Massive MIMO for Ray-Based Channels

Li, Shuang ; Smith, Peter J; Dmochowski, Pawel A; Tataria, Harsh; Matthaiou, Michail; Yin, Jing Wei

Published in:
IEEE International Conference on Communications (ICC) 2019, Proceedings

DOI:
10.1109/ICC.2019.8761057

2019

Document Version:
Early version, also known as pre-print

Massive MIMO for Ray-Based Channels

Shuang Li*, Peter J. Smith†, Pawel A. Dmochowski*, Harsh Tataria‡, Michail Matthaiou‡, and Jingwei Yin§
*School of Engineering and Computer Science, Victoria University of Wellington, Wellington, New Zealand
†School of Mathematics and Statistics, Victoria University of Wellington, Wellington, New Zealand
‡Institute of Electronics, Communications and Information Technology (ECIT), Queens University Belfast, Belfast, U.K.
§College of Underwater Acoustic Engineering, Harbin Engineering University, Harbin, P. R. China
email:{lishua,peter.smith,pawel.dmochowski}@ecs.vuw.ac.nz,{h.tataria,m.matthaiou}@qub.ac.uk, yinjingwei@hrbeu.edu.cn

***Abstract*—Favorable propagation (FP) and channel hardening are desired properties in massive multiple-input and multiple-output (MIMO) systems, where nearly optimal performance is achieved with linear processing techniques, such as maximal-ratio combining. To date, these properties have primarily been analyzed for classical *statistical* channel models, or ray-based models with very specific angular parameters and distributions. This paper presents a thorough mathematical analysis of the asymptotic system behaviour for ray-based channels with *arbitrary* ray distributions and a uniform linear array at the base station. In addition to FP and channel hardening, we analyze the *large system potential* (LSP) which measures the asymptotic ratio of the power in the desired channel to the total interference power when both the antenna and user numbers grow. LSP is said to hold when this ratio converges to a positive constant. The results demonstrate that while FP is guaranteed in ray-based channels, channel hardening may or may not occur depending on the nature of the model. Furthermore, we demonstrate that LSP will not normally hold as the interference power grows logarithmically relative to the power in the desired channel as the system size increases. Nevertheless, we identify some fundamental and attractive properties of massive MIMO in this limiting regime.**

I. Introduction

Two key principles behind the success of massive MIMO are favourable propagation (FP) [1], [2], and channel hardening [3], meaning that the normalized inter-user interference power converges to zero, and the normalized power in the desired channel becomes constant. With FP, the use of large numbers of antennas has an implicit interference reduction mechanism which boosts the achievable rates and enables the use of low complexity signal processing algorithms [2], [4], [5].

The bulk of the theoretical work on FP and channel hardening has employed classical *statistical* channel models. Here, the existence of FP has been demonstrated for channel models of increasing complexity, progressing from independent and identically distributed (i.i.d.) Rayleigh [1], [6], pure line-of-sight [1], [6], correlated Rayleigh [7], [8], and independent Ricean [9] to correlated Ricean channels [10], [11]. In parallel, with the theory, real channel measurements have demonstrated that a large fraction of the theoretical gains due to FP can be obtained [12]–[14].

This work is now mature, but incomplete in the sense that accurate modeling of large dimensional channels requires a strong link to the propagation environment. This is usually obtained through ray-based models which have been extensively validated by measurements and, for this reason, have made their way to the 3GPP standardization exercise [15]. These models are more physically based, have a closer link to the array architecture and are widely used irrespective of the type of frequency band [15]–[17]. The physical nature of the ray-based models also has some advantages for analysis. For example, FP was considered in recent work [10] for very general heterogeneous, correlated Ricean channels. This work gives wide ranging results on FP, but the inherent nature of these models meant that the conclusions relied on various assumptions concerning the correlation structure, line-of-sight direction, etc. In contrast, we are able to prove FP for ray-based models with the most assumptions concerning physical phenomena rather than statistical parameters.

Variations of such models have a proliferation of names including directional, spatial, geometric and Saleh-Valenzulela (SV) type [18] channel models. We prefer the phrase *ray-based* as the main requirements for our work are that the statistical distributions of individual rays can be identified. This is possible for a wide range of such channels. Important work has begun in this area demonstrating the existence of FP with specific ray-based models for a variety of antenna topologies such as the uniform linear array (ULA), uniform rectangular array (URA), and uniform circular array (UCA) [19], [20]. However, these results rely on two very special cases for the rays: an arbitrary ray must arrive with an azimuth angle, ϕ, which satisfies $\phi \sim U[0, 2\pi]$ [19], [20] or $\sin\phi \sim U[0, 2\pi]$ [1], [21]. Hence, a general analysis of FP for ray-based models with arbitrary ray distributions is almost entirely lacking. Further, while FP is an excellent property for a communication system, it only implies that a finite number of users can be served by increasing the number of antenna elements. We refer to this as *single-sided* massive MIMO. Ideally, as you grow the number of BS antennas you would also serve more users leading to a system that becomes large both in users and antennas, i.e., *double-sided* massive MIMO. Hence, we define large system potential (LSP) as the property that the fundamental ratio which measures the power in the desired channel relative to the total interference power converges to a positive constant as both the number of users (K) and the number of antennas (N) grow to infinity, with $N/K \to \alpha$ as $N \to \infty$. This is equivalent to requiring that

The work of M. Matthaiou was supported by EPSRC, UK, under grant EP/P000673/1.

the signal-to-interference ratio (SIR) of maximal ratio combing (MRC) with perfect channel state information (CSI) converges under the same limiting regime. Very little is known about LSP, except for the analysis in [5] for i.i.d. Rayleigh fading. To the best of our knowledge, nothing is known about LSP for ray-based channel models. In this paper, we address the gaps identified in the literature. In particular, for a broad class of ray-based models and a ULA[1] at the base-station, we make the following contributions.

- We show that channel hardening may or may not occur depending on the nature of the model.
- We show that FP is guaranteed for all models where the ray angles are continuous random variables (as assumed by all models to date).
- For LSP, we derive a remarkably simple expression which relates the asymptotic interference behaviour to system size, antenna spacing and the ray distribution. This result highlights the important role played by the end-fire direction in interference growth and leads to the following two contributions.
- We demonstrate that LSP will not normally hold as the interference power grows logarithmically relative to the power of the desired channel as the system size increases.
- Despite the lack of LSP, the implications for massive MIMO are excellent. Although the interference eventually dominates the desired channel, the growth is very slow and is further attenuated by practical factors such as the likely propagation environment and the typical array patterns employed.

Notation. Boldface lower and upper case symbols denote vectors and matrices. Complex conjugation and Hermitian transpose operations are denoted by $(\cdot)^*$ and $(\cdot)^{\mathrm{H}}$, while $\mathcal{CN}(\mathbf{m}, \mathbf{R})$ denotes the circular symmetric complex Gaussian distribution with mean $\mathbf{m}$ and covariance matrix $\mathbf{R}$, while $U[a, b]$ denotes a uniform distribution on $[a, b]$. $\mathbb{E}[\cdot]$ denotes statistical expectation, $\sim$ denotes asymptotic equivalence defined in [22, p. 15], $\xrightarrow{a.s.}$ denotes almost sure convergence, and $\mathcal{F}(\cdot)$ denotes the Fourier transform. Finally, $J_0(\cdot)$ and $I_0(\cdot)$ denote the zero order Bessel and modified Bessel functions of the first kind.

II. Channel Model and System Metrics

We consider an uplink massive MIMO system with N co-located antennas at one BS simultaneously serving K single antenna users, where $N \gg K$. We assume a narrowband flat fading channel model such that the $N \times 1$ channel vector for user i can be written as $\mathbf{h}_i$, defined below, and the composite $N \times K$ channel matrix is denoted by $\mathbf{H} = [\mathbf{h}_1 \mathbf{h}_2 ... \mathbf{h}_K]$.

A. Ray-based Channel Model

In general, the propagation channel to user i can be described as the superposition of many individual rays possibly arriving in clusters from a set of far-field scatterers. In simple terms, the channel is broken down into P incident rays at the BS.[2] Hence, for a ULA we have:

$$\mathbf{h}_i = \sum_{r=1}^{P} \gamma_{ir} \boldsymbol{a}(\phi_{ir}), \tag{1}$$

where ϕ_{ir} is the azimuth angle of the r^{th} ray, γ_{ir} is a complex scaling factor for the magnitude and phase of the ray, and $\boldsymbol{a}(\phi_{ir})$ is the steering vector. In azimuth, the antenna array broadside is at $\phi_{ir} = 0$, and end-fire is $\phi_{ir} = \pm\frac{\pi}{2}$. Common models for the scaling factor include random phase models [15], where $\gamma_{ir} = \sqrt{\beta_{ir}}\exp(j\Phi_{ir})$, β_{ir} is the power of the r^{th} ray and Φ_{ir} are i.i.d. $U[0, 2\pi]$ phase offsets. Hence, $\beta_i = \sum_{r=1}^{P} \beta_{ir}$ is the total link gain for user i. Also, complex Gaussian models have been proposed in [17], where $\gamma_{ir} = \sqrt{\beta_{ir}} u_{ir}$ and $u_{ir} \sim \mathcal{CN}(0, 1)$. For both models, we note that $\mathbb{E}[\gamma_{ir}] = 0$, $\mathbb{E}[|\gamma_{ir}|^2] = \beta_{ir}$ and $\mathbb{E}[\gamma_{ir}^* \gamma_{js}] = 0$ for all pairs $(i, r) \neq (j, s)$. For a ULA with normalized inter-element spacing d, the steering vectors are:

$$\boldsymbol{a}(\phi_{ir}) = \left[1, e^{2\pi j d \sin\phi_{ir}}, e^{2\pi j 2d \sin\phi_{ir}}, \ldots, e^{2\pi j(N-1)d\sin\phi_{ir}}\right]^{\mathrm{T}}.$$

B. FP, Channel Hardening and Large System Potential

Here, FP denotes asymptotic FP where $\mathbf{h}_i^{\mathrm{H}}\mathbf{h}_j/N \to 0$ as $N \to \infty$ [1]. Channel hardening refers to the property that $\mathbf{h}_i^{\mathrm{H}}\mathbf{h}_i/N \to \beta_i$ as $N \to \infty$, which is equivalent in our case to the definition [3]. In addition, we define, LSP where both the number of antennas and users grow at the same rate. LSP is said to hold when the following ratio

$$\zeta_{\text{LSP}} = \frac{|\mathbf{h}_i^{\mathrm{H}}\mathbf{h}_i/N|^2}{\sum_{j\neq i} |\mathbf{h}_i^{\mathrm{H}}\mathbf{h}_j/N|^2}, \tag{2}$$

converges to a positive constant as $N \to \infty$ and $N/K \to \alpha$. Since $|\mathbf{h}_i^{\mathrm{H}}\mathbf{h}_i|^2$ and $|\mathbf{h}_i^{H}\mathbf{h}_j|^2$ relate to the power in the desired channel and to the inherent interference power between two users, ζ_{LSP} is a fundamental performance metric, measuring the ratio of the desired channel power to interference.[3] Hence, the condition that ζ_{LSP} converges to a positive constant ensures that the power of signals carried on the desired channel is never dominated by the total interference, and also indicates that the simplest linear processing (MRC) has the potential to handle large systems. Note that the limiting regime used for LSP, which supports double sided massive MIMO, is far more challenging than traditional massive MIMO. In practice, the number of users will never grow without bound but the asymptotics are still useful in identifying the key properties of systems, which are large in both N and K.

III. Channel Hardening and FP

In this section, channel hardening and FP are considered for ray-based channel models. For ease of notation, let $\mathbf{a}_{ir}$ be the

[1] We extend the analysis to a URA in the forthcoming journal version of this work.

[2] For ease of notation we do not specifically itemize clusters but the P paths include any clustered rays.

[3] The expression in (2) can also be interpreted as the SIR of MRC processing with perfect CSI.

steering vector for user i, path r, so that $\mathbf{a}_{ir} = \mathbf{a}(\phi_{ir})$. Then,

$$\begin{aligned}\frac{\mathbf{h}_i^{\mathrm{H}}\mathbf{h}_i}{N} &= \frac{1}{N}\sum_{r=1}^{P}\gamma_{ir}^{*}\mathbf{a}_{ir}^{\mathrm{H}}\sum_{s=1}^{P}\gamma_{is}\mathbf{a}_{is} \\ &= \sum_{r=1}^{P}|\gamma_{ir}|^2 + \frac{1}{N}\sum_{r=1,r\neq s}^{P}\sum_{s=1}^{P}\gamma_{ir}^{*}\gamma_{is}\mathbf{a}_{ir}^{\mathrm{H}}\mathbf{a}_{is} \\ &= X_i + E_i, \end{aligned} \tag{3}$$

where $X_i = \sum_{r=1}^{P}|\gamma_{ir}|^2$ is independent of N. Thus the limiting value depends entirely on $\lim_{N\to\infty} E_i$, which in turn depends on $\lim_{N\to\infty}\mathbf{a}_{ir}^{\mathrm{H}}\mathbf{a}_{is}/N$, where $r \neq s$. Now,

$$\begin{aligned}\left|\frac{\mathbf{a}_{ir}^{\mathrm{H}}\mathbf{a}_{is}}{N}\right| &= \frac{1}{N}\sum_{n=0}^{N-1} e^{-j2\pi dn\sin\phi_{ir}}e^{j2\pi dn\sin\phi_{is}} \\ &= \frac{1}{N}\left|\frac{\sin\left(N\iota/2\right)}{\sin\left(\iota/2\right)}\right| \xrightarrow{a.s.} 0, \end{aligned} \tag{4}$$

where $\iota = 2\pi d[\sin\phi_{is} - \sin\phi_{ir}]$ using simple results on geometric series. Almost sure convergence follows from the fact that convergence is guaranteed unless $\sin\phi_{ir} = \sin\phi_{is}$, an event with probability zero for continuous angular variables. Thus we have $\mathbf{h}_i^{\mathrm{H}}\mathbf{h}_i/N \xrightarrow{a.s.} X_i$ as $N \to \infty$. Note that for random phase models, $X_i = \beta_i$ and traditional channel hardening occurs where $\mathbf{h}_i^{\mathrm{H}}\mathbf{h}_i/N \xrightarrow{a.s.} \beta_i$, a deterministic limit. In contrast, for complex Gaussian models, $|\gamma_{ir}|^2 = \beta_{ir}|u_{ir}|^2$, which gives a random limit, as $X_i = \sum_{r=1}^{P}|\gamma_{ir}|^2$ is a weighted sum of exponential variables. Hence, we see that the existence of channel hardening depends on the nature of the model for the ray coefficients.

In terms of FP, results are simple following the same methodology as for channel hardening. First, we write

$$\frac{\mathbf{h}_i^{\mathrm{H}}\mathbf{h}_j}{N} = \frac{1}{N}\sum_{r=1}^{P}\sum_{s=1}^{P}\gamma_{ir}^{*}\gamma_{js}\mathbf{a}_{ir}^{\mathrm{H}}\mathbf{a}_{js},$$

and then we use (4) to show that

$$\left|\frac{\mathbf{a}_{ir}^{\mathrm{H}}\mathbf{a}_{js}}{N}\right| = \frac{1}{N}\left|\frac{\sin\left(N\tau/2\right)}{\sin\left(\tau/2\right)}\right| \xrightarrow{a.s.} 0, \tag{5}$$

as $N \to \infty$, where $\tau = 2\pi d[\sin\phi_{js} - \sin\phi_{ir}]$. Hence, FP is proven very simply for all ray-based models where $\sin\phi_{ir} = \sin\phi_{is}$ has probability zero. A simple condition for this to hold is that the angles are continuous random variables, a property held by all proposed models. Therefore, FP, the key property enabling single sided massive MIMO, holds for all ray distributions considered to date.

IV. LARGE SYSTEM POTENTIAL

In this section, we analyse the LSP of ray-based channel models in the limiting regime where $K \to \infty$, $N \to \infty$ and $N/K \to \alpha$. The ratio ζ_{LSP} in (2) has a numerator satisfying $|\mathbf{h}_i^{\mathrm{H}}\mathbf{h}_i|^2/N^2 \xrightarrow{as} X_i^2$ from Sec. III. Hence, LSP depends on the asymptotic properties of the denominator, $\eta = \sum_{j=1,j\neq i}^{K}|\mathbf{h}_i^{\mathrm{H}}\mathbf{h}_j|^2/N^2$. For the denominator, we focus on $\mathbb{E}[\eta]$ since LSP would require a finite value of $\mathbb{E}[\eta]$.

$$\begin{aligned}\mathbb{E}\left[\eta\right] &= \mathbb{E}\left[\frac{1}{N}\sum_{j=1,j\neq i}^{K}\frac{1}{N}\left|\sum_{r=1}^{P}\sum_{s=1}^{P}\gamma_{ir}^{*}\gamma_{js}\boldsymbol{a}_{ir}^{\mathrm{H}}\boldsymbol{a}_{js}\right|^2\right] \\ &= \frac{1}{N}\sum_{j=1,j\neq i}^{K}\sum_{r=1}^{P}\sum_{s=1}^{P}\mathbb{E}\left[|\gamma_{ir}|^2\right]\mathbb{E}\left[|\gamma_{js}|^2\right]\frac{1}{N}\mathbb{E}\left[\left|\boldsymbol{a}_{ir}^{\mathrm{H}}\boldsymbol{a}_{js}\right|^2\right] \\ &= \beta_i\left(\frac{\sum_{j=1,j\neq i}^{K}\beta_j}{N}\right)\mu(N,d), \end{aligned} \tag{6}$$

using the basic properties of the γ_{ir} terms and the notation

$$\begin{aligned}\mu(N,d) &= \frac{1}{N}\mathbb{E}\left[\left|\boldsymbol{a}_{ir}^{\mathrm{H}}\boldsymbol{a}_{js}\right|^2\right] \\ &= \frac{1}{N}\sum_{n=0}^{N-1}\sum_{m=0}^{N-1}\mathbb{E}\left[e^{j2\pi d(m-n)\sin\phi_{ir}}\right]\mathbb{E}\left[e^{j2\pi d(n-m)\sin\phi_{js}}\right] \\ &= \frac{1}{N}\sum_{n=0}^{N-1}\sum_{m=0}^{N-1}\left|\mathbb{E}\left[e^{j2\pi d(m-n)\sin\phi_{ir}}\right]\right|^2. \end{aligned} \tag{7}$$

Now we set $\hat{\phi}_{ir} = 2\pi d\sin\phi_{ir}$, and rewrite (7) as

$$\mu(N,d) = 1 + 2\sum_{s=1}^{N-1}\left(1 - \frac{s}{N}\right)|\mathbb{E}[e^{-js\hat{\phi}_{ir}}]|^2. \tag{8}$$

Assuming that $\sum_{j=1}^{K}\beta_j/K$ converges to $\bar{\beta}$ as $K \to \infty$, where $\bar{\beta}$ is a finite mean power, we have $\lim_{N\to\infty}\mathbb{E}\left[\eta\right] = \beta_i\bar{\beta}\lim_{N\to\infty}\{\mu(N,d)\}$. Hence, the asymptotic behaviour of $\mathbb{E}\left[\eta\right]$ depends on $\mu(N,d)$ which in turn depends on how quickly $|\mathbb{E}[e^{-js\hat{\phi}_{ir}}]|^2$ decays. In the following theorem, we present a general answer to this question.

Theorem 1. *The term $\mathbb{E}[e^{-js\hat{\phi}_{ir}}]$ decays as $s^{-\frac{1}{2}}$ as $s \to \infty$ with the asymptotic representation:*

$$\begin{aligned}&\mathbb{E}[e^{-js\hat{\phi}_{ir}}] \\ &\sim \frac{1}{\sqrt{ds}}\left(f_\phi\left(\frac{-\pi}{2}\right)e^{j\left(2\pi ds - \frac{\pi}{4}\right)} + f_\phi\left(\frac{\pi}{2}\right)e^{-j\left(2\pi ds - \frac{\pi}{4}\right)}\right), \end{aligned} \tag{9}$$

where $f_\phi(\cdot)$ is the probability density function (PDF) of ϕ_{ir}.

Proof. The proof is given in Appendix A. ∎

A. Implications of Theorem 1

Equation (9) in Theorem 1 is a remarkable result with a simple and intuitive interpretation, wide generality and important implications for massive MIMO.

- In terms of generality, (9) only requires the angular PDF, $f_\phi(\cdot)$, not to have singularities which are worse than $O(x^{-1/2})$ at $x = 0$. This covers all proposed models.
- Interpreting (9) we see that if the end-fire direction has no energy, $f_\phi(\pm\pi/2) = 0$, then $\mathbb{E}[e^{-js\hat{\phi}_{ir}}] = 0$. Alternatively, if some end-fire radiation occurs then $f_\phi(\pm\pi/2) > 0$ and $\mathbb{E}[e^{-js\hat{\phi}_{ir}}] = O(s^{-1/2})$.

- From the above, it follows that if there is no end-fire radiation, $\mu(N,d)$ is finite and the mean interference cannot dominate the power of the desired channel.
- Further, if there is end-fire radiation, then LSP does not hold as $\mu(N,d) \to \infty$. This conclusion holds by inspection of (8). When $\mathbb{E}[e^{-js\hat{\phi}_{ir}}]$ is $O(s^{-1/2})$ then $\sum_{s=1}^{N} |\mathbb{E}[e^{-js\hat{\phi}_{ir}}]|^2$ is $O(\log N)$ using well known properties of the series $\sum_{s=1}^{N} \frac{1}{s}$. Also, $\sum_{s=1}^{N} \frac{s}{N} |\mathbb{E}[e^{-js\hat{\phi}_{ir}}]|^2$ is finite so that $\mu(N,d)$ grows to infinity, but at a very slow logarithmic rate.
- The importance of the end-fire direction can be understood in the following way. For a ULA, it is not the proximity of two incoming ray angles that drives the interference, but the difference in the sines of the angles (see τ in (5)). For angles close to broadside the difference in sines is largest and for angles near end-fire the difference is smallest, resulting in greater interference.
- Overall, the result in (9) is extremely positive for double sided massive MIMO. We have shown that in the challenging scenario where both K and N grow large, the interference, relative to the power of the desired channel, grows very slowly (logarithmically). Also, the scaling of this growth factor is very small, since a large amount of end-fire radiation is unlikely and practical deployments employ array patterns which focus on a given sector and massively attenuate the end-fire direction.

Given the power of these results, it is useful to validate the conclusions with some closed form special cases.

B. Special Cases: Uniform Distribution

When $\phi_{ir} \sim U[0, 2\pi]$, $f_\phi(x) = \dfrac{1}{2\pi}$ for $-\pi \le x \le \pi$ and (9) becomes

$$\mathbb{E}[e^{-js\hat{\phi}_{ir}}] \sim \frac{1}{\pi\sqrt{ds}} \cos\left(2\pi ds - \frac{\pi}{4}\right). \tag{10}$$

This limiting value is verified in the uniform case where the exact solution is known as $\mathbb{E}[e^{-js\hat{\phi}_{ir}}] = \mathbb{E}[e^{-js2\pi d \sin\phi_{ir}}] = J_0(2\pi ds)$ [22, p. 375]. For large values of s, $J_0(2\pi ds) \sim \cos(2\pi ds - \frac{\pi}{4})/\pi\sqrt{ds}$ [22, p. 364], which agrees with (9). Hence, the general asymptotic analysis in (9) is supported and the exact value of $\mathbb{E}[e^{-js\hat{\phi}_{ir}}]$ can be used in (8) to give the exact value of $\mu(N,d)$.

C. Special Cases: Von-Mises Distribution

The von-Mises (VM) distribution has also been used in angular modelling [23] and has the PDF given by

$$f_\phi(x) = \frac{e^{\kappa \cos(x-\mu)}}{2\pi I_0(\kappa)}, \ -\pi \le x \le \pi, \tag{11}$$

where μ is a measure of location and κ is a measure of concentration. Substituting the VM PDF into (9) we obtain

$$\mathbb{E}[e^{-js\hat{\phi}_{ir}}] \sim \frac{e^{j2\pi ds}}{\sqrt{ds}} \left(\frac{e^{\kappa\cos(-\frac{\pi}{2}-\mu)}}{2\pi I_0(\kappa)} e^{-j\frac{\pi}{4}} + \frac{e^{\kappa\cos(\frac{\pi}{2}-\mu)}}{2\pi I_0(\kappa)} e^{j\frac{\pi}{4}} \right). \tag{12}$$

The exact solution can be found by integration, giving

$$\mathbb{E}[e^{-js\hat{\phi}_{ir}}] = \frac{I_0\left(\sqrt{\kappa^2\cos^2(\mu) + (\kappa\sin(\mu) - j2\pi ds)^2}\right)}{I_0(\kappa)}. \tag{13}$$

Some further analysis, omitted here for reasons of space, shows that (13) is asymptotically equal to (12). Hence, for the VM case also, we have verified (9) and given an exact solution for $\mathbb{E}[e^{-js\hat{\phi}_{ir}}]$.

V. Numerical Results

In Fig. 1 we demonstrate the channel hardening and FP results discussed in Sec. III for $K = 2$ and an increasing number of antennas. We adopt the non-line-of-sight (NLOS) 3GPP angular and cluster parameters in [15]. The number of clusters is $C = 20$, and the number of subpaths per cluster is $L = 20$. Referring to the channel model in (1), $P = CL$. Each subpath angle of arrival (AoA) is modeled by a central cluster angle with a Gaussian distribution (zero mean and a standard deviation of $76.5°$) along with a subray offset angle which is Laplacian with a standard deviation of $15°$. We assume subrays with equal powers.[4] From the upper plot of Fig. 1, we can see that the normalized power in the desired channel, $S = |\mathbf{h}_i^{\mathrm{H}}\mathbf{h}_i|/N \approx 1$ for large numbers of antennas. Similarly, the lower plot shows the interference term, $I = |\mathbf{h}_i^{\mathrm{H}}\mathbf{h}_j|/N$ decreasing to zero as $N \to \infty$. Note that Fig.1 plots $\mathbb{E}[I]$ for both the ray-based model (via simulation) and i.i.d. Rayleigh fading (via analysis) so that the variations don't obscure the trend. As expected, the convergence to FP is slower for the ray-based model but the initial rate of convergence is similar for both channels. Hence, both channel hardening and FP are shown to occur for a typical parameter set as predicted by the analysis. Fig. 1 shows channel hardening and FP occurring for a clustered channel model with wrapped Gaussian central cluster angles and Laplacian offsets. This numerical example is useful as it verifies the analysis for a commonly used ray-based model structure. The analysis goes much further and proves the existence/non-existence of channel hardening and the existence of FP for all ray-based of the form in (1) for a comprehensive range of ray distributions.

In Fig. 2, we show that the power of the desired channel will either converge to a constant or a random variable, verifying the analysis in Sec. III. We assume the same model as in Fig. 1 but with two possibilities for the ray coefficients, γ_{ir}. The Akdeniz model [15] uses a complex Gaussian variable for γ_{ir}, while the 3GPP model [15] uses a random phase. As shown in Fig. 2, as the number of antennas grows, the cumulative distribution function (CDF) of the normalized desired channel power, $S = \mathbf{h}_i^{\mathrm{H}}\mathbf{h}_i/N$, with the Akdeniz model remains almost the same, indicating convergence to a random variable. In contrast, with the random phase model the CDF converges to a step function indicating that S converges to a constant.

[4] Equal ray powers are adopted for simplicity in Fig. 1 and Fig. 2 for initial verification of the FP and channel hardening results, $\beta_{ir} = 1/CL$, and phases are uniformly distributed, $\phi_{ir} \sim U[0, 2\pi]$.

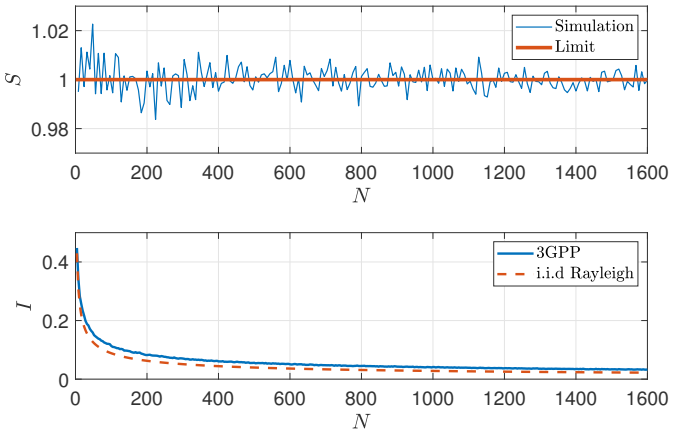

Fig. 1. Channel hardening and FP (3GPP angular parameters).

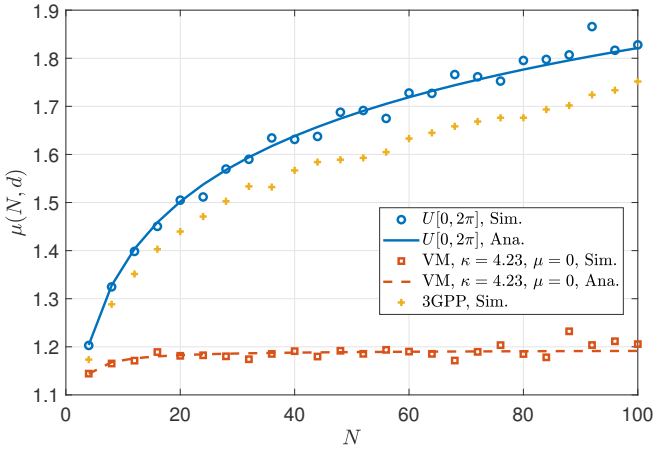

Fig. 3. $\mu(N,d)$ vs N for three different angular distributions.

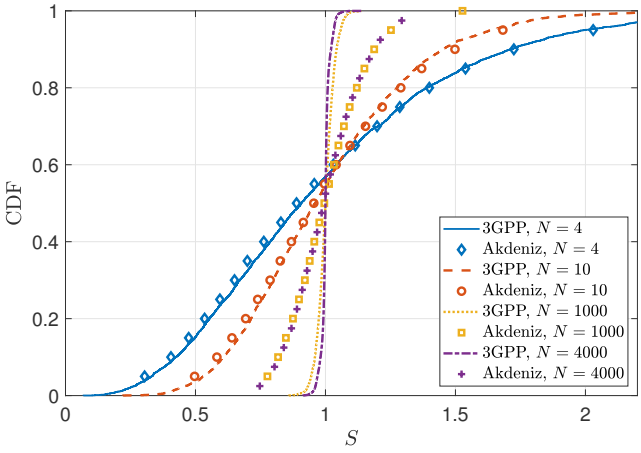

Fig. 2. Channel hardening for two types of channel models.

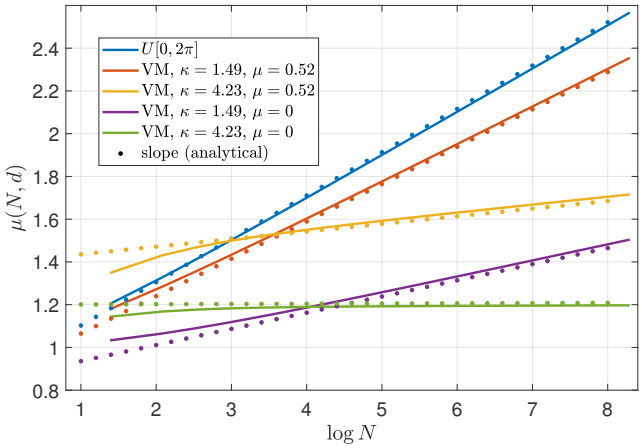

Fig. 4. Logarithmic growth of $\mu(N,d)$.

Hence, as shown in Sec. III, channel hardening can occur for ray-based models but this depends on the models employed for the ray coefficients.

In Fig. 3, we show both the simulated and analytical results for $\mu(N,d)$ with uniform and VM distributions using the results in Secs. IV-B and IV-C. We also show simulated values of $\mu(N,d)$ by adopting the angular parameters of the 3GPP model in [15] as in Fig. 1. The number of antennas and users are growing at the same ratio $N/K = \alpha = 2$, while $\phi_{ir} \sim U[0,2\pi]$ for the uniform model and $\kappa = 4.23$ (for 30° angle spread) and $\mu = 0$ for VM. From Fig. 3 we can see that the analysis agrees well with simulation for both uniform and VM models. We also notice that the growth rates of $\mu(N,d)$ are different for all three models, due to the differences in the AoA distributions. In the following figure, we give more details of the growth rate with regard to angular distributions.

In Fig. 4, we demonstrate the logarithmic growth rate of $\mu(N,d)$ against the number of antennas, N, for VM and uniform models with different parameters. Note that although the analysis in Theorem 1 predicted logarithmic growth, it is hard to verify from Fig. 3. Hence, we substitute (9) into (8) and identify the dominant component of $\mu(N,d)$ giving $\mu(N,d) \sim m_{\text{slope}}\log(N) + C_0$, where C_0 is a constant and

$$m_{\text{slope}} = \frac{2(f_\phi^2(\frac{\pi}{2}) + f_\phi^2(-\frac{\pi}{2}))}{d}. \tag{14}$$

Hence, m_{slope} determines how quickly $\mu(N,d)$ will grow. The uniform distribution has the highest interference growth rate, which is $m_{\text{slope}}^{\text{uniform}} = (\pi^2 d)^{-1}$. For the VM model, the slope depends on κ and μ. In Fig. 4 we observe that $\mu(N,d)$ is clearly logarithmic in N as predicted and that the slope is correctly identified by (9) as shown by the dotted lines which have slope m_{slope}. The analytical results in Secs. IV-B and IV-C were used in generating the results.

As well as verifying the logarithmic growth, Fig. 4 demonstrates some interesting angular properties. For both $\kappa = 4.23$ (angle spread $= 30°$) and $\kappa = 1.49$ (angle spread $= 60°$), $\mu(N,d)$ decreases as μ is reduced from $\mu = 0.52$ (30°) to $\mu = 0$. This is because shifting the mean towards broadside reduces the interference inflation that occurs near end-fire. Secondly, for both $\mu = 0$ and $\mu = 0.52$ there is a cross-over as N increases. For small N, increased angle spread is

beneficial as it spreads the rays and reduces the chance of high interference caused by rays in close proximity. However, for high N the higher angle spread puts more probability near end-fire and this begins to dominate and causes higher interference.

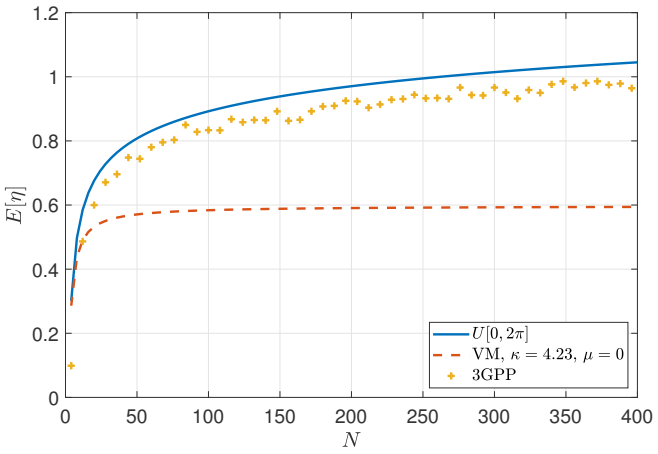

Fig. 5. $\mathbb{E}[\eta]$ vs N for three angular distributions.

Finally, in Fig. 5, we confirm via simulation for the 3GPP parameters and via analysis for the uniform and VM models that the mean global interference term, $\mathbb{E}[\eta]$, grows logarithmically as predicted by the analysis. For the uniform case, $\phi_{ir} \sim U[0, 2\pi]$, for VM, $\kappa = 4.23$, $\mu = 0$, and for 3GPP we use the parameters considered in Fig. 1. For the uniform and VM models all user link gains and ray powers are equal, $\beta_{ir} = (CL)^{-1}$. For the 3GPP parameters, we also consider unequal ray powers and unequal user link gains. To avoid the substantial extra variation caused by shadowing models with large arrays we employ a simple deterministic model for these powers. The link gains decay exponentially from user 1 to user K such that $\beta_K = \frac{1}{10}\beta_1$ and the cluster powers behave similarly. The desired users is then randomly allocated one of the K distinct link gains. The levels are then adjusted to give the same total power as in the uniform and VM models and subrays in a particular cluster all have the same power as assumed in [24]. Fig. 5 shows the same logarithmic growth as Fig. 3, confirming the analysis.

VI. Conclusion

The fundamental properties of massive MIMO have been identified with great generality for a broad class of ray-based models with a ULA at the BS. The generality and insight possible is considerably more than can be achieved with statistical channel models. In particular, we show that channel hardening may or may not occur depending on the model used and FP is guaranteed for all continuous angular distributions. Although LSP will not normally hold, as the interference grows logarithmically relative to the desired channel, the implications for massive MIMO are excellent. As the number of users grows, the interference does grow relative to the desired channel but extremely slowly and this is further reduced by practical considerations, such as the attenuation of end-fire radiation caused by typical array patterns.

Appendix A
Proof of Theorem 1

We note that

$$\mathbb{E}[e^{-js\hat{\phi}_{ir}}] = \int_{-2\pi d}^{2\pi d} e^{-jsx} f_{\hat{\phi}}(x)dx, \tag{15}$$

where $f_{\hat{\phi}}(\cdot)$ is PDF of $\hat{\phi}_{ir}$. Now, $\hat{\phi_{ir}} = 2\pi d \sin\phi_{ir}$ is a non one-to-one transformation of ϕ_{ir}. Using standard transformation theory, we obtain

$$f_{\hat{\phi}}(x) = \frac{p(x)}{\sqrt{l^2 - x^2}}, \ -l \le x \le l, \tag{16}$$

where $l = 2\pi d$ and

$$p(x) = \begin{cases} f_\phi\left(\sin^{-1}\left(\frac{x}{l}\right)\right) + f_\phi\left(\pi - \sin^{-1}\left(\frac{x}{l}\right)\right), x \ge 0 \\ f_\phi\left(\sin^{-1}\left(\frac{x}{l}\right)\right) + f_\phi\left(-\pi - \sin^{-1}\left(\frac{x}{l}\right)\right), x < 0. \end{cases} \tag{17}$$

Hence, (15) is rewritten as

$$\mathbb{E}[e^{-js\hat{\phi}_{ir}}] = \int_{-l}^{l} e^{-jsx} \frac{p(x)}{\sqrt{l^2 - x^2}} dx. \tag{18}$$

Using the notation in [25, Eq.1, p. 15], the Fourier transform (FT) of a function $f(x)$ can be written as

$$g(y) = \int_{-\infty}^{\infty} f(x) e^{-j2\pi xy} dx. \tag{19}$$

If we set $y = \dfrac{s}{2\pi}$, then

$$g\left(\frac{s}{2\pi}\right) = \int_{-\infty}^{\infty} f(x) e^{-jsx} dx. \tag{20}$$

Using the Heaviside function, $H(x)$, we can write (18) as a FT in the same format as (20) as follows,

$$\mathbb{E}[e^{-js\hat{\phi}_{ir}}] = \int_{-\infty}^{\infty} e^{-jsx} \left(H(x+l) - H(x-l)\right) \frac{p(x)}{\sqrt{l^2 - x^2}} dx. \tag{21}$$

Hence, defining $f(x) = \left(H(x+l) - H(x-l)\right) \dfrac{p(x)}{\sqrt{l^2 - x^2}}$, allows $\mathbb{E}[e^{-js\hat{\phi}_{ir}}]$ to be computed as the FT of $f(x)$.

This formulation is particularly useful as we can now leverage known results on the asymptotics of FTs as $s \to \infty$ [25]. These results depend on the singularities of $f(x)$ so we first discuss the nature of these singularities. Clearly, $f(x)$ has singularities at $x = \pm l$ and at any singularities of $p(x)$. Note that the singularities at $x = \pm l$ are infinite discontinuities (indicating that the value of $f(x)$ will grow infinitely large as x approaches $\pm l$). In contrast, the singularities of $p(x)$ are never infinite discontinuities for any proposed, practical angular distribution models. Models such as the wrapped Gaussian have no singularities inside $(-l, l)$ while the Laplacian has only a non-differentiable point at the peak. Hence, the singularities at $x = \pm l$ are the worst. The general principle presented in [25,

p. 55] is that the 'worst' singularity[5] of a function contributes the leading term to the asymptotic expression for its FT. Thus in our case, we only need to consider the two singularities at $x = \pm l$. Near $x_1 = -l$, $f(x)$ behaves like $F_1(x) = H(x+l)p(-l)(2l(l+x))^{-1/2}$ and similarly near $x_2 = l$, $f(x)$ behaves like $F_2(x) = (1 - H(x-l))p(l)(2l(l-x))^{-1/2}$. Rewriting, we obtain

$$F_1(x) = \frac{H(x+l)p(-l)}{\sqrt{2l}}|x+l|^{-\frac{1}{2}}, \tag{22}$$

$$F_2(x) = \frac{p(l)}{\sqrt{2l}}|x-l|^{-\frac{1}{2}} - \frac{H(x-l)p(l)}{\sqrt{2l}}|x-l|^{-\frac{1}{2}}. \tag{23}$$

From [25, Theorem 19, p. 52], we know that if a generalised function, $f(x)$, has a finite number of singularities at $\{x = x_1, x_2, x_3, ..., x_m\}$, and for each of them $f(x) - F_m(x)$ has absolutely integrable N^{th} order derivatives in an interval including x_m, where $F_m(x)$ is a linear combination of functions of type $|x - x_m|^\beta$, $|x - x_m|^\beta \text{sgn}(x - x_m)$, $|x-x_m|^\beta \log|x-x_m|$, $|x-x_m|^\beta \log|x-x_m| \text{sgn}(x-x_m)$, and if $f^{(N)}(x)$ is well behaved at infinity, then $g(y)$, the FT of $f(x)$, satisfies $g(y) = \sum_{m=1}^{M} G_m(y) + o(|y|^{-N})$, as $|y| \longrightarrow \infty$, where $G_m(y)$ is the FT of $F_m(x)$. Using this, we have

$$g\left(\frac{s}{2\pi}\right) \sim G_1\left(\frac{s}{2\pi}\right) + G_2\left(\frac{s}{2\pi}\right), \tag{24}$$

where G_1 and G_2 are the FTs of $F_1(x)$ and $F_2(x)$ in (22) and (23) and $\sim$ denotes asymptotic equivalence defined in [22, p. 15]. From [25, Table 1, p. 43], the FTs required are

$$\begin{aligned} \mathcal{F}(|x-l|^{-\frac{1}{2}}) &= e^{-2\pi jly}|y|^{-\frac{1}{2}}, \\ \mathcal{F}(H(x+l)|x+l|^{-\frac{1}{2}}) &= e^{2\pi jly - \frac{1}{4}j\pi \text{sgn}(y)}|2y|^{-\frac{1}{2}}, \\ \mathcal{F}(H(x-l)|x-l|^{-\frac{1}{2}}) &= e^{-2\pi jly - \frac{1}{4}j\pi \text{sgn}(y)}|2y|^{-\frac{1}{2}}. \end{aligned} \tag{25}$$

Using (25), we obtain

$$\begin{aligned} &g\left(\frac{s}{2\pi}\right) \sim G_1\left(\frac{s}{2\pi}\right) + G_2\left(\frac{s}{2\pi}\right), \\ &= \sqrt{\frac{\pi}{ls}}\left(\frac{p(-l)}{\sqrt{2}}e^{j(ls-\frac{\pi}{4})} + p(l)e^{-jls} - \frac{p(l)}{\sqrt{2}}\sqrt{\frac{\pi}{s}}e^{-j(ls+\frac{\pi}{4})}\right). \end{aligned} \tag{26}$$

Substituting $p(l) = 2f_\phi(\frac{\pi}{2})$, $p(-l) = 2f_\phi(\frac{-\pi}{2})$ and $l = 2\pi d$ into (26), and after some simplification we obtain the result in Theorem 1.

[5]The singularity $x = x_m$ of a function, $f(x)$, is worst if $f(x)$ is of order $|x - x_m|^\beta$ near x_m and β is the smallest value for all singularities. [25, p. 55].

REFERENCES

[1] H. Q. Ngo *et al.*, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.

[2] E. G. Larsson *et al.*, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[3] H. Q. Ngo and E. G. Larsson, "No downlink pilots are needed in TDD massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2921–2935, May 2017.

[4] L. Liang *et al.*, "Low-complexity hybrid precoding in massive multiuser MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 3, no. 6, pp. 653–656, Dec. 2014.

[5] F. Rusek *et al.*, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[6] E. Björnson *et al.*, "Massive MIMO: Ten myths and one critical question," *IEEE Commun. Mag.*, vol. 54, no. 2, pp. 114–123, Feb. 2016.

[7] ——, "A random access protocol for pilot allocation in crowded massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 4, pp. 2220–2234, Apr. 2017.

[8] R. Couillet and M. Debbah, *Random Matrix Methods for Wireless Communications*. Cambridge University Press, 2011.

[9] Q. Zhang *et al.*, "Power scaling of uplink massive MIMO systems with arbitrary-rank channel means," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 966–981, Oct. 2014.

[10] M. Matthaiou *et al.*, "Does massive MIMO fail in Ricean channels?" *IEEE Wireless Commun. Lett.*, 2018.

[11] H. Tataria *et al.*, "Uplink analysis of large MU-MIMO systems with space-constrained arrays in Ricean fading," in *Proc. IEEE ICC*, May 2017, pp. 1–7.

[12] J. Li and Y. Zhao, "Measurement-based asymptotic user orthogonality analysis and modelling for massive MIMO," *IEEE Commun. Lett.*, vol. 21, no. 12, pp. 2762–2765, Dec. 2017.

[13] J. Hoydis *et al.*, "Channel measurements for large antenna arrays," in *Proc. IEEE ISWCS*, Aug. 2012, pp. 811–815.

[14] X. Gao *et al.*, "Massive MIMO performance evaluation based on measured propagation data." *IEEE Trans. Wireless Commun.*, vol. 14, no. 7, pp. 3899–3911, Jul. 2015.

[15] C. T. Neil *et al.*, "Impact of microwave and mmWave channel models on 5G systems performance," *IEEE Trans. Antennas Propag.*, vol. 65, no. 12, pp. 6505–6520, Dec. 2017.

[16] S. Sangodoyin *et al.*, "Cluster characterization of 3D MIMO propagation channel in an urban macrocellular environment," *IEEE Trans. Wireless Commun.*, 2018.

[17] M. R. Akdeniz *et al.*, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1164–1179, Jun. 2014.

[18] A. A. Saleh and R. Valenzuela, "A statistical model for indoor multipath propagation," *IEEE J. Sel. Areas Commun.*, vol. 5, no. 2, pp. 128–137, Feb. 1987.

[19] X. Wu *et al.*, "On favorable propagation in massive MIMO systems and different antenna configurations," *IEEE Access*, vol. 5, pp. 5578–5593, 2017.

[20] Z. Gao *et al.*, "Asymptotic orthogonality analysis of time-domain sparse massive MIMO channels," *IEEE Commun. Lett.*, vol. 19, no. 10, pp. 1826–1829, Oct. 2015.

[21] H. Q. Ngo *et al.*, "Aspects of favorable propagation in massive MIMO," in *Proc. IEEE EUSIPCO*, Sep. 2014, pp. 76–80.

[22] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables*. Courier Corporation, 1964, vol. 55.

[23] S. Wang *et al.*, "Performance characterization of AOA geolocation systems using the Von Mises distribution," in *Proc. IEEE Veh. Technol. Conf. (VTC Fall)*, Sep. 2012, pp. 1–5.

[24] 3GPP, "Study on channel model for frequencies from 0.5 to 100 GHz," 3rd Generation Partnership Project (3GPP), Tech. Rep. TR 38.901 (V14.0.0), Mar. 2017. [Online]. Available: http://www.3gpp.org/

[25] M. J. Lighthill, *An Introduction to Fourier Analysis and Generalised Functions*. Cambridge University Press, 1958.