On Spatial and Temporal Variations in Ultra Dense Wireless Networks

Pranav Madadi, François Baccelli and Gustavo de Veciana
Department of Electrical and Computer Engineering, The University of Texas at Austin.

Abstract—Ultra densification along with the use of wider bands at higher frequencies are likely to be key elements towards meeting the throughput/coverage objectives of 5G wireless networks. In addition to increased parallelism, densification leads to improved, but eventually bounded, benefits from proximity of users to base stations, while resulting in increased aggregate interference. Such networks are expected to be interference limited, and in higher frequency regimes, the interference is expected to become spatially variable due to the increased sensitivity of propagation to obstructions and the proximity of active interferers. This paper studies the characteristics of the spatial random fields associated with interference and Shannon capacity in ultra-dense limiting regimes. They rely on the theory of Gaussian random fields which arise as natural limits under densification. Our models show how densification and operation at higher frequencies, could lead to increasingly rough temporal variations in the interference process. This is characterized by the Hölder exponent of the interference field. We show that these fluctuations make it more difficult for mobile users to adapt modulation and coding. We further study how the spatial correlations in users’ rates impact backhaul dimensioning. Therefore, this paper identifies and quantifies challenges associated with densification in terms of the resulting unpredictability and the correlation of interference on the achievable rates.

I. INTRODUCTION

Traffic from mobile devices has significantly increased over the last decade mainly due to growth in the number of smart wireless devices and bandwidth-demanding applications. It is expected to increase 1000 fold in the next decade, and fifth generation (5G) wireless networks should be able to support this growth. While millimeter wave and massive MIMO technologies have been proposed to improve spectral efficiency and exploit wider bandwidths, it is expected that Ultra Dense Networks (UDNs) will be a key element towards boosting capacity and enhancing coverage with low cost and a power-efficient infrastructure [1].

With the increases in density, the distance between the users and their associated base stations reduces, leading to an increase in the signal strength, but there is also an increase in interference. Together this effectively reduces the impact of thermal noise. Thus, ultra dense networks should be interference limited networks 1. Further, for these short range inter-site distances, the unbounded path loss models often used in the literature are clearly no longer physically relevant as they are singular at the origin, see for e.g., [2]. More realistic and practical models would be based on bounded path loss functions. The focus of this paper will be on studying scaling limits for ultra dense networks for such more realistic models.

In particular, we develop a spatial characterization of the interference and the Shannon rate fields, which in turn helps to better understand the fundamental characteristics underlying densification of wireless systems. For example, a major issue underlying densification is the provisioning of shared backhaul resources. One of the interesting questions is to study how the backhaul capacity requirement per base station scales with densification. Another major issue is to understand the rate variability that mobile users would see in dense networks. Assuming such users are able to immediately connect to the closest or best nearby base station, their signal path should be good, yet they will still experience variability in the interference. It is essential to study the variations in interference over a time period, since such variability can be a challenge to techniques such as adaptive modulation and coding which rely on the predictability of SINR over time periods.

The aim in this paper is to study the characteristics of spatial fields associated with ultra dense wireless networks, and link them to such basic engineering questions.

Related Work. Cellular network performance has been extensively studied by modeling the network using stochastic geometry [3], [4]. With the help of scaling limits for the interference, network performance has been evaluated under densification for various modeling assumptions ([5], [6], [7], [8], [9], [10]). Coverage probability and area spectral efficiency analysis have mostly been used as the main performance metrics. The findings are sometimes conflicting and suggest that densification may eventually stop delivering significant throughput gains.

Most prior work focuses on either studying the scaling limits of the SINR at a typical location, or on two-point correlations in interference and shadowing. [11], [12], [13]. Although, [14] studies the scaling limit of the interference field with singular power law path loss and Rayleigh fading, to our knowledge, a spatial characterization of the limiting interference field for bounded path loss models is lacking.

There is a broad body of relevant work in the field of mathematics of shot noise and Gaussian fields. These are of interest to this class of problems since the interference fields in large wireless systems can be viewed as shot noise fields, where the path loss function is equivalent to the kernel function of the shot noise field. Further, the infinite divisibility property of Poisson point processes allows one to establish convergence.
of the shot noise field to a Gaussian field as the intensity increases ([15], [16]). Precise sample path properties, especially level crossings of shot noise fields have been extensively studied in, e.g., [17], [18], [19], [20]. General results are known concerning the level crossings of smooth Gaussian processes ([17], [21]). However, with closest base station association, the interference is not a shot noise field but a protected shot noise field. Thus, most of the results in the literature are not directly applicable.

Given the importance of backhauling for 5G small cell networks, many researchers have studied centralized and distributed architectures for the backhauling gateways, see, e.g., [22]. Simulation results suggest that a distributed wireless backhaul network architecture is more suitable for future 5G networks employing massive MIMO and/or millimeter wave communication technologies. Millimeter wave communication has been considered as the wireless backhaul solution for small cell networks in 5G communication systems. However, most studies on millimeter wave backhaul technologies focus on the design of the antenna array and radio frequency (RF) components of transceivers, such as beamforming and modulation schemes [23], [24]. To our knowledge, a system level investigation of ultra-dense cellular networks backhaul requirements such as that in this paper is novel.

Adaptive modulation and coding is a critical technique for adapting to time varying channels resulting from, say fading, path loss and interference, see, e.g., [25], [26] and the references therein. Most of the work focuses on channel quality estimation or various adaptive modulation techniques. In this paper, we characterize the time periods over which we can predict SINR with some success and study their dependence on various system parameters.

Contributions. In this paper, we study basic models allowing us to to glean some of the salient characteristics of densification that are yet to be explored. To begin with, we establish a scaling limit for the interference field under bounded, non-negative and integrable path loss functions. This is then used to approximate the interference field in dense networks by a stationary Gaussian field which captures the underlying spatial variations.

Since the interference field primarily depends on the distribution of the base station locations and the path loss, we classify various existing path loss models in the literature such as the dual slope models ([27], [28], [29]) and observe the impact they have on the various sample path properties of the limiting field, e.g., continuity and differentiability.

By transforming or taking functionals of the interference field, one can study various additional properties of such systems. In particular, with an appropriate bandwidth scaling, we can also model the Shannon rate field as a stationary Gaussian field. Further, relevant functions of the Shannon rate field enable the study of the variability in the spatial average rate, and backhaul capacity requirements for ultra dense wireless networks.

We then study how the spatial variability of the interference and Shannon rate fields enables a characterization of the temporal landscape a mobile user would see. For certain path loss models, we show the interference process is nowhere differentiable, which implies that mobiles experience high fluctuations in Shannon rate, which could make it difficult to implement adaptive modulation and coding techniques. To better understand these fluctuations, we quantify their variation via Hölder exponents and leverage these to bound the time scales over which adaptive modulation and coding could be performed. Finally we provide a characterization of the level crossing characteristics of, e.g., interference, and of rate.

Paper Organization: The paper is organized as follows. We describe the system model and the classification of path loss models in Section II. The convergence of the scaled interference field to a stationary Gaussian field and its sample path properties are gathered in Section III. The characterization of the Shannon rate field and various functions of the field, along with the problem of backhaul capacity dimensioning are discussed in Section IV. In Section V, we focus on the temporal characterization of the rate seen by a mobile user, where we quantify the variations in the rate and study their impact in the context of adaptive modulation coding. We then study the mean level crossings of the process.

II. System Model

Consider a cellular network where the base stations are distributed according to a homogeneous Poisson Point Process (PPP) \( \Phi = \{X_1, X_2, \ldots\} \) in \( \mathbb{R}^2 \) of intensity \( \lambda_b \). Let \( l: \mathbb{R}^2 \rightarrow \mathbb{R}^+ \) be a deterministic non-negative function. Consider downlink transmissions and assume all base stations transmit at a fixed power \( p \). Then, the total power received from all base stations at a location \( y \) is referred to as a shot noise field [30], given by

\[
J_{\lambda_b}(y) = \sum_{X_i \in \Phi \setminus X_{\lambda_b}(y)} p l(X_i - y),
\]

where \( X_{\lambda_b}(y) \) denotes the closest base station in \( \Phi \) to location \( y \). Given the interference field, one can study the SINR field and Shannon rate field experienced by a user at a given location. In the absence of fading and shadowing, the SIR field (\( \text{SIR}_{\lambda_b}(y), y \in \mathbb{R}^2 \)) is given by

\[
\text{SIR}_{\lambda_b}(y) = \frac{p \ast l(y - X_{\lambda_b}(y))}{J_{\lambda_b}(y)}.
\]

In the interference limited regime, the Shannon rate field, \( (S_{\lambda_b}(y), y \in \mathbb{R}^2) \) can be defined from the SIR field through

\[
S_{\lambda_b}(y) = w \log(1 + \text{SIR}_{\lambda_b}(y)),
\]

where, \( w \) is the wireless system bandwidth.

In the sequel we will also consider a tagged user moving at a fixed unit velocity along a straight line starting from the origin at time \( t = 0 \). We will denote the interference and the Shannon rate experienced by the mobile user by the stochastic processes \( (J_{\lambda_b}(t), t > 0) \) and \( (S_{\lambda_b}(t), t > 0) \), respectively.
Note that $J_{\lambda b}$ and $I_{\lambda b}$ correspond to spatial processes or fields while $J_{\lambda b}$ and $I_{\lambda b}$ are temporal processes. Table I summarizes the notation used in the paper.

With densification and operation at higher frequencies, the nature of the path loss changes. We shall consider path loss functions that belong to a specific functional space. Recall that the space of continuous functions is in $D_1$. Similarly, if $T$ is the unit square $[0,1]^2$, then the functional space $D_2$ is the uniform closure, in the space of all bounded functions from $T$ to $\mathbb{R}$, of the vector subspace of simple functions which are coordinate wise $D_1$ [31]. The reason for this choice is that weak convergence of stochastic processes can be studied in this functional space under the S-topology [31]. We further classify the path loss functions in the $D_2$ functional space as follows:

1) Class-1 Functions : functions that are smooth, integrable and once differentiable i.e., $l \in C^2(\mathbb{R}^2)$, $l \in L^2(\mathbb{R}^2)$, $l', l'' \in L^1(\mathbb{R}^2)$. The stretched exponential path loss function, $l_1(y) = e^{-a_1 |y|^s}$, [32] for a constant $a_1$, is an example of Class-1 function.

2) Class-2 Functions : functions that are continuous, piece-wise $C^2$ on $\mathbb{R}^2$, with discontinuities in their first derivative. Multi-slope path loss functions are examples of Class-2 [7]:

$$l_2(y) = \begin{cases} 
1 & \text{for } ||y|| \leq r_0, \\
1 - a_1/||y||^2 & \text{for } r_0 < ||y|| \leq r_1, \\
a_2/||y||^2 & \text{for } ||y|| > r_1.
\end{cases}$$

where, $a_1$ and $a_2$ are constants such that the function is continuous.

3) Class-3 Functions : functions with discontinuities in $D_2$. Out-of-sight path loss models studied in [33], where there is a sudden drop in the power due to blockages in urban/sub-urban areas with buildings provide examples for this class of functions. An example of such a path loss function is:

$$l_3(y) = \begin{cases} 
1 & \text{for } -r_0 \leq y_1 < r_0, -r_0 \leq y_2 < r_0, \\
0 & \text{for otherwise}.
\end{cases}$$

A discussion of the approximation of radial discontinuous functions by functions that belong to the $D_2$ class of functions is given in the Appendix A of [34].

III. SCALING LIMIT OF THE INTERFERENCE FIELD

Under our modeling assumptions we have $\mathbb{E}[I_{\lambda b}(0)] = \lambda_b \int_{\mathbb{R}^2} l(y)dy < \infty$, so we can consider the following re-scaling of the interference field:

$$I_{\lambda b}^c(y) = \frac{1}{\sqrt{\lambda_b}}(I_{\lambda b}(y) - \mathbb{E}[I_{\lambda b}(0)]).$$

It is well known that as $\lambda_b \to \infty$, the scaled field $I_{\lambda b}^c$ converges to a stationary Gaussian random field $\hat{I}^c$. Given two locations $z_1, z_2 \in \mathbb{R}^2$, the covariance kernel $c(z_1, z_2)$ depends only on the Euclidean distance $t = ||z_1 - z_2||$ and is given by:

$$c(t) = \mathbb{E}[\hat{I}^c(z_1)\hat{I}^c(z_2)] = \int_{\mathbb{R}^2} p^2 l(y) - z_1 l(y - z_2)dy. \quad (7)$$

**Theorem 1.** For bounded path loss functions in the $D_2$ functional space as defined in [31], consider the re-scaling of the interference field, $J_{\lambda b}$ given by:

$$J_{\lambda b}^c(y) = \frac{1}{\sqrt{\lambda_b}}(J_{\lambda b}(y) - \mathbb{E}[J_{\lambda b}(0)]). \quad (8)$$

Then, as $\lambda_b \to \infty$, $J_{\lambda b}^c$ converges weakly to a stationary Gaussian random field, $\hat{J}^c$. Further, in the limit the expectation of the interference field scales with $\lambda_b$ as

$$\mathbb{E}[J_{\lambda b}(0)] = \lambda_b \kappa, \quad (9)$$

where

$$\kappa = p \int_{\mathbb{R}^2} l(y)dy. \quad (10)$$

In particular, we have convergence in the sense of finite dimensional distribution, i.e., for $z_1, z_2, ..., z_n \in \mathbb{R}^2$, the random vector $(J_{\lambda b}^c(z_1), J_{\lambda b}^c(z_2), ..., J_{\lambda b}^c(z_n))$ converges to a centered Gaussian vector $(\hat{J}^c(z_1), J^c(z_2), ..., J^c(z_n))$ for any $n \geq 1$ with covariance kernel, $c(t)$ given by (7). In addition, the tightness condition as given in [31] holds. Proof is given in Appendix B of [34].

We would like to characterize the interference and the Shannon rate fields. For this we use the above scaling to approximate the interference field. Given the central limit Gaussian field $(\hat{J}^c(y), y \in \mathbb{R}^2)$, for a large values of $\lambda_b$, from (8), the interference field at location $y$, $J_{\lambda b}(y)$, can be approximated as follows:

$$J_{\lambda b}(y) \sim \sqrt{\lambda_b} \hat{J}^c(y) + \mathbb{E}[J_{\lambda b}(0)] + o(\sqrt{\lambda_b}). \quad (11)$$

Since the expectation of the interference scales linearly with $\lambda_b$ as seen in Theorem 1, the Gaussian approximation for the interference field $\hat{J}_{\lambda b} = (\hat{J}_{\lambda b}(y), y \in \mathbb{R}^2)$ is given by:

$$\hat{J}_{\lambda b}(y) = \sqrt{\lambda_b} \hat{J}^c(y) + \lambda_b \kappa. \quad (12)$$

In Appendix G of [34], we illustrate the convergence of the interference field considering a dual-slope path-loss model of Class-2 (4) and for $\lambda_b = 10^4$ and use the Kolmogorov-Smirnov test (K-S test) to compare the marginal empirical CDF with the Gaussian cumulative distribution.

Now, we focus on certain fundamental questions about the Gaussian field $(\hat{J}^c(y), y \in \mathbb{R}^2)$, such as its continuity and differentiability. Recall that for a Gaussian field, these are determined by the mean and covariance kernel given in (7). We leverage some well known results regarding the continuity and differentiability of Gaussian fields as stated in the lemmas given in Appendix C of [34]. The following theorem states the result for the path loss functions of Class-1.
Table 1

<table>
<thead>
<tr>
<th>System Parameters</th>
<th>Model</th>
<th>Gaussian Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \lambda_b )</td>
<td>( J_{\lambda_b} ); ( J_{\lambda_b}^c ) (re-scaled)</td>
<td>( \hat{J}<em>{\lambda_b} ); ( \hat{J}</em>{\lambda_b}^c ) (re-scaled)</td>
</tr>
<tr>
<td>( \rho ) Transmit power</td>
<td>( S_{\lambda_b} )</td>
<td>( \hat{S}_{\lambda_b} )</td>
</tr>
<tr>
<td>( w ) System bandwidth</td>
<td>( l(y) ) Path loss</td>
<td></td>
</tr>
</tbody>
</table>

### Theorem 2

For Class-1 radial path loss functions, if for some \( \eta > 0 \) and all \( t \in [-\eta, \eta] \), the following integral in polar coordinates

\[
\int_0^{2\pi} \int_0^{\infty} l(t)l'(\sqrt{r^2 + t^2 - 2rt\cos(\theta)})
\times \left( \frac{t - r\cos(\theta)}{\sqrt{r^2 + t^2 - 2rt\cos(\theta)}} \right) rdrd\theta
\]

is uniformly convergent, and in addition the covariance kernel \( c(t) \), defined in (7), is convergent, then, \( c(t) \) is continuous and twice differentiable. Thus, the limiting Gaussian field, \( (\hat{J}^c(y), y \in \mathbb{R}^2) \) is mean square and almost surely continuous and differentiable.

The proof is given in Appendix D of [34].
Now based on our characterization of the interference limit as in (12) we can approximate the Shannon rate field as:

\[ S_{\lambda_b}(y) \sim \frac{wp}{\sqrt{\lambda_b J^c(y)}} + \lambda_b \kappa, \]

where \( \kappa \) is defined in (10).

Using a Taylor series expansion, one obtains

\[ S_{\lambda_b}(y) \sim \frac{wp}{\lambda_b} - \frac{wp\sqrt{\lambda_b J^c(y)}}{\kappa^2 \lambda_b^2} + o(1/\lambda_b^{3/2}). \]  

(16)

Thus to compensate for the increase in interference, the Shannon rate must be scaled by a system bandwidth \( w \) scaling linearly in \( \lambda_b \). We let \( w = a \lambda_b \) and define our Gaussian model for the Shannon rate field as follows \( \hat{S}_{\lambda_b} = (\hat{S}_{\lambda_b}(y), y \in \mathbb{R}^2) \) where

\[ \hat{S}_{\lambda_b}(y) = \frac{ap}{\kappa} - \frac{ap\sqrt{\lambda_b J^c(y)}}{\kappa^2 \lambda_b^2}, \]  

(17)

with mean the covariance kernels given by:

\[ E[\hat{S}_{\lambda_b}(y)] = \mu = \frac{ap}{\kappa}, \]

\[ \text{Cov}(\hat{S}_{\lambda_b}(y), \hat{S}_{\lambda_b}(x)) = \frac{\mu^2}{\lambda_b \kappa^2} c(||y - x||), \]

where \( c(.) \) is given in (7).

Note that the variability of the field decreases with the intensity of base stations \( \lambda_b \). See Appendix G of [34] for the plot exhibiting the marginal empirical CDF of the Shannon rate field along with that of the Gaussian approximation.

**B. Variability of the Spatial Average Rate (SAR)**

Given a Gaussian field such as \( \hat{S}_{\lambda_b} = (\hat{S}_{\lambda_b}(y), y \in \mathbb{R}^2) \), one can now consider various relevant functions of the spatial process. For example, below we define the spatial average rate over a fixed region.

**Definition 1.** The spatial average of the Shannon rate field \( \hat{S}_{\lambda_b} \) over a set \( A \subset \mathbb{R}^2 \) is defined by

\[ X_{\lambda_b}(A) = \frac{1}{|A|} \int_A \hat{S}_{\lambda_b}(y) dy, \]

where \( |A| \) denotes the area of \( A \).

It follows immediately from the properties of Gaussian processes that \( X_{\lambda_b}(A) \) is Gaussian such that

\[ E[X_{\lambda_b}(A)] = \mu, \]

\[ \text{Var}(X_{\lambda_b}(A)) = \frac{\mu^2}{\kappa^2 \lambda_b |A|^2} \int_A \int_A c(||y - z||) dy dz, \]

where \( \kappa = \int_{\mathbb{R}^2} l(y) dy \).

For a fixed region \( A \) one might ask how densification will impact variability in the spatial average rate. Our analysis suggests the standard deviation is inversely proportional to \( \sqrt{\lambda_b} \), i.e., leads to concentration in the rates users will see. Fig 3 illustrates such decreases in variability for Class 2 and 3 path loss models. Perhaps as expected, scenarios with more discontinuous path loss characteristics see higher variability but still similar decays.

**C. Backhaul Capacity Dimensioning**

The cost and provisioning of backhauling resources is one of the key issues associated with deploying dense networks. In this subsection we shall study how densification impacts the cost of backhauling, leveraging again functionals of our Gaussian model for the Shannon rate field.

We consider a simple backhauling infrastructure based on a grid tessellation, where each cell (square) is associated with a gateway which provides backhauling for the users in its cell. We assume that the spatial density of users the network serves grows along with the density of base stations. We consider this setting as one of the aims of densification is to provide individual users high throughput which is achieved by cell-splitting gain. Although more general models could be considered, we shall assume that backhauling technology is such that each gateway can handle roughly a fixed number of users say \( m \). Hence, as we density, the cells’ area \( |A_{\lambda_b}| \) decreases inversely proportional to the base station density i.e., \( |A_{\lambda_b}| = \frac{\lambda}{2 \lambda_b} \). We assume that neither the link from the base station to the gateway nor the backhaul to the Internet is a bottleneck. That is, base stations are connected to the gateway via high capacity links, e.g., mmWave, links. Here the key question is about the capacity that should be provisioned from the gateway to the Internet.

There are various sources of variability which impact the provisioning of the gateway to the Internet backhaul capacity: (1) variability in the peak rate of users, (2) correlations amongst users’ peak rates, (3) variability in the number of active users, and (4) sharing of base station resources by one or more users (which limits their peak rate).

Let us first ignore the impact of variability in the number of users. To that end, we shall consider user locations \( \Phi_g \) corresponding to a grid with density \( \lambda_u = \lambda_b \). This scenario might correspond to a deterministic deployment e.g., a video surveillance system with a fixed set of active users. Further, ignoring the sharing of base station resources, we model the aggregate peak rate requirement at a typical gateway cell for a base station density \( \lambda_b \) as:

\[ R_{\lambda_b}(\Phi_g) = \sum_{j \in \Phi_g \cap A_{\lambda_b}} \hat{S}_{\lambda_b}(Y_j). \]
Lemma 1. The gateway capacity $\rho(\delta)$ capable of serving the aggregate peak rate $R_{\lambda_b}(\Phi_g)$ with overflow probability $\delta$ is given by

$$\rho(\delta) = \arg\min_\rho\{\rho \mid P(R_{\lambda_b}(\Phi_u) \geq \rho) \leq \delta\}, \quad (18)$$

where $\rho(\delta) = \bar{\mu} + Q^{-1}(\delta)\tilde{\sigma}$, $Q^{-1}(x) = \sqrt{2}erf^{-1}(1-2x)$, $\bar{\mu} = E[R_{\lambda_b}(\Phi_g)] = \mu m$, and $\tilde{\sigma}^2 = Var(R_{\lambda_b}(\Phi_g))$ which is given by

$$\frac{\mu^2 m}{\kappa^2 \lambda_b} c(0) + \frac{\mu^2 m}{\kappa^2 \lambda_b} \sum_{i,j=1, i \neq j}^m c(||g_{\lambda_b}^{(i)} - g_{\lambda_b}^{(j)}||),$$

with $g_{\lambda_b}^{(j)}$, $j = 1, \ldots, m$ corresponding to the $m$ grid points in the gateway’s square cell.

Proof. In our grid model, the aggregate peak rate, $R_{\lambda_b}(\Phi_g)$, is the sum of $m$ jointly Gaussian (positively correlated) random variables associated with the grid locations $\Phi_g$ in the gateway cell. Thus, $R_{\lambda_b}(\Phi_u)$ is Gaussian $(\bar{\mu}, \tilde{\sigma}^2)$ and the result in the Lemma follows.

Note that the first term in the variance captures the spatial variations in the users’ peak rate while the second term captures correlations amongst the users peak rates. See the Appendix F of [34] for a derivation of the above variance formula. \[\blacksquare\]

Remark 1. (Variability in number of users.) We shall consider the scenario where users’ locations $\Phi_r$ correspond to a Poisson process with intensity $\lambda_u = \lambda_b$, which model variations in both the number and locations of active users. Let $R_{\lambda_b}(\Phi_r)$ correspond to a random sum of random variables corresponding to a Poisson distributed number of users in the gateway cell, having Gaussian peak rates which are correlated. Such a mixture of Gaussian random variables is no longer Gaussian but is reasonably well approximated. So we propose to approximate $R_{\lambda_b}(\Phi_r)$ as a Gaussian Random variable with the same mean as $R_{\lambda_b}(\Phi_g)$ and with

$$Var(R_{\lambda_b}(\Phi_r)) = \mu^2 m + \frac{\mu^2 m}{\kappa^2 \lambda_b} c(0) + \frac{\mu^2 \lambda_b}{\kappa^2} \int_{A_{\lambda_b}} \int_{A_{\lambda_b}} c(||y - z||)dydz.$$ 

See the Appendix F of [34] for a derivation of the above variance formula as well as an example of validation of this approximation. The first term in the above variance captures variability in the number of users, the second is associated with variability in user’s peak rate while the third again captures correlations amongst the users peak rates.

In addition to determining the required backhaul capacity for the above two scenarios, Grid and Random, we can also determine the capacity one would provision if one ignored the terms in their variance corresponding to positive spatial correlations in users peak rates. We refer to the latter as Grid-simple and Random-simple.

We evaluated the required backhaul capacity according to Lemma 1, for $\delta = 0.01, \lambda_b |A_{\lambda_b}| = m = 20$ and dual slope path loss with $r_0 = 100m$ for all the above mentioned scenarios and Fig 4 exhibits the comparison of the required backhaul capacity with increasing density, $\lambda_b$. We have the following observations:

- Since a single gateway serves approximately a fixed number of base stations, $m$, the capacity can be viewed as the required backhaul capacity per unit base station. Thus, the cost of providing backhaul capacity decreases with densification due to decrease in the variance of the Shannon rate.
- As seen in Fig.4, the positive rate correlations impact the capacity requirements differently for random and grid users. The relative increase in the capacity ranges from 22% - 19.2% for grid users and 8.1% - 4.3% for random users. The user variability dominates the variability due to correlations. Thus the relative increase in the required capacity is higher for grid users.
- The positive rate correlations also impact differently for various path loss models and Table II states the values of the relative increase in the required backhaul capacity for various classes of path loss functions.
- The required capacity is higher for random users due to the additional contribution of user variability.

<table>
<thead>
<tr>
<th>$\lambda_b$</th>
<th>Class-1</th>
<th>Class-2</th>
<th>Class-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 $\Phi_g$</td>
<td>8.9%</td>
<td>22%</td>
<td>19.2%</td>
</tr>
<tr>
<td>100 $\Phi_r$</td>
<td>0.8%</td>
<td>8.1%</td>
<td>9.6%</td>
</tr>
<tr>
<td>300 $\Phi_g$</td>
<td>5.3%</td>
<td>21.2%</td>
<td>22.9%</td>
</tr>
<tr>
<td>300 $\Phi_r$</td>
<td>0.2%</td>
<td>5.6%</td>
<td>8.5%</td>
</tr>
<tr>
<td>500 $\Phi_g$</td>
<td>4.1%</td>
<td>19.2%</td>
<td>22.7%</td>
</tr>
<tr>
<td>500 $\Phi_r$</td>
<td>0.1%</td>
<td>4.3%</td>
<td>7.3%</td>
</tr>
</tbody>
</table>

Table II

V. Temporal Characteristics of the Shannon Rate Process

Definition 2. Given a stationary spatial field, such as the interference field $J_{\lambda_b}$, the temporal stochastic process seen by a mobile moving at a constant velocity, $v$ along a straight line, $(J_{\lambda_b}(t), t > 0)$ is defined as:

$$J_{\lambda_b}(t) = J_{\lambda_b}(y_v(t)),$$  \hspace{1cm} (19)

where $y_v(t) = (vt, 0)$. 

Fig. 4. Required backhaul capacity with increasing density of base stations.
Similarly, we can define the temporal Shannon rate process, $(\hat{S}\lambda(t), t > 0)$. Given the fields are stationary and isotropic, without loss of generality, we can assume that the user is moving along the $x$-axis starting from the origin at time $t = 0$. In a first step, consider that the mobile user is moving at a fixed unit velocity.

Given the asymptotic characterization of the interference and Shannon rate fields as Gaussian fields, one can asymptotically characterize the above stochastic processes as stationary Gaussian processes, $(J\lambda(t), t > 0)$ and $(\hat{S}\lambda(t), t > 0)$. The continuity and differentiability properties of the processes follow immediately from those of the fields.

Note that, as the mobile moves through space, the variations in the rate it experiences depend on the path loss functions. For Class-1 functions, we have smooth variations in the interference and rate i.e., differentiable processes. For Class-3 functions we have no-where differentiable processes with high variations as illustrated in Fig 5. To analyze the high temporal variations in the rate, we characterize them with the help of Hölder exponents [35].

![Fig. 5. Sample path of Interference processes for various path loss functions.](image)

**Remark 2.** For a radial path loss function of Class-3, the interference experienced by the mobile user moving with unit velocity along a straight line is a stochastic process which is equivalent to the number of users in a $M/GI/\infty$ queue whose limiting processes has a Hölder exponent of $1/2$. See Appendix H of [34] for details.

For certain Class-2 and Class-3 path loss functions of interest, we could verify that they satisfy the above condition by numerically evaluating the covariance kernel.

**B. Time Scales in Adaptive Modulation and Coding**

In our environment, i.e., dense networks with the bounded path loss functions considered here and no fading, the signal power is asymptotically a constant, and the variability in the Shannon rate is primarily due to variations in the interference. In this section, we discuss ways to cope with such interference variations in the context of adaptive modulation and coding.

Adaptive modulation and coding is a technique used to adapt to variations in signal quality (SINR), where it dynamically selects the best modulation and coding scheme (MCS) based on estimates of current conditions. For simplicity, our model, we assume uniform binning of the Shannon rate itself with bin size $\Delta$. Let $b_0, b_1, \ldots, b_n$ be the discrete rates defining the bins, where bin $i$ is associated with rate $b_i$ and $|b_i - b_{i-1}| = \Delta$. This is a simplified model, as in practice, the range of the estimated channel state information (CSI) (e.g., SINR) is divided into non-uniform bins, each corresponding to a MCS, which are then mapped to a transmission rate by a non-linear function.

We now explore the timescales on which the adaptive modulation and coding (AMC) should operate. Slower rate of adaptation leads to difficulties in keeping up with the local variations. The user may experience conditions worse than those required for the selected rate, which should be avoided. At the same time a significant amount of overhead is involved in estimating the interference power and selecting a new rate. Thus one should try to limit the rate of adaptations.

We use our model to provide an understanding of the rate of adaptation and its dependence on various system parameters. Assume that adaptive coding takes place periodically every $h$ seconds. Namely, every $h$ seconds, the transmitter selects a particular rate, $b_\sigma \in \{b_1, \ldots, b_n\}$, i.e., $b_\sigma = b_{i-1} + \Delta$. Then we have the following theorem which gives a way to choose $h$ such that the chance that the selected modulation and coding rate is fine for the next $h$ seconds.

In order to cope with variations, we consider a conservative approach where, if the current rate $\sigma$ belongs to bin $i$, then we pick a code rate corresponding to bin $(i-1)^+, i.e., b_\sigma = b_{i-1} + \Delta/2$. Then we have the following theorem which gives a way to choose $h$ such that the chance that the selected modulation and coding rate is fine for the next $h$ seconds.

**Theorem 2.** For our adaptive modulation and coding model, if the centered Gaussian process, $(\hat{J}\xi(t), t \geq 0)$ satisfies the condition given in Lemma 2, then for all $|h| \leq g_{\lambda_0}(\alpha, k)$ with:

$$g_{\lambda_0}(\alpha, k) = \left(\frac{\Delta^2 \sqrt{\lambda_0}}{apk}\right)^{1/\alpha},$$

where $\lambda_0$ is the variance of the interference process.
\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|}
\hline
$\lambda_b$ & v m/s & Class-2 & Class-3 \\
\hline
300 & 1 & $\sim$ ms & $\sim$ 10$^{-1}$ ms \\
& 10 & $\sim$ 10$^{-1}$ ms & $\sim$ ms \\
1000 & 1 & $\sim$ 10$^{-4}$ s & $\sim$ 1 ms \\
& 10 & $\sim$ 10$^{-2}$ s & $\sim$ 10$^{-1}$ ms \\
\hline
\end{tabular}
\caption{Time scales determined by Theorem 3, Equation (23).}
\end{table}

and $\hat{k}$ any number such that $\hat{k} > \sqrt{2}k$, where $k_0 = \frac{\alpha}{\hat{k}^2}$, almost surely, all the sample paths satisfy $|\hat{S}_{\lambda_b}(t+h) - \hat{S}_{\lambda_b}(t)| < \Delta$, where $\hat{S}_{\lambda_b}$ is the Gaussian approximation of the Shannon rate process.

Proof. For $h$ small, from (17), $|\hat{S}_{\lambda_b}(t+h) - \hat{S}_{\lambda_b}(t)| < \Delta$ if and only if $|\hat{J}^c(t+h) - \hat{J}^c(t)| < \frac{\Delta e^2 \sqrt{N_0}}{\hat{k} \sigma^2}$. If the limiting Gaussian interference process, $J^c(t)$, satisfies the conditions of Lemma 2, then for all sufficiently small $h$ and $\hat{k} > \sqrt{2k}/k_0$, almost surely, all sample paths satisfy $|J^c(t+h) - J^c(t)| < k|\hat{k}|^\alpha$. Thus, for all $|h| \leq g_{\lambda_b}(\alpha, k)$, almost surely for all sample paths, $|\hat{S}_{\lambda_b}(t+h) - \hat{S}_{\lambda_b}(t)| < \Delta$.

Role of velocity. Let us now assume that the mobile user is moving with constant velocity $v$ instead of unit velocity. From Definition 2, we have

$$|\hat{J}^c(t+h) - \hat{J}^c(t)| = |\hat{J}^c(y_v(t+h)) - \hat{J}^c(y_v(t))|,$$ (22)

where $y_v(t) = (vt, 0)$. Further, since bin size has to be in the same order of magnitude as the Shannon rate, it makes sense to assume that the bin size is a fraction of the mean Shannon rate, i.e., $\Delta = \frac{1}{\eta}E[|S_{\lambda_b}(0)|]$ for some $\eta \in \mathbb{R}^+$ e.g., $\eta = 10$. Then, the function $g_{\lambda_b}(\alpha, k, v)$ is now given by:

$$g_{\lambda_b}(\alpha, k) = \frac{1}{v} \left( \frac{\kappa \sqrt{N_0}}{\eta k} \right)^{1/\alpha}.$$ (23)

The time period at which AMC should operate at primarily depends on: (1) the intensity of base stations ($\lambda_b$), (2) the velocity ($v$), and (3) the path loss models through $\kappa, \hat{k}, \alpha$. We study these various dependencies by considering a specific set of values for parameters: Intensity of base stations, $\lambda_b = 300, 1000$ per Km$^2$; Bandwidth, $w = 900$ MHz; transmitted power, $p = 1$ Watt and velocity, $v = $ from 1 to 10 m/s.

Further, we consider the dual slope path loss function of Class-2 and a discontinuous path loss function. We then numerically estimate the Hölder exponents to be 1 and 0.5 with constant values $k$ of 5 and 50 respectively. Then, for $\hat{k} = \frac{\kappa}{\eta k}$, the function $g_{\lambda_b}(\alpha, k)$ is given in Table III.

From the numerical evaluation and (23), we have the following observations:

> We get a lower rate of adaptation when increasing the density of base stations, $\lambda_b$, since the variance of the process decreases as studied in the previous section. The magnitude of the rate is due to the fact that function $g$ in (23) is polynomial in $\lambda_b$ which is in per Km$^2$.

> Increasing the velocity increases the rate of adaptation. For the same set of parameters, if one considers unit velocity we have that $g_{\lambda_b}(\alpha, k) \sim ms$ for Class-2 path loss models.

- Higher variability in the Shannon rate, i.e., lower values of $\alpha$ in case of discontinuous path loss models, leads to higher rates of adaptation.

Table III illustrates the time scales determined by Theorem 3, equation (23), at which adaptive modulation and coding should operate for the above mentioned network parameters. We validated this result with the help of simulations, by considering uniform binning of a fixed bin size $\Delta$ of the Shannon rate seen by a user moving at a constant velocity $v = 1$ m/s with $\lambda_b = 300$ and the network parameters as above. We considered three different time scales, 10$\mu$s, 1ms and 10$^{-1}$s, for adaptive modulation and evaluated the fraction of the time periods that are in error when applying the technique defined above. The simulation values are in agreement with our numerical result given by the Hölder exponent analysis, since the fraction of error is considerably lower for Class-2 path loss functions at 1ms and for Class-3 path loss functions at 10$\mu$s as can be seen in Table IV. The fraction of error is higher if one considers time scales greater than the values determined by (23).

Thus, this provides an understanding of how $h$ scales with different system parameters like the environment through the path loss $l(r)$, density of base stations, $\lambda_b$ and velocity $v$.

C. Level-crossings of the Scaled Interference Process

In this subsection, we study the expected number of upcrossings in an interval of the stationary differentiable Gaussian process $(\hat{J}^c(t), t \geq 0)$. We assume that this process is a.s. differentiable. This also gives the expected up-crossings for the approximated interference, $(\hat{J}_{\lambda_b}(t), t \geq 0)$ and Shannon rate processes, $(\hat{S}_{\lambda_b}(t), t \geq 0)$.

Given a threshold $u$, define the number of up-crossings in an interval $[0, T]$ as $N_u^+[0,T] = \# \{ t \in [0,T] : J^c(t) = u, \hat{J}^c(t) > 0 \}$. Then, the Rice formula for the expected number of upcrossings in the interval is given by (36)

$$E[N_u^+[0,T]] = \frac{\sqrt{\omega_2 T e^{-u^2/2(\sigma^2(0))}}}{2\pi \sqrt{\sigma(0)}},$$ (24)

where $\omega_2$ is the second spectral moment. Since the Gaussian process is mean square differentiable, the second spectral moment is given by $\omega_2 = \frac{1}{\hat{k}}$.

Using simulations, we estimated the expected number of up-crossings in an interval for various sample paths of the scaled interference process, $\hat{J}^c_{\lambda_b}(t)$. The aim is to compare this with the above result for the limiting Gaussian process. Thus, we evaluated Rice formula by calculating the second spectral moment numerically. The simulated mean value is within a 5 percent error margin from the numerically value. Thus, one can expect to use the results for the limiting Gaussian process to work for the original processes. We can also characterize the coverage and outage times i.e., the level crossings of the
approximated Gaussian Shannon rate process, \((\tilde{S}_{\lambda}(t), t \geq 0)\) and its asymptotics using the existing results as in [36], [37].

VI. CONCLUSION

By properly rescaling the interference and Shannon rate fields we have characterized their corresponding limiting Gaussian fields in ultra dense settings. This opens the opportunity to apply the rich set of tools and results for Gaussian fields to study dense wireless networks. Their characteristics depend primarily on the path loss. By taking functions of these fields, one can also shed light on fundamental engineering questions in ultra dense networks such as (1) the role of spatial correlation on backhaul dimensioning and (2) the characteristics of temporal variations mobile users would see and their impact on adaptive modulation and coding. Overall this provides a new approach for the assessment of the fundamental characteristics of densification.

VII. ACKNOWLEDGMENTS

This work is supported in part by the National Science Foundation under Grant No. NSF-CCF-1218338, NSF Grant CNS-1343383 and an award from the Simons Foundation (No.197982), all to the University of Texas at Austin.

REFERENCES


