Estimation of Interference Correlation in mmWave Cellular Systems

Stefano Tomasin, *Senior Member, IEEE* Raphael Hasler, Antonia M. Tulino, *Fellow, IEEE*, and Matilde Sánchez-Fernández, *Senior Member, IEEE*

***Abstract*—We consider a cellular network, where the uplink transmissions to a base station (BS) are interferenced by other devices, a condition that may occur, e.g., in cell-free networks or when using non-orthogonal multiple access (NOMA) techniques. Assuming that the BS treats this interference as additional noise, we focus on the problem of estimating the interference correlation matrix from received signal samples. We consider a BS equipped with multiple antennas and operating in the millimeter-wave (mmWave) bands and propose techniques exploiting the fact that channels comprise only a few reflections at these frequencies. This yields a specific structure of the interference correlation matrix that can be decomposed into three matrices, two rectangular depending on the angle of arrival (AoA) of the interference and the third square with smaller dimensions. We resort to gridless approaches to estimate the AoAs and then project the least square estimate of the interference correlation matrix into a subspace with a smaller dimension, thus reducing the estimation error. Moreover, we derive two simplified estimators, still based on the gridless angle estimation that turns out to be convenient when estimating the interference over a larger number of samples.**

***Index Terms*—Angle of arrival, Cell-free networks, Interference estimation, Non-orthogonal multiple access.**

I. Introduction

THE evolution of cellular networks is pushing towards non-orthogonal multiple access (NOMA) schemes, wherein the wireless resource is not exclusively assigned to a user (as in orthogonal multiple access) but is shared among several devices that interfere. Such an approach may be applied either within each cell (in what is more specifically NOMA, [1]) or among different cells, e.g., in dynamic time-division duplexing (TDD) networks [2] and is particularly attractive in systems with partial or distributed coordination, such as in device-to-device (D2D) communications [3] or cell-free networks [4]. In all these cases, the interference among devices becomes challenging and is addressed by several techniques that can be classified according to the underlying interference model: when interference is seen as an additional data signal with a suitable structure (e.g., with known modulation and coding parameters) successive interference cancellation (SIC) or multiuser decoding solutions are applied; when interference is seen as additional noise, it is handled by suitable signal processing at the transmitter or the receiver, e.g., by beamforming and combining signals over multiple antennas in multiple-input multiple-output (MIMO) systems. The first approach, which is preferable under strong interference coming from a few users, requires the estimation of the interference channels, and yields intensive computations. The latter approach is preferable when several low-power interferers are present, as they cannot be demodulated without errors; this approach requires the estimation of the statistics of the interference and entails fewer computations. However, interference can always be treated as additional noise, even under strong interference, thus our approach has a very wide application. This is particularly true when either strong interferers cannot be decoded, e.g., because the receiver either does not know the used modulation or coding or does not have enough computational capabilities. In all these cases, the estimation of the interference correlation matrix turns out to be very useful.

We focus on the latter scenario, where interference is treated as additional noise, and consider the problem of estimating its correlation matrix when the receiver is equipped with multiple antennas. The estimation of the interference statistics is not only useful for the design of beamformers but also to improve the channel estimation [5], estimate the phase noise and frequency offset [6], and perform link adaptation [7]. Note however that we aim at estimating the correlation matrix of the interference averaging over the signal transmitted by the interferers, but assuming their channel as time-invariant: this makes our problem significantly different from the covariance matrix estimation of the channel, where the average is taken with respect to the channel fading (see for example [8]).

Several works in the literature have addressed the estimation of the noise variance, under the assumption of independent identically distributed noise samples on each receive antenna. Among other works, we mention [9], where the noise variance estimation is considered in the context of massive MIMO assuming uncorrelated noise at the antennas and under practical impairments such as pilot contamination. Although a different noise power is assumed at each antenna, noise samples at different antennas are still considered uncorrelated. In this paper, we instead focus on the case wherein interference is correlated among the various antennas. In this respect, in the literature we can find exponential models for the correlation of the noise among antennas, see [10] and references therein. These

Corresponding author: Stefano Tomasin.
S. Tomasin is with the University of Padova, Italy, and also with the Consorzio Nazionale Interuniversitario per le Telecomunicazioni (CNIT) (email: stefano.tomasin@unipd.it).
R. Hasler at the time of writing was with the University of Padova, Italy.
A. Tulino is with the Universityà degli Studi di Napoli Federico II, Naples, Italy (email: antoniamaria.tulino@unina.it)
M. Sánchez-Fernández is with the Universidad Carlos III de Madrid, Spain (email: mati@tsc.uc3m.es). The work of A. M. Tulino has been supported in part by projects FRACTAL - Project number: E59C20001020005 and RESTART - Project number: E63C2200204000. The work of M. Sánchez-Fernández has been partly funded by the Spanish Government through project IRENE-EARTH (PID2020-115323RB-C33/ AEI / 10.13039/501100011033) (Ministerio de Ciencia e Innovación/AEI/FEDER, UE).

models neither exemplify any particular propagation scenario nor are suited to capture the underlying composite interference signals. In fact, they are typically proposed as the simplest analytical model that can be easily adapted to empirically derived covariances [11]. Nevertheless, exponential correlation modeling has some drawbacks: on one hand, the fitting may not be straightforward given that they are not linear functions, on the other, they may miss relevant components, e.g., at high frequencies. In [12], [13] the training sequence for channel estimation in MIMO systems with colored noise is designed, but no specific solution is proposed for the estimation of the noise correlation matrix. In [14] a generic covariance matrix of the interference is considered, and the problem of its joint estimation with the channel is addressed and solved by a low-rank approximation of the channel. However, no particular structure of the interference correlation matrix is assumed. The estimation of colored interference is considered in [2] and a least square (LS) technique is proposed. The noise correlation matrix estimation is included in the estimation of the channel correlation in [15], and a modified version of the LS estimator is derived, taking into account the channel statistics.

However, to our knowledge, there are no specific solutions for interference correlation matrix estimation when operating at millimeter-wave (mmWave) or Terahertz (THz) bands. Indeed, at those frequencies, the channels exhibit only a few propagation paths, due to the strong attenuation incurred by radio signals. This feature has been extensively exploited to obtain accurate channel estimators [16]. Still, it remains not used for the estimation of the interference correlation matrix in systems operating at mmWave or THz bands. Note that, although the presence of few paths could make the beamforming to the interfering users more accurate, the strong attenuations combined to more flexible architectures (including D2D, cell-free, and dynamic-TDD communications) may make the detection and cancellation of *all* interfering users more problematic. Thus, also with mmWaves the estimation of the interference correlation matrix is still relevant.

In this paper, we propose three techniques to estimate the interference correlation matrix in mmWave or THz systems with different complexity and performance. All solutions exploit the structure of the interference channels, and the objective is to obtain more accurate estimations using fewer samples than those needed by the LS approach. In particular, we consider a receiver equipped with a linear array of antennas and first observe that the interference correlation matrix can be decomposed into the product of three matrices, where the two external depend only on the angles of arrival (AoAs) of the interference signals. At the same time, the internal one represents the correlation among the interference signals. The internal matrix has a much lower rank than the interference correlation matrix since it only provides the correlation of signals with the (few) different AoAs. Then, we propose a procedure by which we estimate a) the AoAs, b) the internal correlation matrix, and c) the interference matrix. Lastly, we reconstruct the interference correlation matrix.

In particular, for the AoA estimation, we adopt a gridless angle estimation (GAE) approach [17], where the estimation is described as a gridless identification of several multi-dimensional frequency vectors, each corresponding to a different AoA. Although this estimator shows an improvement over the LS estimator already when using a few samples for each antenna, it becomes infeasible when applied to several samples. Therefore, we introduce two alternative approaches, still based on GAE: the subspace gridless estimation (SGE) and the gridless estimation and clustering (GEC). SGE first estimates the subspace of the LS correlation estimate and then applies the GAE method on the subspace basis vectors, with a complexity that does not increase with the number of observations. GEC instead performs a separate GAE estimation on each new interference observation and then fuses the estimates by clustering the estimated AoAs. In this case, the complexity increases with the number of observations, since a larger number of points must be clustered; however, the resulting complexity is still less than that of the original GAE technique. For comparison purposes, we also consider the multiple signal classification (MUSIC) algorithm for the AoA estimation.

The rest of the paper is organized as follows. Section II introduces the considered interference scenario, the mmWave channel model with few reflections, and the LS estimate of the interference correlation matrix. Section III describes the proposed projection-based correlation estimate and the procedure that first estimates the AoAs and then provides a more accurate estimate of the interference correlation matrix. The channel interference matrix estimators based on GAE techniques are described in Section IV. The proposed approaches for the interference correlation estimate are then compared with the LS estimate in a typical cellular communication scenario in Section V before the main conclusions are outlined in Section VI.

Notation: Matrices and vectors are denoted in boldface, with uppercase and lowercase letters, respectively. $\boldsymbol{A}^T$ and $\boldsymbol{A}^{\mathrm{H}}$ denote the transpose and Hermitian operators on matrix $\boldsymbol{A}$, respectively. $\boldsymbol{A} \succcurlyeq 0$ indicates that matrix $\boldsymbol{A}$ is semidefinite positive. $\boldsymbol{I}_L$ is the identity matrix of size $L \times L$. $\boldsymbol{B} = \mathrm{diag}\{\boldsymbol{b}\}$ is a diagonal matrix having on the diagonal elements of vector $\boldsymbol{b}$. $\mathrm{atan2}(\boldsymbol{p})$ is the 2-argument arc-tangent of point $\boldsymbol{p} = [x, y]$. $\boldsymbol{A}^\dagger$ denotes the More-Penrose pseudoinverse of matrix $\boldsymbol{A}$.

II. System Model

In a communication system, we consider one base station (BS) with N antennas (denoted *main* BS) receiving signals in uplink from several served user equipments (UEs). In the surroundings, a set of users transmit to *their own* serving BSs and interfere at the main BS. In this system, we set L as the total number of interferers to the *main* BS, each interferer is equipped with N_I antennas, and interferers transmit statistically independent signals. We consider one BS (denoted *main* BS) of a cellular communication system receiving signals in uplink from several served UEs through its N antennas. This scenario fits cellular systems dynamic TDD networks (where interferes are in other cells) and cell-free networks.

Let us indicate with $\boldsymbol{J}^{(\ell)}(t) \in \mathbb{C}^{N_I \times 1}$, $\ell = 1, \ldots, L$, the column vector of symbols transmitted by interferer ℓ at symbol time t. Assuming time-invariant narrowband channels, the matrices of the channels from the L interferers to the main BS are $\boldsymbol{G}(\ell) \in \mathbb{C}^{N \times N_I}$, $\ell = 1, \ldots, L$. We consider here that only a few interferers are present, i.e., L is a small number. Due to the specific propagation conditions in mmWave, such as severe large-scale fading and blockage, we consider here that only a few interferers are present, i.e., L is a small number. The baseband equivalent interference and noise signal received at sample time t by the main BS can then be written as the N-size column vector

$$\boldsymbol{N}(t) = \boldsymbol{Y}(t) + \boldsymbol{Z}(t) = \sum_{\ell=1}^{L} \boldsymbol{G}(\ell)\boldsymbol{J}^{(\ell)}(t) + \boldsymbol{Z}(t), \quad (1)$$

where $\boldsymbol{Z}(t)$ is the vector of additive white Gaussian noise (AWGN), with independent entries having zero mean and variance σ^2, which is supposed to be known at the main BS.

We assume that $\boldsymbol{J}^{(\ell)}(t)$ and $\boldsymbol{Z}(t)$ are independent complex Gaussian vectors with zero mean. The cross-correlation matrices of vectors $\boldsymbol{J}^{(\ell)}(t)$ and $\boldsymbol{Z}(t)$ are $\boldsymbol{R}_J = \mathbb{E}[\boldsymbol{J}^{(\ell)}(t)\boldsymbol{J}^{(\ell)H}(t)]$, and $\boldsymbol{R}_Z = \mathbb{E}[\boldsymbol{Z}(t)\boldsymbol{Z}^{\mathrm{H}}(t)] = \sigma^2 \boldsymbol{I}_N$, respectively. Note that the cross-correlation matrix of $\boldsymbol{J}^{(\ell)}(t)$, in general, is not an identity matrix as each interferer may use a transmit beamforming to convey one data stream with its multiple antennas. The cross-correlation of $\boldsymbol{N}(t)$ is therefore

$$\begin{aligned}\boldsymbol{R} =& \mathbb{E}\left[\boldsymbol{N}(t)\boldsymbol{N}^{\mathrm{H}}(t)\right] \\ =& \mathbb{E}\left[\left(\sum_{\ell=1}^{L} \boldsymbol{G}(\ell)\boldsymbol{J}^{(\ell)}(t) + \boldsymbol{Z}(t)\right) \times \right. \\ & \left. \left(\sum_{\ell=1}^{L} \boldsymbol{G}(\ell)\boldsymbol{J}^{(\ell)}(t) + \boldsymbol{Z}(t)\right)^{\mathrm{H}}\right] \\ =& \sum_{\ell=1}^{L} \boldsymbol{G}(\ell)\boldsymbol{R}_J\boldsymbol{G}^{\mathrm{H}}(\ell) + \sigma^2 \boldsymbol{I}_N.\end{aligned} \quad (2)$$

In this paper, we aim at estimating matrix $\boldsymbol{R}$ from T samples of the interference and noise signal, i.e., from $\boldsymbol{N}(t)$, $t = 1, \ldots, T$. The main BS can obtain such samples for example in correspondence of pilot signals transmitted by the UE canceling the received signals (and also some other strong interference signals). The estimated interference correlation matrix $\hat{\boldsymbol{R}}$ will then be used to process the forthcoming received signals containing uplink data and mitigate the effects of interference. As mentioned in the introduction, we consider a scenario wherein the interference is treated as additional noise, as an example in the presence of several low-power interferes whose signals cannot be properly detected and canceled.

A. Channel Model

All devices are equipped with linear antenna arrays. Signals transmitted by interferer ℓ, with $\ell = 1, \ldots, L$, propagate through N_g scatterers and we indicate with $\alpha_{i,\ell}^{(\mathrm{Tx})}$ and $\alpha_{i,\ell}^{(\mathrm{Rx})}$ the angle of departure (AoD) and AoA of the i-reflection, for $i = 0, \ldots, N_g - 1$. Let us also define the transmit and receive phase shifts as

$$\begin{aligned}\gamma_{i,\ell} &= \frac{D_{\min}^{(\mathrm{Tx})}}{\lambda} \sin \alpha_{i,\ell}^{(\mathrm{Tx})}, \\ \beta_{i,\ell} &= \frac{D_{\min}^{(\mathrm{Rx})}}{\lambda} \sin \alpha_{i,\ell}^{(\mathrm{Rx})},\end{aligned} \quad (3)$$

where $D_{\min}^{(\mathrm{Tx})}$ and $D_{\min}^{(\mathrm{Rx})}$ are the distances between the elements in the transmit and receive linear antenna arrays, respectively, and λ is the wavelength of the carrier signal. Defining the steering vector function

$$\boldsymbol{a}_N(\beta) = [1, e^{-2\pi \mathtt{i}\beta}, \ldots, e^{-2\pi \mathtt{i}\beta(N-1)}]^{\mathrm{T}}, \quad (4)$$

with $\mathtt{i} = \sqrt{-1}$, we can write the interference channel matrices as

$$\boldsymbol{G}(\ell) = \sum_{i=0}^{N_g-1} v_{i,\ell}\boldsymbol{a}_N(\beta_{i,\ell})\boldsymbol{a}_{N_I}^{\mathrm{H}}(\gamma_{i,\ell}), \quad (5)$$

where $v_{i,\ell}$ is the complex gain of signal i of the ℓth interferer, with power given by the path-loss ($P_{i,\ell}$). In mmWave, the attenuation is high, due to the high operating frequencies and only a few rays participate in the formation of the useful signal (typically $N_g \in \{1, 2, 3\}$). In the following, we assume that the BS knows the value of N_g, although solutions are available to estimate its value [18], [19]. Note that our model may easily include the presence of blockage, a typical phenomenon of mmWaves, by considering a higher path-loss $P_{i,\ell}$, [20], [21]. Moreover, we do not have any specific assumption on the distribution of the interferes that will affect in general the statistics of $v_{i,\ell}$, $\beta_{i,\ell}$, and $\gamma_{i,\ell}$.

B. LS Interference Correlation Matrix Estimate

A basic estimate of the interference correlation matrix $\boldsymbol{R}$ from the T samples is provided by the LS estimate (or sample covariance matrix)

$$\hat{\boldsymbol{R}}_{\mathrm{LS}}(T) = \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{N}(t)\boldsymbol{N}^{\mathrm{H}}(t). \quad (6)$$

Although for $T \to \infty$ we have $\hat{\boldsymbol{R}}_{\mathrm{LS}} \to \boldsymbol{R}$, the convergence of this estimator alone is slow, thus we consider techniques having a faster convergence. We will also integrate the LS estimate into advanced solutions that, starting from a still rough estimate, are able to refine it significantly.

III. Projection-based Correlation Estimate

In this section, we propose an estimator of the interference correlation matrix that exploits the structure of the interference channel given by (1). From (5), we immediately have

$$\boldsymbol{Y}^{(\ell)}(t) = \boldsymbol{G}(\ell)\boldsymbol{J}^{(\ell)}(t) = \sum_{i=0}^{N_g-1} \boldsymbol{a}_N(\beta_{i,\ell}) \underbrace{[v_{i,\ell}\boldsymbol{a}_{N_I}^{\mathrm{H}}(\gamma_{i,\ell})\boldsymbol{J}^{(\ell)}(t)]}_{x_{i,\ell}(t)}, \quad (7)$$

where (for fixed $v_{i,\ell}$) $\{x_{i,\ell}(t)\}$, $i = 0, \ldots, N_g - 1$, are zero-mean complex Gaussian variables (since $\boldsymbol{J}^{(\ell)}$ is assumed to be Gaussian) with cross-correlations

$$\begin{aligned}\mathbb{E}[x_{i,\ell}(t)x_{j,\ell}^*(t)] &= v_{i,\ell}v_{j,\ell}^*\mathbb{E}\left[\boldsymbol{a}_{N_I}^{\mathrm{H}}(\gamma_{i,\ell})\boldsymbol{J}^{(\ell)}(t)\boldsymbol{J}^{(\ell)H}(t)\boldsymbol{a}_{N_I}(\gamma_{j,\ell})\right] \\ &= v_{i,\ell}v_{j,\ell}^*\boldsymbol{a}_{N_I}^{\mathrm{T}}(\gamma_{i,\ell})\boldsymbol{R}_J\boldsymbol{a}_{N_I}\gamma_{j,\ell}).\end{aligned} \quad (8)$$

Therefore, from (1) and (7) we have

$$\boldsymbol{Y}(t) = \sum_{i=0}^{N_g-1}\sum_{\ell=1}^{L} \boldsymbol{a}_N(\beta_{i,\ell})x_{i,\ell}(t). \quad (9)$$

Note that we assume that both L and N_g are small numbers and in particular $N_gL << N$.

To obtain a more compact model, we collect the interference channels into the matrix $\boldsymbol{G} = [\boldsymbol{G}(1), \dots, \boldsymbol{G}(L)]$ and analogously we collect the transmitted interference symbol vectors into the column vector $\boldsymbol{J}(t) = [\boldsymbol{J}^{(1)\mathrm{T}}(t), \dots, \boldsymbol{J}^{(L)\mathrm{T}}(t)]^T$. Now, by defining

$$\begin{aligned} \boldsymbol{x}(\ell, t) = [x_{1,\ell}(t), x_{2,\ell}(t), \dots, x_{N_g-1,\ell}(t)]^T, \\ \boldsymbol{x}(t) = [\boldsymbol{x}^T(1,t), \dots, \boldsymbol{x}^T(L,t)]^T, \end{aligned} \tag{10}$$

we can rewrite the interference vector at sample time t as

$$\boldsymbol{N}(t) = \boldsymbol{G}\boldsymbol{J}(t) + \boldsymbol{Z}(t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{Z}(t), \tag{11}$$

with $\boldsymbol{A} = [\boldsymbol{A}(1), \dots, \boldsymbol{A}(L)]$ and $\boldsymbol{A}(\ell) = [\boldsymbol{a}_N(\beta_{0,\ell}), \dots, \boldsymbol{a}_N(\beta_{N_g-1,\ell})]$. From (11) we can write the interference vector as the sum of LN_g correlated Gaussian sources, arriving at the BS with receive phase shifts $\beta_{i,\ell}$, for $i = 0, \dots, N_g - 1$, and $\ell = 1, \dots, L$.

Thus, inserting (11) into (2), the interference correlation matrix can be written as

$$\boldsymbol{R} = \boldsymbol{A}\boldsymbol{R}_{\boldsymbol{x}}\boldsymbol{A}^{\mathrm{H}} + \sigma^2\boldsymbol{I}_N, \tag{12}$$

where $\boldsymbol{R}_{\boldsymbol{x}} = \mathbb{E}[\boldsymbol{x}(t)\boldsymbol{x}^{\mathrm{H}}(t)]$ is block-diagonal with L blocks, as we assumed that the signals transmitted by the L interferers are independent; entries of block ℓ are given by (8). From (12), we observe that the interference correlation matrix (neglecting noise) has a particular structure, as it can be decomposed into the product of three matrices: the two external matrices depend only on the AoAs of the interference signals, while the inner matrix depends only on the correlation of the interference signals coming from the different angles.

As we consider a scenario with few paths and few interferers, the number of rows (columns) N_gL of $\boldsymbol{R}_{\boldsymbol{x}}$ is much smaller than the number of rows (columns) N of $\boldsymbol{R}$. Moreover, although matrix $\boldsymbol{A}$ may have a large number of rows (N), it actually depends again on a small number of parameters. Therefore, we propose the projection-based correlation estimation (PBCE) technique that splits the correlation estimation into two sub-problems: a) the estimation of the receive phase shifts (from which an estimate of matrix $\boldsymbol{A}$ is obtained), and b) the estimate of the inner correlation matrix $\boldsymbol{R}_{\boldsymbol{x}}$.

In particular, PBCE works as follows. Once T received vectors $\boldsymbol{N}(t)$, $t = 1, \dots, T$, have been collected:

1) estimate the receive phase shifts as $\hat{\beta}_{i,\ell}$, $i = 0, \dots, N_g - 1$, $\ell = 1, \dots, L$, from $\boldsymbol{N}(t)$, $t = 1, \dots, T$;
2) obtain the LS estimate of the interference correlation matrix $\hat{\boldsymbol{R}}_{\mathrm{LS}}(T)$ as in (6);
3) remove the noise contribution from the LS estimate of the correlation matrix to obtain an estimate of the interference correlation
$$\hat{\boldsymbol{R}}'_{\mathrm{LS}}(T) = \hat{\boldsymbol{R}}_{\mathrm{LS}}(T) - \sigma^2\boldsymbol{I}_N; \tag{13}$$
4) project $\hat{\boldsymbol{R}}'_{\mathrm{LS}}(T)$ into the subspace defined by $\hat{\boldsymbol{A}}(\ell) = [\boldsymbol{a}_N(\hat{\beta}_{0,\ell}), \dots, \boldsymbol{a}_N(\hat{\beta}_{N_g-1,\ell})]$ and $\hat{\boldsymbol{A}} = [\hat{\boldsymbol{A}}(1), \dots, \hat{\boldsymbol{A}}(L)]$ to obtain an estimate of $\boldsymbol{R}_{\boldsymbol{x}}$ as
$$\hat{\boldsymbol{R}}_{\boldsymbol{x}}(T) = \hat{\boldsymbol{A}}^{\dagger}\hat{\boldsymbol{R}}'_{\mathrm{LS}}(T)(\hat{\boldsymbol{A}}^{\dagger})^{\mathrm{H}}; \tag{14}$$
5) obtain the new estimate of the interference correlation matrix as (see (12))
$$\hat{\boldsymbol{R}}_{\mathrm{PBCE}}(T) = \hat{\boldsymbol{A}}\hat{\boldsymbol{R}}_{\boldsymbol{x}}(T)\hat{\boldsymbol{A}}^{\mathrm{H}} + \sigma^2\boldsymbol{I}_N. \tag{15}$$

This is the explanation of the various points:

Point 1) refers to the estimate of the phase shifts and will be detailed in the next Section IV, where several techniques will be considered.

Point 2) provides the starting estimate of the interference correlation matrix, from which the inner correlation matrix $\boldsymbol{R}_x$ will be estimated, given the estimate of the AoA (thus the estimate of $\boldsymbol{A}$) obtained in point 1).

Point 3) elaborates on (12), first removing the AWGN noise contribution from the LS estimate of the interference correlation matrix.

Point 4) is obtained by observing that the LS estimate of $\boldsymbol{x}(t)$ from $\boldsymbol{N}(t)$, given the estimate $\hat{\boldsymbol{A}}$, is obtained as (see (11))

$$\hat{\boldsymbol{x}}_{\mathrm{LS}}(t) = \hat{\boldsymbol{A}}^{\dagger}\boldsymbol{N}(t). \tag{16}$$

Assuming that the angle estimates are correct ($\hat{\boldsymbol{A}} = \boldsymbol{A}$), the correlation matrix of $\hat{\boldsymbol{x}}_{\mathrm{LS}}(t)$ is

$$\mathbb{E}[\hat{\boldsymbol{x}}_{\mathrm{LS}}(t)\boldsymbol{x}^{\mathrm{H}}_{\mathrm{LS}}(t)] = \boldsymbol{A}^{\dagger}\boldsymbol{R}\boldsymbol{A}^{\dagger H} = \boldsymbol{R}_{\boldsymbol{x}} + \sigma^2\boldsymbol{A}^{\dagger}\boldsymbol{A}^{\dagger H}, \tag{17}$$

and its LS estimate is

$$\begin{aligned} \frac{1}{T}\sum_{t=1}^{T}\hat{\boldsymbol{x}}_{\mathrm{LS}}(t)\hat{\boldsymbol{x}}^{\mathrm{H}}_{\mathrm{LS}}(t) = \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{A}^{\dagger}\boldsymbol{N}(t)\boldsymbol{N}^{\mathrm{H}}(t)\boldsymbol{A}^{\dagger H} \\ = \boldsymbol{A}^{\dagger}\hat{\boldsymbol{R}}_{\mathrm{LS}}(T)\boldsymbol{A}^{\dagger H}. \end{aligned} \tag{18}$$

Then, considering (18) as an estimate of (17) and removing the contribution of noise, we obtain the estimate (14) for $\boldsymbol{R}_{\boldsymbol{x}}$.

Lastly, point 5) follows from (12), and the estimated matrices replace the true matrices. The resulting PBCE estimate of $\boldsymbol{R}$ can also be written as

$$\hat{\boldsymbol{R}}_{\mathrm{PBCE}}(T) = \hat{\boldsymbol{A}}\left[\hat{\boldsymbol{A}}^{\dagger}\hat{\boldsymbol{R}}_{\mathrm{LS}}(T)\hat{\boldsymbol{A}}^{\dagger H} - \sigma^2\hat{\boldsymbol{A}}^{\dagger}\hat{\boldsymbol{A}}^{\dagger H}\right]\hat{\boldsymbol{A}}^{\mathrm{H}} + \sigma^2\boldsymbol{I}_N. \tag{19}$$

A. MSE For Correct Phase Shift Estimates

We now compute the mean square error (MSE) of the interference correlation matrix estimate under the hypothesis of correct receive phase-shift estimates and compare it with that obtained with the LS estimate. The MSE is defined as

$$\Gamma = \frac{1}{N^2}\sum_{m,n}\mathbb{E}[|\hat{\boldsymbol{R}}_{n,m} - \boldsymbol{R}_{n,m}|^2]. \tag{20}$$

First, note that both the LS and the PBCE estimates (under correct angle estimation) are unbiased, i.e., $\mathbb{E}[\hat{\boldsymbol{R}}_{\mathrm{LS}}] = \mathbb{E}[\hat{\boldsymbol{R}}_{\mathrm{PBCE}}] = \boldsymbol{R}$. Then, for the LS estimate we have

$$\Gamma_{\mathrm{LS}}(T) = \frac{1}{N^2}\sum_{m,n}\mathbb{E}\left[\left|\frac{1}{T}\sum_{t=1}^{T}[\boldsymbol{N}(t)\boldsymbol{N}^{\mathrm{H}}(t)]_{n,m} - [\boldsymbol{R}]_{n,m}\right|^2\right]. \tag{21}$$

Now, considering that $[\boldsymbol{N}(t)\boldsymbol{N}^{\mathrm{H}}(t)]_{n,m}$ are zero-mean and independent in t, we have

$$\Gamma_{\mathrm{LS}}(T) = \frac{1}{N^2}\sum_{m,n}\mathbb{E}[|[\boldsymbol{N}(t)\boldsymbol{N}^{\mathrm{H}}(t)]_{n,m} - [\boldsymbol{R}]_{n,m}|^2], \tag{22}$$

and from the results of the Appendix, we have

$$\begin{aligned}\Gamma_{\rm LS}(T) =&\frac{1}{TN^2}\sum_{m,n}[\boldsymbol{R}]_{n,n}[\boldsymbol{R}]_{m,m}\\ =&\frac{1}{TN^2}\mathrm{trace}^2(\boldsymbol{R}).\end{aligned} \tag{23}$$

For the PBCE, from (19) and assuming $\hat{\boldsymbol{A}} = \boldsymbol{A}$, we have

$$\begin{aligned}\Gamma_{\rm PBCE}(T) =&\frac{1}{N^2}\sum_{m,n}\mathbb{E}[|[\boldsymbol{R}_{\rm PBCE}]_{n,m} - [\boldsymbol{R}]_{n,m}|^2]\\ =&\frac{1}{N^2}\sum_{m,n}\mathbb{E}\left[\left|\frac{1}{T}\sum_{t=1}^{T}\left[\boldsymbol{A}\left[\boldsymbol{A}^{\dagger}\boldsymbol{N}(t)\boldsymbol{N}^{\rm H}(t)\boldsymbol{A}^{\dagger H}\right.\right.\right.\right.\\ &\left.\left.\left.\left.-\sigma^2\boldsymbol{A}^{\dagger}\boldsymbol{A}^{\dagger H}\right]\boldsymbol{A}^{\rm H} + \sigma^2\boldsymbol{I}_N\right]_{n,m} - [\boldsymbol{R}]_{n,m}\right|^2\right].\end{aligned} \tag{24}$$

Since each term of the summation is zero-mean and independent also in this case, we have

$$\begin{aligned}\Gamma_{\rm PBCE}(T) =& \frac{1}{TN^2}\sum_{m,n}\mathbb{E}\Big[\Big|\left[\boldsymbol{A}(\boldsymbol{A}^{\dagger}\boldsymbol{N}(t)\boldsymbol{N}^{\rm H}(t)\boldsymbol{A}^{\dagger H} - \right.\\ &\left.\sigma^2\boldsymbol{A}^{\dagger}\boldsymbol{A}^{\dagger H})\boldsymbol{A}^{\rm H} + \sigma^2\boldsymbol{I}_N\right]_{n,m} - [\boldsymbol{R}]_{n,m}\Big|^2\Big]\\ =&\frac{1}{TN^2}\sum_{m,n}\mathbb{E}\Big[\,|[\boldsymbol{A}\boldsymbol{A}^{\dagger}\boldsymbol{N}(t)\boldsymbol{N}^{\rm H}(t)\boldsymbol{A}^{\dagger H}\boldsymbol{A}^{\rm H}]_{n,m} -\\ &\left[\sigma^2\boldsymbol{A}\boldsymbol{A}^{\dagger}\boldsymbol{A}^{\dagger H}\boldsymbol{A}^{\rm H} - \sigma^2\boldsymbol{I}_N + \boldsymbol{R}\right]_{n,m}\Big|^2\Big]\end{aligned} \tag{25}$$

and again from the results of the Appendix, we have

$$\begin{aligned}\Gamma_{\rm PBCE}(T) =&\frac{1}{TN^2}\sum_{m,n}[\boldsymbol{A}\boldsymbol{A}^{\dagger}\boldsymbol{R}\boldsymbol{A}^{\dagger H}\boldsymbol{A}^{\rm H}]_{n,n}[\boldsymbol{A}\boldsymbol{A}^{\dagger}\boldsymbol{R}\boldsymbol{A}^{\dagger H}\boldsymbol{A}^{\rm H}]_{m,m}\\ =&\frac{1}{TN^2}\mathrm{trace}^2[\boldsymbol{A}\boldsymbol{A}^{\dagger}\boldsymbol{R}\boldsymbol{A}^{\dagger H}\boldsymbol{A}^{\rm H}].\end{aligned} \tag{26}$$

Since $\boldsymbol{A}^{\dagger}\boldsymbol{A} = \boldsymbol{I}_{LN_g}$ and using (12), we have

$$\Gamma_{\rm PBCE}(T) =\frac{1}{TN^2}\mathrm{trace}^2[\boldsymbol{R} - \sigma^2(\boldsymbol{I}_N - \boldsymbol{A}\boldsymbol{A}^{\dagger}\boldsymbol{R}\boldsymbol{A}^{\dagger H}\boldsymbol{A}^{\rm H})]. \tag{27}$$

IV. RECEIVE PHASE-SHIFT ESTIMATION

For the estimation of the receive phase shifts $\hat{\beta}_{i,\ell}$, $i = 0,\ldots,N_g-1$, $\ell = 1,\ldots,L$, we exploit the assumptions that interferers' channels have only a few paths and there are only a few interferers. This yields a sparse Fourier transform of the correlation matrix, i.e., its sparse representation in the *angular* domain. We consider three solutions, based on the gridless approach of [17], which has been proven to provide accurate estimates even from a single or a few samples (small T). Still, as T grows, its complexity becomes prohibitive, and we propose also two simplified approaches. In particular, first, we propose the gridless angle estimation (GAE) based on [17], [22]; in the second approach, denoted as subspace gridless estimation (SGE), the receive phase shifts are estimated from the LS-estimated interference correlation matrix; lastly, in the third approach, denoted gridless estimation and clustering (GEC), GAE is performed on windows of samples and the estimated phase shifts are clustered.

A. Gridless Angle Estimation

First, we consider a solution where the AoAs are not forced to be on a grid but can take continuous values, in what is called a *gridless* estimation. In particular, we resort to the approach of [17], [22], which leverages the equivalence between the computation of the ℓ_0 atomic norm (AN) and a rank minimization problem restricted to the set of positive semidefinite canonical multi-level Toeplitz matrices.

In the GAE algorithm, the receive phase shifts are estimated by imposing a sparsity (ℓ_0 norm) on the estimated vectors $\hat{\boldsymbol{Y}}(t)$, $t = 1,\ldots,T$. First, we recall that the ℓ_0 AN of vectors $\{\hat{\boldsymbol{Y}}(t)\}$ in the atomic set $\mathcal{E} = \{\boldsymbol{a}_N(\beta), \beta \in [0, 2\pi)\}$, is defined as

$$||\{\hat{\boldsymbol{Y}}(t)\}||_{\mathcal{E},0} = \inf_{\{\zeta_p\},\{q_p\}}\left\{P : \hat{\boldsymbol{Y}}(t) = \sum_{p=1}^{P} q_p(t)\boldsymbol{a}_N(\zeta_p),\ \forall t\right\}. \tag{28}$$

Note that this definition is consistent with (9) when $\{\zeta_p\}$ are the estimated receive phase shifts. Then, we aim at minimizing the MSE between $\boldsymbol{N}(t)$ and $\hat{\boldsymbol{Y}}(t)$, under the sparsity constraint on $\hat{\boldsymbol{Y}}(t)$, with $t = 1,\ldots,T$, i.e.,

$$\min_{\{\hat{\boldsymbol{Y}}(t)\}} \frac{1-\eta}{T}\sum_{t=1}^{T}||\boldsymbol{N}(t) - \hat{\boldsymbol{Y}}(t)||_2^2 + \eta||\{\hat{\boldsymbol{Y}}(t)\}||_{\mathcal{E},0}, \tag{29}$$

where $\eta \in (0, 1)$ is a parameter ruling the sparsity of the obtained solution. The solution of (29) provides $\hat{\boldsymbol{Y}}(t)$, $t = 1,\ldots,T$, which in turn provide the estimated receive phase shifts $\hat{\beta}_{i,\ell}$, $i = 0,\ldots,N_g-1$, $\ell = 1,\ldots,L$, through (28).

Now, the function (29) is non-convex, making its minimization hard. Therefore, we consider its l_1-AN relaxation to obtain a simpler problem. Let $\mathcal{T}_{\rm N}$ be the set of positive semidefinite canonical 1-Level Toeplitz matrices with dimension N (see [17] for its definition and properties). The relaxation of (29) leads to the following problem [17]

$$\min_{\tau,\{\boldsymbol{q}(t)\},\boldsymbol{Q}\in\mathcal{T}_{\rm N}} \frac{1-\eta}{T}\sum_{t=1}^{T}||\boldsymbol{N}(t) - \hat{\boldsymbol{Y}}(t)||_2^2 + \frac{\eta}{2}[\tau + \mathrm{trace}(\boldsymbol{Q})], \tag{30a}$$

s.t.

$$\begin{bmatrix}\boldsymbol{Q} & \hat{\boldsymbol{Y}}(t)\\ \hat{\boldsymbol{Y}}^{\rm H}(t) & \tau\end{bmatrix} \succcurlyeq 0, \quad t = 1,\ldots,T. \tag{30b}$$

This is a convex problem that can be solved with standard methods. The new estimates of the receive phase shifts $\hat{\beta}_{i,\ell}$, $i = 0,\ldots,N_g-1$, $\ell = 1,\ldots,L$, are then obtained by the Vandermonde decomposition of matrix $\boldsymbol{Q}$, as from the [17, Algorithm 1] (see also [22, Lemma 1]).

Complexity Analysis: As shown in [17], [23], solving (30) requires

$$\mathcal{N}(N,T,\epsilon) = \mathcal{O}(T(N^3 + TN^2 + T^2)\sqrt{N}\log(1/\epsilon)), \tag{31}$$

arithmetic operations, where ϵ is the accuracy parameter. We note that as T grows, the number of constraints in (30) also grows linearly with T, and the complexity of solving the optimization problem grows as $\mathcal{O}(T^3)$, as shown in [17]. Hence, we now consider two simplified solutions that are adequate when the number of samples grows large.

B. Subspace Gridless Estimation

The second solution is based on the estimation of the receive phase shifts from the LS estimate of the interference matrix, rather than on samples $\boldsymbol{N}(t)$, $t = 1, \ldots, T$. This approach holds when the number of interfering signals is smaller than the number of BS antennas, i.e.,

$$S = LN_g < N. \tag{32}$$

First, we observe that we can remove the contribution of the noise by subtracting the noise correlation as in (13), obtaining $\hat{\boldsymbol{R}}'_{\rm LS}(T) = \hat{\boldsymbol{R}}_{\rm LS}(T) - \sigma^2 \boldsymbol{I}_N$. Then, we can improve the LS estimate by imposing the Hermitian symmetry, obtaining

$$\hat{\boldsymbol{R}}'_{\rm LS}(T) = \frac{1}{2}\left[\hat{\boldsymbol{R}}'_{\rm LS}(T) + \hat{\boldsymbol{R}}'^{H}_{\rm LS}(T)\right]. \tag{33}$$

Now, from (12), we observe that the columns of the true interference-only correlation matrix

$$\boldsymbol{R}' = \boldsymbol{R} - \sigma \boldsymbol{I}_N \tag{34}$$

lay in a sub-space of size S. Therefore, we can decompose the LS estimate of the interference-only correlation by singular value decomposition (SVD) as

$$\hat{\boldsymbol{R}}'_{\rm LS}(T) = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{\rm H}, \tag{35}$$

where $\boldsymbol{V}$ is unitary and $\boldsymbol{\Lambda}$ is diagonal with real non-negative entries $\lambda_1 \geq \ldots \geq \lambda_N$, denoted as singular values.

To partially remove the estimation error, we define the tall $N \times S$ sub-matrix $\boldsymbol{\Lambda}'$ of $\boldsymbol{\Lambda}$ containing the S largest singular values, i.e., $[\boldsymbol{\Lambda}']_{s,s} = \lambda_s$, while all other entries of $\boldsymbol{\Lambda}'$ are zero, and compute the new correlation estimate

$$\hat{\boldsymbol{R}}''_{\rm LS}(T) = \boldsymbol{V}\boldsymbol{\Lambda}'\boldsymbol{V}^{\rm H}. \tag{36}$$

Until now we have reduced the estimation error in the LS estimate by exploiting the properties of the SVD of the interference correlation matrix. Now, by comparing (36) with (12), we note that in the columns of $\boldsymbol{V}$ we have the steering vectors with the receive phase shifts. Thus, the SGE algorithm exploits this property. To this end, we first define the *square root* matrix of $\hat{\boldsymbol{R}}''_{\rm LS}(T)$ as

$$\boldsymbol{S} = \boldsymbol{V}\,{\rm diag}\{\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_S}\}, \tag{37}$$

and then we apply the GAE method on the columns of $\boldsymbol{S}$. In particular, we solve (30) with $T = S$ and $\boldsymbol{N}(t)$ replaced by the columns of $\boldsymbol{S}$.

Complexity Analysis: The number of operations required by this approach is

$$\mathcal{N}(N, S, \epsilon) = \mathcal{O}(N^3 + S(N^3 + SN^2 + S^2)\sqrt{N}\log(1/\epsilon)), \tag{38}$$

where, with respect to (31) we added also the operations required by the SVD (N^3). Note that, in this case, regardless of the value of T, the number of constraints is always S, thus limiting the computational complexity of this solution.

C. Gridless Estimation and Clustering

We propose also the GEC algorithm, wherein GAE is performed on windows of T_0 samples, and the receive phase shifts are estimated on several windows. They are then suitably clustered and the new receive phase shift estimates are the cluster heads. The obtained solution is denoted as GEC.

In detail, let us first define the subset of observations obtained on window n of T_0 samples as

$$\mathcal{N}(n) = \{\boldsymbol{N}(t), t = nT_0 + 1, \ldots, nT_0 + T_0\}. \tag{39}$$

Then, the GEC algorithm works as follows. At time nT_0, for $n \geq 1$:

1) apply the GAE on $\mathcal{N}(n)$ to obtain the estimated LN_g receive phase shifts $\{\hat{\beta}_{1,1}(n), \ldots, \hat{\beta}_{N_g-1,L}(n)\}$;
2) cluster the n previously estimated receive phase shifts $\{\hat{\beta}_{1,1}(1), \ldots, \hat{\beta}_{N_g-1,L}(n)\}$ into S clusters, as detailed in the following;
3) obtain the estimate of the receive phase shifts as the centroids of the S clusters.

We now detail the clustering procedure and the computation of the centroids. First, note that the k-means algorithm [21] cannot be immediately applied on the estimated receive phase shifts, since it does not ensure that the centroids are phase shifts in the interval $[0, 2\pi)$. Therefore, we first convert each estimated receive phase shift $\hat{\beta}_{i,\ell}(n)$ into the two-dimensional vector

$$\hat{\boldsymbol{\beta}}_{i,\ell}(n) = [\cos(2\pi\hat{\beta}_{i,\ell}(n)), \sin(2\pi\hat{\beta}_{i,\ell}(n))]. \tag{40}$$

Hence, the set of all vectors at time nT_0 is

$$\bar{\mathcal{B}} = \{\hat{\boldsymbol{\beta}}_{0,1}(1), \ldots, \hat{\boldsymbol{\beta}}_{N_g-1,L}(1), \ldots, \hat{\boldsymbol{\beta}}_{0,1}(n), \ldots, \hat{\boldsymbol{\beta}}_{N_g-1,L}(n)\}. \tag{41}$$

The S clusters are sets $\mathcal{B}_1, \ldots, \mathcal{B}_S$, with $|\mathcal{B}_s| = n/S$: these sets constitute a partition of $\bar{\mathcal{B}}$, i.e., $\bigcup_s \mathcal{B}_s = \bar{\mathcal{B}}$ and $\mathcal{B}_{s_1} \cap \mathcal{B}_{s_2}$ for $s_1 \neq s_2$. We use the MSE metric for clustering and finding the cluster head

$$\boldsymbol{\beta}_s = {\rm argmin}_{\boldsymbol{\zeta}} \sum_{\boldsymbol{\beta}\in\mathcal{B}_s} ||\boldsymbol{\zeta} - \boldsymbol{\beta}||_2^2 = \sum_{\boldsymbol{\beta}\in\mathcal{B}_s} \boldsymbol{\beta} \tag{42}$$

Lastly, we convert $\boldsymbol{\beta}_s$ into the estimated phase shift as

$$\beta_s = {\rm atan2}(\boldsymbol{\beta}_s). \tag{43}$$

Complexity Analysis: To obtain equal-size clusters of observations, we can resort for example to the algorithm of [24], having complexity $\mathcal{O}(T^{1.7})$, thus the number of operations required by this approach is

$$\mathcal{N}(N, T, \epsilon) = \mathcal{O}(T^{1.7} + T_0(N^3 + T_0N^2 + T_0^2)\sqrt{N}\log(1/\epsilon)). \tag{44}$$

D. MUSIC Angle Estimation

For comparison purposes, we consider the MUSIC algorithm [25] to estimate the receive phase shift. This algorithm is based on the SVD of the LS estimate; as for SGE, we consider the LS-estimated interference-only correlation matrix $\hat{\boldsymbol{R}}'_{\rm LS}(T)$ and its SVD (35). Then, the receive phase shifts $\{\beta_{i,\ell}\}$ are

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2023.3291917

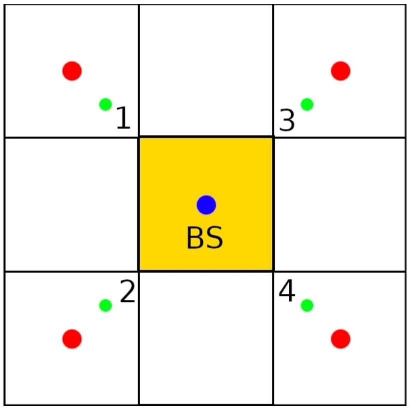


Fig. 1. Interference environment with the main BS (blue dot), neighborhood cells (white squares) with their BSs (red dots), and interferers (green dots).

estimated by finding the LN_g values of θ corresponding to the largest peaks of the frequency estimation function

$$P_{\rm MU}(\theta) = \frac{1}{\boldsymbol{a}_N^{\rm H}(\theta)\boldsymbol{V}\boldsymbol{V}^{\rm H}\boldsymbol{a}_N(\theta)}. \tag{45}$$

Note that the peak is searched in a continuous space; however, in practice, θ is sampled in a discrete space over which the maxima are searched, thus a grid approach is obtained.

Complexity Analysis: The complexity of MUSIC then depends on the number of points of the grid $N_{\rm G}$, and can be written as [26]

$$\mathcal{O}(N^2(LN_g + T + N_{\rm G})). \tag{46}$$

V. Numerical Results

We assess now the performance of the proposed interference correlation matrix in the scenario of Fig. 1, where the *main* BS (blue point) is at the center of its square cell (in yellow), and the cell is surrounded by 4 neighboring cells (in white), in which $L = 4$ interferers (green points) communicate with their own secondary BSs (red points). The interferers are placed along the segments between the main and the secondary BSs, which define the average AoA for channels $\boldsymbol{G}(\ell)$. For simulation purposes, we draw ray amplitudes $v_{i,\ell}$ from a complex Gaussian distribution with zero mean and unitary variance, i.e., $P_{i,\ell} = 1$, modeling interferers in no-line-of-sight (nLOS) condition with respect to the main BS. Note that such a scenario is the worst-case for interference estimation since all channel paths are weak and thus particularly challenging to be estimated.

The main BS is equipped with $N = 32$ antennas, and the channel is modeled as in Section II.A, with $N_g = 3$ rays. AoAs $\{\alpha_{i,\ell}^{(\rm Rx)}\}$ are uniformly distributed with mean given by the angle between the transmitting interferer and the main BS and support size $\frac{\pi}{6}$, which corresponds to a cell with 12 sectors (a higher-order sectorization, see [27]), while AoDs $\{\alpha_{i,\ell}^{(\rm Tx)}\}$ are uniformly distributed in the interval $[0, \pi]$. The antenna spacings are $D_{\min}^{(\rm Tx)} = D_{\min}^{(\rm Rx)} = \frac{\lambda}{2}$, for a carrier frequency $f_c = 28$ GHz.

The performance will be shown as a function of the ROT, defined as the ratio between the interference power and the noise power, i.e.,

$$\text{ROT} = \frac{\frac{1}{N}\text{trace}\boldsymbol{R} - \sigma^2}{\sigma^2} = \frac{\frac{1}{N}\text{trace}\boldsymbol{R}}{\sigma^2} - 1. \tag{47}$$

In particular, in the following we vary the noise power σ^2 while leaving unchanged the interference power. Therefore, from (47), we note by increasing the noise power we decrease the ROT.

We focus on the case of single-antenna users and interferes ($N_I = 1$) and the received vector signal at sample t (including also the data signal, not only noise and interference) is

$$\boldsymbol{r}(t) = \boldsymbol{h}s(t) + \boldsymbol{N}(t), \tag{48}$$

with $\boldsymbol{h}$ the column channel vector between the transmitter in the main cell and the BS and $s(t)$ the transmitted useful data symbol at time t, assumed to be Gaussian zero-mean with variance σ_x^2. We assume here that the BS perfectly knows $\boldsymbol{h}$.

We assume that the BS requests from the transmitter a data rate R to be used for uplink transmission. Since the BS does not have a perfect estimate of the interference correlation matrix, its estimate $\hat{C}$ of the achievable rate in the uplink channel is affected by errors. Thus, the BS will transmit at rate $R = \delta\hat{C}$, with $\delta \in [0, 1]$, suitably chosen, as detailed in the following. Note that even with this choice, *outages* may occur, when $R > C$, and C is the true achievable rate of the uplink channel.

A. Performance Metrics

Performance has been compared in terms of a) MSE of the estimate of the interference correlation matrix, b) achievable data rate, and c) effective throughput (taking into account outages).

MSE: The MSE of the estimate of the interference correlation matrix is defined in (20).

Achievable Rate: To define the achievable rate and the throughput, we first revise the operations performed at the BS upon reception of a message. We recall that we treat the interference as additional noise, but we explicitly exploit the knowledge of its correlation matrix, estimated with the techniques presented in this paper. Thus, the BS first whitens the interference on $\boldsymbol{r}(t)$. This is achieved by decomposing the estimated interference correlation matrix $\hat{\boldsymbol{R}}$ by computing its SVD $\hat{\boldsymbol{R}} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{\rm H}$ and then whitening the signal to obtain

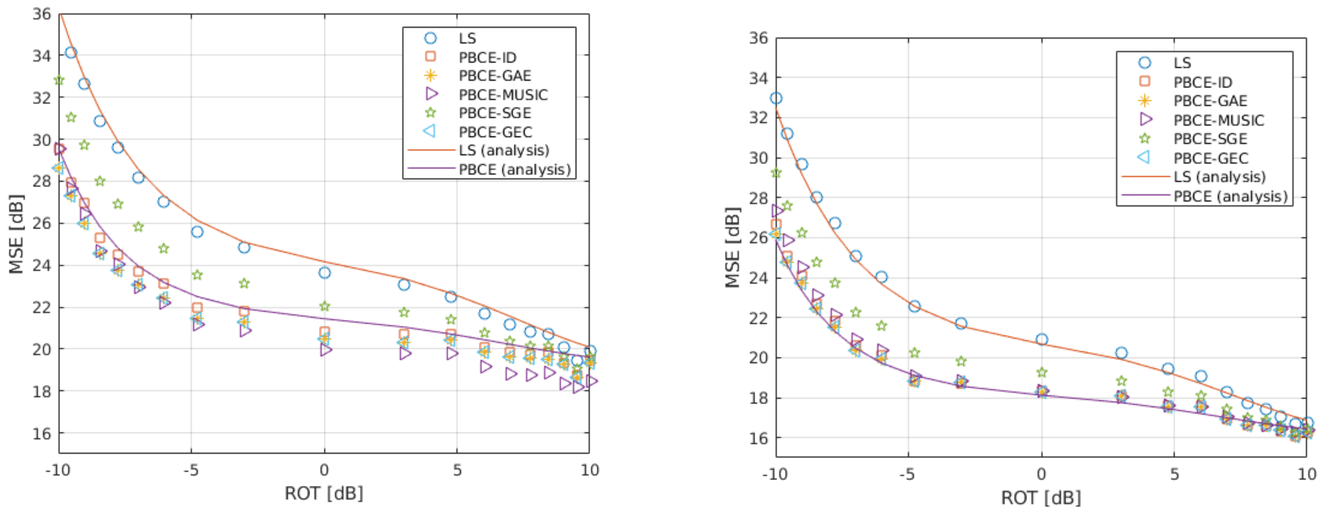
$$\boldsymbol{r}'(t) = \boldsymbol{\Lambda}^{-1/2}\boldsymbol{U}^{\rm H}\boldsymbol{r}(t) = \boldsymbol{g}s(t) + \boldsymbol{N}'(t), \tag{49}$$

where

$$\boldsymbol{g} = \boldsymbol{\Lambda}^{-1/2}\boldsymbol{U}^{\rm H}\boldsymbol{h} \tag{50}$$

is the equivalent channel and $\boldsymbol{N}'(t)$ is the noise term (that includes the noise-like interference). Note that if the estimation of the interference correlation matrix is perfect, i.e., $\hat{\boldsymbol{R}} = \boldsymbol{R}$, then $\boldsymbol{N}'(t)$ has an identity correlation matrix. Otherwise, it is still colored with correlation matrix

$$\boldsymbol{R}_{\boldsymbol{N}'} = \boldsymbol{\Lambda}^{-1/2}\boldsymbol{U}^{\rm H}\boldsymbol{R}\boldsymbol{U}\boldsymbol{\Lambda}^{-1/2}. \tag{51}$$



a) Case $T = 2$ samples.

b) Case $T = 4$ samples.

Fig. 2. MSE vs raise-over-thermal (ROT) for $T = 2$ and $T = 4$ samples and various interference correlation matrix estimators.

However, the BS is not aware of the residual correlation and cannot exploit it to compute the achievable data rate. Assuming for the sake of simplicity a single user transmitting in the uplink to the main BS (i.e., a resource block is assigned to a single user), the BS then applies a maximal ratio combiner (MRC) to obtain

$$\boldsymbol{r}''(t) = \boldsymbol{g}^{\rm H}\boldsymbol{r}'(t) = \boldsymbol{g}^{\rm H}\boldsymbol{g}s(t) + \boldsymbol{g}^{\rm H}\boldsymbol{N}'(t), \tag{52}$$

where now the correlation of the noise becomes $\boldsymbol{g}^{\rm H}\boldsymbol{R}_{\boldsymbol{N}'}\boldsymbol{g}$. Hence, and the true achievable data rate at the BS is

$$C = \log_2\left(1 + \frac{(\boldsymbol{g}^{\rm H}\boldsymbol{g})^2\sigma_x^2}{\boldsymbol{g}^{\rm H}\boldsymbol{R}_{\boldsymbol{N}'}\boldsymbol{g}}\right). \tag{53}$$

Throughput: The throughput takes into account outages in the uplink transmission. The estimated achievable rate at the BS is

$$\hat{C} = \log_2\left(1 + \boldsymbol{g}^{\rm H}\boldsymbol{g}\sigma_x^2\right). \tag{54}$$

When $\hat{\boldsymbol{R}} = \boldsymbol{R}$, $\hat{C} = C$ and this is also the spectral efficiency of the system, denoted as $C_{\rm opt}$: it provides an upper bound to the achievable rate $\hat{C}$. The throughput is then defined as

$$\rho = \begin{cases} 0, & \delta\hat{C} > C, \\ \delta\hat{C}, & \text{otherwise}. \end{cases} \tag{55}$$

We choose δ to maximize $\mathbb{E}[\rho]$: this is obtained by numerical methods.

B. Compared Solutions

We compare the following techniques for the interference correlation matrix:

- LS: the LS estimate of Section II-B;
- PBCE-GAE: PBCE using the GAE algorithm of Section IV-A for the estimate of the receive phase shifts;
- PBCE-SGE: PBCE using the SGE algorithm of Section IV-B for the estimate of the receive phase shifts;
- PBCE-GEC: PBCE using the GEC algorithm of Section IV-C for the estimate of the receive phase shifts;
- PBCE-MUSIC: PBCE using the MUSIC algorithm of Section IV-D for the estimate of the receive phase shifts;
- PBCE-ideal (ID): PBCE with the perfect estimate of the receive phase shifts.

We have considered $T_0 = 1$ and $N_{\rm G} = 10^3$.

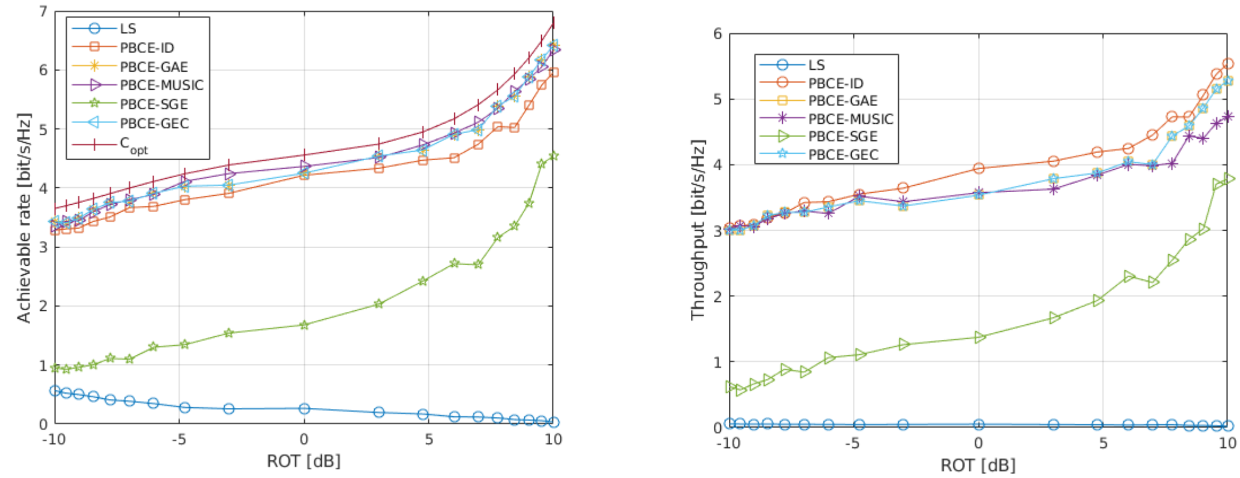
C. Performance Results

First, we compare the performance in terms of MSE of the interference correlation matrix estimate.

Fig. 2 shows the interference correlation matrix MSE (20) as a function of the ROT for the considered estimators and two values of the number of observed samples, namely $T = 2$ and 4. Note that we consider a very small number of samples to better assess the performance of the estimators based on the AoAs. In the following, we will also consider larger values of T. We first observe that as the noise increases, the ROT decreases, and the advantage of the proposed techniques becomes more evident since they distinguish noise from interference better than the LS technique. When comparing the various PBCE algorithms, we note that the PBCE-SGE provides only a partial improvement with respect to the LS estimator, while other approaches further reduce the MSE and achieve similar performance. Fig. 2 shows also the MSE obtained with the analysis of Section III-A, which well matches the simulation results.

We now consider both the achievable rate and the throughput as performance metrics.

Fig. 3 shows both the achievable rate and the throughput as a function of the ROT for $T = 2$ samples and various correlation estimation methods. We observe that the LS estimate for such a small value of T provides a very low achievable rate, much lower than that achieved when using PBCE techniques for the interference correlation matrix estimation. Indeed, all PBCE methods achieve a rate close to the capacity $C_{\rm opt}$, all



a) Achievable rate b) Throughput

Fig. 3. Achievable rate (a) and throughput (b) vs ROT using $T = 2$ samples and various interference correlation matrix estimators.

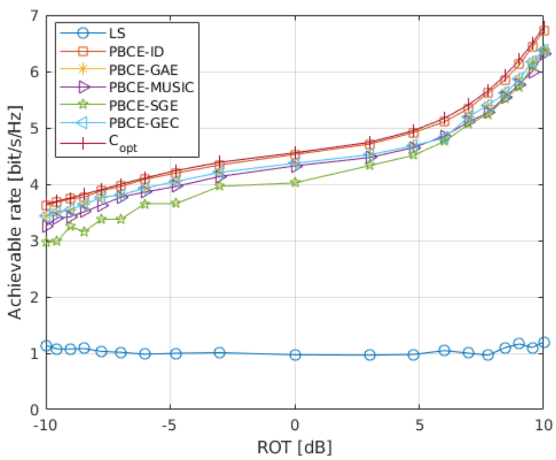


Fig. 4. Achievable rate vs ROT with $T = 5$ samples and various interference correlation matrix estimators.

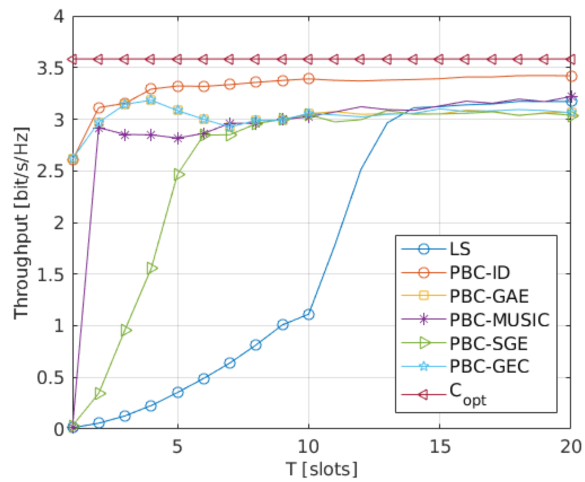


Fig. 5. Throughput vs T with $ROT = -10$ dB and various interference correlation matrix estimators.

with similar performance, except for the PBCE-SGE method, showing a higher sensitivity to the interference and yielding an achievable rate close to LS for high ROT values. Lastly, note that all PBCE techniques exhibit an achievable rate reduction as the ROT increases, since the overall interference increases. For the LS technique, the estimation is very bad and the slight increase of the rate with the ROT is due to the fact that interference is structured and the whitening becomes a bit more effective. When comparing Fig. 2 and Fig. 3, we conclude that the MSE of the correlation matrix estimate is not indicative of the data rate performance, as the correlation matrix is used for whitening (see (50) and (53)) and the impact of correlation matrix estimation errors is not linear on the data rates.

Fig. 4 shows the achievable rate as a function of the ROT, when $T = 5$ samples, for various correlation estimation methods. The trend is similar to that of Fig. 3, with the PBCE solutions getting closer to the upper bound, and the PBCE-SGE approach significantly increasing the achievable rate and aligning with the other techniques. Even the LS approach yields a higher achievable rate but is still considerably reduced with respect to that of the PBCE techniques. The LS method is still significantly affected by the estimation error, providing a very low achievable rate: indeed, the estimation error is so dominant that it turns out to be almost insensitive to the ROT.

Lastly, Figs. 5 and 6 show the throughput as a function of the number of samples T, for $ROT = -10$ and 0 dB, respectively. As expected, a larger number of samples used for the estimation provides higher throughput. However, the PBCE approaches show a much faster increase of the throughput with the number of samples than the LS approach, which is a significant advantage of the proposed solutions. Indeed, the LS approach suffers from the increased interference and requires more samples to achieve values close to the optimal when the ROT increases, and even when using more than 10

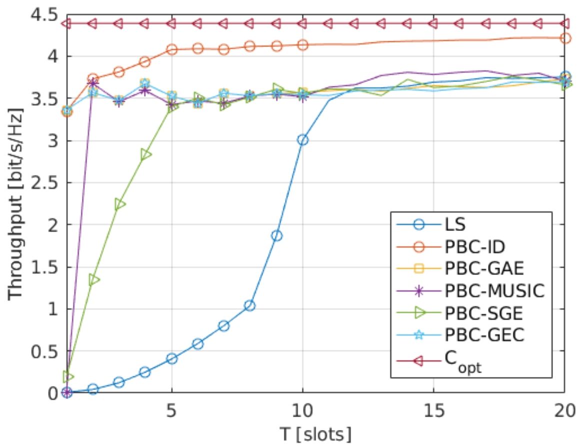


Fig. 6. Throughput vs T with $ROT = 0$ dB and various interference correlation matrix estimators.

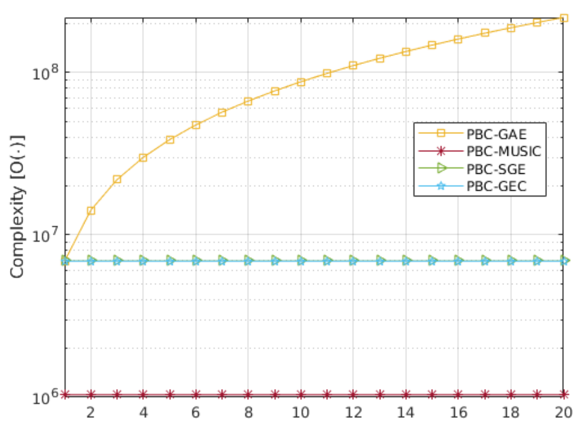


Fig. 7. Complexity comparison among the techniques.

samples, the throughput improvement is very slow. Note also that for high values of ROT, the PBCE method based on GEC outperforms PBCE-MUSIC providing a 10% increase in the throughput.

In summary, we note that the proposed GEC methods outperform both the LS and the MUSIC techniques, at least at certain values of ROT. Moreover, GEC offer a faster convergence than MUSIC.

D. Complexity Comparison

We have also compared the complexity of the various receive phase-shift estimation algorithms, as shown in Fig. 7 that reports the asymptotic ($\mathcal{O}(\cdot)$) number of operations required as a function of T, with all other parameters set as in the other performance figures. We note that PBCE-GAE has the highest complexity, which also grows fast with T, while both PBCE-SGE and PBCE-GEC have a complexity independent of T, as indeed it is related to T_0 and S, which are here kept constant. The MUSIC algorithm exhibits the lowest complexity, as expected. From the figure we first appreciate the reduction of complexity achieved by PBCE-SGE and PBCE-GEC with respect to PBCE-GAE. Then, taking into account the performance (in terms of throughput) achieved by the proposed PBCE techniques, we note that this advantage over MUSIC comes at the cost of higher complexity.

VI. Conclusions

In this paper, we have proposed three techniques for the estimation of the interference correlation matrix in a cellular system. The approaches are based on the estimation of the AoAs of interference signals for channels with few rays. The investigation of several techniques for the estimation of the receive phase shift and their performance comparison in a simulated environment has shown that the GAE algorithm is very effective and that the sub-optimal (and less complex) solution GEC is also well-performing, especially when more symbols are available for the interference correlation estimate. A future study related to our proposed estimators may be devoted to the design of interference correlation matrix estimators when hybrid analog-digital receive structures are employed: in this case, the constraints on the number of radio-frequency chains and thus the accessible signals call for specific solutions that exploit both the geometry of the receive antennas and their grouping.

References

[1] Y. Yuan, S. Wang, Y. Wu, H. V. Poor, Z. Ding, X. You, and L. Hanzo, "NOMA for next-generation massive IoT: Performance potential and technology directions," *IEEE Communications Magazine*, vol. 59, no. 7, pp. 115–121, Sept. 2021.

[2] H. Kim, J. Kim, and D. Hong, "Dynamic TDD systems for 5G and beyond: A survey of cross-link interference mitigation," *IEEE Communications Surveys Tutorials*, vol. 22, no. 4, pp. 2315–2348, Apr. 2020.

[3] L. Tlebaldiyeva, G. Nauryzbayev, S. Arzykulov, and A. M. Eltawil, "Performance of NOMA-based mmwave D2D networks under practical system conditions," *IEEE Access*, vol. 9, pp. 160 958–160 974, 2021.

[4] H. A. Ammar, R. Adve, S. Shahbazpanahi, G. Boudreau, and K. V. Srinivas, "User-centric cell-free massive MIMO networks: A survey of opportunities, challenges and solutions," *IEEE Communications Surveys Tutorials*, vol. 24, no. 1, pp. 611–652, Jan. 2022.

[5] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 264–273, Feb. 2013.

[6] D. D. Lin, R. Pacheco, T. J. Lim, and D. Hatzinakos, "Joint estimation of channel response, frequency offset, and phase noise in OFDM," *IEEE Transactions on Signal Processing*, vol. 54, no. 9, pp. 3542–3554, Sept. 2006.

[7] B. Classon, A. Nimbalker, S. Sesia, and I. Toufik, *Link Adaptation and Channel Coding*. John Wiley & Sons, Ltd, 2011, ch. 10, pp. 215–248. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470978504.ch10

[8] D. Neumann, M. Joham, and W. Utschick, "Covariance matrix estimation in massive MIMO," *IEEE Signal Processing Letters*, vol. 25, no. 6, pp. 863–867, 2018.

[9] J. Iscar, I. Güvenç, S. Dikmese, and N. Rupasinghe, "Efficient noise variance estimation under pilot contamination for massive MIMO systems," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 4, pp. 2982–2996, Apr. 2018.

[10] M. L. Ammari, P. Fortier, and M. El Khaled, "Channel and noise covariance matrix estimation for MIMO systems with optimal training design," in *Proc. IEEE Vehicular Technology Conference (VTC-Fall)*, 2014, pp. 1–5.

[11] T. Schubert, J. Korte, J. M. Brockmann, and W.-D. Schuh, "A generic approach to covariance function estimation using arma-models," *Mathematics*, vol. 8, no. 4, p. 591, Apr 2020. [Online]. Available: http://dx.doi.org/10.3390/math8040591

[12] Y. Liu, T. F. Wong, and W. W. Hager, "Training signal design for estimation of correlated MIMO channels with colored interference," *IEEE Transactions on Signal Processing*, vol. 55, no. 4, pp. 1486–1497, Apr. 2007.

[13] M. Biguesh, S. Gazor, and M. H. Shariat, "Optimal training sequence for MIMO wireless systems in colored environments," *IEEE Transactions on Signal Processing*, vol. 57, no. 8, pp. 3144–3153, Aug. 2009.

[14] G. Soatti, A. Murtada, M. Nicoli, J. Gambini, and U. Spagnolini, "Low-rank channel and interference estimation in mm-Wave massive antenna arrays," in *Proc. European Signal Processing Conference (EUSIPCO)*, 2018, pp. 922–926.

[15] F. A. Dietrich, T. Ivanov, and W. Utschick, "Estimation of channel and noise correlations for MIMO channel estimation," in *Proc. International ITG Workshop on Smart Antennas, Ulm, Germany*, 2006.

[16] K. Hassan, M. Masarra, M. Zwingelstein, and I. Dayoub, "Channel estimation techniques for millimeter-wave communication systems: Achievements and challenges," *IEEE Open Journal of the Communications Society*, vol. 1, pp. 1336–1363, 2020.

[17] A. V. Delgado, M. Sánchez-Fernández, L. Venturino, and A. Tulino, "Super-resolution in automotive pulse radars," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 4, pp. 913–926, Apr. 2021.

[18] H. Krim and J. Cozzens, "A data-based enumeration technique for fully correlated signals," *IEEE Transactions on Signal Processing*, vol. 42, no. 7, pp. 1662–1668, Jul. 1994.

[19] A. Di, "Multiple source location–a matrix decomposition approach," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 5, pp. 1086–1091, May 1985.

[20] T. Bai, V. Desai, and R. W. Heath, "Millimeter wave cellular channel models for system evaluation," in *Proc. International Conference on Computing, Networking and Communications (ICNC)*, 2014, pp. 178–182.

[21] N. Benvenuto, G. Cherubini, and S. Tomasin, *Algorithms for communications systems and their applications*. John Wiley & Sons, 2021.

[22] M. Sánchez-Fernández, V. Jamali, J. Llorca, and A. M. Tulino, "Gridless multidimensional angle-of-arrival estimation for arbitrary 3D antenna arrays," *IEEE Transactions on Wireless Communications*, vol. 20, no. 7, pp. 4748–4764, Jul. 2021.

[23] K. Krishnan and T. Terlaky, *Interior Point and Semidefinite Approaches in Combinatorial Optimization*. Boston, MA: Springer US, 2005, pp. 101–157.

[24] W. Tang, Y. Yang, L. Zeng, and Y. Zhan, "Optimizing MSE for clustering with balanced size constraints," *Symmetry*, vol. 11, no. 3, 2019. [Online]. Available: https://www.mdpi.com/2073-8994/11/3/338

[25] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, Mar. 1986.

[26] F. Yan, M. Jin, and X. Qiao, "Low-complexity DOA estimation based on compressed MUSIC and its performance analysis," *IEEE Transactions on Signal Processing*, vol. 61, no. 8, pp. 1915–1930, 2013.

[27] R. Joyce, D. Morris, S. Brown, D. Vyas, and L. Zhang, "Higher order horizontal sectorization gains for 6, 9, 12 and 15 sectored cell sites in a 3GPP/HSPA+ network," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 5, pp. 3440–3449, 2016.

APPENDIX

Let x and y be zero-mean jointly circularly symmetric complex Gaussian variables, with powers σ_x^2 and σ_y^2, cross-correlation $\mathbb{E}[xy^*] = \xi\sigma_x\sigma_y$, and $\xi = \xi_R + j\xi_I$. We now aim at computing the variance of the random variable xy^*.

Variable y can be written as a function of x as

$$y = \xi^* \frac{\sigma_y}{\sigma_x} x + \sqrt{\sigma_y^2 - |\xi\sigma_y|^2}\, w, \tag{56}$$

with w zero-mean unitary-variance circularly symmetric complex Gaussian variable, independent of x.

Then the variance of the random variable xy^* is

$$\begin{aligned}\mathbb{E}[|xy^* - \xi\sigma_x\sigma_y|^2] = \mathbb{E}[|xy^*|^2) + |\xi\sigma_x\sigma_y|^2 + \\ - 2\mathbb{E}(\Re\{xy^*\xi^*\sigma_x\sigma_y\}].\end{aligned} \tag{57}$$

Since we have

$$\begin{aligned}\mathbb{E}[|xy^*|^2] &= \mathbb{E}[|x|^4]|\xi|^2 \frac{\sigma_y^2}{\sigma_x^2} + (\sigma_y^2 - |\xi\sigma_y|^2)\sigma_x^2 \\ &= 2\sigma_x^4|\xi|^2 \frac{\sigma_y^2}{\sigma_x^2} + (\sigma_y^2 - |\xi\sigma_y|^2)\sigma_x^2,\end{aligned} \tag{58}$$

$$\begin{aligned}&\mathbb{E}\left[\Re\{xy^*\xi^*\sigma_x\sigma_y\}\right] = \sigma_x\sigma_y[\xi_R\Re\{xy^*\} - \xi_I\Im\{xy^*\}] = \\ &\sigma_x^2\sigma_y^2|\xi|^2,\end{aligned} \tag{59}$$

we obtain

$$\begin{aligned}&\mathbb{E}[|xy^* - \xi\sigma_x\sigma_y|^2] = 2|\xi|^2\sigma_x^2\sigma_y^2 \\ &\quad + \sigma_x^2\sigma_y^2(1 - |\xi|^2) - 2\sigma_x^2\sigma_y^2|\xi|^2 + |\xi|^2\sigma_x^2\sigma_y^2 = \sigma_x^2\sigma_y^2.\end{aligned} \tag{60}$$

Stefano Tomasin received the Ph.D. degree from the University of Padova, Italy, in 2003. During his studies he did internships with IBM Research (Switzerland) and Philips Research (Netherlands). He joined the University of Padova where he has been Assistant Professor (2005-2015), Associate Professor (2016-2022), and Full Professor (since 2022). He was visiting faculty at Qualcomm, San Diego (CA) in 2004, the Polytechnic University in Brooklyn (NY) in 2007 and the Mathematical and Algorithmic Sciences Laboratory of Huawei in Paris (France) in 2015. His current research interests include physical layer security, security of global navigation satellite systems, signal processing for wireless communications, synchronization, and scheduling of communication resources. He is a senior member of IEEE since 2011 (member since 1999) and a member of EURASIP since 2011. He is or has been an Editor of the IEEE Transactions on Vehicular Technologies (2011-2016), of the IEEE Transactions on Signal Processing (2017-2020), of the EURASIP Journal of Wireless Communications and Networking (since 2011) and of the IEEE Transactions on Information Forensics and Security (since 2020). He serves also as a Deputy Editor-in-Chief of the IEEE Transactions on Information Forensics and Security since January 2023.

Raphael Hasler received the B.Sc. degree in telecommunication engineering from the Polytechnic University of Milan, Italy, in 2018. In 2022 he received an M.Sc. degree in ICT for Internet and multimedia from the University of Padova, Italy. He graduated with a thesis entitled "Statistics estimation of interference in mmWave channels". In 2022 he became the winner of a research grant, at the department of information engineering (DEI) of the University of Padova. The research was focused on the estimation of the interference correlation matrix.

Antonia M. Tulino received the Ph.D. degree in electrical engineering from Seconda Universita' degli Studi di Napoli, Italy, in 1999. She held research positions with the Center for Wireless Communications in Finland, Princeton University in USA and also with the Universita' degli Studi del Sannio in Italy.. From 2002 to 2016, she has been an Associate Professor with the Universita' degli Studi di Napoli Federico II, where she has been a Full Professor since 2017. Since 2002, she has been collaborating with Nokia Bell Labs. Starting from October 2019, she is, also, Research Professor at the at Dep. of Electrical and Computer Engineering NYU Tandon School of Engineering, NY, 11201, USA. In September 2020 she is appointed as teaching director of the 5G Academy jointly organized by the Università degli Studi di Napoli, Federico II in collaboration with Capgemini, Nokia and TIM. From 2011 to 2013, she has been a member of the Editorial Board of the IEEE Transactions on Information Theory and in 2013, she was elevated to IEEE Fellow. She was selected by the National Academy of Engineering for the Frontiers of Engineering program in 2013. From 2019, she is the chair of the Information Theory society Fellows Committee. She has received several paper awards and among the others the 2009 Stephen O. Rice Prize in the Field of Communications Theory for the best paper published in the IEEE Transactions on Communications in 2008. She was a recipient of the UC3M-Santander Chair of Excellence from 2018 to 2019.

Matilde Sánchez-Fernández received the M.Sc. degree in telecommunications engineering and the Ph.D. degree from Universidad Politécnica de Madrid, Madrid, Spain, in 1996 and 2001. In 2000, she joined Universidad Carlos III de Madrid, Madrid, where she has been an Associate Professor since 2009 teaching several undergraduate and graduate courses (M.Sc. and Ph.D.) related to communication theory and digital communications. Previously, she was a Telecommunication Engineer with Telefónica. She performed several research stays at the Information and Telecommunication Technology Center, The University of Kansas, Lawrence (1998), Bell Laboratories, Crawford Hill, NJ (2003-2006, 2015), Centre Tecnologic de Telecomunicacions de Catalunya, Barcelona, Spain (2007), and Princeton University, Princeton, NJ (2011). Her current research interests are signal processing for wireless communications, multiple-input multiple-output techniques, channel modeling in wireless communications and game theory and machine learning techniques applied to communications. In these fields, she has (co)authored more than 50 contributions to international journals and conferences.