The Robustness of Favorable Propagation in Massive MIMO to Location and Phase Errors

Elham Anarakifirooz and Sergey Loyka, *Senior Member, IEEE*

***Abstract*— The impact of random errors (implementation inaccuracies) in element locations and beamforming phases on favorable propagation (FP) in massive multiple-input multiple-output (MIMO) line-of-sight (LOS) channels is studied. For arbitrary array geometry and under independent (possibly non-Gaussian) errors, the FP property is shown to hold for the perturbed array as long as it holds for the unperturbed one. This means that small errors do not have catastrophic impact on the FP, even for a large number of antennas. The negative impact of random errors is to slow down the convergence to the asymptotic value so that more antennas are needed under random errors to achieve the same low inter-user interference as without errors. For large but finite number of antennas, the distribution and an analytically-tractable approximation of the inter-user interference power are obtained. Practical design guidelines are given that quantify the accuracy level needed to make the impact of random errors negligible. The analytical results are validated via numerical simulations and are in agreement with measurement-based studies.**

***Index Terms*— Massive MIMO, favorable propagation, inter-user interference, robustness, location/phase errors.**

I. Introduction

SINCE the seminal work by Marzetta [1], massive MIMO (mMIMO) has been attracting significant and increasing attention, both in academia [2], [3], [4] and industry, especially for 5/6G applications [5], [6]. Its main advantage is a significant increase in spectral and energy efficiency as well as simplified processing in multi-user environments [2], [3], [4], [5], [6], [7]. This is due to a phenomenon known as "favorable propagation" (FP), whereby different users' channels become orthogonal (or nearly so) to each other when the number of base station antennas increases, thereby substantially reducing inter-user interference (IUI), even with simple linear processing [7]. Although these channels are not exactly orthogonal to each other in practice, they become nearly orthogonal under certain conditions and a significant part of mMIMO benefits can be achieved in this case [14], [15], [16], [17], [18], [19]. Massive MIMO has also significant advantages for single-user applications, including high spectral efficiency and large multiplexing/diversity gains [20], [21].

Manuscript received 20 October 2022; revised 8 February 2023 and 19 April 2023; accepted 21 April 2023. Date of publication 26 May 2023; date of current version 9 January 2024. An earlier version of this paper was presented in part at the IEEE Information Theory Workshop, Saint-Malo, France, in April 2023. The associate editor coordinating the review of this article and approving it for publication was Z. Ding. *(Corresponding author: Sergey Loyka.)*

The authors are with the School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, ON K1N 6N5, Canada (e-mail: sergey.loyka@ieee.org).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TWC.2023.3278490.

Digital Object Identifier 10.1109/TWC.2023.3278490

The FP property has been studied both theoretically [7], [8], [9], [10], [11], [12] and experimentally [14], [15], [16], [17], [18], [19]; antenna array geometry and propagation channel properties were identified as key factors affecting the FP. In particular, it was shown, using the law of large numbers, that the FP holds in i.i.d. fading channels [2], [7] [8], [20]. However, the i.i.d. fading assumption neglects the impact of antenna array geometry and is justified provided that (i) multipath is rich enough (without a single dominant component) and (ii) antenna spacing is large enough. If either of these conditions is violated, e.g. if there is a dominant line-of-sight (LOS) component, then the i.i.d. assumption does not hold anymore. In fact, the LOS environment is extreme opposite of i.i.d. fading and is considered to be "particularly difficult" for users' orthogonality and mMIMO system performance [15], [19]; the law of large numbers is not applicable in this case. Real-world channels are expected to be somewhere in-between of these 2 extremes [7], [11]. mMIMO in LOS environment was studied in [7], [8], [9], [10], [11], and [12] and the FP was shown to hold for uniform linear, planar, circular and cylindrical arrays (ULA, UPA, UCA and UCLA) of fixed element spacing and distinct angles-of-arrival (AoA) of different users (or for uniformly-random users) but it does not hold for a fixed antenna array size, clearly showing the impact of antenna array geometry and related constraints. While increasing the array size improves the performance due to higher spacial resolution, especially in LOS channels with closely-located users, this fails to hold if grating lobes appear (due to large element spacing) and some users align with their directions [13], [18]. In the latter case, non-uniform array design can be used to eliminate this negative effect.

Measurement-based studies with large numbers of antennas [14], [15], [16], [17], [18], [19] confirmed the general tendencies observed in the theoretical studies, i.e. performance improvement with increasing number of antennas and antenna array size, but also revealed some key differences. Namely, this improvement falls behind that predicted by the i.i.d. fading model and, at certain point, it starts saturating, indicating that channel correlation plays a role here. LOS environment with closely-located users was identified as the most difficult one while non-LOS environment is somewhat better for channel orthogonality [15], [19].

The above-mentioned theoretical studies assume perfect channel knowledge/estimation, perfect element location or array calibration, no inaccuracies in beamforming weights etc. In practice, such perfect setting is hardly possible as implementation inaccuracies and tolerances always exist. It is not clear whether the FP property will still hold under such inaccuracies, when the number of antennas increases without bound.

The robustness property, whereby small perturbations/errors in element locations, beamforming weights etc. do not have a catastrophic impact on the performance, is desirable from a practical perspective, especially for mmWave and THz systems, where the impact of implementation inaccuracies becomes more pronounced since the wavelength becomes much smaller; in addition, the presence of phase noise also becomes a major issue affecting the system performance.

The impact of implementation inaccuracies/errors on the traditional antenna arrays has been studied and robust beamforming strategies have been proposed, including the well-known diagonal loading technique and its modifications [28], [29], [30], [31]. While some studies advocate random error model [28], others use deterministic normed or bounded uncertainty models and worst-case performance optimization [29], [30], [31]. A general conclusion is that, under proper design, small errors do not have a catastrophic impact on the performance. Likewise, the impact of channel uncertainty on traditional MIMO system performance was also studied in the classical setting using information-theoretic tools (compound channel capacity) and it was shown that small normed uncertainties do not have large impact under proper transmitter and receiver designs [32], [33]. However, the above results were established in the traditional settings (not mMIMO) and it remains unclear whether they still holds in the mMIMO setting and apply to the FP property as well (especially because increasing the number of antennas to very large values has a potential to "amplify" small per-element errors and generate a large aggregate effect thereby destroying the FP property). In a related line of work, the impact of hardware imperfections on mMIMO energy and spectral efficiencies in Rayleigh-fading channels was studied in [34] using an additive noise model of imperfections. However, their impact on the FP was not studied so that it remains unclear whether the FP holds under random errors, especially in LOS channels, which behave differently to i.i.d. Rayleigh-fading ones and are particularly difficult for the FP.

This paper studies the robustness of the FP in mMIMO LOS channels to random errors in element locations and beamforming phases (due to implementation inaccuracies, quantization effects, imperfect array calibration, component aging, environmental effects, etc.). Both Gaussian and non-Gaussian error distributions are considered and the following three questions are addressed:

1) Does the FP hold under random errors? If so, under what conditions?
2) What is the IUI scaling with the number of antennas under random errors?
3) How accurate does a design need to be for the impact of random errors to be negligible?

We show that, for an arbitrary array geometry (including ULA, UPA, UCA and UCLA as special cases), the FP holds for the perturbed array under random independent errors as long as it holds for the nominal array (i.e. the one without errors). From a practical perspective, this means that small inaccuracies in array implementation do not "blow up" even if the number of array elements is large and that small levels of IUI are achievable under random errors (provided the number of antennas is large enough).

While random errors do not affect the FP asymptotically, they have a profound negative impact on the convergence speed to the asymptotic value as the number N of antennas increases: while the IUI power scales as N^{-2} for the nominal array (no errors), it scales only as N^{-1} for the perturbed array, i.e. the impact of random errors is to slow down the convergence from N^{-2} (20 dB per decade) to N^{-1} (10 dB per decade) so that more antennas are needed to achieve the same low IUI under random errors. For a finite number of antennas, the impact of random errors is shown to be qualitatively different in small and large perturbation regimes. A quantitative condition for this impact to be negligible is given, which also determines the required implementation accuracy for a given number of antennas and IUI level. For large but finite N, the IUI power under random i.i.d. errors is shown to follow the non-central chi-squared distribution with 2 degrees of freedom. Based on this, an analytically-tractable approximation for the IUI power is proposed.

To compare the above IUI scaling under random errors to that without errors, note from [11] that, in the latter case, the IUI power scales as $N^{-2}\ln^2 N$ in the LOS channel with uniformly-random users but only as N^{-1} in the i.i.d. Rayleigh fading channel, i.e. the IUI power decreases with N much faster in the LOS channel when there are no errors. However, as our results above indicate, this IUI scaling in LOS channel slows down from N^{-2} to N^{-1} when random location and phase errors are present, i.e. becomes the same as in the i.i.d. Rayleigh fading channel.

The rest of the paper is organized as follows. Sec. II introduces the system model and the FP property. The impact of random (possibly non-Gaussian) errors is introduced and studied in Sec. III; the main result is in Theorem 1. Sec. IV gives the distribution and an analytically-tractable estimate of the IUI power for large but finite number of antennas. Sec. V illustrates the asymptotic results of Theorem 1 for a finite number of antennas and gives some design guidelines as to what accuracy is needed to make errors negligible. It also compares the finite-N scaling of the IUI power with and without errors. The analytical results are validated via numerical simulations and are in agreement with measurement-based studies. Sec. VI concludes the paper.

Notations: lower-case bold letters denote column vectors while regular letters are scalars; $|\boldsymbol{h}|$, $\boldsymbol{h}'$ and $\boldsymbol{h}^+$ denote Euclidean norm (length), transposition and Hermitian conjugation, respectively, of vector $\boldsymbol{h}$; $\mathbb{E}\{\cdot\}$ is statistical expectation while $\mathrm{Var}\{\cdot\}$ is the variance (with respect to relevant random variables), $\mathrm{Re}\{\cdot\}$ and $\mathrm{Im}\{\cdot\}$ are the real and imaginary parts of a complex number.

II. System Model and FP

We use the standard basedband model [1], [2], [3], [7], [8], [9], [10], [11] whereby M single-antenna users communicate simultaneously to a base station (BS) equipped with an N-element antenna array:

$$\boldsymbol{y}(t) = \boldsymbol{h}_1 x_1(t) + \sum_{i=2}^{M} \boldsymbol{h}_i x_i(t) + \boldsymbol{\xi}(t) \tag{1}$$

where $\boldsymbol{y}(t)$ is the vector signal received by the BS at time t, $\boldsymbol{\xi}(t)$ is a zero-mean white Gaussian circularly-symmetric noise vector of variance σ_0^2 per dimention; $\boldsymbol{h}_i$ and $x_i(t)$ are the channel vector and transmitted signal of user i, $i = 1 \ldots M$, respectively.

To detect user 1 signal $x_1(t)$ (the main user), a linear (low-complexity) beamforming $\boldsymbol{w}_1^+ \boldsymbol{y}(t)$ is used by the BS [7], where $\boldsymbol{w}_1$ is the beamforming vector for detecting user 1, and the other users' contribution $\boldsymbol{w}_1^+ \sum_{i=2}^M \boldsymbol{h}_i x_i(t)$ is treated as interference. Since $x_1(t), .., x_M(t)$ are independent of each other and also of $\boldsymbol{\xi}(t)$, the main user SINR can be expressed as follows:

$$\begin{aligned} \text{SINR} &= \frac{|\boldsymbol{w}_1^+ \boldsymbol{h}_1|^2 \sigma_{x_1}^2}{\sum_{i=2}^M |\boldsymbol{w}_1^+ \boldsymbol{h}_i|^2 \sigma_{x_i}^2 + |\boldsymbol{w}_1|^2 \sigma_0^2} && (2)\\ &= \frac{|\alpha_{11,N}|^2 \gamma_1}{\sum_{i=2}^M |\alpha_{1i,N}|^2 \gamma_i + 1} && (3)\\ &\leq |\alpha_{11,N}|^2 \gamma_1 && (4)\\ &\leq \gamma_1 && (5)\end{aligned}$$

where $\sigma_{xi}^2 = \mathbb{E}\{|x_i(t)|^2\}$ and $\gamma_i = |\boldsymbol{h}_i|^2 \sigma_{xi}^2 / \sigma_0^2$ are the signal power and the single-user SNR of user i, respectively, and

$$\alpha_{1i,N} = \frac{\boldsymbol{w}_1^+ \boldsymbol{h}_i}{|\boldsymbol{w}_1||\boldsymbol{h}_i|} \quad (6)$$

where $|\alpha_{11,N}|^2 \leq 1$ is the normalized channel power gain of the main user (the normalization is with respect to the no-error case, which corresponds to $|\alpha_{11,N}|^2 = 1$) and $|\alpha_{11,N}|^2 \gamma_1$ is its after-beamforming SNR, which may also include beamforming errors that induce $|\alpha_{11,N}|^2 < 1$; $|\alpha_{1i,N}|^2 \leq 1$, $i = 2 \ldots M$, is the IUI power "leakage" factor of user i to the main user and $|\alpha_{1i,N}|^2 \gamma_i$ is the "leaked" IUI power measured in σ_0^2. In fact, $|\alpha_{1i,N}|$ is the measure of the FP (or channel orthogonality) adopted in the reported studies, both theoretical and experimental [2], [3], [7], [8], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18]. The channel is normalized so that $|\boldsymbol{h}_i|^2 = N$ and the propagation path loss is absorbed into the single-user SNR γ_i. While the model in (1) applies directly to the uplink (users-to-BS), the user 1 SINR in (2) also applies to the downlink (BS-to-users) if (i) equal amount of channel state information is available in both cases and (ii) the channel is reciprocal. As a consequence, if the FP holds for uplink, it will also hold for downlink and vice versa.

In the case of no perturbations/errors, the channel is known precisely and the beamforming weights are also set precisely, as is usually assumed in the literature [1], [2], [3], [7], [8], [9], [10], [11], [12]. In this case and for the matched filter beamforming (also known as maximum ratio combining, which maximizes the single-user SNR), $\boldsymbol{w}_1 = \boldsymbol{h}_1$ so that $\alpha_{11,N} = 1$ and hence the inequality in (5) becomes equality. Under this condition, the upper bound in (4) is attained with equality and thus the SINR is maximized achieving its single-user value, if the total IUI power "leakage" vanishes, $\sum_{i=2}^M |\alpha_{1i,N}|^2 = 0$. This favorable condition can be approached, in some cases, by increasing the number of antennas without bound. This is known as "asymptotically favorable propagation" [7], [8], [9], [10], [11], [12]. Since the total IUI is rarely exactly zero for a finite number of antennas, we further omit "asymptotically" and use the term "favorable propagation" (FP) with understanding that $N \to \infty$ is included in its definition. For a finite number of users, $M < \infty$, and for uniformly-bounded per-user SNRs, $\gamma_i \leq A < \infty$ for all i, where A is independent of N (but γ_i may depend on N, as they usually do in practice), the total IUI vanishes and thus the FP holds if

$$\lim_{N\to\infty} \sum_{i=2}^M |\alpha_{1i,N}|^2 = 0 \text{ or } \lim_{N\to\infty} |\alpha_{1i,N}|^2 = 0 \quad \forall i > 1 \quad (7)$$

so that the SINR achieves its maximum, SINR $= \gamma_1$, i.e. its single-user value. Note that, when the FP holds, the performance of matched filtering (MF), also known as maximum-ratio combining (MRC), zero-forcing (ZF) and MMSE receivers is the same where the MF/MRC is attractive due to its simplicity, robustness and suitability for parallel/distributed implementation.[1] While $N \to \infty$ is not possible in practice, an important practical implication of (7), which follows from the limit definition, is that the IUI power can be made as small as desired provided the number of antennas is large enough. This is not the case anymore if the FP does not hold. A metric alternative to the FP measure in (6)(7) was introduced in [19], namely, "angle to interference subspace". It is not difficult to see that these two metrics are equivalent, i.e. if the FP holds for one, it also holds for the other.

In the case of random errors, (7) cannot be used anymore since $\alpha_{1i,N}$ becomes a random sequence and thus the limits in (7) do not exist (in the deterministic sense). Hence, an extension of the FP condition in (7) is needed to accommodate random errors. This is done in the next section.

III. FP Under Location and Phase Errors

Let us consider an N-element antenna array of arbitrary geometry (including ULA, UPA, UCA and UCLA as special cases) where element locations as well as beamforming phases are subject to random errors (perturbations) due to e.g. implementation and array calibration inaccuracies, quantization errors, aging, environmental effects etc. In particular, the actual location vector $\boldsymbol{p}_n$ of n-th antenna array element is

$$\boldsymbol{p}_n = \boldsymbol{p}_n^0 + \Delta\boldsymbol{p}_n, \quad n = 1 \ldots N \quad (8)$$

where $\boldsymbol{p}_n^0$ is the nominal location vector and $\Delta\boldsymbol{p}_n$ is its random offset, all measured in wavelengths. Following [7], [8], [9], [10], [11], [12], we consider the LOS environment, primarily due to the following reasons: (i) it is a particularly difficult environment for the FP to hold and it also represents the extreme opposite of i.i.d. fading; (ii) there are many LOS application scenarios, including cellular, WiFi, wireless backhaul, short-range, UAV and satellite systems [17], [22], [23], [24]; (iii) unlike below-6 GHz systems, the LOS environment is critical for emerging mmWave and THz systems (key technologies for 5/6G, most likely to employ mMIMO) [6], [25], [26], [27]. Under LOS propagation in the far field, the normalized channel vector entries for user i can be expressed as [7] and [28]

$$h_{in} = \exp(j2\pi \boldsymbol{u}_i^+ \boldsymbol{p}_n), \quad i = 1..M, \; n = 1 \ldots N, \quad (9)$$

[1] One should further note that, under the FP, the rate delivered by the MF/MRC is also the same as that of the more-complicated (nonlinear) but capacity-optimal successive interference cancellation receiver [42] in this case (since there is no interference to cancel under the FP).

where i and n are user and element indexes, $\boldsymbol{u}_i$ is the unit direction vector for user i.

When matched filtering is used to detect user 1 signal, the beamforming weights $\boldsymbol{w}_1 = [w_1, \ldots, w_N]'$ are matched to user 1 nominal (rather than actual) channel vector $\boldsymbol{h}_1^0$, whose entries are

$$h_{1n}^0 = \exp(j2\pi \boldsymbol{u}_1^+ \boldsymbol{p}_n^0), \tag{10}$$

and are perturbed in phase as well, typically due to imperfect phase shifters and quantization errors, so that

$$w_n = \exp(j2\pi \boldsymbol{u}_1^+ \boldsymbol{p}_n^0 + j\Delta\phi_n) \tag{11}$$

where $\Delta\phi_n$ are beamforming phase errors; following the widely-accepted approach [28], they are modeled as zero-mean i.i.d. random variables. Likewise, $\Delta \boldsymbol{p}_n$ are also modeled as zero-mean i.i.d. random vectors. Unlike the existing studies, we do not assume here that they are Gaussian, so that non-Gaussian distributions are included as well.

Using (6) and (8)-(11), $\alpha_{1i,N}$ can be expressed as

$$\alpha_{1i,N} = \frac{1}{N}\sum_{n=1}^{N} w_n^* h_{in} = \frac{1}{N}\sum_{n=1}^{N} e^{j\Psi_{in}}, \tag{12}$$

$$\Psi_{in} = \Psi_{in}^0 + \Delta\Psi_{in} \tag{13}$$

$$\Psi_{in}^0 = 2\pi(\boldsymbol{u}_i - \boldsymbol{u}_1)^+ \boldsymbol{p}_n^0 \tag{14}$$

$$\Delta\Psi_{in} = \Delta\psi_{in} - \Delta\phi_n \tag{15}$$

where Ψ_{in}^0 represents phase differences of i-th and 1st users' signals for the nominal array; $\Delta\psi_{in} = 2\pi \boldsymbol{u}_i^+ \Delta \boldsymbol{p}_n$ represents uncompensated phase shifts due to element location errors. The corresponding IUI leakage factor of the nominal array is

$$\alpha_{1i,N}^0 = \frac{1}{N}\sum_{n=1}^{N} e^{j\Psi_{in}^0} \tag{16}$$

Since, under random errors above, $|\alpha_{1i,N}|^2$ is a random sequence (indexed by N), (7) cannot be used since the respective limits do not exist in the deterministic sense. Therefore, one has to use a notion of stochastic convergence. The following definition gives three such notions widely used in many areas of stochastic analysis and applications; this is a slight extension of the definitions in [35, p. 306] and [36, Ch. 6].

Definition 1: A sequence $z_1, z_2, \ldots$ of random variables converges to a deterministic sequence $a_1, a_2, \ldots$ in the mean square sense (mse), $z_N \xrightarrow{mse} a_N$, if

$$\lim_{N\to\infty} \mathbb{E}\{|z_N - a_N|^2\} = 0 \tag{17}$$

z_N converges in probability to a_N, $z_N \xrightarrow{\Pr} a_N$, if, for any $\epsilon > 0$,

$$\lim_{N\to\infty} \Pr\{|z_N - a_N| > \epsilon\} = 0 \tag{18}$$

z_N converges to a_N almost surely (a.s.) or with probability one, $z_N \xrightarrow{a.s.} a_N$, if

$$\Pr\left\{\lim_{N\to\infty} (z_N - a_N) = 0\right\} = 1 \tag{19}$$

where the probability applies to the set of all events where the limit exists and equals to 0.

We will further use $\to$ (without a superscript) to denote stochastic convergence in all three senses above as well as the regular (deterministic) convergence when all related sequences are deterministic. Note that convergence in the mean square sense and almost sure convergence imply convergence in probability but the converse is not true in general, see e.g. [35, p. 310] [36, p. 130-132] for examples and further discussion.

A few remarks are in order as to why these three different modes of convergence are needed here. First, mean square error is a time-tested tool in many areas of communications, signal processing and stochastic control, including robust beamforming [31], so its use is appropriate here. It ensures that random IUI z_N converges to its mean value a_N in the MSE sense. Second, $\Pr\{|z_N - a_N| > \epsilon\}$ is the probability that the deviation of random IUI from its mean is not small. This is akin to outage probability widely used in may areas of wireless communications. In fact, it becomes the outage probability if one designs a system based on the mean IUI and the actual random IUI deviates significantly from this design. Lastly, (19) is needed since, even if (18) holds, it does not guarantee that $|z_N - a_N|$ cannot become arbitrary large (i.e., large deviation) for infinitely-many N [37, p. 237]. The guarantee that $|z_N - a_N|$ becomes and *stays* small as N increases is provided by (19) (the set of all events where this does not hold has a combined probability measure of zero, i.e. extremely unlikely to be encountered in the real world) [36, p. 315]. It comes the closest to the deterministic convergence and guarantees that once a mMIMO design is acceptable for a given (large) number N_0 of antennas, it will also *remain* acceptable for any $N > N_0$. In fact, the results we establish below hold for all three modes of convergence, which also ensures that they are not an artifact of a particular definition used.

Next, we replace the deterministic limits in the FP definition in (7) by the stochastic convergence modes in Definition 1. Upon this replacement, the FP holds under random perturbations if, as $N \to \infty$,

$$|\alpha_{1i,N}|^2 \to 0 \quad \forall i > 1 \tag{20}$$

The following Theorem shows that the FP property holds for a perturbed array under i.i.d. perturbations defined above if it holds for the respective nominal array, i.e. i.i.d. random errors do not have a catastrophic impact on the FP. This becomes "if and only if" (iff) under certain conditions and holds for all three convergence definitions above and, hence, is insensitive to a particular definition used.

Theorem 1: Under i.i.d. perturbations of finite variance, the FP property of the channel in (8)-(11) can be characterized as follows:

1) *If perturbations are Gaussian: the FP holds for a perturbed array if and only if it holds for a nominal (unperturbed) one, i.e., for any $i > 1$,*

$$|\alpha_{1i,N}|^2 \to 0 \quad \textit{iff} \quad \lim_{N\to\infty} |\alpha_{1i,N}^0|^2 = 0 \tag{21}$$

2) *If perturbations are non-Gaussian: (21) holds if $c_i \triangleq \mathbb{E}\{e^{j\Delta\Psi_{in}}\} \neq 0$; if $c_i = 0$, then the FP always holds for a perturbed array and 2nd condition in (21) is not necessary.*

3) *The following convergence holds in all considered cases for $i \geq 1$:*

$$|\alpha_{1i,N}|^2 \to \mathbb{E}\{|\alpha_{1i,N}|^2\} \tag{22}$$

$$= |c_i|^2|\alpha^0_{1i,N}|^2 + (1-|c_i|^2)N^{-1} \quad (23)$$
$$\to |c_i|^2|\alpha^0_{1i,N}|^2 \quad (24)$$

Proof: First, we establish (22)-(24), from which (21) will follow. We begin with a proof of the above claims for convergence in the MSE sense. Then, convergence in probability and almost sure convergence are established. To simplify notations, we further use $\alpha_N = \alpha_{1i,N}$, $\alpha^0_N = \alpha^0_{1i,N}$ for $i > 1$.

A. Convergence in the MSE Sense

The definition in (17) with $z_N = |\alpha_N|^2$ and $a_N = \mathbb{E}\{|\alpha_N|^2\}$ is equivalent to

$$\begin{aligned}\mathrm{Var}\{|\alpha_N|^2\} &= \mathbb{E}\{(|\alpha_N|^2 - \mathbb{E}\{|\alpha_N|^2\})^2\} \\ &= \mathbb{E}\{|\alpha_N|^4\} - (\mathbb{E}\{|\alpha_N|^2\})^2 \to 0 \end{aligned} \quad (25)$$

Finding the variance in (25) is rather involved since the fourth moment analysis is complicated. To circumvent this difficulty, we present an upper bound and prove that it converges to zero as $N \to \infty$. To this end, the following Lemma shows that $\mathbb{E}\{|\alpha_N|^4\}$ can be "sandwiched" via $|\mathbb{E}\{\alpha_N\}|^4$.

Lemma 1: $\mathbb{E}\{|\alpha_N|^4\}$ can be bounded as follows:

$$|\mathbb{E}\{\alpha_N\}|^4 \le \mathbb{E}\{|\alpha_N|^4\} \le |\mathbb{E}\{\alpha_N\}|^4 + 12N^{-1} \quad (26)$$

Proof: See Appendix A. □

It follows from (26) that $\mathbb{E}\{|\alpha_N|^4\}$ and $|\mathbb{E}\{\alpha_N\}|^4$ behave similarly as $N \to \infty$:

$$\mathbb{E}\{|\alpha_N|^4\} \to 0 \ \text{ iff } \ |\mathbb{E}\{\alpha_N\}|^4 \to 0 \quad (27)$$

The next Lemma establishes the needed upper bound.

Lemma 2: $\mathrm{Var}\{|\alpha_N|^2\}$ can be upper bounded, for any N, as follows:

$$\mathrm{Var}\{|\alpha_N|^2\} \le 12N^{-1} \quad (28)$$

Proof: Since $|\cdot|^2$ is convex, it follows from Jensen's inequality [40] that

$$\mathbb{E}\{|\alpha_N|^2\} \ge |\mathbb{E}\{\alpha_N\}|^2 \quad (29)$$

Combining this with the upper bound in (26), one obtains

$$\begin{aligned}\mathrm{Var}\{|\alpha_N|^2\} &= \mathbb{E}\{|\alpha_N|^4\} - (\mathbb{E}\{|\alpha_N|^2\})^2 \\ &\le |\mathbb{E}\{\alpha_N\}|^4 + 12N^{-1} - |\mathbb{E}\{\alpha_N\}|^4 = 12N^{-1}\end{aligned}$$

as required. While this is not the tightest possible bound and, in many cases, the variance decreases faster with N (e.g. as N^{-2} - see Fig. 2), it is sufficient for our purpose here. □

Now, using (28) and (25),

$$\mathrm{Var}\{|\alpha_N|^2\} = \mathbb{E}\{(|\alpha_N|^2 - \mathbb{E}\{|\alpha_N|^2\})^2\} \le 12N^{-1} \to 0$$

which establishes the MSE convergence,

$$|\alpha_N|^2 \xrightarrow{mse} \mathbb{E}\{|\alpha_N|^2\} \quad (30)$$

i.e. (22) in the MSE sense. The next Lemma establishes (23).

Lemma 3: The following holds for $\alpha_N = \alpha_{1i,N}$ as defined in (12):

$$\mathbb{E}\{\alpha_N\} = c_i\alpha^0_N, \quad c_i = \mathbb{E}\{e^{j\Delta\Psi_{in}}\} \quad (31)$$
$$\mathbb{E}\{|\alpha_N|^2\} = (1-|c_i|^2)N^{-1} + |c_i|^2|\alpha^0_N|^2 \quad (32)$$

Proof: Let us first evaluate $\mathbb{E}\{|\alpha_N|^2\}$ using (12):

$$\begin{aligned}\mathbb{E}\{|\alpha_N|^2\} &= \frac{1}{N^2}\sum_{n,m}\mathbb{E}\{e^{j(\Psi_{in}-\Psi_{im})}\} \\ &= \frac{1}{N} + \frac{1}{N^2}\sum_{m\neq n}\mathbb{E}\{e^{j\Psi_{in}}\}\mathbb{E}\{e^{-j\Psi_{im}}\} \quad (33) \\ &= \frac{1}{N} + \frac{1}{N^2}\sum_{n}\mathbb{E}\{e^{j\Psi_{in}}\}\sum_{m}\mathbb{E}\{e^{-j\Psi_{im}}\} \\ &\quad - \frac{1}{N^2}\sum_{n}\mathbb{E}\{e^{j\Psi_{in}}\}\mathbb{E}\{e^{-j\Psi_{in}}\} \\ &= \frac{1}{N} + |\mathbb{E}\{\alpha_N\}|^2 - \frac{1}{N^2}\sum_{n}|\mathbb{E}\{e^{j\Psi_{in}}\}|^2 \end{aligned} \quad (34)$$

where (33) follows from independence of Ψ_{in} and Ψ_{im} for $n \neq m$. Next, using (12) and (16), one obtains:

$$\mathbb{E}\{\alpha_N\} = \frac{1}{N}\sum_{n=1}^{N}\mathbb{E}\{e^{j\Psi_{in}}\} = \frac{c_i}{N}\sum_{n=1}^{N}e^{j\Psi^0_{in}} = c_i\alpha^0_N \quad (35)$$

where

$$\mathbb{E}\{e^{j\Psi_{in}}\} = \mathbb{E}\{e^{j(\Psi^0_{in}+\Delta\Psi_{in})}\} = c_i e^{j\Psi^0_{in}} \quad (36)$$

Using (34)-(35), (32) follows:

$$\mathbb{E}\{|\alpha_N|^2\} = N^{-1} + |c_i|^2|\alpha^0_N|^2 - N^{-1}|c_i|^2 \quad (37)$$

This completes the proof of Lemma 3. □

Finally, note that (32) implies the following convergence as $N \to \infty$:

$$\mathbb{E}\{|\alpha_N|^2\} \to |c_i|^2|\alpha^0_N|^2 \quad (38)$$

This establishes (22)-(24), where perturbations do not have to be Gaussian (no such assumption was made in the above derivations). For Gaussian perturbations, from (56), $c_i = \mathbb{E}\{e^{j\Delta\Psi_{in}}\} = e^{-\delta^2/2} \neq 0$ and therefore, from (38),

$$\lim_{N\to\infty}\mathbb{E}\{|\alpha_N|^2\} = 0 \quad \text{iff} \quad \lim_{N\to\infty}|\alpha^0_N|^2 = 0 \quad (39)$$

and, since $|\alpha_N|^2 \xrightarrow{mse} \mathbb{E}\{|\alpha_N|^2\}$ as established in (30) above, (21) follows (in the MSE sense):

$$|\alpha_N|^2 \xrightarrow{mse} 0 \quad \text{iff} \quad \lim_{N\to\infty}|\alpha^0_N|^2 = 0 \quad (40)$$

For non-Gaussian perturbations, $c_i = 0$ is possible, in which case, from (30) and (32),

$$|\alpha_N|^2 \xrightarrow{mse} \mathbb{E}\{|\alpha_N|^2\} \to 0 \quad (41)$$

even if $\lim_{N\to\infty}|\alpha^0_N|^2 \neq 0$, i.e. the condition $\lim_{N\to\infty}|\alpha^0_N|^2 = 0$ is not necessary. This completes the proof for the MSE convergence.

B. Convergence in Probability

To establish (22) with convergence in probability, we use (18) with $z_N = |\alpha_N|^2$, $a_N = \mathbb{E}\{|\alpha_N|^2\}$ and Chebyshev inequality [35, p. 46]:

$$\Pr\{|z_N - \mathbb{E}\{z_N\}| > \epsilon\} \le \mathrm{Var}\{z_N\}\epsilon^{-2} \quad (42)$$

which holds for any $\epsilon > 0$. Using (42) and (28), one obtains:

$$\begin{aligned}\lim_{N\to\infty} \Pr\{||\alpha_N|^2 - \mathbb{E}\{|\alpha_N|^2\}| > \epsilon\} \\ \leq \lim_{N\to\infty} \mathrm{Var}\{|\alpha_N|^2\}\epsilon^{-2} \\ \leq \lim_{N\to\infty} 12(N\epsilon^2)^{-1} = 0 \end{aligned} \quad (43)$$

This establishes (22) with convergence in probability,

$$|\alpha_N|^2 \xrightarrow{\Pr} \mathbb{E}\{|\alpha_N|^2\} \quad (44)$$

Combining it with (23) and (24) established above (note that they apply to the deterministic sequences and therefore do not depend on the stochastic convergence definitions), we finally obtain (21) with convergence in probability.

C. Almost Sure Convergence

To establish (22) with almost sure convergence,[2] $|\alpha_N|^2 \xrightarrow{a.s.} \mathbb{E}\{|\alpha_N|^2\}$, we need the following technical Lemma, which is an extension of [38, Theorem 1] to two-dimensional sequences with non-zero means.

Lemma 4: Consider the empirical mean z_N of N^2 uniformly-bounded random variables z_{nm},

$$z_N = \frac{1}{N^2}\sum_{n,m=1}^{N} z_{nm} \quad (45)$$

Then, $z_N \xrightarrow{a.s.} \mathbb{E}\{z_N\}$ if:

$$\text{(a) } |z_{nm}| \leq 1 \text{ and (b) } \sum_{N=1}^{\infty}\frac{1}{N}\mathrm{Var}\{z_N\} < \infty \quad (46)$$

Proof: See Appendix B. □

Now, use Lemma 4 with $z_N = |\alpha_N|^2$,

$$z_N = |\alpha_N|^2 = \frac{1}{N^2}\sum_{n,m} e^{j(\Psi_{in}-\Psi_{im})} \quad (47)$$

which is the empirical mean of N^2 random variables $z_{nm} = \exp\{j\Psi_{in} - j\Psi_{im}\}$ as in (45) and observe that $|z_{nm}| \leq 1$, which satisfies (46)(a). Using (28), $\mathrm{Var}\{z_N\} \leq 12N^{-1}$ so that

$$\sum_{N\geq 1} N^{-1}\mathrm{Var}\{z_N\} \leq 12\sum_{N\geq 1} N^{-2} = 2\pi^2 < \infty \quad (48)$$

i.e. (46)(b) is also satisfied. Therefore, $|\alpha_N|^2 \xrightarrow{a.s.} \mathbb{E}\{|\alpha_N|^2\}$, as required. This establishes (22) with almost sure convergence. Combining it with (23) and (24) established above, (21) follows with almost sure convergence. □

IV. The Distribution and a Bound for $|\alpha_N|^2$

Eq. (22) implies that, for large N, $\mathbb{E}\{|\alpha_N|^2\}$ can be used as an estimate of the actual IUI $|\alpha_N|^2 = |\alpha_{1i,N}|^2$. However, as Fig. 3 below shows, random fluctuations of $|\alpha_N|^2$ do exists for finite N and, hence, $|\alpha_N|^2$ can exceed $\mathbb{E}\{|\alpha_N|^2\}$. This possibility should be taken into account for a more reliable design with finite N. A simple way to accomplish this is via the following approximate upper bound, which accounts for random fluctuations and holds with high probability for large N and sufficiently-large m,

$$|\alpha_N|^2 \lesssim |\alpha_N^{up}|^2 = \mathbb{E}\{|\alpha_N|^2\} + m\sigma_{|\alpha_N|^2} \quad (49)$$

where $\sigma^2_{|\alpha_N|^2} = \mathrm{Var}\{|\alpha_N|^2\}$ is the variance, $m = 1\ldots 3$ controls the outage probability (i.e. the probability that $|\alpha_N|^2$ exceeds the bound $|\alpha_N^{up}|^2$) and the design is based on $|\alpha_N^{up}|^2$. As Fig. 3 below shows, using $m = 0$ is not sufficient and $m = 1\ldots 3$ provides more reliable design, with larger m corresponding to smaller outage probability (note that $m = 3$ corresponds to the well-known three-sigma rule).

The following Proposition gives finite-N approximations for the distribution of $|\alpha_N|^2$ and its variance, which can be used in (49).

Proposition 1: For large N and if the FP holds for the nominal (no error) array at double the element spacing, i.e.

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} e^{j2\Psi^0_{in}} = 0, \quad (50)$$

$|\alpha_N|^2$ is distributed as a non-central chi-squared random variable $\chi^2_2(\lambda)$ with 2 degrees of freedom,

$$|\alpha_N|^2 \sim \frac{1}{2}\sigma_N^2\chi_2^2(\lambda) \quad (51)$$

where $\lambda = 2\sigma_N^{-2}|\,\mathbb{E}\{\alpha_N\}|^2$ is the noncentrality parameter, and

$$\lambda = 2\, N(1-|c_i|^2)^{-1}|c_i|^2|\alpha_N^0|^2 \quad (52)$$

$$\sigma_N^2 = \mathrm{Var}\{\alpha_N\} = N^{-1}(1-|c_i|^2) \quad (53)$$

$$\sigma^2_{|\alpha_N|^2} = \mathrm{Var}\{|\alpha_N|^2\} \approx \sigma_N^2(\sigma_N^2 + 2|c_i|^2|\alpha_N^0|^2) \quad (54)$$

Proof: See Appendix C. □

The accuracy of these approximations is examined in the next section.

V. Examples and Discussion

To illustrate and validate the above results, let us consider a ULA with the nominal element spacing $d = 0.5$ measured in wavelengths, as illustrated in Fig. 1. The perturbed array has both location and phaseshift errors, which are zero-mean i.i.d Gaussian of variances σ_p^2 and σ_ϕ^2, respectively. It follows that $\Delta\Psi_{in}$ is also zero mean Gaussian of variance

$$\delta^2 = 4\pi^2\sigma_p^2 + \sigma_\phi^2 \quad (55)$$

and therefore (see e.g. [28, p. 68])

$$c_i = \mathbb{E}\{e^{j\Delta\Psi_{in}}\} = e^{-\delta^2/2} \quad (56)$$

so that $0 < c_i < 1$ if $0 < \delta < \infty$.

First, we validate the key part of the Theorem 1 proof, the asymptotic behaviour $\mathrm{Var}\{|\alpha_N|^2\} \to 0$ and also the approximation in (54). Fig. 2 illustrates the behaviour of $\mathrm{Var}\{|\alpha_N|^2\}$ as N increases for the ULA as in Fig. 1, confirming that indeed $\mathrm{Var}\{|\alpha_N|^2\} \to 0$, as expected. This also implies, from Chebyshev inequality, convergence in probability in (21). Note also that the empirical (Monte-Carlo (MC) simulated) variance and its estimate in (54) agree well with each other over the whole range of N.

The qualitatively-different behaviour of $\mathrm{Var}\{|\alpha_N|^2\}$ for smaller and larger $\sigma_{p,\phi}$ can be explained using (54) as follows.

[2]note that this convergence mode is not implied by the two convergence modes above [35, p. 310] [36, p. 130-132].

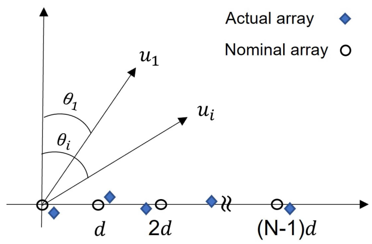

Fig. 1. N-element ULA under perturbed element locations. θ_i is the AoA of user i, where $i = 1$ corresponds to the main user, all measured from the array broadside.

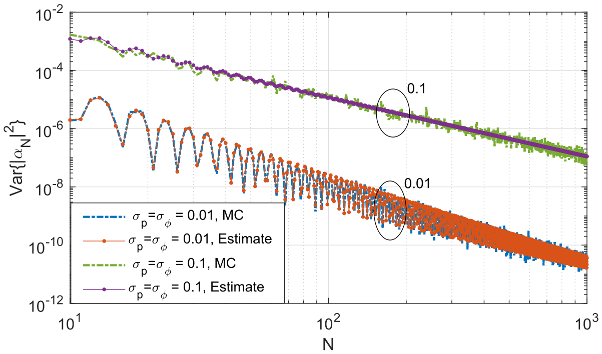

Fig. 2. Empirical Monte-Carlo (MC) simulated $\mathrm{Var}\{|\alpha_N|^2\}$ and its estimate in (54) in the presence of zero-mean i.i.d. Gaussian errors of variances σ_p^2 and σ_ϕ^2, for the ULA with $d = 0.5$, the user angles-of-arrival are $\theta_1 = 0$, $\theta_i = \pi/8$ (with respect to the array broadside). The MC variance was evaluated over 100 trials generated independently for each N.

In the large error regime $\sigma_{p,\phi} = 0.1$, σ_N^2 is larger and hence the first term in (54) dominates, $\mathrm{Var}\{|\alpha_N|^2\} \approx \sigma_N^4$, which decreases monotonically with N, as in (53). In the small error regime $\sigma_{p,\phi} = 0.01$, the 2nd term in (54) dominates until about $N = 100$ and hence $\mathrm{Var}\{|\alpha_N|^2\} \approx 2|c_i|^2|\alpha_N^0|^2\sigma_N^2$, which exhibits an oscillatory decrease with N due to $|\alpha_N^0|$.

A. Impact of Errors: Small and Large Perturbation Regimes

1) Large Perturbation Regime: Fig. 3 illustrates the behaviour of the IUI factors as N increases in the large perturbation regime. By comparing $|\alpha_N|^2$ and $|\alpha_N^0|^2$, note that perturbations have a significant impact on the IUI in this regime and cannot be neglected for any finite N and, thus, $|\alpha_N^0|^2$ is not a good approximation of $|\alpha_N|^2$. While all three IUI factors, i.e. the random IUI $|\alpha_N|^2$, its mean $\mathbb{E}\{|\alpha_N|^2\}$ and nominal $|\alpha_N^0|^2$ values, decrease with N, in agreement with Theorem 1, the 1st one exhibit the slowest decrease and also statistical fluctuations due to random location and phase errors while the nominal IUI shows the fastest decrease. In general, there is significant difference between all three IUI factors. Using the mean IUI, as done in some studies, can significantly underestimate the actual IUI level, since the former ignores statistical fluctuations around the mean and their impact is significant in this regime; the approximate upper bound $|\alpha_N^{up}|^2$ in (49) is a more reliable measure of IUI in this regime. Likewise, using the nominal IUI factor $|\alpha_N^0|^2$ will also significantly underestimate the actual IUI.

2) Small Perturbation Regime: Fig. 4 illustrates the IUI factors in the small perturbation regime with $\sigma_p = \sigma_\phi = 0.01$. Note that here, in a stark contrast to Fig. 3, all three factors, i.e. the random IUI, its mean and nominal values, behave similarly until about $N = 100$, making the impact of random errors

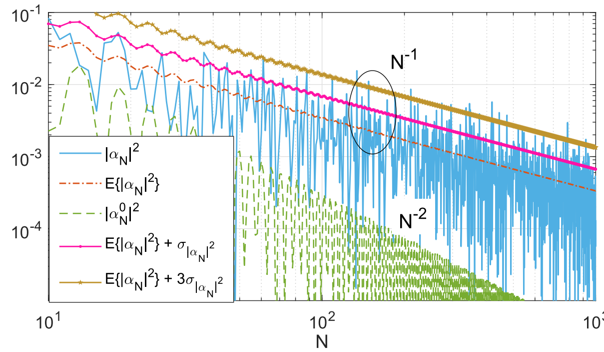

Fig. 3. IUI factor $|\alpha_N|^2$, its mean and nominal values for the N-element ULA with nominal element spacing $d = 0.5$, under zero-mean i.i.d. Gaussian perturbations with $\sigma_p = \sigma_\phi = 0.1$ (generated independently for each N); $\theta_1 = 0$, $\theta_i = \pi/8$. While all three decrease with N, $|\alpha_N|^2$ fluctuates due to random errors and $|\alpha_N^0|^2$ exhibits fastest decrease. Note that perturbations have non-negligible impact for any finite N so that $|\alpha_N^0|^2$ is not a good approximation of $|\alpha_N|^2$.

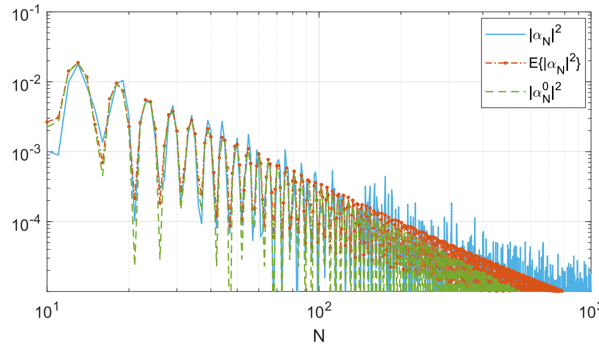

Fig. 4. The IUI factors as in Fig. 3 but for $\sigma_p = \sigma_\phi = 0.01$ (small perturbation regime). Note that, unlike Fig. 3, all three factors behave similarly until about $N = 100$, the impact of random errors is negligible and $|\alpha_N^0|^2$ is a good approximation of $|\alpha_N|^2$.

almost negligible. This can be explained via (22) and (23), whereby 1st term of (23) dominates in the small perturbation regime, for which $|c_i| \approx 1$ and therefore

$$(1 - |c_i|^2)N^{-1} \ll |c_i|^2|\alpha_N^0|^2 \tag{57}$$

so that, from (22),

$$|\alpha_N|^2 \approx \mathbb{E}\{|\alpha_N|^2\} \approx |\alpha_N^0|^2 \tag{58}$$

i.e. the random and nominal IUI leakage factors are almost the same, making the impact of random perturbations negligible. Using (56) for Gaussian perturbations, (57) is equivalent to

$$\delta^2 \ll \ln(1 + N|\alpha_N^0|^2) \tag{59}$$

which quantifies the notion of small perturbation regime, where the impact of location and phase errors is negligible and (58) holds, i.e. all 3 IUI leakage factors are approximately the same. We caution the reader not to interpret (59) as that large errors are tolerable for larger N (and especially that arbitrarily-large errors are allowed as $N \to \infty$). The reason is that α_N^0 also depends on N and, in many cases, $|\alpha_N^0|^2 \sim N^{-2}$ so that the overall scaling of the upper bound in (59) is as $\ln(1 + N^{-1}) \sim N^{-1}$, i.e. just the opposite of what naive interpretation would suggest.

Note also that random errors affect not only $|\alpha_N|^2$ but also its mean value $\mathbb{E}\{|\alpha_N|^2\}$, as evidenced by (58) as well as Fig. 3 and 4: while $\mathbb{E}\{|\alpha_N|^2\} \approx |\alpha_N^0|^2$ in the small perturbation regime (Fig. 4 until about $N = 100$), this is no

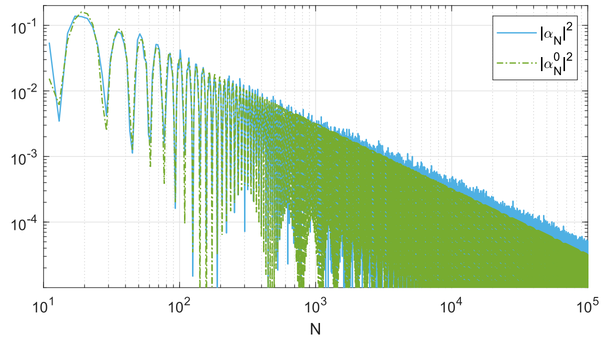

Fig. 5. IUI factor $|\alpha_N|^2$ and its nominal values for the N-element UCA and all other settings as in Fig. 3. Note the same tendencies as in Fig. 4 and significantly smaller impact of errors compared to Fig. 3.

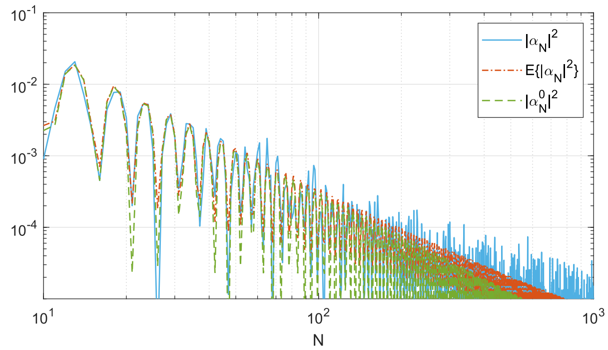

Fig. 6. IUI factor $|\alpha_N|^2$, its mean and nominal values for the N-element ULA as in Fig. 4 under uniform perturbations with $\sigma_p = \sigma_\phi = 0.01$. Note the similarity to Fig. 4, where perturbations are Gaussian.

longer the case in the large perturbation regime when, as in Fig. 3,

$$\mathbb{E}\{|\alpha_N|^2\} \approx (1-|c_i|^2)N^{-1} \tag{60}$$

The above observations are not limited to ULA geometry but also hold for other array geometries. Fig. 5 illustrates this for uniform circular array (UCA) and all other settings as in Fig. 3: while the actual values of IUI are different from those for the ULA, the general tendencies are the same as in Fig. 4 and the FP also holds under random perturbations since it holds for the nominal array, as was established in [12]. This is consistent with Theorem 1, which holds for any array geometry. Comparing Fig. 5 to Fig. 3, where the error standard deviations are the same in both cases, we conclude that the UCA is more robust to random errors (their impact is almost negligible) compared to the ULA. This can be explained by the fact that while $|\alpha_N^0|^2 \sim N^{-2}$ for the ULA, the scaling is much slower for the UCA, where $|\alpha_N^0|^2 \sim N^{-1}$ and hence the impact of random errors is not that visible for the latter.

It should be noted that the above observations also hold for non-Gaussian errors. Fig. 6 shows the IUI factors under uniformly-distributed errors with all other parameters being the same as in Fig. 4. As the comparison of these figures shows, the impact of errors is almost the same in both cases. Hence, we conclude that, under the same error variance, the shape of its distribution plays only a minor role. This is also consistent with Theorem 1 and Proposition 1, which hold for any error distribution (provided the errors are i.i.d. and of finite variance).

B. IUI Scaling With N and Design Guidelines

It follows from Theorem 1 that, if the FP holds for the nominal array, then it also holds for the perturbed one and all three IUI factors converge to zero as $N \to \infty$,

$$|\alpha_N|^2, \quad \mathbb{E}\{|\alpha_N|^2\}, \ |\alpha_N^0|^2 \to 0 \tag{61}$$

Note, however, that while their convergence point is the same, the convergence speed is significantly different: while for the nominal array in many cases (e.g. an ULA with fixed element spacing, distinct AoAs and no grating lobes)

$$|\alpha_N^0|^2 \sim N^{-2} \tag{62}$$

i.e. 20 dB per decade, for the perturbed one

$$|\alpha_N|^2, \quad \mathbb{E}\{|\alpha_N|^2\} \sim N^{-1} \tag{63}$$

i.e. 10 dB per decade, so that the impact of random errors, even if the FP holds, is to slow down the convergence from N^{-2} to N^{-1} and, hence, more antennas are needed to achieve the same low IUI leakage under random errors. As a further confirmation of this result, the $N=100$ point in Fig. 3, where the IUI is about -20 dB under random errors, agrees well with the respective experimental results in [17, Fig. 11].

The above scalings should be compared to the ones in [11], where it was observed that, without errors, the IUI power scales as $N^{-2}\ln^2 N \approx N^{-2}$ in the LOS channel with uniformly-random users (located on a sphere) but only as N^{-1} in the i.i.d. Rayleigh fading channel. However, when random location and phase errors are present, (63) indicates that, even in the LOS channel, the scaling slows down to N^{-1}, i.e. the same as in the fading channel.

We note that (59) may be difficult to use in practice since $|\alpha_N^0|^2$ depends on user AoAs, which may be unknown or known only imprecisely. On the other hand, a sensible practical design is one which can tolerate *any* IUI leakage not exceeding a certain threshold ε, i.e. any $|\alpha_N|^2 \leq \varepsilon$ is acceptable, regardless of its actual value. In this case, one can use (59) at this target IUI threshold to obtain

$$\delta^2 \ll \ln(1+N\varepsilon) \tag{64}$$

which would ensure that the threshold is not exceeded under random errors (provided that it is not exceeded without errors). To illustrate this, let $N=100$ and $\varepsilon = 10^{-3}$ (i.e. -30 dB) for Fig. 4, so that, from (64), $\delta^2 \ll 0.1$ makes the impact of random errors negligible for this design, which is consistent with $\delta^2 = 4\pi^2\sigma_p^2 + \sigma_\phi^2 \approx 0.004$ in Fig. 4 but not with $\delta^2 \approx 0.4$ in Fig. 3. Using (62) and (63), one concludes that the number of antennas needed to achieve a low IUI threshold $\varepsilon \ll 1$ scales as $1/\varepsilon$ with errors and only as $\sqrt{1/\varepsilon}$ without errors, i.e. more antennas are needed in the former case for small ε.

In summary, while random errors do affect the IUI leakage factor, their impact is not catastrophic at all and small $|\alpha_N|^2$ can still be attained, even under random errors, provided that N is sufficiently large (or, equivalently, if δ^2 is sufficiently small). This confirms the asymptotic result in Theorem 1.

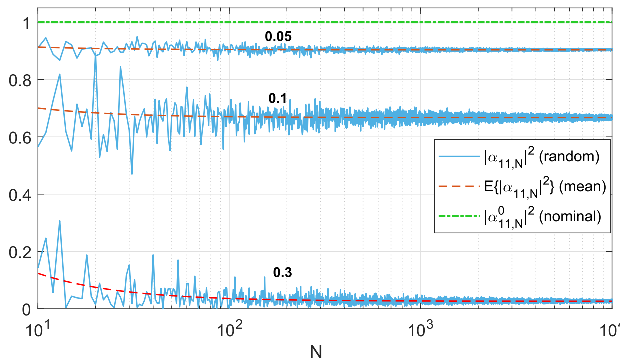

Fig. 7. The main user normalized power gain $|\alpha_{11,N}|^2$, its mean and nominal values for the N-element ULA with $d = 0.5$ and $\theta_1 = 0$, under zero-mean i.i.d. Gaussian errors with $\sigma_p = \sigma_\phi = 0.05, 0.1, 0.3$ (generated independently for each value of N). While power loss is not large for $\sigma_p, \sigma_\phi \leq 0.1$, it quickly increases otherwise but $|\alpha^0_{11,N}|^2 = 1$ for the nominal (no errors) array.

C. Impact of Errors on the Main User

Finally, we examine the impact of random errors on the main user via its normalized power gain $|\alpha_{11,N}|^2$ (equal to 1 if there are no errors). Using (22)-(24) in Theorem 1, one obtains:

$$\begin{aligned} |\alpha_{11,N}|^2 &\rightarrow \mathbb{E}\{|\alpha_{11,N}|^2\} \\ &= |c_1|^2|\alpha^0_{11,N}|^2 + (1-|c_1|^2)N^{-1} \\ &= |c_1|^2 + (1-|c_1|^2)N^{-1} \\ &\rightarrow |c_1|^2 = e^{-\delta^2} \end{aligned} \quad (65)$$

where the last equality holds for Gaussian errors while the others hold for non-Gaussian errors as well; (65) is due to $|\alpha^0_{11,N}|^2 = 1$, which follows from (14) and (16) with $\Psi^0_{1n} = 0$. $|c_1|^2 \leq 1$ is a measure of the average normalized power received by the main user under random errors, which can be evaluated from (56), where 1st equality holds for non-Gaussian errors as well. Note that random errors do induce power/SNR loss for the main user and this loss is not large as long as δ^2 is not large:

$$|\alpha_{11,N}|^2 \approx 1 \quad \text{if } \delta^2 \ll 1 \quad (66)$$

This condition is less demanding than the one for the low IUI leakage in (59) if $N|\alpha^0_N|^2 < e - 1$. Fig. 7 shows $|\alpha_{11,N}|^2$ as well as its mean and nominal values for N-element ULA. Note that power loss is not large for $\sigma_p, \sigma_\phi \leq 0.1$ but it quickly increases above that threshold. While random fluctuations of $|\alpha_{11,N}|^2$ induced by random errors do diminish with N (especially above $N = 100$), the significant power/SNR loss remains, regardless of N, for $\sigma_p, \sigma_\phi > 0.1$. Therefore, the latter regime should be avoided in practice.

VI. Conclusion

While the existing studies analysed favorable propagation in massive MIMO assuming no implementation inaccuracies, this paper studies the impact of random errors in array element location and beamforming phases on the FP. In particular, the FP property is rigorously shown to hold asymptotically for the perturbed array if it holds for the unperturbed one and, hence, random errors do not have a catastrophic impact, even when the number of antennas increases without bound. Even though random errors do not affect the FP asymptotically, they do slow down considerably the convergence to the asymptotic value and, thus, more antennas are needed under random errors to achieve the same low IUI as without errors. These results are general enough to include arbitrary array geometry as well as non-Gaussian error distributions. Practical guidelines as to what accuracy is needed to make the impact of errors negligible for a finite number of antenna are given. The analytical results are validated via numerical simulations and are in agreement with measurement-based studies.

Appendix

A. Proof of Lemma 1

First, we prove the upper bound. Using (12), one obtains

$$\begin{aligned} &\mathbb{E}\{|\alpha_N|^4\} \\ &\quad = \mathbb{E}\{\alpha_N\alpha_N^*\alpha_N\alpha_N^*\} \\ &\quad = \frac{1}{N^4}\mathbb{E}\left\{\sum_{n_1=1}^N e^{j\Psi_{in_1}} \sum_{n_2=1}^N e^{-j\Psi_{in_2}} \sum_{n_3=1}^N e^{j\Psi_{in_3}} \right. \\ &\qquad\qquad \left. \times \sum_{n_4=1}^N e^{-j\Psi_{in_4}}\right\} \\ &\quad = \frac{1}{N^4}\sum_{n_1,n_2,n_3,n_4} \mathbb{E}\{e^{j(\Psi_{in_1}-\Psi_{in_2}+\Psi_{in_3}-\Psi_{in_4})}\} \end{aligned} \quad (67)$$

where the last summation is over $1 \leq n_k \leq N$, $k = 1\ldots 4$. Since $e^{j\Psi_{in_k}}$ are independent for different n_k, we divide the total set S_t of $\{n_1, \ldots, n_4\}$ into the set of distinct indices S_d and its complementary set $S^c_d = S_t - S_d$,

$$S_t = \{\{n_1, n_2, n_3, n_4\}, 1 \leq n_k \leq N\} \quad (68)$$

$$S_d = \{\{n_1, n_2, n_3, n_4\}, n_i \neq n_j, \forall i \neq j\} \quad (69)$$

Next, we determine the cardinalities of these sets. First, note that the cardinality of S_t is $|S_t| = N^4$ and, likewise,

$$|S_d| = N(N-1)(N-2)(N-3) \quad (70)$$

(the latter follows from the number $(N)_k$ of all possible ordered selections of k distinct items out of a set of N distinct items, $(N)_k = \frac{N!}{(N-k)!}$ [39, p. 745]). Hence, the cardinality of the complementary set S^c_d is

$$|S^c_d| = N^4 - |S_d| = 6N^3 - 11N^2 + 6N \quad (71)$$

To simplify the derivations, define

$$\beta_{i,n_1..n_4} = \prod_{k=1}^4 \mathbb{E}\{e^{j(-1)^{k+1}\Psi_{in_k}}\} \quad (72)$$

$$\beta'_{i,n_1..n_4} = \mathbb{E}\{\prod_{k=1}^4 e^{j(-1)^{k+1}\Psi_{in_k}}\} \quad (73)$$

Since $e^{j\Psi_{in_k}}$ are independent for different n_k,

$$\beta_{i,n_1..n_4} = \beta'_{i,n_1..n_4} \quad \text{if } \{n_1..n_4\} \in S_d \quad (74)$$

Using (67) and (74),

$$\mathbb{E}\{|\alpha_N|^4\} = \frac{1}{N^4}\sum_{n_1..n_4}\beta'_{i,n_1..n_4} \quad (75)$$

$$= \frac{1}{N^4}\{\sum_{S_d}\beta_{i,n_1..n_4} + \sum_{S_d^c}\beta'_{i,n_1..n_4}\} \tag{76}$$

$$= \frac{1}{N^4}\{\sum_{S_d}\beta_{i,n_1..n_4} + \sum_{S_d^c}\beta_{i,n_1..n_4}\} + \frac{1}{N^4}\{\sum_{S_d^c}\beta'_{i,n_1..n_4} - \sum_{S_d^c}\beta_{i,n_1..n_4}\} \tag{77}$$

$$= \frac{1}{N^4}\{\sum_{n_1..n_4}\beta_{i,n_1..n_4} + \sum_{S_d^c}(\beta'_{i,n_1..n_4} - \beta_{i,n_1..n_4})\} \tag{78}$$

$$\leq |\mathbb{E}\{\alpha_N\}|^4 + N^{-4}\sum_{S_d^c}|\beta'_{i,n_1..n_4} - \beta_{i,n_1..n_4}| \tag{79}$$

$$\leq |\mathbb{E}\{\alpha_N\}|^4 + 12N^{-1} - 22N^{-2} + 12N^{-3} \tag{80}$$

$$\leq |\mathbb{E}\{\alpha_N\}|^4 + 12N^{-1} \tag{81}$$

where (76) is due to (74), (79) is due to the triangle inequality and

$$|\mathbb{E}\{\alpha_N\}|^4 = \frac{1}{N^4}\sum_{n_1..n_4}\beta_{i,n_1..n_4} \tag{82}$$

and (80) is due to (71) and

$$|\beta'_{i,n_1..n_4} - \beta_{i,n_1..n_4}| \leq 2 \tag{83}$$

which follows from $|\beta'_{i,n_1..n_4}|, |\beta_{i,n_1..n_4}| \leq 1$; (81) is due to

$$-22N^{-2} + 12N^{-3} < 0, \quad N \geq 1 \tag{84}$$

This establishes the upper bound. The lower bound in (26) follows from Jensen's inequality, since $|\cdot|^4$ is a convex function, see e.g. [40].

B. Proof of Lemma 4

We will need the following technical Lemmas.

Lemma 5 [38]: Let $\{a_N\}_{N=1}^{\infty}$ be a sequence of real numbers such that

$$a_N \geq 0, \quad \sum_{N=1}^{\infty}\frac{a_N}{N} < \infty \tag{85}$$

Then, there exists an increasing sequence of integers $\{N_k\}_{k=1}^{\infty}$ such that $N_{k+1}/N_k \to 1$ as $k \to \infty$ and

$$\sum_{k=1}^{\infty} a_{N_k} < \infty \tag{86}$$

Lemma 6 [38]: If y_n are random variables such that $\sum_{n=1}^{\infty}\mathbb{E}\{|y_n|^2\} < \infty$, then $y_n \xrightarrow{a.s.} 0$ as $n \to \infty$.

Now, use Lemma 5 with

$$a_N = \mathrm{Var}\{z_N\} = \mathbb{E}\{|z_N - \mathbb{E}\{z_N\}|^2\} \tag{87}$$

which satisfy (85), to conclude that there exits an increasing sequence $\{N_k\}$ such that $N_{k+1}/N_k \to 1$ and

$$\sum_{k=1}^{\infty}\mathbb{E}\{|z_{N_k} - \mathbb{E}\{z_{N_k}\}|^2\} < \infty \tag{88}$$

Now, using Lemma 6 with $y_k = z_{N_k} - \mathbb{E}\{z_{N_k}\}$,

$$z_{N_k} \xrightarrow{a.s.} \mathbb{E}\{z_{N_k}\} \text{ as } k \to \infty \tag{89}$$

i.e. almost sure convergence holds for the subsequence $\{N_k\}$. It remains to prove that it also holds for any N in-between, $N_k < N < N_{k+1}$. To this end, let $y_{nm} = z_{nm} - \mathbb{E}\{z_{nm}\}$, define the following sets I_1, I_2 of indexes (n, m):

$$\begin{aligned} I_1 &= \{(n,m) : 1 \leq n \leq N_k,\ N_k < m \leq N\} \\ I_2 &= \{(n,m) : N_k < n \leq N,\ 1 \leq m \leq N\} \end{aligned} \tag{90}$$

and note that

$$|y_N| = |z_N - \mathbb{E}\{z_N\}| \tag{91}$$

$$= \frac{1}{N^2}\Big|\sum_{n,m\leq N} y_{nm}\Big| \tag{92}$$

$$\leq \frac{1}{N_k^2}\Big|\sum_{n,m\leq N} y_{nm}\Big| \tag{93}$$

$$= \frac{1}{N_k^2}\Big|S_k + \sum_{(n,m)\in I_1} y_{nm} + \sum_{(n,m)\in I_2} y_{nm}\Big| \tag{94}$$

$$\leq \frac{|S_k|}{N_k^2} + \sum_{(n,m)\in I_1}\frac{|y_{nm}|}{N_k^2} + \sum_{(n,m)\in I_2}\frac{|y_{nm}|}{N_k^2} \tag{95}$$

$$\leq \frac{|S_k|}{N_k^2} + \sum_{\substack{n\leq N_k \\ N_k<m\leq N_{k+1}}}\frac{|y_{nm}|}{N_k^2} + \sum_{\substack{m\leq N_{k+1} \\ N_k<n\leq N_{k+1}}}\frac{|y_{nm}|}{N_k^2} \tag{96}$$

$$\leq \frac{|S_k|}{N_k^2} + \frac{2(N_{k+1} - N_k)}{N_k} + \frac{2N_{k+1}(N_{k+1} - N_k)}{N_k^2} \tag{97}$$

$$\xrightarrow{a.s.} 0 \tag{98}$$

where $S_k = \sum_{n,m\leq N_k} y_{nm}$; (93) is due to $N > N_k$, (96) is due to $N_{k+1} > N_k$ and all summation terms being positive, (97) is due to

$$|y_{nm}| = |z_{nm} - \mathbb{E}\{z_{nm}\}| \leq |z_{nm}| + \mathbb{E}\{|z_{nm}|\} \leq 2 \tag{99}$$

since $|z_{nm}| \leq 1$; (98) follows since, from (89),

$$\frac{S_k}{N_k^2} = z_{N_k} - \mathbb{E}\{z_{N_k}\} \xrightarrow{a.s.} 0 \tag{100}$$

and $N_{k+1}/N_k \to 1$. Therefore, from (98), $z_N \xrightarrow{a.s.} \mathbb{E}\{z_N\}$, as required.

C. Proof of Proposition 1

To estimate $|\alpha_N|^2$ for large N, we first obtain the asymptotic distribution of $\alpha_N = \alpha_{N1} + j\alpha_{N2}$. To this end, observe from (12) that it is an empirical average of i.i.d random variables with finite variance. Therefore, from the central limit theorem [35, p. 406], its real α_{N1} and imaginary α_{N2} parts are asymptotically Gaussian,

$$\alpha_{Nk} \sim \mathcal{N}(\mathbb{E}\{\alpha_{Nk}\}, \sigma_{Nk}^2), \quad k = 1, 2 \tag{101}$$

where $\sigma_{Nk}^2 = \text{Var}\{\alpha_{Nk}\}$ and, using (35), $\mathbb{E}\{\alpha_N\} = c_i\alpha_N^0$ from which $\mathbb{E}\{\alpha_{Nk}\}$ can be found. Next, σ_{N1}^2 can be evaluated as follows:

$$\begin{aligned}\sigma_{N1}^2 &= \text{Var}\left\{\frac{1}{N}\sum_{n=1}^{N}\cos(\Psi_{in})\right\} &(102)\\ &= \frac{1}{N^2}\sum_{n=1}^{N}\text{Var}\{\cos(\Psi_{in})\} &(103)\\ &= \frac{1}{N^2}\sum_{n=1}^{N}(\mathbb{E}\{\cos^2(\Psi_{in})\} - (\mathbb{E}\{\cos(\Psi_{in})\})^2) &(104)\\ &= \frac{1}{N^2}\sum_{n=1}^{N}\left(\frac{1}{2}+\frac{1}{4}\mathbb{E}\{e^{j2\Psi_{in}}+e^{-j2\Psi_{in}}\}\right)\\ &\quad - \frac{1}{4N^2}\sum_{n=1}^{N}(\mathbb{E}\{e^{j\Psi_{in}}+e^{-j\Psi_{in}}\})^2 &(105)\\ &= \frac{1}{N^2}\sum_{n=1}^{N}\left(\frac{1}{2}+\frac{1}{4}e^{j2\Psi_{in}^0}(c_i'-c_i^2)\right)\\ &\quad + \frac{1}{N^2}\sum_{n=1}^{N}\left(\frac{1}{4}e^{-j2\Psi_{in}^0}(c_i'-c_i^2)^* - \frac{1}{2}|c_i|^2\right) &(106)\\ &= \frac{1}{2N}(1-|c_i|^2)+\frac{1}{N^2}\\ &\quad \times\sum_{n=1}^{N}\left(\frac{1}{4}e^{j2\Psi_{in}^0}(c_i'-c_i^2)+\frac{1}{4}e^{-j2\Psi_{in}^0}(c_i'-c_i^2)^*\right) &(107)\\ &= \frac{1}{2N}(1-|c_i|^2)+\frac{1}{2N}\text{Re}\{(c_i'-c_i^2)\beta_N^0\} &(108)\\ &= \frac{1}{2N}(1-|c_i|^2)(1+o(1)) &(109)\end{aligned}$$

where (103) is due to Ψ_{in} being independent for different n; (106) is due to (13) and the following expectations:

$$\begin{aligned}\mathbb{E}\{e^{j2\Psi_{in}}\} &= e^{j2\Psi_{in}^0}c_i', \quad c_i' = \mathbb{E}\{e^{j2\Delta\Psi_{in}}\} &(110)\\ \mathbb{E}\{e^{j\Psi_{in}}\} &= e^{j\Psi_{in}^0}c_i, \quad c_i = \mathbb{E}\{e^{j\Delta\Psi_{in}}\} &(111)\end{aligned}$$

(107) follows by noting that 1st and last summation terms are independent of n; (108) follows from $\beta_N^0 = N^{-1}\sum_{n=1}^{N}e^{j2\Psi_n^0}$, c_i, c_i' being independent of n and $z+z^* = 2\text{Re}\{z\}$; (109) follows from (50), where $\beta_N^0 = o(1) \to 0$ as $N\to\infty$. Following the steps in (102)-(109) adopted for α_{N2}, one obtains:

$$\begin{aligned}\sigma_{Nk}^2 &= \frac{1}{2}\sigma_N^2 - \frac{(-1)^k}{2N}\text{Re}\{(c_i'-c_i^2)\beta_N^o\}, \quad k=1,2\\ &= \frac{1}{2}\sigma_N^2(1+o(1)) \approx \frac{1}{2}\sigma_N^2 &(112)\end{aligned}$$

where, using (31) and (32),

$$\begin{aligned}\sigma_N^2 &= \text{Var}\{\alpha_N\} = \mathbb{E}\{|\alpha_N|^2\} - |\mathbb{E}\{\alpha_N\}|^2 &(113)\\ &= N^{-1}(1-|c_i|^2)\end{aligned}$$

It can be further shown that $\sigma_{Nk} > 0$ unless $\Delta\Psi_{in}$ is a deterministic constant with probability one, i.e. the errors are deterministic (biases) rather then random, - a degenerate case not considered here.

To obtain the asymptotic distribution of $|\alpha_N|^2$, we will now show that α_{N1} and α_{N2} are asymptotically independent. To this end, we use the normalized random variables α_{Nk}', $k=1,2$,

$$\alpha_{Nk}' = \sigma_{Nk}^{-1}\alpha_{Nk} \sim \mathcal{N}(\mathbb{E}\{\alpha_{Nk}\},1) \quad (114)$$

and evaluate their covariance r_N:

$$\begin{aligned}r_N &= \mathbb{E}\{\alpha_{N1}'\alpha_{N2}'\} - \mathbb{E}\{\alpha_{N1}'\}\,\mathbb{E}\{\alpha_{N2}'\} &(115)\\ &= (\sigma_{N1}\sigma_{N2})^{-1}(\mathbb{E}\{\alpha_{N1}\alpha_{N2}\} - \mathbb{E}\{\alpha_{N1}\}\,\mathbb{E}\{\alpha_{N2}\})\\ &= \sum_n\sum_m\frac{\mathbb{E}\{\cos(\Psi_{in})\sin(\Psi_{im})\}}{\sigma_{N1}\sigma_{N2}N^2}\\ &\quad - \sum_n\sum_m\frac{\mathbb{E}\{\cos(\Psi_{in})\}\,\mathbb{E}\{\sin(\Psi_{im})\}}{\sigma_{N1}\sigma_{N2}N^2} &(116)\\ &= \sum_n\frac{\mathbb{E}\{\sin(2\Psi_{in})\} - 2\,\mathbb{E}\{\cos(\Psi_{in})\}\,\mathbb{E}\{\sin(\Psi_{in})\}}{2\sigma_{N1}\sigma_{N2}N^2} &(117)\\ &= \sum_n\frac{(c_i'-c_i^2)e^{j2\Psi_{in}^0} - (c_i'-c_i^2)^*e^{-j2\Psi_{in}^0}}{4j\sigma_{N1}\sigma_{N2}N^2} &(118)\\ &= \frac{1}{2\sigma_{N1}\sigma_{N2}N}\text{Im}\{(c_i'-c_i^2)\beta_N^0\} &(119)\\ &= \text{Im}\{(c_i'-c_i^2)\beta_N^0\}\\ &\quad\times\left[(1-|c_i|^2)^2 - (\text{Re}\{(c_i'-c_i^2)\beta_N^0\})^2\right]^{-1/2} &(120)\end{aligned}$$

where (116) is obtained by substituting the real and imaginary parts of α_N, defined in (12); the double sum in (116) reduces to the single sum in (117) since Ψ_{in} is independent of Ψ_{im} for $n \neq m$ and so $\mathbb{E}\{\cos(\Psi_{in})\sin(\Psi_{im})\} = \mathbb{E}\{\cos(\Psi_{in})\}\,\mathbb{E}\{\sin(\Psi_{im}\}$; (118) follows from (110) and (111); (119) follows from $z - z^* = 2j\text{Im}\{z\}$ and c_i, c_i' being independent of n; (120) follows from (113) and (112). Once can further show that $|c_i| < 1$ unless $\Delta\Psi_{in}$ is a deterministic constant (no random errors). Thus, using (120) and (50):

$$\lim_{N\to\infty} r_N = \lim_{n\to\infty}\frac{\text{Im}\{(c_i'-c_i^2)\beta_N^0\}}{\sqrt{(1-|c_i|^2)^2 - (\text{Re}\{(c_i'-c_i^2)\beta_N^0\})^2}} = 0$$

This shows that α_{N1}' and α_{N2}' are asymptotically uncorrelated and, since they are asymptotically Gaussian, also independent of each other. Having established this, we use the following technical Lemma [41, p. 447] to establish the asymptotic distribution of $|\alpha_N|^2$.

Lemma 7: Let $(x_1, x_2, \ldots, x_k)$ be k independent, normally distributed random variables (RV) with means μ_i and unit variances. Then $y = \sum_{i=1}^{k}x_i^2$ is the noncentral chi-squared RV with k degrees of freedom and noncentrality parameter λ,

$$y \sim \chi_k^2(\lambda), \quad \lambda = \sum_{i=1}^{k}\mu_i^2, \quad \text{Var}\{y\} = 2(k+2\lambda) \quad (121)$$

Since α_{N1}' and α_{N2}' are two unit variance and asymptotically independent Gaussian random variables with means $\mathbb{E}\{\alpha_{N1}\}$ and $\mathbb{E}\{\alpha_{N2}\}$, one obtains from Lemma 7, (114) and (112):

$$2\alpha_{N1}^2\sigma_N^{-2} + 2\alpha_{N2}^2\sigma_N^{-2} \sim \chi_2^2(\lambda) \quad (122)$$

Using (31) and (121),

$$\lambda = 2(\mathbb{E}\{\alpha_{N1}\})^2\sigma_N^{-2} + 2(\mathbb{E}\{\alpha_{N2}\})^2\sigma_N^{-2} = 2|\,\mathbb{E}\{\alpha_N\}|^2\sigma_N^{-2} = 2\sigma_N^{-2}|c_i|^2|\alpha_N^0|^2 \quad (123)$$

Using (121) with $k = 2$ and λ in (123),

$$\mathrm{Var}\{2\alpha_{N1}^2\sigma_N^{-2} + 2\alpha_{N2}^2\sigma_N^{-2}\} = 4\sigma_N^{-4}\,\mathrm{Var}\{|\alpha_N|^2\} \approx 2(2 + 4\sigma_N^{-2}|c_i|^2|\alpha_N^0|^2) \quad (124)$$

Hence, using (122), the IUI is asymptotically distributed as

$$|\alpha_N|^2 = \alpha_{N1}^2 + \alpha_{N2}^2 \sim \frac{1}{2}\sigma_N^2\chi_2^2(\lambda) \quad (125)$$

and, using (124), $\sigma^2_{|\alpha_N|^2} = \mathrm{Var}\{|\alpha_N|^2\}$ can be asymptotically approximated as

$$\sigma^2_{|\alpha_N|^2} \approx \sigma_N^4 + 2|c_i|^2|\alpha_N^0|^2\sigma_N^2 \quad (126)$$

as required.

References

[1] 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.

[2] H. Quoc 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] L. Lu, G. Ye Li, A. Lee Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Oct. 2014.

[4] 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.

[5] M. Shafi et al., "5G: A tutorial overview of standards, trials, challenges, deployment, and practice," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 6, pp. 1201–1221, Jun. 2017.

[6] H. Tataria, M. Shafi, A. F. Molisch, M. Dohler, H. Sjöland, and F. Tufvesson, "6G wireless systems: Vision, requirements, challenges, insights, and opportunities," *Proc. IEEE*, vol. 109, no. 7, pp. 1166–1199, Jul. 2021.

[7] T. L. Marzetta, *Fundamentals of Massive MIMO*. Cambridge, U.K.: Cambridge Univ. Press, 2016.

[8] H. Quoc Ngo, E. G. Larsson, and T. L. Marzetta, "Aspects of favorable propagation in massive MIMO," in *Proc. 22nd Eur. Signal Process. Conf. (EUSIPCO)*, Sep. 2014, pp. 76–80.

[9] J. H. Chen, "When does asymptotic orthogonality exist for very large arrays?" in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Atlanta, GA, USA, Nov. 2013, pp. 4146–4150.

[10] 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.

[11] H. Yang and T. L. Marzetta, "Massive MIMO with max-min power control in line-of-sight propagation environment," *IEEE Trans. Commun.*, vol. 65, no. 11, pp. 4685–4693, Nov. 2017.

[12] E. Anarakifirooz and S. Loyka, "Favorable propagation for massive MIMO with circular and cylindrical antenna arrays," *IEEE Wireless Commun. Lett.*, vol. 11, no. 3, pp. 458–462, Mar. 2022.

[13] E. Anarakifirooz and S. Loyka, "Structural design of non-uniform linear arrays for favorable propagation in massive MIMO," *IEEE Commun. Lett.*, vol. 27, no. 1, pp. 367–371, Jan. 2023.

[14] J. Hoydis, C. Hoek, T. Wild, and S. T. Brink, "Channel measurements for large antenna arrays," in *Proc. Int. Symp. Wireless Commun. Syst. (ISWCS)*, Paris, France, Aug. 2012, pp. 811–815.

[15] X. Gao, O. Edfors, F. Rusek, and F. Tufvesson, "Massive MIMO performance evaluation based on measured propagation data," *IEEE Trans. Wireless Commun.*, vol. 14, no. 7, pp. 3899–3911, Jul. 2015.

[16] M. Gauger, J. Hoydis, C. Hoek, H. Schlesinger, A. Pascht, and S. T. Brink, "Channel measurements with different antenna array geometries for massive MIMO systems," in *Proc. 10th Int. ITG Conf. Syst., Commun. Coding*, Feb. 2015, pp. 1–6.

[17] P. Harris et al., "Performance characterization of a real-time massive MIMO system with LOS mobile channels," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 6, pp. 1244–1253, Jun. 2017.

[18] À. O. Martínez, J. Ø. Nielsen, E. De Carvalho, and P. Popovski, "An experimental study of massive MIMO properties in 5G scenarios," *IEEE Trans. Antennas Propag.*, vol. 66, no. 12, pp. 7206–7215, Dec. 2018.

[19] J. Flordelis, F. Rusek, X. Gao, G. Dahman, O. Edfors, and F. Tufvesson, "Spatial separation of closely-located users in measured massive MIMO channels," *IEEE Access*, vol. 6, pp. 40253–40266, 2018.

[20] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Dec. 1999.

[21] S. Loyka and G. Levin, "Finite-SNR diversity-multiplexing tradeoff via asymptotic analysis of large MIMO systems," *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 4781–4792, Oct. 2010.

[22] F. Bohagen, P. Orten, and G. E. Oien, "Design of optimal high-rank line-of-sight MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 6, no. 4, pp. 1420–1425, Apr. 2007.

[23] P. Wang, Y. Li, X. Yuan, L. Song, and B. Vucetic, "Tens of gigabits wireless communications over E-band LoS MIMO channels with uniform linear antenna arrays," *IEEE Trans. Wireless Commun.*, vol. 13, no. 7, pp. 3791–3805, Jul. 2014.

[24] H. Do, N. Lee, and A. Lozano, "Reconfigurable ULAs for line-of-sight MIMO transmission," *IEEE Trans. Wireless Commun.*, vol. 20, no. 5, pp. 2933–2947, May 2021.

[25] E. G. Larsson, T. L. Marzetta, H. Q. Ngo, and H. Yang, "Antenna count for massive MIMO: 1.9 GHz vs. 60 GHz," *IEEE Commun. Mag.*, vol. 56, no. 9, pp. 132–137, Sep. 2018.

[26] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proc. IEEE*, vol. 102, no. 3, pp. 366–385, Mar. 2014.

[27] T. S. Rappaport et al., "Wireless communications and applications above 100 GHz: Opportunities and challenges for 6G and beyond," *IEEE Access*, vol. 7, pp. 78729–78757, 2019.

[28] H. L. Van Trees, *Optimum Array Processing*. Hoboken, NJ, USA: Wiley, 2002.

[29] S. A. Vorobyov, A. B. Gershman, and Z.-Q. Luo, "Robust adaptive beamforming using worst-case performance optimization: A solution to the signal mismatch problem," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 313–324, Feb. 2003.

[30] R. G. Lorenz and S. P. Boyd, "Robust minimum variance beamforming," *IEEE Trans. Signal Process.*, vol. 53, no. 5, pp. 1684–1696, May 2005.

[31] J. Li and P. Stoica, *Robust Adaptive Beamforming*. New York, NY, USA: Wiley, 2006.

[32] S. Loyka and C. D. Charalambous, "On the compound capacity of a class of MIMO channels subject to normed uncertainty," *IEEE Trans. Inf. Theory*, vol. 58, no. 4, pp. 2048–2063, Apr. 2012.

[33] S. Loyka and C. D. Charalambous, "Novel matrix singular value inequalities and their applications to uncertain MIMO channels," *IEEE Trans. Inf. Theory*, vol. 61, no. 12, pp. 6623–6634, Dec. 2015.

[34] E. Björnson, J. Hoydis, M. Kountouris, and M. Debbah, "Massive MIMO systems with non-ideal hardware: Energy efficiency, estimation, and capacity limits," *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 7112–7139, Nov. 2014.

[35] A. N. Shiryaev, *Probability*, 3rd ed. New York, NY, USA: Springer-Verlag, 2015.

[36] A. A. Borovkov, *Probability Theory*, 5th ed. London, U.K.: Springer, 2013.

[37] W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. 2. Hoboken, NJ, USA: Wiley, 1971.

[38] R. Lyons, "Strong laws of large numbers for weakly correlated random variables," *Michigan Math. J.*, vol. 35, no. 3, pp. 353–359, Jan. 1988.

[39] I. N. Bronshtein and K. A. Semendyayev, *Handbook of Mathematics*, 5th ed. New York, NY, USA: Springer, 2007.

[40] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[41] N. L. Johnson, *Continuous Univariate Distributions*, 2nd ed. New York, NY, USA: Wiley, 1995.

[42] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.