\mathbb{P}|}$ at Tx and $|\cdot|$ denotes the cardinality of \mathbb{P} . A particular P is chosen based on the received $Z(t)$ estimate from Rx. Each $Z(t)$ is assigned to a region $[Z_{j-1}, Z_j]$, $j = 1, \dots, |\mathbb{P}|$ with $Z_0 = 0$ and $Z_{|\mathbb{P}|} = \infty$. Boundaries Z_j ($j = 1, \dots, |\mathbb{P}| - 1$) and $|\mathbb{P}|$ are chosen depending on the application; the better we want to make use of channel, larger is the value of $|\mathbb{P}|$, and vice-versa.

Power level P_j is selected at Tx if $Z(t) \in [Z_{j-1}, Z_j]$. Tx then continues to transmit with P_j for the next ζ slots without any further pilot signal transmission. Accordingly, the energy efficiency $\eta^{(\max)}$ is obtained for a given $\rho^{(l)}$ or $\sigma_e^{(l)}$.

Fig. 10(a) shows the effect of the proposed characterization in context of adaptive transmission power control. It can be noted that, in terms of energy efficiency the proposed GM characterization-based strategy offers a consistent gain of more than 3.15 times over and above the per-slot feedback based power control, irrespective of the value of $Z(t)$. Sudden rise in η are noted at $Z(t) = 8$ dB and $Z(t) = 14$ dB, which are because these are the thresholds at which P changes.

From (27) we get $\eta^{(\zeta)}(Z(t), \zeta) = \frac{\Lambda_{\text{prop}}(Z(t), \zeta)}{P[\zeta - \delta]T_p + 2P_{\text{pilot}}T_f}$, from (28) we get $\eta^{(l)}(Z(t)) = \frac{\Lambda_{\text{ext}}(Z(t))}{P[1 - \delta]T_p + 2P_{\text{pilot}}T_f}$, and pilot signal duration $T_f = \delta T_p$. Thus, the gain of the proposed characterization is also dependent on δ . Therefore, although the proposed characterization consistently offers a gain in terms of energy efficiency, the gain margin depends on the pilot signal duration. The shorter the pilot signal duration, the more is the time for data transmission, and vice-versa.

The corresponding throughput performance plot in Fig. 10(b) demonstrates that the proposed transmission strategy does not incur any loss in terms of throughput; we also observe a consistent gap that exists between Λ_{prop} and Λ_{ext} irrespective of $Z(t)$. Thus, for a baseline throughput performance, the proposed characterization offers significant gain in energy efficiency, as it does not require pilot signal transmission in each time slot. Lastly, we also observe the $\varphi \neq 0$ scenario, where Λ follows the general trend, though in a rugged manner. This behavior can be attributed to the randomness of $e \sim \mathcal{CN}(0, 1)$, which has been explained earlier in context of Fig. 5.

Unlike transmission power adaptive communication system, when a rate-adaptive multi-level modulation based communication system is used, the analytical relation (22) of $C_{\text{est}}(Z(t) = Z_0, \zeta)$ with $Z(t)$ for a given ζ ceases to apply. In this case the effective throughput depends on the assigned modulation m , chosen based on $Z(t)$. However, here also the concept of $\zeta^{(\max)}$ helps to improve the energy efficiency. The corresponding discussion is presented in Appendix B.

Significance of the proposed approach

Channel-aware resource allocation approaches in the existing literature [20]–[22] consider the wireless channel state to be static over the coherence time duration $t_{\text{ch}} = \frac{0.423}{f_D}$ [50]. To this end, the study in [12] demonstrated that this assumption is a good rule-of-thumb for systems that focus on average performance, for example traditional cellular or WiFi-type systems. However, in scenarios, where worst-case performance guarantee is required, such as in URLLC, this assumption falls short. At the same time, deciding transmission strategy based on CSI feedback in every time slot involves significant communication and energy consumption overheads, as observed in Fig. 10(a). In this context, the definition of $\zeta^{(\max)}$ serves the purpose; depending on the user-defined acceptable reliability limit $\rho^{(l)}$ or standard error of estimate $\sigma_e^{(l)}$, the inter-feedback interval $\zeta^{(\max)}$ is estimated to assess the amount of data that can be transferred in this time interval without any additional CSI feedback from the Rx.

Fig. 11 demonstrates the advantage of the proposed characterization in terms effectively tracking the temporally-varying channel. In this figure we calculate error $\Delta^{\text{err}}(i) = P_{\text{act}}(i) - P_{\text{est}}(i)$, where $P_{\text{act}}(i)$ and $P_{\text{est}}(i)$ respectively denote the actual and estimated transmission power in the i^{th} slot. For simplicity, the following mapping is done: $\mathbb{P} = \{-20, -10, 0\}$ dBm $\rightarrow \{1, 2, 3\}$, to calculate Δ^{err} . $\Delta^{\text{err}} = 0$ implies correct choice of transmission power in the i^{th} slot. We observe from the figure that $\Delta^{\text{err}} \neq 0$ over $\sim 13.4\%$ of transmission slots in the proposed characterization, whereas it is $\sim 38.8\%$ in t_{ch} -based

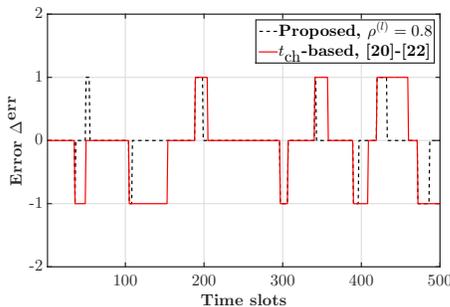


Fig. 11. Channel utilization. System parameters: $v = 20$ kmph, $T_p = 500$ μ s, $\varphi = 0$, $K = 4$, $\frac{P}{\sigma_o^2} = 10$ dB, $\mathbb{P} = \{-20, -10, 0\}$ dBm, thresholds = $\{8, 14\}$ dB, $\rho^{(l)} = 0.8$.

scheme. This justifies our claim, and also in a way it reaffirms the claim in [12].

Lastly, from Fig. 9 we observe that higher $\rho^{(l)}$ results in lower energy efficiency. At the same time, higher $\rho^{(l)}$ results in better channel utilization, as $\zeta^{(\max)}$ decreases with increasing $\rho^{(l)}$ (observed in Fig. 7(b)). Based on this trade-off, user can specify the acceptable $\rho^{(l)}$ and accordingly $\zeta^{(\max)}$ is obtained by using the proposed characterization. An analogous argument can also be presented with $\sigma_e^{(l)}$.

VII. CONCLUDING REMARKS

In this paper, we characterized the temporal variation of the W.S.S wireless channel based on its current state. In the proposed generalized model, the temporal variations of channel is modeled as a finite GM distribution. We validated the proposed model through analysis and extensive Monte-Carlo simulations. The time interval over which this characterization is done depends on the fading scenario as well as the application at hand. Accordingly, we proposed an upper bound on this time interval based on the fading scenario and the limit of acceptable error in estimation. We noted that there exists a non-trivial trade-off between this upper bound and accuracy of the proposed channel characterization.

On the basis of the proposed characterization, we introduced definitions for *short term outage probability* and *short term channel capacity* that accurately capture the short-term variations of the wireless channel. The convergence of these definitions with the existing classical definitions of outage probability and channel capacity was also shown. Results demonstrated that, when the proposed characterization is used in the context of channel adaptive communication, energy efficiency is as high as 2.94 times over its nearest approach.

We believe that the proposed channel characterization will provide a benchmark for characterizing the temporally varying wireless channel in context of URLLC. We also believe that it will be a yardstick for estimating the inter-feedback duration for the realization of large amount of data transfer in low power and energy-constrained scenarios like WSNs and IoT.

APPENDIX A

ERROR OF ESTIMATE BASED CHOICE OF $\zeta^{(\max)}$

From [51] we define *standard error of estimate* σ_e in terms of ζ as:

$$\sigma_e(\zeta) = \dot{\sigma}_{Z\zeta} \sqrt{1 - \rho^2(\zeta)}, \quad (\text{A-1})$$

where $\dot{\sigma}_{Z\zeta}^2 = \zeta \dot{\sigma}_{Z_1}^2$. The smaller the value of $\sigma_e(\zeta)$, the better is the estimate of ζ . Hence we can see that, for a given set of f_D , T_p , $\frac{P}{\sigma_o^2}$, and K , $\sigma_e(\zeta)$ is a function of the temporal spacing $\zeta \cdot T_p$ between the samples under consideration. It should be noted here that, like $\rho^{(l)}$, $\sigma_e^{(l)}$ is also an application specific parameter and it is given as input by the user, on the basis of which corresponding $\zeta^{(\max)}$ is calculated. Like in Section V-A1, here also we consider the domain of $[0, 2.4048]$ to deal with the oscillatory nature of $J_0(\cdot)$. We obtain $\zeta^{(\max)}$ in a way that $\sigma_e(\zeta^{(\max)})$ does not go above an acceptable threshold $\sigma_e^{(l)}$.

It can be seen from Fig. 7(a) that, $\rho(\zeta)$ decreases monotonically in the range $0 \leq \zeta \leq \frac{0.3827}{f_D T_p}$ for a given set of f_D and T_p . Further, for a given set of f_D , T_p , $\frac{P}{\sigma_o^2}$, and Rice factor K , $\dot{\sigma}_{Z\zeta}$ is an increasing function of ζ . Hence we conclude that for a given set of system parameters, $\sigma_e(\zeta)$ is a monotonically increasing function of ζ in $[0, \frac{0.3827}{f_D T_p}]$. Taking into account the monotonic nature of $\sigma_e(\zeta)$ in $[0, \frac{0.3827}{f_D T_p}]$, we obtain $\zeta^{(\max)}$ for a given $\sigma_e^{(l)}$ as:

$$(\text{P1}) : \zeta^{(\max)} = \left\{ \zeta \left| \underset{0 \leq \zeta \leq \frac{0.3827}{f_D T_p}}{\operatorname{argmin}} \left[\dot{\sigma}_{Z\zeta} \sqrt{1 - \rho^2(\zeta)} - \sigma_e^{(l)} \right]^2 \right. \right\}. \quad (\text{A-2})$$

Considering the nature of $\dot{\sigma}_{Z\zeta} \sqrt{1 - \rho^2(\zeta)}$ in $[0, \frac{0.3827}{f_D T_p}]$, we can observe that the objective function of (A-2) is unimodal in this region with the minimum being at $\dot{\sigma}_{Z\zeta} \sqrt{1 - \rho^2(\zeta)} = \sigma_e^{(l)}$. Accordingly we propose Algorithm 1 using Golden section based line search [52] to solve (A-2).

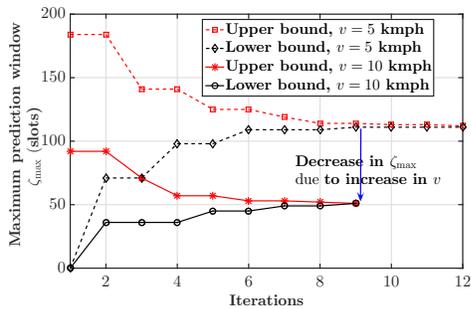
Algorithm 1 Algorithm to find $\zeta^{(\max)}$

Require: f_D , T_p , $\frac{P}{\sigma_o^2}$, Rician K factor, $\sigma_e^{(l)}$, and $\xi \geq 0$

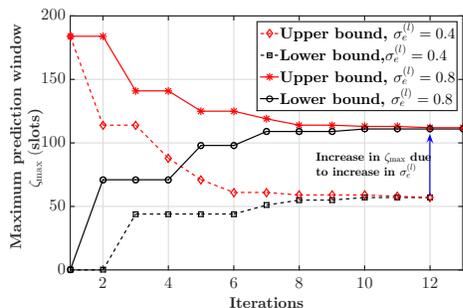
Ensure: $\zeta^{(\max)}$

- 1: Define $\zeta^{lb} = 0$ and $\zeta^{ub} = \left\lceil \frac{0.3827}{f_D T_p} \right\rceil$
 - 2: Define $r(\zeta) = \left[\dot{\sigma}_{Z\zeta} \sqrt{1 - \rho^2(\zeta)} - \sigma_e^{(l)} \right]^2$
 - 3: Set $a = 0$
 - 4: Calculate $\zeta_p = \zeta^{ub} - 0.618(\zeta^{ub} - \zeta^{lb})$
 - 5: Calculate $\zeta_q = \zeta^{lb} + 0.618(\zeta^{ub} - \zeta^{lb})$
 - 6: Calculate $r(\zeta_p)$ and $r(\zeta_q)$
 - 7: Set $\Delta_g = \zeta^{ub} - \zeta^{lb}$
 - 8: **while** $\Delta_g > \xi$ **do**
 - 9: **if** $r(\zeta_p) \leq r(\zeta_q)$ **then**
 - 10: | Set $\zeta^{ub} = \zeta_q$, $\zeta_q = \zeta_p$, and $\zeta_p = \zeta^{ub} - 0.618(\zeta^{ub} - \zeta^{lb})$
 - 11: **else**
 - 12: | Set $\zeta^{lb} = \zeta_p$, $\zeta_p = \zeta_q$, and $\zeta_q = \zeta^{lb} + 0.618(\zeta^{ub} - \zeta^{lb})$
 - 13: **end if**
 - 14: Set $a = a + 1$
 - 15: Calculate $r(\zeta_p)$ and $r(\zeta_q)$
 - 16: Set $\Delta_g = \zeta^{ub} - \zeta^{lb}$
 - 17: **end while**
 - 18: **if** $r(\zeta_p) < r(\zeta_q)$ **then**
 - 19: Set $\zeta^{(\max)} = \zeta_p$
 - 20: **else**
 - 21: Set $\zeta^{(\max)} = \zeta_q$
 - 22: **end if**
-

Convergence of Algorithm 1: The proposed algorithm estimates $\zeta^{(\max)}$ by reducing the search space $[\zeta^{lb}, \zeta^{ub}]$ in successive iterations. Due to the unimodal nature of the objective function of (P1), the reduced search space always contains the optimal ζ . Each iteration reduces the search



(a) Effect of v on the convergence of Algorithm 1. System parameters: Rice factor $K = 2$, $T_p = 500 \mu\text{s}$, $\frac{P}{\sigma_o^2} = 10 \text{ dB}$, $v = 5, 10 \text{ kmph}$, and $\sigma_e^{(l)} = 1.0$.



(b) Effect of $\sigma_e^{(l)}$ on the convergence of Algorithm 1. System parameters: Rice factor $K = 4$, $T_p = 500 \mu\text{s}$, $\frac{P}{\sigma_o^2} = 10 \text{ dB}$, $v = 5 \text{ kmph}$, and $\sigma_e^{(l)} = 0.4, 0.8$.

Fig. 12. Illustration of convergence of Algorithm 1.

space by a factor of 0.618. For golden section method it is known that the algorithm terminates after N iterations if $(\zeta^{ub} - \zeta^{lb}) 0.618^N \leq \xi$, where ξ is the tolerance level [52]. In other words, for a given set of ζ^{ub} , ζ^{lb} , and ξ ,

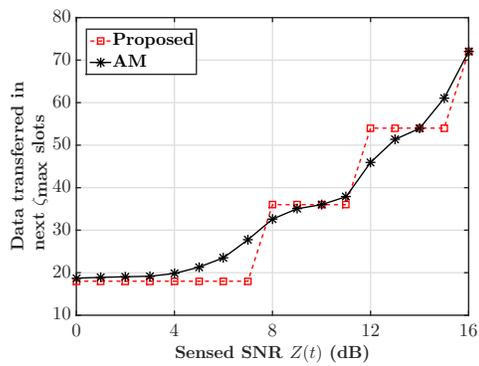
$$N \leq \frac{\ln\left(\frac{\xi}{\zeta^{ub} - \zeta^{lb}}\right)}{\ln(0.618)} \approx 2 \ln\left(\frac{\zeta^{ub} - \zeta^{lb}}{\xi}\right). \quad (\text{A-3})$$

Fig. 12 shows the convergence of Algorithm 1 with increasing iterations. It can be observed that the convergence is faster in the initial stages as the gap between ζ^{lb} and ζ^{ub} reduces very sharply in the first few iterations.

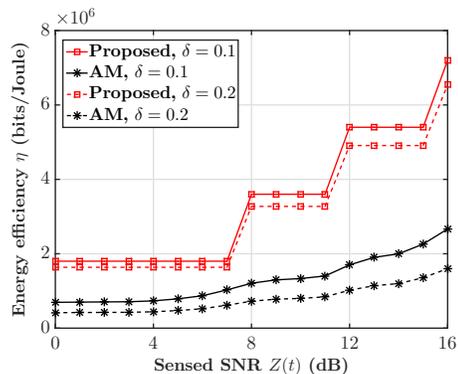
From Fig. 12(a) it can be seen that $\zeta^{(\max)}$ is higher for lower node velocity and vice-versa, which reaffirms the observation on Fig. 7(b). Fig. 12(b) shows the dependence of $\zeta^{(\max)}$ on $\sigma_e^{(l)}$; $\zeta^{(\max)}$ increases with increase in $\sigma_e^{(l)}$ and vice-versa. This is also intuitive, i.e., we can estimate the channel state for a large number of time slots if we increase the limit of acceptable error. This shows that there exists a trade-off between quality of estimation and $\zeta^{(\max)}$, which had also been observed in Fig. 7(b).

APPENDIX B RATE ADAPTATION

In adaptive modulation (AM), thresholds are chosen according to the desired bit error rate (BER). BER is given by $P_b = a_1 \exp\left(-\frac{a_2 Z(t)}{m^{a_3 - a_4}}\right)$, where a_1, \dots, a_4 are modulation specific real constants and m is the size of the chosen con-



(a) Variation of $D^{(\text{ST})}$ in next $\zeta^{(\max)}$ slots with $Z(t)$



(b) Variation of η with $Z(t)$

Fig. 13. Effect of currently sensed SNR on AM. System parameters: node velocity = 20 kmph, slot duration $T_p = 500 \mu\text{s}$, Rice factor $K = 4$, $P = 0 \text{ dBm}$, $P_{\text{pilot}} = 10 \text{ dBm}$, $\rho^{(l)} = 0.8$, and $P_b = 10^{-3}$.

stellation [35]. For example, for M-ary quadrature amplitude modulation (M-QAM), $a_1 = 2$, $a_2 = 1.5$, $a_3 = 1$, and $a_4 = 1$.

By *classical* AM, we imply that the modulation scheme m is changed in each time slot based on the regular CSI feedback. In contrast, the proposed adaptive modulation scheme receives feedback after every $\zeta^{(\max)}$ slots, based on minimum acceptable reliability limit $\rho^{(l)}$ or acceptable standard error of estimate threshold $\sigma_e^{(l)}$. Modulation scheme is chosen based on the received feedback and this chosen scheme is used for next consecutive $\zeta^{(\max)}$ slots, after which a feedback is again available at Tx from the Rx end.

Fig. 13 shows the effect of the proposed characterization on adaptive modulation, where we illustrate our claim over the following constellations: BPSK, 4-, 8-, 16-, and 32-QAM. Fig. 13(a) shows that the proposed characterization does not provide an appreciable gain in terms of data transferred in $\zeta^{(\max)}$ slots with respect to the classical AM. But its advantage can be observed from Fig. 13(b), which shows the energy efficiency to be as high as nearly 1.6 times above the energy efficiency obtained in AM irrespective of δ .

REFERENCES

- [1] "The power of 5G," Ericsson Mobility Report, Nov. 2018.
- [2] G. Y. Li, Z. Xu, C. Xiong, C. Yang, S. Zhang, Y. Chen, and S. Xu, "Energy-efficient wireless communications: Tutorial, survey, and open issues," *IEEE Wireless Commun. Mag.*, vol. 18, no. 6, pp. 28–35, Dec. 2011.
- [3] C.-L. I, C. Rowell, S. Han, Z. Xu, G. Li, and Z. Pan, "Toward green and soft: A 5G perspective," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 66–73, Feb. 2014.

- [4] P. Popovski, J. J. Nielsen, C. Stefanovic, E. d. Carvalho, E. Strom, K. F. Trillingsgaard, A. S. Bana, D. M. Kim, R. Kotaba, J. Park, and R. B. Sorensen, "Wireless access for ultra-reliable low-latency communication: Principles and building blocks," *IEEE Netw.*, vol. 32, no. 2, pp. 16–23, Mar. 2018.
- [5] A. Duel-Hallen, "Fading channel prediction for mobile radio adaptive transmission systems," *Proc. IEEE*, vol. 95, no. 12, pp. 2299–2313, Dec. 2007.
- [6] A. Duel-Hallen, S. Hu, and H. Hallen, "Long-range prediction of fading signals," *IEEE Signal Process. Mag.*, vol. 17, no. 3, pp. 62–75, May 2000.
- [7] T. Jia, A. Duel-Hallen, and H. Hallen, "Data-aided noise reduction for long-range fading prediction in adaptive modulation systems," *IEEE Trans. Veh. Technol.*, vol. 62, no. 5, pp. 2358–2362, June 2013.
- [8] M. Chen, T. Ekman, and M. Viberg, "New approaches for channel prediction based on sinusoidal modeling," *EURASIP J. Adv. Signal Process.*, vol. 2007, no. 1, pp. 1–13, 2006.
- [9] A. Heidari, A. K. Khandani, and D. Mcavoy, "Adaptive modelling and long-range prediction of mobile fading channels," *IET Commun.*, vol. 4, no. 1, pp. 39–50, Jan. 2010.
- [10] Z. Shen, J. G. Andrews, and B. L. Evans, "Short range wireless channel prediction using local information," in *Proc. IEEE ACSSC*, Pacific Grove, CA, USA, Nov. 2003, pp. 1147–1151.
- [11] R. Atawia, H. Abou-zeid, H. S. Hassanein, and A. Noureldin, "Joint chance-constrained predictive resource allocation for energy-efficient video streaming," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 5, pp. 1389–1404, May 2016.
- [12] V. N. Swamy, P. Rigge, G. Ranade, B. Nikolić, and A. Sahai, "Wireless channel dynamics and robustness for ultra-reliable low-latency communications," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 4, pp. 705–720, Apr. 2019.
- [13] P. Mukherjee, D. Mishra, and S. De, "Exploiting temporal correlation in wireless channel for energy-efficient communication," *IEEE Trans. Green Commun. and Netw.*, vol. 1, no. 4, pp. 381–394, Dec. 2017.
- [14] P. Mukherjee and S. De, "cDIP: Channel-aware dynamic window protocol for energy-efficient IoT communications," *IEEE Internet Things J.*, vol. 5, no. 6, pp. 4474–4485, Dec. 2018.
- [15] A. Chelli, E. Zedini, M. Alouini, M. Pätzold, and I. Balasingham, "Throughput and delay analysis of HARQ with code combining over double rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 67, no. 5, pp. 4233–4247, May 2018.
- [16] Z. Shi, S. Ma, G. Yang, and M. Alouini, "Energy-efficient optimization for HARQ schemes over time-correlated fading channels," *IEEE Trans. Veh. Technol.*, vol. 67, no. 6, pp. 4939–4953, June 2018.
- [17] A. Mukherjee, "Energy efficiency and delay in 5G ultra-reliable low-latency communications system architectures," *IEEE Netw.*, vol. 32, no. 2, pp. 55–61, Mar. 2018.
- [18] L. Tong and Q. Zhao, "Joint order detection and blind channel estimation by least squares smoothing," *IEEE Trans. Signal Process.*, vol. 47, no. 9, pp. 2345–2355, Sep. 1999.
- [19] S. Han, S. Ahn, E. Oh, and D. Hong, "Effect of channel-estimation error on BER performance in cooperative transmission," *IEEE Trans. Veh. Technol.*, vol. 58, no. 4, pp. 2083–2088, May 2009.
- [20] Chen Xu, Leiming Zhang, Lingyang Song, Yuping Zhao, and Bingli Jiao, "Feedback compression for time-correlated MIMO block-fading channels using Huffman coding," in *Proc. IEEE Int. Conf. Commun. Tech.*, Jinan, China, Sept. 2011, pp. 411–415.
- [21] B. Makki and T. Eriksson, "Feedback subsampling in temporally-correlated slowly-fading channels using quantized CSI," *IEEE Trans. Commun.*, vol. 61, no. 6, pp. 2282–2294, June 2013.
- [22] P. N. Alevizos, X. Fu, N. D. Sidiropoulos, Y. Yang, and A. Bletsas, "Limited feedback channel estimation in massive mimo with non-uniform directional dictionaries," *IEEE Trans. Signal Process.*, vol. 66, no. 19, pp. 5127–5141, Oct. 2018.
- [23] C. Wen, S. Jin, K. Wong, J. Chen, and P. Ting, "Channel estimation for massive MIMO using gaussian-mixture bayesian learning," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1356–1368, Mar. 2015.
- [24] B. Selim, O. Alhussein, S. Muhaidat, G. K. Karagiannidis, and J. Liang, "Modeling and analysis of wireless channels via the mixture of gaussian distribution," *IEEE Trans. Veh. Technol.*, vol. 65, no. 10, pp. 8309–8321, Oct. 2016.
- [25] J. Mo, P. Schniter, and R. W. Heath, "Channel estimation in broadband millimeter wave MIMO systems with few-bit ADCs," *IEEE Trans. Signal Process.*, vol. 66, no. 5, pp. 1141–1154, Mar. 2018.
- [26] C. Snow, L. Lampe, and R. Schober, "Error rate analysis for coded multicarrier systems over quasi-static fading channels," *IEEE Trans. Commun.*, vol. 55, no. 9, pp. 1736–1746, Sep. 2007.
- [27] B. Makki and T. Eriksson, "On hybrid ARQ and quantized CSI feedback schemes in quasi-static fading channels," *IEEE Trans. Commun.*, vol. 60, no. 4, pp. 986–997, Apr. 2012.
- [28] H. Holma and A. Toskala, *LTE for UMTS: OFDMA and SC-FDMA based radio access*. Wiley, 2009.
- [29] C. Tepedelenioglu, A. Abdi, G. B. Giannakis, and M. Kaveh, "Estimation of doppler spread and signal strength in mobile communications with applications to handoff and adaptive transmission," *Wireless Commun. Mob. Comput.*, vol. 1, no. 2, pp. 221–242, 2001.
- [30] M. Zorzi, R. R. Rao, and L. B. Milstein, "ARQ error control for fading mobile radio channels," *IEEE Trans. Veh. Technol.*, vol. 46, no. 2, pp. 445–455, May 1997.
- [31] C. Xiao, Y. R. Zheng, and N. C. Beaulieu, "Novel sum-of-sinusoids simulation models for Rayleigh and Rician fading channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3667–3679, Dec. 2006.
- [32] M. K. Simon and M. Alouini, "A unified approach to the performance analysis of digital communication over generalized fading channels," *Proc. IEEE*, vol. 86, no. 9, pp. 1860–1877, Sep. 1998.
- [33] C. Tepedelenioglu, A. Abdi, and G. B. Giannakis, "The Ricean K factor: estimation and performance analysis," *IEEE Trans. Wireless Commun.*, vol. 2, no. 4, pp. 799–810, July 2003.
- [34] S. L. Cotton, "Second-order statistics of $\kappa - \mu$ shadowed fading channels," *IEEE Trans. Veh. Technol.*, vol. 65, no. 10, pp. 8715–8720, Oct. 2016.
- [35] A. Goldsmith, *Wireless communications*. Cambridge University Press, 2005.
- [36] M. D. Yacoub, J. E. V. Bautista, and L. G. de Rezende Guedes, "On higher order statistics of the Nakagami-m distribution," *IEEE Trans. Veh. Technol.*, vol. 48, no. 3, pp. 790–794, May 1999.
- [37] K. T. Wong, Y. I. Wu, and M. Abdulla, "Landmobile radiowave multipaths' DOA-distribution: Assessing geometric models by the open literature's empirical datasets," *IEEE Trans. Antennas Propag.*, vol. 58, no. 3, pp. 946–958, Mar. 2010.
- [38] C. Ziolkowski and J. M. Kelner, "Empirical models of the azimuthal reception angle—part I: Comparative analysis of empirical models for different propagation environments," *Wireless Pers. Commun.*, vol. 91, no. 2, pp. 771–791, Nov. 2016.
- [39] S. J. Nawaz, A. Ahmed, S. Wyne, K. Cumanan, and Z. Ding, "A tunable 3-D statistical channel model for spatio-temporal characteristics of wireless communication networks," *Trans. Emerg. Telecommun. Technol.*, vol. 28, no. 12, p. e3213, Dec. 2017.
- [40] —, "3-D spatial modeling of network interference in two-tier heterogeneous networks," *IEEE Access*, vol. 5, pp. 24 040–24 053, 2017.
- [41] C. Ziolkowski and J. M. Kelner, "Statistical evaluation of the azimuth and elevation angles seen at the output of the receiving antenna," *IEEE Trans. Antennas Propag.*, vol. 66, no. 4, pp. 2165–2169, Apr. 2018.
- [42] H. Sifaou, A. Kammoun, L. Sanguinetti, M. Debbah, and M. Alouini, "Max-Min SINR in large-scale single-cell MU-MIMO: Asymptotic analysis and low-complexity transceivers," *IEEE Trans. Signal Process.*, vol. 65, no. 7, pp. 1841–1854, Apr. 2017.
- [43] B. S. Everitt, *Finite Mixture Distributions*. Wiley Online Library, 1981.
- [44] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Royal Statistical Soc.*, vol. 39 (Series B), pp. 1–38, 1977.
- [45] A. Borhani and M. Pätzold, "Correlation and spectral properties of vehicle-to-vehicle channels in the presence of moving scatterers," *IEEE Trans. Veh. Technol.*, vol. 62, no. 9, pp. 4228–4239, Nov 2013.
- [46] D. B. da Costa and S. Aissa, "Capacity analysis of cooperative systems with relay selection in Nakagami-m fading," *IEEE Commun. Lett.*, vol. 13, no. 9, pp. 637–639, Sep. 2009.
- [47] M. Zorzi and R. R. Rao, "Error control and energy consumption in communications for nomadic computing," *IEEE Trans. Comput.*, vol. 46, no. 3, pp. 279–289, Mar. 1997.
- [48] R. K. Mungara, G. George, and A. Lozano, "Overhead and spectral efficiency of pilot-assisted interference alignment in time-selective fading channels," *IEEE Trans. Wireless Commun.*, vol. 13, no. 9, pp. 4884–4895, Sep. 2014.
- [49] R. Yan, H. Sun, and Y. Qian, "Energy-aware sensor node design with its application in wireless sensor networks," *IEEE Trans. Instrum. Meas.*, vol. 62, no. 5, pp. 1183–1191, May 2013.
- [50] T. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. Upper Saddle River, NJ, USA: Prentice Hall PTR, 2001.
- [51] B. M. Ayyub and R. H. McCuen, *Probability, Statistics, and Reliability for Engineers and Scientists*. CRC press, 2011.
- [52] A. D. Belegundu and T. R. Chandrupatla, *Optimization Concepts and Applications in Engineering*. Cambridge University Press, 2011.