Massive MIMO with a Generalized Channel Model: Fundamental Aspects

Massive MIMO with a Generalized Channel Model: Fundamental Aspects

Michail Matthaiou*, Hien Quoc Ngo*, Peter J. Smith†, Harsh Tataria‡, and Shi Jin§
*Institute of Electronics, Communications and Information Technology (ECIT), Queen's University Belfast, Belfast, U.K.
†School of Mathematics and Statistics, Victoria University of Wellington, Wellington, New Zealand
‡Department of Electrical and Information Technology, Lund University, Lund, Sweden
§National Mobile Communications Research Laboratory, Southeast University, Nanjing, P. R. China
Email: {m.matthaiou, hien.ngo}@qub.ac.uk, peter.smith@vuw.ac.nz, harsh.tataria@eit.lth.se, jinshi@seu.edu.cn

***Abstract*—Massive multiple-input multiple-output (MIMO) is becoming a mature technology, and has been approved for standardization in the 5G space. Although there are many papers on the theoretical analysis of massive MIMO, the majority of relevant work assumes the simplified, yet overly idealistic, Kronecker-type model for spatial correlation. Motivated by the deficiencies of the Kronecker model, we invoke a naturally generalized spatial correlation model, that is the Weichselberger model. For this model, we pursue a comprehensive analysis of massive MIMO performance in terms of channel hardening and favorable propagation (FP). We identify a number of scenarios under which massive MIMO may fail and discuss their relevance from a practical perspective.**

I. Introduction

Massive MIMO is now a mature technology that is making its way into 5G trials and standards [1], [2]. The biggest body of related literature is based on Kronecker-type models, which offer analytical tractability, thereby, facilitating performance analysis and transceiver design. Yet, it is a well-known fact from the early days of conventional MIMO, that Kronecker-based models enforce the spatial correlation properties at both ends to be separable [3]; that is all directions-of-arrival (DoAs) are treated as completely independent from the directions-of-departure (DoDs), and vice versa. A far more realistic channel model is the so-called Weichselberger model [4], which alleviates the deficiencies of the Kronecker model by considering the joint correlation structure of both ends; therefore, the average coupling between the spatial subchannels is effectively modeled. We also note that the Weichselberger model includes the Kronecker model and the virtual channel representation (VCR) [5] as special cases. Despite its importance, there is a dearth of literature on the performance analysis of conventional multi-user MIMO (MU-MIMO) and massive MIMO with this generalized Gaussian fading model. This can be partially attributed to the increased number of parameters that have to be specified for the Weichselberger model compared to the Kronecker model and the VCR. In the MU-MIMO literature, we point out the work of [6], which investigated the capacity-achieving input covariance matrix for a single-user Weichselberger Ricean fading MIMO channel and [7] which extended [6] to the MU case but only in the low-power regime. More recently, [8] investigated the asymptotic sum-rate of the MU Weichselberger Ricean fading MIMO channel using the replica method. The massive MIMO literature is even more scarce with respect to the Weichselberger model. The work in [9] compared the accuracy of the Kronecker and Weichselberger models against a set of measurement data at 2.6 GHz and concluded that the latter indeed provides more accurate modeling. Also, [10] analyzed both centralized and distributed massive MIMO using the Weichselberger model but neglected any line-of-sight components.

This paper moves away from the state-of-art and analyzes, for the first-time, the theoretical performance of massive MIMO using the Weichselberger model. After introducing the new system model and discussing some basic statistical properties, our analysis targets the two fundamental performance metrics of any massive MIMO communication system, namely *channel hardening* and *favorable propagation*. By leveraging tools of Gaussian theory, we derive mathematical conditions under which these two concepts become valid. We also identify analytically scenarios under which these two concepts break down and corroborate them with a set of numerical results. Our work complements and extends some recent theoretical papers on massive MIMO with pure LoS fading [11], [12], i.i.d Rayleigh fading [11], semi-correlated Rayleigh fading [13], and semi-correlated Ricean fading [14], [15].

Notation: A complex normal vector with mean $\mathbf{b}$ and covariance $\mathbf{\Sigma}$ reads as $\mathcal{CN}(\mathbf{b}, \mathbf{\Sigma})$. The expectation of a random variable is denoted as $\mathbb{E}[\cdot]$, while the matrix trace by $\mathrm{tr}(\cdot)$. The symbols $(\cdot)^*$, $(\cdot)^T$ and $(\cdot)^H$ represent the conjugate, transpose and Hermitian transpose of a matrix. The notation $\xrightarrow{\text{a.s.}}$ implies almost sure convergence, $\xrightarrow{\text{P}}$ denotes convergence in probability, while $\odot$ denotes the element-wise (Hadamard) multiplication between two matrices (or vectors).

II. System model

Consider the massive MIMO uplink, where the BS is equipped with M antennas and serves $L \ll M$ single-antenna users. The $M \times 1$ channel from the k-th user to the BS is [4]

$$\mathbf{h}_k = \eta_k \bar{\mathbf{h}}_k + \gamma_k \underbrace{\mathbf{U}_k \left(\widetilde{\boldsymbol{\omega}}_k \odot \mathbf{h}_{\mathtt{iid}} \right)}_{\triangleq \widetilde{\mathbf{h}}_k} \tag{1}$$

The work of M. Matthaiou was supported by EPSRC, UK, under grant EP/P000673/1 and by the RAEng/The Leverhulme Trust Senior Research Fellowship LTSRF1718\14\2.

where K_k is the Ricean K-factor seen by the k-th user, while $\bar{\mathbf{h}}_k$ is the LoS component, $\eta_k \triangleq (K_k/(K_k+1))^{1/2}$ and $\gamma_k \triangleq (1/(K_k+1))^{1/2}$. Moreover, $\widetilde{\boldsymbol{\omega}}_k$ is the element-wise square root of $\boldsymbol{\omega}_k = [\omega_{k,1}, \ldots \omega_{k,M}]^T \in \mathbb{R}^{M\times 1}$ (which will be defined shortly), while $\mathbf{h}_{\mathtt{iid}}$ is an $M \times 1$ vector with i.i.d. $\mathcal{CN}(0,1)$ entries. We also define the small-scale fading matrix $\mathbf{H} \triangleq [\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_L] \in \mathbb{C}^{M\times L}$.[1] The spatial structure of the random component in (1) can be efficiently captured by the one-sided correlation matrix

$$\begin{aligned}\mathbf{Q}_k &= \mathbb{E}\left[\widetilde{\mathbf{h}}_k\widetilde{\mathbf{h}}_k^H\right] = \mathbf{U}_k\mathbb{E}\left[\left(\widetilde{\boldsymbol{\omega}}_k \odot \mathbf{h}_{\mathtt{iid}}\right)\left(\widetilde{\boldsymbol{\omega}}_k \odot \mathbf{h}_{\mathtt{iid}}\right)^H\right]\mathbf{U}_k^H \\ &= \mathbf{U}_k\left(\widetilde{\boldsymbol{\omega}}_k\widetilde{\boldsymbol{\omega}}_k^T\right) \odot \mathbb{E}\left[\mathbf{h}_{\mathtt{iid}}\mathbf{h}_{\mathtt{iid}}^H\right]\mathbf{U}_k^H = \mathbf{U}_k\boldsymbol{\Lambda}_k\mathbf{U}_k^H \qquad (2)\end{aligned}$$

where $\boldsymbol{\Lambda}_k = \mathrm{diag}\left(\boldsymbol{\omega}_k\right)$ and we have leveraged the following property $(\mathbf{a}\odot\mathbf{b})(\mathbf{c}\odot\mathbf{d})^H = \left(\mathbf{a}\mathbf{c}^H\right)\odot\left(\mathbf{b}\mathbf{d}^H\right)$. Looking at (2), we infer that $\mathbf{U}_k$ contains the eigenbases at the BS side related with user k. The diagonal matrix $\boldsymbol{\Lambda}_k$ contains the eigenvalues of $\mathbf{Q}_k$ which are also equal to the real coefficients $\omega_{k,1}, \ldots \omega_{k,M}$. As was articulated in [4], $\omega_{k,m}$ represents the coupling coefficients which specify the mean amount of energy coupled from the m-th user to the m-th receive eigenvector. Referring to (1), the following constraints should be satisfied: $\bar{\mathbf{h}}_k^H\bar{\mathbf{h}}_k = M$ and $\mathrm{tr}(\boldsymbol{\Lambda}_k) = M, \forall k = 1, \ldots, L$ [8], [14].

III. Channel hardening and FP

Since the pioneering work [16], the theoretical advancement of massive MIMO has been based on the concepts of *channel hardening* and *FP*. In this section, we will elaborate on them and identify scenarios under which they break down. To keep our notation clean, we focus on the perfect CSI case, though very similar results can be obtained for the imperfect CSI case.

Proposition 1. *For the channel model in* (1)*, we have that*

$$\mathbb{E}\left[||\mathbf{h}_k||^4\right] = \gamma_k^4\mathrm{tr}\left(\boldsymbol{\Lambda}_k^2\right) + M^2 + 2\eta_k^2\gamma_k^2\left(\mathbf{v}_k^H\boldsymbol{\Lambda}_k\mathbf{v}_k\right) \qquad (3)$$

$$\mathbb{E}\left[||\mathbf{h}_k||^2\right] = M, \qquad (4)$$

where $\mathbf{v}_k = \mathbf{U}_k^H\bar{\mathbf{h}}_k = [\nu_{k,1}, \ldots, \nu_{k,M}]^T$.

Proof. By expanding the left-hand side of (3), we get:

$$\begin{aligned}&\mathbb{E}\left[||\mathbf{h}_k||^4\right] = \mathbb{E}\left[\mathbf{h}_k^H\mathbf{h}_k\mathbf{h}_k^H\mathbf{h}_k\right] \\ &= \mathbb{E}\left[\left(\eta_k^2\bar{\mathbf{h}}_k^H\bar{\mathbf{h}}_k + \eta_k\gamma_k\bar{\mathbf{h}}_k^H\widetilde{\mathbf{h}}_k + \eta_k\gamma_k\widetilde{\mathbf{h}}_k^H\bar{\mathbf{h}}_k + \gamma_k^2\widetilde{\mathbf{h}}_k^H\widetilde{\mathbf{h}}_k\right)^2\right] \\ &= \mathbb{E}\left[\left(\eta_k^2 M + \eta_k\gamma_k\bar{\mathbf{h}}_k^H\widetilde{\mathbf{h}}_k + \eta_k\gamma_k\widetilde{\mathbf{h}}_k^H\bar{\mathbf{h}}_k + \gamma_k^2||\widetilde{\mathbf{h}}_k||^2\right)^2\right] \\ &= \eta_k^4M^2 + \eta_k^2\gamma_k^2M\mathbb{E}\left[||\widetilde{\mathbf{h}}_k||^2\right] \\ &+ 2\eta_k^2\gamma_k^2\mathbb{E}\left[\bar{\mathbf{h}}_k^H\widetilde{\mathbf{h}}_k\widetilde{\mathbf{h}}_k^H\bar{\mathbf{h}}_k^H\right] + \gamma_k^4\mathbb{E}\left[||\widetilde{\mathbf{h}}_k||^4\right]. \qquad (5)\end{aligned}$$

The proof then concludes by appropriate simplifications and by noticing that $\mathbb{E}\left[||\widetilde{\mathbf{h}}_k||^2\right] = \mathrm{tr}(\boldsymbol{\Lambda}_k) = M$ and that $\mathbb{E}\left[||\widetilde{\mathbf{h}}_k||^4\right] = \left(\mathrm{tr}(\boldsymbol{\Lambda}_k^2) + (\mathrm{tr}\left(\boldsymbol{\Lambda}_k\right))^2\right)$ [14]. The proof of (4) is trivial. □

Proposition 2. *For the channel model in* (1)*, we have that*

$$\begin{aligned}\mathbb{E}\left[\left|\mathbf{h}_k^H\mathbf{h}_\ell\right|^2\right] &= \eta_k^2\eta_\ell^2\left|\bar{\mathbf{h}}_k^H\bar{\mathbf{h}}_\ell\right|^2 + \gamma_k^2\gamma_\ell^2\mathrm{tr}\Big(\mathbf{Q}_k\mathbf{Q}_\ell\Big) \\ &+ \eta_k^2\gamma_\ell^2\left(\bar{\mathbf{h}}_k^H\mathbf{Q}_\ell\bar{\mathbf{h}}_k\right) + \gamma_k^2\eta_\ell^2\left(\bar{\mathbf{h}}_\ell^H\mathbf{Q}_k\bar{\mathbf{h}}_\ell\right). \qquad (6)\end{aligned}$$

Proof. The proof follows after some basic algebra. □

Note that the results in Propositions 1 and 2 can generalize our results in [14] for the case of semi-correlated Ricean fading with Kronecker-type of spatial correlation.[2] The following result will be particularly useful in our asymptotic analysis:

Lemma 1. *Chebyshev's theorem: Let* $X_1, X_2, \ldots, X_n$ *be independent RVs with* $\mathbb{E}[X_i] = \mu_i$ *and* $\mathrm{Var}(X_i) \leq z < \infty, \forall i \in 1, 2, \ldots, n$*. Then, as* $n \to \infty$

$$\frac{1}{n}\left(X_1 + X_2 + \ldots + X_n\right) - \frac{1}{n}\left(\mu_1 + \mu_2 + \ldots + \mu_n\right) \xrightarrow{\mathrm{P}} 0.$$

A. Channel hardening

Channel hardening arises whenever the randomness of a fading channel is averaged, converting the normalized channel power into a deterministic quantity. The authors of [17] provided the following definition of asymptotic channel hardening

$$\frac{||\mathbf{h}_k||^2}{\mathbb{E}\left[||\mathbf{h}_k||^2\right]} \xrightarrow{\text{a.s.}} 1, \text{ as } M \to \infty. \qquad (7)$$

We now introduce the notation $\underline{\mathbf{X}} \triangleq \mathbf{X} \odot \mathbf{X}^*$ and also invoke $||\mathbf{x}||_\infty = \max\left(|x_1|, |x_2|, \ldots\right)$ and $||\mathbf{X}||_{\max} = \max_{ij}|x_{ij}|$.

Assumption 1. *As* $M \to \infty$*, for every* $k = 1, \ldots, L$*,* $\limsup_M ||\underline{\boldsymbol{\omega}_k}||_\infty < \infty$ *and* $\limsup_M ||\underline{\mathbf{v}_k}||_\infty < \infty$.

The physical interpretation of Assumption 1 is that no $\omega_{k,m}, |\nu_{k,m}|, m = 1, \ldots, M$ grows without bound as M increases. We will later on see what happens when this constraint is not satisfied.

Corollary 1. *For the channel model in* (1) *and provided that Assumption 1 is fulfilled, we have*

$$\frac{||\mathbf{h}_k||^2}{\mathbb{E}\left[||\mathbf{h}_k||^2\right]} \xrightarrow{\mathrm{P}} 1, \text{ as } M \to \infty. \qquad (8)$$

Proof. The proof begins by evaluating the asymptotic behavior of $\widetilde{\mathbf{h}}_k^H\widetilde{\mathbf{h}}_k$ which arises in the expansion of $||\mathbf{h}_k||^2$:

$$\begin{aligned}\frac{1}{M}\widetilde{\mathbf{h}}_k^H\widetilde{\mathbf{h}}_k &= \frac{1}{M}\mathrm{tr}\left(\left(\widetilde{\boldsymbol{\omega}}_k\widetilde{\boldsymbol{\omega}}_k^T\right)\odot\left(\mathbf{h}_{\mathtt{iid}}\mathbf{h}_{\mathtt{iid}}^H\right)\right) \\ &= \frac{1}{M}\sum_{m=1}^M\omega_{k,m}|X_m|^2 \xrightarrow{\mathrm{P}} \frac{1}{M}\sum_{m=1}^M\omega_{k,m} = 1 \qquad (9)\end{aligned}$$

[1] In this paper, we deliberately neglect the impact of large-scale fading as our main focus in on modeling the small-scale fading variations. Besides, such an extension is trivial by following the standard approach in [1], [2].

[2] For the case of single-antenna users, the Weichselberger and Kronecker models are mathematically equivalent through a linear transformation [4]. The beauty of the former is that it enables an insightful second-order characterization through the power coupling coefficients.

where X_m is an i.i.d. complex Gaussian RV with zero mean and unit variance. Note that in the last step of (9), we have utilized Lemma 1. In a very similar manner, we can show that $\frac{1}{M}\widetilde{\mathbf{h}}_k^H\bar{\mathbf{h}}_k \xrightarrow{\text{P}} 0$ and $\frac{1}{M}\bar{\mathbf{h}}_k^H\widetilde{\mathbf{h}}_k \xrightarrow{\text{P}} 0$. The proof then concludes after some basic algebra. □

Discussion: By inspection of (9), we can infer that channel hardening will break down if some $\omega_{k,m}, 1 \leq n \leq M$ scale as $\mathcal{O}(M^\epsilon)$ with $0 < \epsilon \leq 1$. In this case, $\frac{1}{M}\widetilde{\mathbf{h}}_k^H\widetilde{\mathbf{h}}_k$ will asymptotically start to behave as a linear combination of $|X_m|^2$ terms, which are exponential variates, thereby experiencing random fluctuations. This scenario kicks in if the energy from the k-th user couples only into a small number of receive eigenvectors. This can happen when there is a limited number of resolvable multipath components from the k-th user impinging on the BS. Similar conclusions can be drawn for the elements of $\mathbf{v}_k$ that appear in the expansion of $\frac{1}{M}\widetilde{\mathbf{h}}_k^H\bar{\mathbf{h}}_k$ and $\frac{1}{M}\bar{\mathbf{h}}_k^H\widetilde{\mathbf{h}}_k$. Note that similar observations were made in the earlier work [18] for conventional point-to-point MIMO systems.

Mathematically speaking, the convergence in probability in Corollary 1 is a weaker condition compared to almost sure convergence. Nevertheless, in a practical system such a difference will cause little performance variation (if any). Most important though is the deviation from these "asymptotically-optimal" conditions in the finite number of antennas regime; this can be quantified by the scaled second-order moment of the channel gain, defined as below

$$\text{Var}\left(\frac{||\mathbf{h}_k||^2}{\mathbb{E}\left[||\mathbf{h}_k||^2\right]}\right) = \frac{\gamma_k^2}{M^2}\Big(\underbrace{2\eta_k^2\mathbf{v}_k^H\mathbf{\Lambda}_k\mathbf{v}_k}_{\text{LoS contribution}} + \underbrace{\gamma_k^2\text{tr}\left(\mathbf{\Lambda}_k^2\right)}_{\text{NLoS contribution}}\Big). \tag{10}$$

The above expression decouples nicely the NLoS and LoS contributions; in fact, the scaled variance of the signal power is an increasing function of γ_k; for $\gamma_k = 0$ (pure LoS propagation), this variance becomes exactly zero regardless of the value of M, whereas for $\gamma_k = 1$ (correlated Rayleigh fading conditions), the right-hand side of (10) becomes equal to $\text{tr}\left(\mathbf{\Lambda}_k^2\right)/M^2$. This value agrees with [17, Eq. (2.17)]. For an arbitrary and fixed value of $\gamma_k \in [0, 1]$, we have:

- From the Rayleigh-Ritz theorem, the LoS component in (10) is upper and lower bounded as follows:

$$\frac{\min(\boldsymbol{\omega}_k)}{M} \leq \frac{1}{M^2}\mathbf{v}_k^H\mathbf{\Lambda}_k\mathbf{v}_k \leq \frac{\max(\boldsymbol{\omega}_k)}{M} \leq 1. \tag{11}$$

  Interestingly, the lower limit above does not go asymptotically to zero when $\bar{\mathbf{h}}_k^H$ is aligned with the weakest eigenvector of $\mathbf{Q}_k$ and the corresponding eigenvalue scales as $\mathcal{O}(M)$. More generally, by definition $\bar{\mathbf{h}}_k$ can be expressed as a linear combination of the linearly independent eigenvectors of $\mathbf{Q}_k$ (i.e. the columns of $\mathbf{U}_k$). If the corresponding eigenvalues scale as $\mathcal{O}(M)$, then $\left(\mathbf{v}_k^H\mathbf{\Lambda}_k\mathbf{v}_k\right)/M^2$ does not vanish as $M \to \infty$. As such, whenever we have non-vanishing alignment of a LoS response of the k-th user with the dominant eigenvectors of the corresponding covariance matrix $\mathbf{Q}_k$, asymptotic channel hardening becomes challenging.
- The NLoS component in (10) is maximized when $\mathbf{Q}_k$ is rank-1 [17], in which case $\text{tr}\left(\mathbf{\Lambda}_k^2\right) = M^2$. Then,

$$\text{Var}\left(\frac{||\mathbf{h}_k||^2}{\mathbb{E}\left[||\mathbf{h}_k||^2\right]}\right) = \frac{\gamma_k^2}{M^2}\left(2\eta_k^2 M|\bar{\mathbf{h}}_k^H\mathbf{u}_{k,1}|^2 + \gamma_k^2 M^2\right)$$

  where $\mathbf{u}_{k,1}$ is the dominant eigenvector of $\mathbf{Q}_k$. On the other hand, the variance is minimized when $\mathbf{Q}_k = \mathbf{I}_M$ in which case $\text{tr}\left(\mathbf{\Lambda}_k^2\right) = M$ and (10) becomes

$$\text{Var}\left(\frac{||\mathbf{h}_k||^2}{\mathbb{E}\left[||\mathbf{h}_k||^2\right]}\right) = \frac{\gamma_k^2}{M}\left(1 + \eta_k^2\right),$$

  which goes smoothly to zero as $M \to \infty$.

We will now investigate how the variance in (10) behaves for a range of scenarios. Assuming a uniform linear array at the BS, the LoS channel response can be expressed as

$$\bar{\mathbf{h}}_k = \left[1, e^{j2\pi d\cos(\phi_k)}, \ldots, e^{j2\pi d(M-1)\cos(\phi_k)}\right]^T \tag{12}$$

where d is the equidistant inter-element antenna spacing normalized by the carrier wavelength and ϕ_k is the angle-of-arrival from the k-user. We will now consider three different scenarios for the coupling vector $\boldsymbol{\omega}_k$ of the k-th user:

Scenario 1: $\boldsymbol{\omega}_k = [1, 1, \ldots, 1]^T$

Scenario 2: $\boldsymbol{\omega}_k = [M/2, M/(2M-2), \ldots, M/(2M-2)]^T$

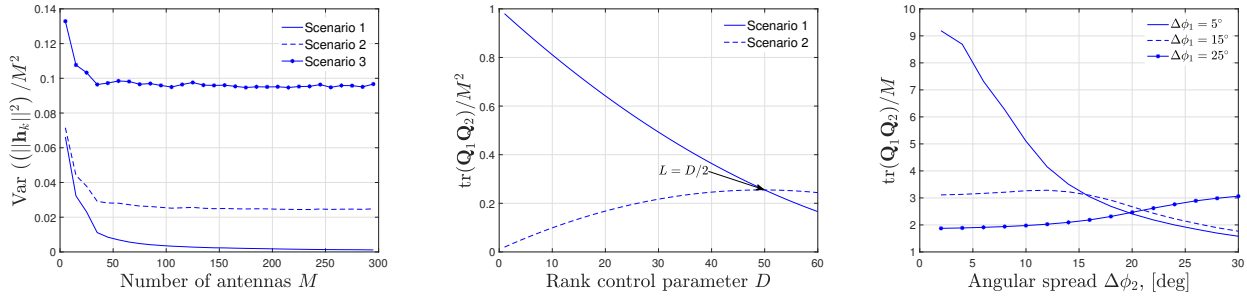
Scenario 3: $\boldsymbol{\omega}_k = [M, 0, \ldots, 0]^T$.

The first scenario represents an equal distribution of power across all antenna elements and resembles that of i.i.d. Rayleigh fading; in the second and third scenarios, there are entries $\omega_{k,m}$ that scale as $\mathcal{O}(M)$. The last scenario represents an extreme case where only a single antenna can capture energy from the k-th user. In Fig. 1(a), we plot $\text{Var}\left(||\mathbf{h}_k||^2/\mathbb{E}\left[||\mathbf{h}_k||^2\right]\right) = \text{Var}\left(||\mathbf{h}_k||^2\right)/M^2$ against M. Note that the unitary matrices $\mathbf{U}_k$ have been randomly generated. The figure validates that for isotropic fading conditions (as those studied in the early papers on massive MIMO [2], [11]), the normalized variance converges smoothly to zero. This is because all entries $\omega_{k,m}$ scale as $\mathcal{O}(1)$. On the other hand, for Scenarios 2 and 3, the findings are substantially different, since the scaled variance converges to non-zero limits, thereby indicating the deviation from the channel hardening regime. The situation is exacerbated for Scenario 3, which corroborates our theoretical analysis that rank-1 matrices maximize the NLoS contribution in (10).

B. Favorable propagation

FP has been identified as the key feature of massive MIMO that enables successful inter-user interference cancellation [11]–[13]. Mathematically speaking, two channel vectors offer asymptotic FP if they satisfy the following relationship [17]

$$\frac{\mathbf{h}_k^H\mathbf{h}_\ell}{\sqrt{\mathbb{E}\left[||\mathbf{h}_k||^2\right]\mathbb{E}\left[||\mathbf{h}_\ell||^2\right]}} \xrightarrow{\text{a.s.}} 0, \text{ as } M \to \infty. \tag{13}$$



(a) (b) (c)

Fig. 1: (a): Var $\left(||\mathbf{h}_k||^2\right)/M^2$ against the number of antennas, M ($K_k = 0.5, \phi_k = \pi/3$); (b): Inter-user covariance interference term, $\mathrm{tr}(\mathbf{Q}_1\mathbf{Q}_2)/M^2$ against the rank control parameter ($M = 100$); (c): Inter-user covariance interference term, $\mathrm{tr}(\mathbf{Q}_1\mathbf{Q}_2)/M$ against the angular spread $\Delta\phi_2$ ($M = 100, \phi_0^1 = \pi/4, \phi_0^2 = 3\pi/4, d = 1/2$).

Let us now define the following matrices: $\mathbf{\Omega}_{\ell,k} = \boldsymbol{\omega}_\ell\boldsymbol{\omega}_k^T$ and $\mathbf{R}_{k,\ell} = \mathbf{U}_k^H\mathbf{U}_\ell$ and the vector $\boldsymbol{\nu}_{k,\ell} = \bar{\mathbf{h}}_k^H\mathbf{U}_\ell$.

Assumption 2. *As* $M \to \infty$, *for every* $k, \ell = 1, \ldots, L$, $\limsup_M ||\underline{\mathbf{\Omega}_{\ell,k}}||_{max} < \infty$ *and* $\limsup_M ||\underline{\mathbf{R}_{k,\ell}}||_{max} < \infty$.

Assumption 3. *As* $M \to \infty$, *for every* $k, \ell = 1, \ldots, L$, $\limsup_M ||\underline{\boldsymbol{\nu}_{k,\ell}}||_\infty < \infty$.

Assumption 4. *As* $M \to \infty$, *for every* $k, \ell = 1, \ldots, L$, $\limsup_M |\bar{\mathbf{h}}_k^H\bar{\mathbf{h}}_\ell| < \infty$.

As with Assumption 1, Assumption 2-4 bound every individual entry of the involved vectors/matrices to remain finite, as the number of antennas, M, grows large. We will later on see what happens if any of these Assumptions are violated.

Corollary 2. *For the channel model in* (1) *and provided that Assumptions 1-4 are fulfilled, we have*

$$\frac{\mathbf{h}_k^H\mathbf{h}_\ell}{\sqrt{\mathbb{E}\left[||\mathbf{h}_k||^2\right]\mathbb{E}\left[||\mathbf{h}_\ell||^2\right]}} \xrightarrow{\mathrm{P}} 0, \text{ as } M \to \infty. \tag{14}$$

Proof. By expanding the left-hand side of (13) we obtain

$$\frac{\mathbf{h}_k^H\mathbf{h}_\ell}{\sqrt{\mathbb{E}\left[||\mathbf{h}_k||^2\right]\mathbb{E}\left[||\mathbf{h}_\ell||^2\right]}} = \frac{1}{M}\left(\eta_k\bar{\mathbf{h}}_k^H + \gamma_k\widetilde{\mathbf{h}}_k^H\right)\left(\eta_\ell\bar{\mathbf{h}}_\ell + \gamma_\ell\widetilde{\mathbf{h}}_\ell\right)$$
$$= \frac{1}{M}\left(\eta_k\eta_\ell\bar{\mathbf{h}}_k^H\bar{\mathbf{h}}_\ell + \eta_k\gamma_\ell\bar{\mathbf{h}}_k^H\widetilde{\mathbf{h}}_\ell + \gamma_k\eta_\ell\widetilde{\mathbf{h}}_k^H\bar{\mathbf{h}}_\ell + \gamma_k\gamma_\ell\widetilde{\mathbf{h}}_k^H\widetilde{\mathbf{h}}_\ell\right).$$

Now we can apply Lemma 1 on each random term of the above equation, which requires Assumptions 1-3 to be fulfilled. The deterministic quantity $\frac{1}{M}\bar{\mathbf{h}}_k^H\bar{\mathbf{h}}_\ell$ remains bounded if and only if Assumption 4 is satisfied. This concludes the proof. □

An important measure of performance is how orthogonal the channel vectors are for a practical number of antennas; this, in turn, will also indicate the actual amount of inter-user interference. This can be quantified by the variance of (13):

$$\mathrm{Var}\left(\frac{\mathbf{h}_k^H\mathbf{h}_\ell}{\sqrt{\mathbb{E}\left[||\mathbf{h}_k||^2\right]\mathbb{E}\left[||\mathbf{h}_\ell||^2\right]}}\right) = \frac{1}{M^2}\mathbb{E}\left[\left|\mathbf{h}_k^H\mathbf{h}_\ell\right|^2\right]. \tag{15}$$

Recall that the right-hand side of (15) has already been derived in closed-form in (6). In some of our recent work [15], we identified scenarios under which terms that have the form of (6) do not asymptotically go to zero. In a nutshell, this happens whenever we have strong alignment of two distinct LoS responses and/or non-vanishing alignment of a LoS response of the k-th user with the eigenvectors of the covariance matrix of the ℓ-th user, i.e. $\mathbf{Q}_\ell$, whose eigenvalues scale as $\mathcal{O}(M)$.

Let us see this via a toy example and focus on the term $\bar{\mathbf{h}}_\ell^H\mathbf{Q}_k\bar{\mathbf{h}}_\ell$. The unitary eigenvector matrix of $\mathbf{Q}_k$ can be expressed through its column eigenvectors as follows $\mathbf{U}_k = \left[\mathbf{u}_1^{(k)}, \mathbf{u}_2^{(k)}, \ldots, \mathbf{u}_M^{(k)}\right]$. Now, we consider the case where $\bar{\mathbf{h}}_\ell$ is a linear combination of the two principal eigenvectors of $\mathbf{U}_k$:

$$\bar{\mathbf{h}}_\ell = \sqrt{\frac{M}{2}}\mathbf{u}_1^{(k)} + \sqrt{\frac{M}{2}}\mathbf{u}_2^{(k)}.$$

We can easily show after some basic algebra that

$$\frac{1}{M^2}\bar{\mathbf{h}}_\ell^H\mathbf{Q}_k\bar{\mathbf{h}}_\ell = \frac{\omega_{k,1} + \omega_{k,2}}{2M}. \tag{16}$$

Thus, if any of the eigenvalues $\omega_{k,1}, \omega_{k,2}$ scales as $\mathcal{O}(M)$, FP will break down. We can now investigate further the behavior of the *inter-user covariance interference* term $\mathrm{tr}\Big(\mathbf{Q}_k\mathbf{Q}_\ell\Big)$ in (6). Our analysis begins by defining $\mathbf{V}_{k\ell} = \mathbf{U}_\ell^H\mathbf{U}_k$ and then:

$$\begin{aligned}\mathrm{tr}\Big(\mathbf{Q}_k\mathbf{Q}_\ell\Big) &= \mathrm{tr}\Big(\mathbf{\Lambda}_\ell\mathbf{V}_{k\ell}\mathbf{\Lambda}_k\mathbf{V}_{k\ell}^H\Big)\\ &= \sum_{m=1}^{M}\sum_{n=1}^{M}\omega_{\ell,m}\omega_{k,n}|\mathbf{V}_{k\ell}(m,n)|^2\\ &\leq ||\mathbf{V}_{k\ell}||_{\max}^2\sum_{m=1}^{M}\sum_{n=1}^{M}\omega_{\ell,m}\omega_{k,n}\\ &\leq M^2||\mathbf{V}_{k\ell}||_{\max}^2,\end{aligned} \tag{17}$$

where the upper bound in (17) is attained when both $\mathbf{Q}_k, \mathbf{Q}_\ell$ are fully-aligned rank-1 matrices. Since, by definition, $||\mathbf{V}_{k\ell}||_{\max} \leq 1$, we can infer that at most $\mathrm{tr}\Big(\mathbf{Q}_k\mathbf{Q}_\ell\Big) = M^2$. We now consider the ideal case of no inter-user interference, that is $\mathrm{tr}\Big(\mathbf{Q}_k\mathbf{Q}_\ell\Big) = 0$, which requires the covariance matrices

to have orthogonal support, i.e. $\mathbf{Q}_k\mathbf{Q}_\ell = \mathbf{0}$. This condition was utilized in a stream of papers [19], [20] to eliminate the effects of pilot contamination, by allocating pilots to users whose covariance matrices have nearly orthogonal support. However, as was articulated in [21], the orthogonality support condition is very unlikely in practice. We now consider two intuitive scenarios for $\mathbf{Q}_1$, $\mathbf{Q}_2$ to examine the effect of the matrix rank on the amount of inter-user covariance interference.

$$\text{Scenario 1: } \mathbf{Q}_1 = \begin{bmatrix} \mathbf{1}_{M-D,M-D} & \mathbf{0}_{M-D,D} \\ \mathbf{0}_{D,M-D} & \mathbf{I}_D \end{bmatrix}, \quad \mathbf{Q}_2 = \mathbf{1}_M$$

$$\text{Scenario 2: } \mathbf{Q}_1 = \begin{bmatrix} \mathbf{1}_{D,M-D} & \mathbf{0}_{D,D} \\ \mathbf{0}_{M-D,D} & \mathbf{I}_{M-D} \end{bmatrix}, \quad \mathbf{Q}_2 = \mathbf{1}_M,$$

where $\mathbf{1}_{m,n}$ is a $m \times n$ matrix full of 1's, and D is a rank control parameter. In Fig. 1(b), we validate that for Scenario 1 and low values of D, $\mathbf{Q}_1$ is rank-deficient and most, importantly, aligned with $\mathbf{Q}_2$. This is a catastrophic scenario that makes the inter-user covariance interference term scale with M^2. On the other extreme, for Scenario 2 and $D = 1$, we have that $\mathbf{Q}_1$ is full-rank and this makes $\mathrm{tr}(\mathbf{Q}_1\mathbf{Q}_2)/M^2$ approach zero. As D increases, $\mathbf{Q}_1$ becomes more and more rank-deficient, yet, it never becomes fully aligned with $\mathbf{Q}_2$.

We will now investigate the case when $\mathrm{tr}\Big(\mathbf{Q}_k\mathbf{Q}_\ell\Big)$ scales as $\mathcal{O}(M)$, which lies in the intersection of the previously investigated scenarios. To this end, we leverage Chebyshev's sum inequality and obtain the following lower bound:

$$\mathrm{tr}\Big(\mathbf{Q}_k\mathbf{Q}_\ell\Big) \geq \sum_{m=1}^{M}\sum_{n=1}^{M} |\mathbf{V}_{k\ell}(m,n)|^2 = ||\mathbf{V}_{k\ell}||_F^2 = M. \quad (18)$$

Note that the lower bound in (18) becomes exact if $\omega_{\ell,m}, \omega_{k,n}$ are all equal, such that $\omega_{\ell,m} = 1, \forall m = 1, \ldots, M$ and $\omega_{k,n} = 1, \forall n = 1, \ldots, M$. To visualize this case, we now consider the one-ring model, for which, the (m,n)-th entry of $\mathbf{Q}_k$ is:

$$[\mathbf{Q}_k]_{m,n} = \frac{1}{2\Delta\phi_k}\int_{-\Delta\phi_k+\phi_0^k}^{\Delta\phi_k+\phi_0^k} e^{-j2\pi d(n-m)\sin(\phi_k)} d\phi_k, \quad (19)$$

where $\Delta\phi_k$ is the azimuth angular spread corresponding to the k-th user, ϕ_0^k is the nominal direction-of-arrival, while d is the normalized antenna spacing as in (12). Figure 1(c) compares the term, $\mathrm{tr}(\mathbf{Q}_1\mathbf{Q}_2)/M$, against the angular spread $\Delta\phi_2$ for different values of $\Delta\phi_1$. We see that when both $\mathbf{Q}_1, \mathbf{Q}_2$ are very rank-deficient (i.e. they have small $\Delta\phi_k$), the interference term is boosted. Surprisingly, when the angular spreads are high, e.g., $\Delta\phi_k > 20^\circ$, this is not the ideal scenario since we notice a steady increase of $\mathrm{tr}(\mathbf{Q}_1\mathbf{Q}_2)/M$. This implies that, from the perspective of minimum inter-user covariance interference, the best scenario is when one covariance matrix is full-rank and the other one is very rank-deficient.

IV. Conclusion

The main motivation behind this work has been the inherent deficiencies of the Kronecker-type models. For this reason, we invoked the generalized Weichselberger spatial correlation model, which is able to effectively capture the average coupling between the spatial subchannels. Our analysis began by recasting the standard point-to-point Weichselberger model to a massive MIMO setup. We then derived closed-form expressions for the average desired signal and interference powers and examined the two critical performance measures of any massive MIMO system, these are channel hardening and FP. Our conclusions articulate that if the coupling vector has at least one entry that scales as $\mathcal{O}(M)$, channel hardening breaks down. Moreover, FP between two users is more pronounced whenever one user's covariance matrix is full-rank and the other one is very rank-deficient.

References

[1] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*. Cambridge, UK: Cambridge University Press, 2016.

[2] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp.1436–1449, Apr. 2013.

[3] H. Özcelik, M. Herdin, W. Weichselberger, J. Wallace, and E. Bonek, "Deficiencies of the Kronecker MIMO radio channel model," *IEE Elect. Lett.*, vol. 39, no. 16, pp. 1209–1210, 2003.

[4] W. Weichselberger, M. Herdin, H. Özcelik, and E. Bonek, "A stochastic MIMO channel model with joint correlation at both link ends," *IEEE Trans. Wireless Commun.*, vol. 5, no. 1. pp. 90–100, Jan. 2006.

[5] A. M. Sayeed, "Deconstructing multi-antenna fading channels," *IEEE Trans. Signal Process.*, vol. 50, pp. 2563–2579, Oct. 2002.

[6] X. Gao, *et al.*, "Statistical eigenmode transmission over jointly-correlated MIMO channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 8, Aug. 2009.

[7] P. Memmolo, M. Lops, A. M. Tulino, and R. A. Valenzuela, "Up-link multi-user MIMO capacity in low-power regime," in *Proc. IEEE ISIT,* June 2010, pp. 2308–2312.

[8] C.-K. Wen, S. Jin, and K.-K. Wong, "On the sum-rate of multiuser MIMO uplink channels with jointly-correlated Rician fading," *IEEE Trans. Commun.*, vol. 59, pp. 2883-2895, Oct. 2011.

[9] X. You and Y. Wang, *Massive MIMO Channel Modeling*, Master Thesis, Department of Electrical & Information Technology, Lund University, 2015.

[10] G. N. Kamga, M. Xia, and S. Aïssa, "Spectral-efficiency analysis of massive MIMO systems in centralized and distributed schemes," *IEEE Trans. Commun.* vol. 64, no. 5, pp. 1930–1941, May 2016.

[11] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Aspects of favorable propagation in massive MIMO," in *Proc. EUSIPCO*, Sept. 2014.

[12] C. Masouros and M. Matthaiou, "Space-constrained massive MIMO: Hitting the wall of favorable propagation," *IEEE Commun. Lett.*, vol. 19, no. 5, pp. 771–774, May 2015.

[13] J. H. Chen, "When does asymptotic orthogonality exist for very large arrays?" in *Proc. IEEE GLOBECOM*, Nov. 2013, pp. 6–10.

[14] H. Tataria, *et al.*, "Impact of line-of-sight and unequal spatial correlation in uplink MU-MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 6, no. 5, pp. 634–637, Oct. 2017.

[15] M. Matthaiou, P. J. Smith, H. Q. Ngo, and H. Tataria, "Does massive MIMO fail in Ricean channels?," *IEEE Wireless Commun. Lett.*, vol. 8, no. 1, pp. 61–64, Feb. 2019.

[16] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[17] E. Björnson, J. Hoydis, and L. Sanguinetti, *Massive MIMO Networks: Spectral, Energy, and Hardware Efficiency*, Foundations and Trends in Signal Processing: vol. 11, no. 3–4, pp. 154–655, 2017.

[18] V. Raghavan and A. M. Sayeed, "Sublinear capacity scaling laws for sparse MIMO channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 345–364, Jan. 2011.

[19] A. Adhikary, J. Nam, J. Y. Ahn, and G. Caire, "Joint spatial division and multiplexing The large-scale array regime," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 64416463, Oct. 2013.

[20] L. You, X. Gao, X. G. Xia, N. Ma, and Y. Peng, "Pilot reuse for massive MIMO transmission over spatially correlated Rayleigh fading channels," *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3352–3366, Jun. 2015.

[21] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO has unlimited capacity," *IEEE Trans. Wireless Commun.*, vol. 17, no. 1, pp. 574–590, Jan. 2018.