Corvaja, R. & Armada, A. G. (2020). Analysis of SVD-Based Hybrid Schemes for Massive MIMO With Phase Noise and Imperfect Channel Estimation. *IEEE Transactions on Vehicular Technology*, 69(7), pp. 7325–7338.

DOI: 10.1109/tvt.2020.2990351

Analysis of SVD-Based Hybrid Schemes for Massive MIMO with Phase Noise and Imperfect Channel Estimation

Roberto Corvaja, *Senior Member, IEEE*, Ana García Armada, *Senior Member, IEEE*

***Abstract*—In hybrid analog-digital schemes, proposed to reduce the number of RF chains especially at millimeter waves, the precoding at the transmitter and the combining at the receiver are split into digital and analog parts. We analyze the sensitivity of hybrid schemes to phase noise and channel estimation errors and we compare them to a full-digital approach. The scheme adopted for the analog part employs fixed phase shifters, then the digital part is optimized by a singular-value decomposition. We derive analytical expressions for the interference and the SNR degradation arising from the imperfect decomposition due to phase noise and the channel estimation error, for typical millimter-wave massive MIMO channels. In particular we show that when the channel estimation is made in the beam-space, this hybrid scheme is more robust to the phase noise and to the channel estimation errors than a full-digital approach.**

***Index Terms*—Massive MIMO, hybrid precoding and combining, millimeter-wave MIMO, phase noise.**

I. Introduction

The use of millimeter waves (mmWave) for the deployment of massive multiple input – multiple output (MIMO) systems is pushed forward by the larger available bandwidth and the reduced size of the devices. However the shift towards mmWave determines a major complexity and a serious increase of the power consumption, especially of devices such as analog-to-digital (ADC) and digital-to-analog (DAC) converters. In particular all-digital architectures become very difficult due to the extremely high sampling rate required, so that hybrid analog-digital solutions become a need and are widely studied [1]–[3]. Massive MIMO systems at mmWave may be deployed at the backhaul/fronthaul and at the radio access, for different types of networks, from HetNet [1] to wearable devices [4], mainly at the base station, but also at the user equipment if the reduction of the antenna size, facilitated by the migration to mmW, is enough. In particular one of the first applications of massive MIMO at mmWave is for Fixed Wireless Acess [5]. Many previous studies consider the environment with single-antenna receivers [6], [7], which corresponds to a massive antenna deployment only at the BS. However, here we consider a more general scenario, where a large number of antennas may be possible at both the transmitter and receiver. At mmWave the frequency stability of oscillators becomes a challenging issue and the effect of phase noise (PN) must be considered. Here we study the design of hybrid analog-digital precoding and combining at transmitter and receiver, and we analyze the degradation introduced by the PN of the RF chains and by the channel estimation error. The hybrid combination of the analog and digital precoding and combining causes a reduction of the achievable rate compared to the full-digital approach, unless complex joint optimization procedures for the analog and digital part are adopted [3], [8], [9]. Here a simplified approach is explored, by fixing the analog part and optimizing the digital part, according to a singular-value decomposition (SVD) approach. Although in principle these schemes are sub-optimal and can achieve a lower rate than a full-digital approach, it is shown that in the presence of PN or channel estimation errors, they can achieve rates close to a full-digital scheme. In particular we consider a fixed analog matrix based on the Discrete Fourier Transform (DFT), for which we derive a bound, based on the singular values of the reduced-dimensionality equivalent channel seen at the RF chains. Note that the choice of an analog matrix based on the DFT is similar to solutions adopted in codebook beamforming [10], [11] or in lens-based systems [12], where the selection of the spatial frequencies is typically done by switches. The performance metric analyzed here is the achievable rate, which is strictly related to the signal to interference plus noise ratio (SINR) per spatial data stream. [13], [14] address the characterization of the SINR in a multi-user scenario, where different links determine the large scale SINR at the receiver input. Here, we consider the SINR which includes the precoding and combining, in a point-to-point link, with SVD-based precoding and combining [3] for the massive MIMO scenario at mmWave, characterized by its specific sparse channel. We derive the statistics of the channel singular values, and we characterize the SINR, in terms of probability density function, providing an analytical expression of the SINR mean and variance. Moreover, we show that the probability density function (PDF) of the SINR, in a log-scale, fits perfectly a Gaussian distribution. The results can be applied to both full-digital and hybrid analog-digital schemes. In fact, in the latter, with the SVD-based scheme, the optimization of the analog matrices corresponds to the selection of the largest singular values for the transmission spatial modes associated to the streams.

We extend the analysis of the SINR to the presence of PN, with a degradation due to the fact that the actual channel to which precoding and combining are applied differs from the channel used in the SVD. A similar degradation arises

This work has been supported by the Spanish National Project TERESA-ADA (TEC2017-90093-C3-2-R) through MINECO/AEI/FEDER, UE.

Roberto Corvaja is with the Department of Information Engineering, University of Padova, Padova, Italy. (e-mail: corvaja@dei.unipd.it).

Ana García Armada is with the Department of Signal Theory and Communications, Universidad Carlos III de Madrid, 28911 Leganés, Spain. (e-mail: agarcia@tsc.uc3m.es).

by considering the effect of the channel estimation error. The effects of an imperfect channel estimation (and PN) can be considered as the result of *channel aging* [15]–[17], i.e. the evolution of both the channel and the PN between the instant of estimation and their actual effect on the detection process. With respect to previous works, here we consider spatial multiplexing of several streams for a single transmitter to a single receiver both with multiple antennas, rather than the downlink to multiple single-antenna users. We do not address specifically the channel aging, since both the channel estimation error and PN derive not only from the channel aging, but also from the accuracy and the specific algorithms for channel estimation and phase synchronization. Hence we consider a general approach taking as parameters the PN variance and the channel estimation MSE. Our objective is to provide analytical expressions for the degradation due to PN and channel estimation errors, comparing the sensitivity of full-digital and hybrid schemes. In detail the main contributions of this work are the following:

- The statistics of the SINR and the analytical expressions of its mean and variance are derived from the channel representation in the beam-space, by the joint PDF of the largest singular values of the channel matrix. Together with the Gaussian statistical model of the SINR, this allows the theoretical evaluation of the achievable rate.
- In the presence of PN, we derive the analytical expression of the interference arising from the non-perfect SVD. From this result one can evaluate the SINR degradation due to PN.
- Also in the presence of channel estimation errors, the interference term and the SINR degradation is analytically derived. Two estimation models are evaluated: the direct estimation of the channel matrix and the estimation of the beam-space parameters.

II. System Model

We consider single-carrier spatial multiplexing of N_s parallel data streams from a transmitter equipped with N_t antennas to a receiver with N_r antennas. The data streams are pre-coded by a baseband $L_t \times N_s$ matrix $\boldsymbol{F}_{BB}$ into L_t RF chains and are DAC converted. Each RF chain at the transmitter comprises a DAC, a mixer to RF where PN is introduced, and a filter/amplifier. Then the signal is precoded by the analog $N_t \times L_t$ matrix $\boldsymbol{F}_{RF}$ and sent by N_t transmit antennas. The hybrid analog-digital and the full-digital schemes are shown in Fig. 1. We consider

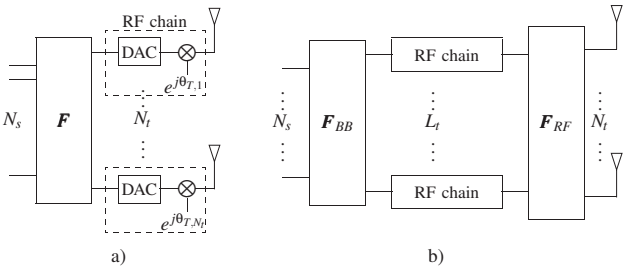


Fig. 1. Precoding schemes: a) full-digital and b) hybrid digital-analog.

a narrow-band flat fading channel represented by a $N_r \times N_t$ matrix $\boldsymbol{H}$. In mmWave, the delay spread is small [18], due to the effect of high directional beamforming with a large number of antennas, which makes the channel become almost flat. The flat channel is a common simplification that allows to do some analysis that would be otherwise very difficult and is adopted in several other studies on the evaluation of the SINR (see for example [19]). Some considerations on the effect of the delay spread of the channel can be found for example in [20].

The received signal by N_r antennas is combined by the $L_r \times N_r$ analog matrix $\boldsymbol{W}_{RF}^H$ into L_r RF chains. After ADC a baseband combining is performed by the $N_s \times L_r$ matrix $\boldsymbol{W}_{BB}^H$ back into N_s signal streams. In the hybrid analog-digital architecture the number of RF chains is reduced with respect to the number of antennas, $N_s \le L_t \le N_t$ ($N_s \le L_r \le N_r$ at the receiver). In some solutions, proposed to further reduce the complexity, the antennas are grouped into sub-arrays: only some RF chains are connected to each sub-array, instead of connecting all the RF chains to all the antennas [3], [21], [22], [23]. In the absence of PN, the received vector $\boldsymbol{y}$ at the detection point, is given by

$$\boldsymbol{y} = \boldsymbol{W}^H \boldsymbol{H} \boldsymbol{F} \boldsymbol{s} + \boldsymbol{W}^H \boldsymbol{n} \tag{1}$$

where $\boldsymbol{s}$ is the vector of zero-mean independent transmitted symbols and the vector $\boldsymbol{n}$ denotes the AWGN contribution. We define as $\boldsymbol{H}_{bb}$ the effective $N_s \times N_s$ channel between the input and output streams. In the absence of PN it is

$$\boldsymbol{H}_{bb} = \boldsymbol{W}^H \boldsymbol{H} \boldsymbol{F} \tag{2}$$

The precoding matrix is $\boldsymbol{F} = \boldsymbol{F}_{RF} \boldsymbol{F}_{BB}$. At the receiver side, similar considerations can be made for the combining matrix $\boldsymbol{W} = \boldsymbol{W}_{RF} \boldsymbol{W}_{BB}$.

A. mmWave Channel Model

In the mmWave frequency range, the channel shows high directivity, with a sparse matrix of coefficients between the transmit and receive antennas. This fact leads to a representation of the channel in the beam-space [9], [24] with a limited number of scatterers. According to this model, a flat fading channel is described by N_p paths or scatterers associated to their corresponding transmit and receive angles. A pictorial representation of the mmWave channel model is given in Fig. 2, for $N_p = 3$ scatterers. The channel matrix in the beam-

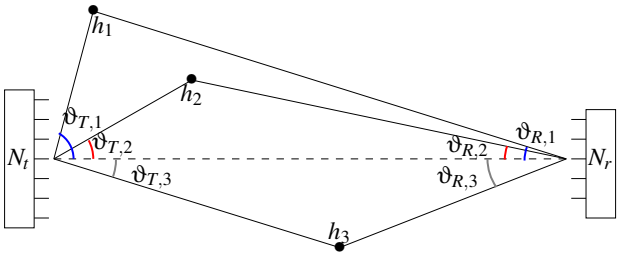


Fig. 2. Millimeter-wave channel: example with $N_p = 3$ scatterers.

space is expressed as

$$\boldsymbol{H} = \sum_{p=1}^{N_p} \alpha_p \boldsymbol{a}_R(\vartheta_{R,p}) \boldsymbol{a}_T^H(\vartheta_{T,p}) \ , \tag{3}$$

where the vectors $\boldsymbol{a}_R(\vartheta_{R,p})$ and $\boldsymbol{a}_T(\vartheta_{T,p})$ denote the array response to the angles $\vartheta_{R,p}$ and $\vartheta_{T,p}$ at the transmitter and receiver, respectively. Thus, the channel is characterized by N_p triplets $(\vartheta_{T,p}, \vartheta_{R,p}, \alpha_p)$ of transmit and receiver angles and the corresponding complex gain. The response vector of a uniform linear array (ULA) with spacing d, at the wavelength λ, is

$$\boldsymbol{a}_T(\vartheta) = \left[1, e^{j2\pi\phi}, e^{j4\pi\phi} \dots e^{j2\pi\phi(N_t-1)}\right]^T \tag{4}$$

with $\phi = \frac{d}{\lambda}\sin\vartheta$. The steering vectors corresponding to two different angles $\vartheta_{T,1}$ and $\vartheta_{T,2}$ are not perfectly orthogonal. The inner product between the vectors associated to $\vartheta_{T,1}$ and $\vartheta_{T,2}$ is given by

$$\boldsymbol{a}_T^H(\vartheta_1)\,\boldsymbol{a}_T(\vartheta_2) = \frac{e^{j2\pi\frac{\lambda}{d}\left(\sin\vartheta_{T,2}-\sin\vartheta_{T,1}\right)N_t} - 1}{e^{j2\pi\frac{\lambda}{d}\left(\sin\vartheta_{T,2}-\sin\vartheta_{T,1}\right)} - 1} \tag{5}$$

As a function of the separation angle $\vartheta_{T,1} - \vartheta_{T,2}$, the inner product (5) is sinc-shaped and clearly decreases as the angular separation increases. Moreover, increasing the value of N_t, the inner product diminishes with N_t and the main lobe of the radiation pattern gets narrower. Therefore, in massive MIMO, characterized by a very large number of antennas, the inner product among steering vectors associated to different angles becomes almost zero (almost orthogonal steering vectors), as long as the two angles are different.

The channel coefficients α_p are modeled by independent Gaussian random variables with zero mean and variance normalized to $1/N_p$ [24], [25]. For the angular distribution of the scatterers, the angles in (3) describing the relative position of the scatterers with respect to the transmitter and receiver are considered independent uniform random variables between $-\pi/2$ and $\pi/2$, as in [25]. Other more sophisticated models for mmWave channels [24], [26] consider the scatterers grouped in clusters with independent fading among different clusters and a spread of the angles within each cluster. In the following we compare the results also with channel matrices obtained by a well known software channel simulator [27] based on the model of [26].

1) Spatially white Rayleigh channel: For comparison purposes we present also the *spatially white Rayleigh channel*. In this case the channel matrix elements are independent and identically distributed (iid) Gaussian random variables, with zero-mean and variance normalized to $1/N_p$.

B. Phase noise

The phase noise is modeled by two independent contributions at the transmitter and receiver. In hybrid architectures, the effect is accounted for by a multiplicative term in the RF chains, more precisely, by the multiplication of a diagonal matrix, with entries corresponding to the phase shifts of PN. At the transmitter we have the matrix $\boldsymbol{P}_T$, acting between $\boldsymbol{F}_{BB}$ and $\boldsymbol{F}_{RF}$

$$\boldsymbol{P}_T = \mathrm{diag}\left[e^{j\theta_{T,1}}, \dots, e^{j\theta_{T,L_t}}\right], \tag{6}$$

where $\theta_{T,m}$ are zero-mean Gaussian random variables. Assuming the classic Wiener PN model with independent increments, the PN increment has variance $\sigma^2_{\theta_T}$. Another contribution is introduced at reception at the downconversion, with corresponding diagonal matrix $\boldsymbol{P}_R$ of size $L_r \times L_r$ and variance of the increments $\sigma^2_{\theta_R}$. The PN variance is related to the 3-dB oscillator bandwidth B_θ and to the time elapsed from a perfect phase synchronization T_d (for example to account for the channel aging) by $\sigma^2_{\theta_R} = 4\pi B_\theta T_d$.

Several combinations of PN could be considered depending on the number of independent oscillators used to feed the RF chains. In massive MIMO, even with a reduced number of RF chains, it is almost impossible to drive all the RF chains with a single local oscillator, due to losses in the distribution circuits.

Several considerations on the number of oscillators have been presented for different antenna configurations and up- or down-link [21], [28], [29]. The paper [28] investigates the high-SNR capacity of single-input multiple-output (SIMO) and multiple-output single-input (MISO) phase-noise channels whether the RF circuitries connected to each antenna are driven by separate (independent) local oscillators (SLO) or by a common local oscillator (CLO). For the SIMO case, the SLO configuration provides a diversity gain, which for the case of Wiener phase noise can be of at least $0.5\ln(N_r)$, where N_r is the number of receive antennas. For the MISO, the CLO configuration can obtain a coherent-combining gain through maximum ratio transmission (a.k.a. conjugate beamforming), while this gain is unattainable with the SLO configuration. The conclusion of [28] is that SLO is better for SIMO and CLO is better for MISO. In our case we have MIMO, so both effects are compensating one another and, as it is found also in other papers [29], [30], we end up with the conclusion that CLO is better due to the coherent combining.

In [29] a large-scale analysis is performed with the number of antennas and users tending to infinity and considerations are made on the number of oscillators, with the conclusion that a single oscillator at the base station (BS) achieves the highest signal-to-interference-plus-noise ratio (SINR). They observe that as the number of oscillators increases, the SINR of the ZF precoder degrades as the interference power increases, and the desired signal power decreases. Hence, here we consider the limiting case where each RF chain is driven by an independent oscillator.

III. Hybrid precoding and combining

For the precoding and combining, together with the full-digital scheme, we use a hybrid design where the analog stage is implemented by fixed phase-shifters. Then the digital stage is obtained by a SVD.

A. Full-digital approach

In the full-digital case, in order to maximize the achievable rate [3], the matrices $\boldsymbol{F}$ and $\boldsymbol{W}$ are derived from the SVD of the full $N_r \times N_t$ channel matrix, namely $\boldsymbol{H} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{V}^H$, with square matrices $\boldsymbol{U}$ and $\boldsymbol{V}$, of size $N_r \times N_r$ and $N_t \times N_t$, respectively. The matrix $\boldsymbol{\Lambda}$ has diagonal entries λ_i in position (i,i) and λ_i are ordered in decreasing order $\lambda_1 \geq \lambda_2 \geq \lambda_{\min\{N_r,N_t\}}$.

Then the matrices $\boldsymbol{F}$ and $\boldsymbol{W}$ are the first N_s columns of $\boldsymbol{U}$ and $\boldsymbol{V}$, $\boldsymbol{F} = \boldsymbol{V}_{N_t \times N_s}$, $\boldsymbol{W} = \boldsymbol{U}_{N_r \times N_s}$.

B. Hybrid solution

1) Analog precoding and combiner: For the analog stage, we address the case in which the matrices $\boldsymbol{F}_{RF}$ and $\boldsymbol{W}_{RF}$ are made of phase shifts. They can be designed in several ways [3] and can affect differently the singular values of the equivalent $L_t \times L_r$ channel, seen at the RF chains stage.

To reduce the system design complexity, we consider fixed analog matrices. In particular, in [21] we considered beamforming with equally spaced angles, random angles between $-\pi/2$ and $\pi/2$, and Discrete Fourier Transform (DFT) matrix. We showed that DFT analog matrices exhibit a better robustness to PN [21]. Moreover, the DFT matrix has been proposed in hybrid schemes with codebook beamforming [10], [11] and it is also the result of using a lens for the analog stage [31]. The use of a lens is interesting since it achieves near-optimal performance with low hardware cost and high power efficiency and its use has been proposed also for channel estimation [12]. Note that both give a projection of the channel matrix into the domain of spatial frequencies. A lens-based system often uses switches where after a selection process some of the lens outputs are sent to the RF chains. The scheme that we analyze here corresponds to a fixed reduced-size DFT analog matrix followed by a SVD digital part, so that all the antennas outputs are combined and the selection is left to the SVD. As shown in the following results, this sub-optimal choice has a degradation in terms of rate with respect to a full-digital scheme that becomes noticeable only if the number of antennas is much greater than the number of RF chains.

Then, the analog matrix $\boldsymbol{F}_{RF}$ built on the DFT matrix, has its j-th column equal to the column jN_t/L_t of the DFT matrix of size $N_t \times N_t$, for $j = 0, \ldots, L_t - 1$. The analog combiner $\boldsymbol{W}_{RF}$ is obtained in the same way, by the substitution of the corresponding dimensions L_r and N_r.

2) Digital precoding and combiner: Once the analog matrices are set, the equivalent $L_r \times L_t$ channel $\boldsymbol{H}_{eq}$ is obtained by the cascade of the RF matrices and the actual channel,

$$\boldsymbol{H}_{eq} = \boldsymbol{W}_{RF}^H \boldsymbol{H} \boldsymbol{F}_{RF} \ . \tag{7}$$

Again, in order to maximize the achievable rate, the digital part should be based on the SVD of $\boldsymbol{H}_{eq}$, that is, $\boldsymbol{H}_{eq} = \boldsymbol{U}_{eq} \boldsymbol{\Lambda}_{eq} \boldsymbol{V}_{eq}^H$, where we assume that the elements of $\boldsymbol{\Lambda}_{eq} = \mathrm{diag}(\mu_i)$ are ordered in decreasing order. Again, the matrices $\boldsymbol{F}_{BB}$ and $\boldsymbol{W}_{BB}$ are obtained by the first N_s columns of $\boldsymbol{U}_{eq}$ and $\boldsymbol{V}_{eq}$ corresponding to the largest singular values.

$$\boldsymbol{F}_{BB} = \boldsymbol{V}_{eq, L_t \times N_s} \qquad \boldsymbol{W}_{BB} = \boldsymbol{U}_{eq, L_r \times N_s} \ . \tag{8}$$

We remind that the full-digital case can be seen as a particular hybrid scheme where the number of RF chains is equal to the number of antennas and the analog matrices $\boldsymbol{W}_{RF}$, $\boldsymbol{F}_{RF}$ are identities. Then $\boldsymbol{H}_{eq} = \boldsymbol{H}$, which means that the SVD is performed directly on the channel $\boldsymbol{H}$.

IV. Performance indicators

In the following sections a system performance measure is the mean achievable sum-rate on the N_s streams, namely $R = \sum_{i=1}^{N_s} E[R_i]$. Considering Gaussian signalling and equal power allocation for the data streams, without channel estimation errors and without PN, the rate on the i-th stream is

$$R_i = \log_2(1 + \rho \Lambda_i) \ , \tag{9}$$

where ρ is the signal-to-noise ratio (SNR), $\rho = \frac{E[|\boldsymbol{s}|^2]}{E[|\boldsymbol{n}|^2]}$ and Λ_i is related to the i-th largest singular value of the matrix $\boldsymbol{H}_{eq}$, seen by the RF chains. In particular: in the full-digital scheme $\boldsymbol{H}_{eq} = \boldsymbol{H}$ and $\Lambda_i = |\lambda_i|^2$, where λ_i are the singular values of $\boldsymbol{H}$ in decreasing order. In the hybrid scheme $\boldsymbol{H}_{eq}$ is the $L_r \times L_t$ matrix of (7), $\Lambda_i = |\mu_i|^2$, with μ_i the singular values of $\boldsymbol{H}_{eq}$, again in decreasing order. In the absence of PN and channel estimation errors, we define

$$\mathrm{SINR}_i = \rho \Lambda_i \ , \tag{10}$$

which we denote as signal-to-interference-plus-noise ratio (SINR) although in the absence of PN and estimation errors the contribution of interference in canceled by the SVD precoding and combining. Note that SINR_i is a random variable and its value in decibels $\mathrm{SINR}_{i,dB} = 10 \log_{10} \mathrm{SINR}_i$ has a relevance, since the sum-rate over the N_s streams is related to the average SINR in dB over the streams. In the following we derive the mean and variance for $\mathrm{SINR}_{i,dB}$ and we show that the average

$$\mathrm{SINR}_{dB} = \frac{1}{N_s} \sum_{i=1}^{N_s} (\mathrm{SINR}_i)_{dB} \ , \tag{11}$$

has Gaussian statistics. Note that in the limit case of a single scatterer $N_p = 1$ and a single stream $N_s = 1$, the statistical description of (11) is given by a log-Rayleigh random variable [32], considered in the following for comparison.

On the other hand, with PN or channel estimation errors, the imperfect diagonalization gives rise to interference among the streams, with a different SINR with respect to (10). Then we employ the very commonly used approximations [33], [34] for the rate and equivalent SNR, that are very tight for massive MIMO. The mean rate R_i of stream i is given by

$$\overline{R_i} = E[R_i] \approx \log_2\left(1 + \overline{\mathrm{SINR}_i}\right) \ , \tag{12}$$

where the SINR on the i-th stream is

$$\overline{\mathrm{SINR}_i} \approx \frac{E\left[|\boldsymbol{H}_{bb}(i,i)|^2\right]}{\frac{1}{\rho} + \sum_{j \neq i}^{N_s} E\left[|\boldsymbol{H}_{bb}(i,j)|^2\right]} \ . \tag{13}$$

where $\boldsymbol{H}_{bb}$ is defined in (2).

V. Channel singular values characterization

In order to obtain the joint statistics of the largest channel singular values, we note that (3) gives a spectral decomposition of the channel matrix $\boldsymbol{H}$, if the steering vectors in (3) are orthogonal to each other. Hence the values Λ_i are equal to the squared channel coefficients, $\Lambda_i = |\alpha_i|^2$, considered in decreasing order. This assumption is realistic in massive MIMO, where the inner product between the steering vectors (5) becomes negligible. Moreover, for mmWave channels, the number of scatterers N_p is low [26], [27]. Therefore, it is negligible the probability that the angles of different scatterers are close enough to determine a sensible overlap. In conclusion, the beam-space representation (3) for massive

MIMO mmWave channels provides a very good approximation of an SVD. Thus, we can approximate the N_s largest values Λ_i by the N_s largest channel squared fading coefficients $|\alpha_i|^2$, which are exponential random variables (chi-squared with 2 degrees of freedom). Using the approach of [35], where it is applied to a selection combining, or with the derivation of Appendix A, the average value is

$$E[\Lambda_i] = \frac{(N_p-1)!}{(i-1)!(N_p-i)!}\sum_{k=0}^{N_p-i}\binom{N_p-i}{k}\frac{(-1)^k}{(k+i)^2} \tag{14}$$

For the case of the SINR, following a logarithmic transformation, it is detailed in Appendix A, leading to the results.

Result 1 The first moments of the random variables SINR_i of (11) on each stream are obtained in closed-form by expressing $(\text{SINR}_i)_{dB} = 10\log_{10}\rho + \frac{10}{\ln(10)}y_i$, where the moments of y_i are

$$m_{y_i} = E[y_i] = \frac{N_p!}{(i-1)!\,(N_p-i)!}\sum_{n=0}^{N_p-i}\binom{N_p-i}{n}\frac{(-1)^n}{n+1} \times\left[\ln\frac{1}{\beta(n+1)}-\gamma\right] \tag{15}$$

$$M_{y_i} = E[y_i^2] = \frac{N_p!}{(i-1)!\,(N_p-i)!}\sum_{n=0}^{N_p-i}\binom{N_p-i}{n}\frac{(-1)^n}{i+n+1} \times\left[\frac{\pi^2}{6}+\left(\ln\frac{1}{\beta(n+1)}-\gamma\right)^2\right]. \tag{16}$$

where γ is the Euler-Mascheroni constant.

Result 2 The joint PDF of the singular values on the N_s streams, and consequently of the SINR_i, is given by

$$f_{\boldsymbol{\Lambda}}(a_1,\ldots,a_{N_s}) = \frac{N_p!}{(N_p-N_s)!}\left(1-e^{-a_{N_s}}\right)^{N_p-N_s}\prod_{m=1}^{N_s}e^{-a_m}. \tag{17}$$

for $a_1 \geq a_2 \geq \ldots \geq a_{N_s}$.

From (15)–(16), one can obtain the mean and standard deviation of SINR_{dB} (11) assuming that SINR_i are independent, or within the limits of the numerical complexity one can obtain them form the joint statistics of (17). The results are presented in Section VIII and compared with simulations.

1) Spatially white Rayleigh channel: In the white Rayleigh channel (Section II-A1), considered for comparison, the singular values are the eigenvalues of the Wishart matrix $\boldsymbol{H}\boldsymbol{H}^H$. Their joint PDF has been derived in [36] for the central case and generalized to non-central and correlated cases in [37].

A. Singular values with hybrid analog-digital beamforming

The previous analysis on the equivalence between the channel singular values and the coefficients of the model (3) has been derived for the full-digital case. However, the results apply also to more sophisticated hybrid analog-digital schemes (e.g. [38]) in which the optimization of the analog stage maintains the largest channel singular values of the channel in the equivalent matrix seen by the RF chains.

In the case of using fixed matrices in the analog stage, as considered in this work, the operation performed by these matrices projects the channel on a subspace of size $L_r \times L_t$.

Due to the structure of the matrices $\boldsymbol{W}_{RF}$ and $\boldsymbol{F}_{RF}$ with orthogonal columns obtained from the DFT matrix the operation (7) corresponds to a projection on an *orthogonal quotient* of the channel matrix $\boldsymbol{H}$ for which a generalization of the Cauchy interlacing theorem (see for example Corollary 24 of [39]) provides a bound for the singular values μ_i of $\boldsymbol{H}_{eq}$ with reference to the singular values of $\boldsymbol{H}$. Then, the values μ_i, in decreasing order $\mu_1 \geq \mu_2 \geq \ldots$, are related to the singular values λ_j of $\boldsymbol{H}$ by the following upper and lower bounds

$$\begin{cases}\mu_i \leq \lambda_i & i=1,\ldots,r\\ \mu_{\ell+1-i} \geq \lambda_{r+1-i} & i=1,\ldots,\ell\end{cases}. \tag{18}$$

where $r = \text{rank}(\boldsymbol{H})$ and $\ell = r - L_t - L_r$. Then $r \leq \min\{N_t, N_r\}$ and according to the channel model (3) we have $r = N_p$. According to the order statistic theory [40] described in Appendix A, we can obtain lower and upper bounds for the joint PDF of μ_j. The upper bound coincides with the full-digital case. For the lower bound, we have to consider the ordered (in decreasing order) values up to ℓ of (18). The analysis of Appendix A gives the mean and variance of the SINR. In the most general case one should have $N_p - L_t - L_r \geq N_s$.

Note, however, that in the case of symmetric configurations such as $N_r = N_t$ and $L_r = L_t$, the projection performed at the transmitter is the same as the one at the receiver, in other words $\ell = N_p - L_t$. Also, for many practical implementations, one can have the situation in which $N_r = N_t$ while L_t is a multiple of L_r (or viceversa) then the columns of $\boldsymbol{F}_{RF}$ are a sub-set of columns of $\boldsymbol{W}_{RF}$ and one can consider the projection on $\min\{L_t, L_r\}$.

In most of the cases one should have a rather dense channel matrix with a sufficiently large number of scatterers N_p in order to apply the lower bound (18).

A more useful and general lower bound can be obtained by accounting for the effect of the DFT matrix as an equivalent analog beam-steering on equally spaced angles. With the steering vector response (4) and considering the overall DFT matrix, the array gain of the RF matrix is a collection of L_t (L_r) sinc-shaped gain patterns [41] at equally spaced angles. The reduction of the singular values corresponds to the factors

$$\mu_i \geq F_r F_t \lambda_i \qquad i=1,\ldots,N_s \tag{19}$$

where F_t (F_r) is the attenuation introduced by the RF stage at transmission (reception).

In order to obtain a bound one can approximate the array pattern by a pessimistic, two-level pattern, with unitary attenuation in the main lobes and a constant value between two main lobes, corresponding to the lowest lobe of the sinc-shaped pattern. The smallest lobe between two main lobes for a uniform linear array gives a factor (at the transmitter)

$$F_t \approx \begin{cases}1 & \text{with probability } \frac{L_t}{N_t}\\ F = \frac{4}{(2\frac{N_t}{2L_t}+1)^2\pi^2} & \text{with probability } \left(1-\frac{L_t}{N_t}\right)\end{cases} \tag{20}$$

Then one can apply the theory outlined in Appendix A for the largest N_s values among N_p, where the PDF of each is now the convex combination of $f_{y_i}(a_i)$ and $\frac{1}{F}f_{y_i}\left(\frac{a_i}{F}\right)$ with the probabilities $\frac{L_t}{N_t}$ and $\left(1-\frac{L_t}{N_t}\right)$.

VI. Effect of phase noise

In the presence of PN, the precoding and combiner matrices are derived from the SVD of a channel, but then applied to a different channel, due to the evolution of the PN. In fact, $\boldsymbol{F}$ and $\boldsymbol{W}$ are designed on the basis of the SVD of $\boldsymbol{H}$, but then they are applied to the actual channel corrupted by the PN matrices $\boldsymbol{P}_T$ and $\boldsymbol{P}_R$. The degradation due to PN is measured by the *SNR penalty*, defined as the increase of SNR necessary to reach the same achievable rate as in the absence of PN (at a reference SNR). The overall effective channel, which determines the rate on the N_s streams, is the matrix $\boldsymbol{H}_{bb}$, which, in the presence of PN, is

$$\boldsymbol{H}_{bb} = \left(\boldsymbol{W}_{BB}^H \boldsymbol{P}_R \boldsymbol{W}_{RF}^H\right) \boldsymbol{H} \left(\boldsymbol{F}_{RF} \boldsymbol{P}_T \boldsymbol{F}_{BB}\right) . \tag{21}$$

The effect of the multiplication by the PN matrices in (21) is shown schematically in Fig. 3 for the combiner (receiver side). PN compromises the orthogonality between the rows of $\boldsymbol{W}_{BB}$ and the columns of $\boldsymbol{U}$. The same effect is present at the precoding side, with $\boldsymbol{F}_{BB}$ and $\boldsymbol{V}^H$. In order to analyze

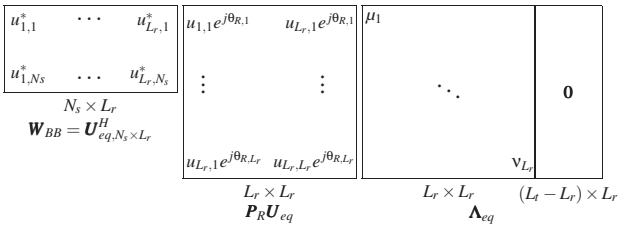


Fig. 3. Effect of PN on the SVD-based combining matrix at the receiver.

the degradation induced by the PN, let us consider a random variable $\Theta(\boldsymbol{a})$ given by a linear combination with coefficients a_n of the PN contributions

$$\Theta(\boldsymbol{a}) = \sum_{n=1}^{N} a_n e^{j\theta_n} , \tag{22}$$

where the iid PN random variables θ_n represent either $\theta_{T,n}$ (transmitter) or $\theta_{R,n}$ (receiver). From $E[e^{j\theta_i}] = e^{-\frac{1}{2}\sigma_\theta^2}$, we have

$$m_y = E[\Theta(\boldsymbol{a})] = e^{-\frac{1}{2}\sigma_\theta^2} \sum_{n=1}^{N} a_n \tag{23}$$

$$M_y = E[\Theta(\boldsymbol{a})\Theta^*(\boldsymbol{a})] = \sum_{n=1}^{N} |a_n|^2 + e^{-\sigma_\theta^2} \sum_{n=1}^{N} \sum_{m \neq n}^{N} a_n a_m^* . \tag{24}$$

We develop the derivation for the hybrid scheme, noting that for the full-digital case it is sufficient to substitute L_t and L_r with N_t, N_r and $\boldsymbol{U}_{eq}$, $\boldsymbol{V}_{eq}$ with $\boldsymbol{U}$, $\boldsymbol{V}$. The multiplication by the SVD-based matrices gives, at the transmitter, the $L_t \times N_s$ matrix $\boldsymbol{\Psi}_T = \boldsymbol{V}_{eq}^H \boldsymbol{P}_T \boldsymbol{F}_{BB}$ and at the receiver the $N_s \times L_r$ matrix $\boldsymbol{\Psi}_R = \boldsymbol{W}_{BB}^H \boldsymbol{P}_R \boldsymbol{U}_{eq}$. As in Fig. 3, the elements of $\boldsymbol{\Psi}_R$ are

$$\boldsymbol{\Psi}_{R,i,i} = \sum_{n=1}^{L_r} e^{j\theta_{R,n}} \left|\boldsymbol{U}_{eq,n,i}\right|^2 \quad \boldsymbol{\Psi}_{R,i,j} = \sum_{n=1}^{L_r} e^{j\theta_{R,n}} \boldsymbol{U}_{eq,i,n}^* \boldsymbol{U}_{eq,n,j} \tag{25}$$

Note that without phase noise $\boldsymbol{\Psi}_{R,i,i} = 1$ and $\boldsymbol{\Psi}_{R,i,j} = 0$, since the columns of $\boldsymbol{U}$ are orthonormal. The elements of $\boldsymbol{\Psi}_R$ can be expressed as (22). In particular, for the diagonal elements we have $\boldsymbol{\Psi}_{R,i,i} = \Theta(\boldsymbol{a})$ with $a_n = \left|\boldsymbol{U}_{eq,n,i}\right|^2$ and off the diagonal $\boldsymbol{\Psi}_{R,i,j} = \Theta(\boldsymbol{a})$ with $a_n = \boldsymbol{U}_{eq,i,n}^H \boldsymbol{U}_{eq,n,j}$. The same applies to $\boldsymbol{\Psi}_T$ at the transmitter, with $\boldsymbol{V}_{eq}$ instead of $\boldsymbol{U}$. By expanding the final effective $N_s \times N_s$ channel $\boldsymbol{H}_{bb}$ as $\boldsymbol{\Psi}_R \boldsymbol{\Lambda}_{eq} \boldsymbol{\Psi}_T$ one can find that the diagonal entries are

$$\boldsymbol{H}_{bb}(i,i) = \mu_i \boldsymbol{\Psi}_{T,i,i} \boldsymbol{\Psi}_{R,i,i} + \sum_{n \neq i}^{\min(L_t,L_r)} \mu_n \boldsymbol{\Psi}_{R,i,n} \boldsymbol{\Psi}_{T,n,i} \tag{26}$$

and off the diagonal we have

$$\boldsymbol{H}_{bb}(i,j) = \sum_{n \neq i,j}^{\min(L_t,L_r)} \mu_n \boldsymbol{\Psi}_{R,i,n} \boldsymbol{\Psi}_{T,n,i} + \mu_i \boldsymbol{\Psi}_{R,i,i} \boldsymbol{\Psi}_{T,i,j} + \mu_j \boldsymbol{\Psi}_{R,i,j} \boldsymbol{\Psi}_{T,j,j} \tag{27}$$

We should note that the elements of $\boldsymbol{\Psi}_T$ and $\boldsymbol{\Psi}_R$ are independent, since the PN samples at the transmitter $\theta_{T,n}$ and receiver $\theta_{R,n}$ are independent.

The approaches used in this paper to evaluate the effect of PN on the terms of (13) are:

Method 1. The conditional values of (26)–(27) are evaluated for each channel realization, then the average is numerically evaluated over different channel realizations. In detail, for each given channel, the matrices $\boldsymbol{U}_{eq}$ and $\boldsymbol{V}_{eq}$ and the values μ_i of the SVD are known. Therefore, both the diagonal terms $|\boldsymbol{H}_{bb}(i,i)|^2$ in (26) and the interference contributions $|\boldsymbol{H}_{bb}(i,j)|^2$ with $i \neq j$ of (27) are a combination with given values of variables of the type (22), for which the expression of the moments (23), (24) allows the evaluation of the expectation (with respect to the phase noise only).

Method 2. Another conditional approach is obtained by modeling the columns of the unitary matrices $\boldsymbol{U}_{eq}$ and $\boldsymbol{V}_{eq}$ as Gaussian vectors. In particular, it is assumed that each column of $\boldsymbol{V}_{eq}$ ($\boldsymbol{U}_{eq}$) is independent of the other columns, to reflect the orthogonality among the columns, and the elements of each column are iid Gaussian random variables with zero mean and variance $1/L_t$ ($1/L_r$), to model unitary vectors. As shown in the following this represents a quite good approximation. The terms of (13) are evaluated for each channel realization, i.e. it is assumed that the singular values μ_i (or λ_i in the full-digital case) are known. From the independence of $\boldsymbol{\Psi}_T$ and $\boldsymbol{\Psi}_R$ and noting that $E[\boldsymbol{\Psi}_{T,i,j}\boldsymbol{\Psi}_{T,m,n}] = E[\boldsymbol{\Psi}_{R,i,j}\boldsymbol{\Psi}_{R,m,n}] = 0$, $\forall m \neq i, n \neq j$ one gets

$$E[|\boldsymbol{\Psi}_{R,i,i}|^2] = e^{-\sigma_\theta^2} \frac{L_r - 1}{L_r} + \frac{3}{L_r^3} \quad E[\left|\boldsymbol{\Psi}_{R,i,j}\right|^2] = \frac{e^{-\sigma_\theta^2}}{L_r} \tag{28}$$

and the same for $\boldsymbol{\Psi}_{T,i,j}$ (with L_t instead of L_r). Finally

$$E\left[|\boldsymbol{H}_{bb}(i,i)|^2 \left|\boldsymbol{\Lambda}_{eq}\right.\right] \approx \Lambda_i \left[e^{-\sigma_\theta^2} \frac{L_r - 1}{L_r} + \frac{3}{L_r^3}\right] \times \left[e^{-\sigma_\theta^2} \frac{L_t - 1}{L_t} + \frac{3}{L_t^3}\right] + \sum_{n \neq i}^{N_s} \Lambda_n \frac{e^{-2\sigma_\theta^2}}{L_r L_t} \tag{29}$$

$$E\left[|\boldsymbol{H}_{bb}(i,j)|^2 \left|\boldsymbol{\Lambda}_{eq}\right.\right] \approx \sum_{n \neq i,j}^{N_s} \Lambda_n \frac{e^{-2\sigma_\theta^2}}{L_r L_t} + \Lambda_i \left[e^{-\sigma_\theta^2} \frac{L_r - 1}{L_r} + \frac{3}{L_r^3}\right] \frac{e^{-\sigma_\theta^2}}{L_t} + \Lambda_j \left[e^{-\sigma_\theta^2} \frac{L_t - 1}{L_t} + \frac{3}{L_t^3}\right] \frac{e^{-\sigma_\theta^2}}{L_r} \tag{30}$$

where it is stressed that the values are conditioned on the knowledge of the singular values of $\boldsymbol{\Lambda}_{eq}$.

Method 3. A complete analytical approximation of the terms in (13) is considered by putting together the results of Section V on the characterization of the matrix singular values with the approximation just outlined in Method 2. The terms of (13) are given by (29)–(30) where the expectation is taken over Λ_j and $E[\Lambda_j]$ is given by (14).

VII. Effect of Channel Estimation Errors

Another performance degradation is introduced by the estimation error in the channel matrix $\boldsymbol{H}$. Considering now only the effect of the channel estimation (without PN), the estimated channel matrix $\boldsymbol{H}_e$ is

$$\boldsymbol{H}_e = \boldsymbol{H} + \delta\boldsymbol{H}\,, \tag{31}$$

where $\delta\boldsymbol{H}$ accounts for the estimation error. Note that in the hybrid schemes the estimation error can be considered on the reduced size channel $\boldsymbol{H}_{eq}$. Then $\boldsymbol{H}_{bb}$ turns out to be

$$\boldsymbol{H}_{bb} = \left(\boldsymbol{W}_{BB,e}^H \boldsymbol{W}_{RF}^H\right) \boldsymbol{H} \left(\boldsymbol{F}_{RF}\, \boldsymbol{F}_{BB,e}\right)\,. \tag{32}$$

where precoding and combining, $\boldsymbol{W}_{BB,e}$ and $\boldsymbol{F}_{BB,e}$, are derived on the basis of $\boldsymbol{H}_e$. The estimation process can be done either directly on the channel matrix by means of pilots, or in a reduced space, like the beam-space, where the channel matrix is estimated from the arrival angles ϑ_p and the coefficients α_p. The first approach finds applications in TDD multi-user scenarios for the uplink, especially when the number of antennas of each user is small. Since both cases have a practical interest, we examine both.

A. Direct estimation of the channel matrix

Assuming an additive noise model for the estimation error, the elements of $\delta\boldsymbol{H}$ in (31) are i.i.d. Gaussian random variables [42] with zero-mean and variance σ^2_{MSE} equal to the mean squared error (MSE). In other words, we can assume that the estimation error is iid over all the channel matrix elements. A perturbation analysis [43] provides the tools to evaluate the degradation and in particular the derivation of the variation in the SVD matrices is outlined in Appendix B. According to this scheme, the contribution of interference is given by

$$\sum_{j\neq i}^{N_s} E\left[\left|\boldsymbol{H}_{bb}(i,j)\right|^2\right] \approx \frac{3(Ns-1)}{L_r L_t}\sigma^2_{\text{MSE}}\,, \tag{33}$$

with $L_t = N_t$ and $L_r = N_r$ in the full-digital case. The goodness of the theoretical expression in terms of SINR degradation is shown in Fig. 14 of Section VIII.

B. Estimation of the beam-space matrix

The channel estimation can be done on a reduced space, due to the channel sparsity in the mmWave range. When the channel is estimated in the beam-space, the estimation error affects directly the channel parameters α_p and ϑ_p of (3). A detailed statistical characterization of the estimation error is out of the scope of this work and can depend on the specific estimation algorithm. Here we model the estimation errors by zero-mean Gaussian random variables with variance $\sigma^2_{MSE,bs}$, which is justified in the case of Gaussian distribution of each angle of arrival [44]. To reduce the number of parameters in the presentation of the results, we assume that the variance of the error is the same on α_p and ϑ_p: note however that the errors on the angles are more detrimental, due to the massive number of antennas, which provide high directivity and a drop of the antenna gain at angles different form the steering angle. In this case the estimation error affects directly the singular values and the columns/rows of the SVD matrices, as outlined in Appendix C. Then again we have interference arising from estimation error, affecting the SINR (13). This can be represented by terms as in (29)–(30) with

$$E[|\boldsymbol{\Psi}_{R,i,i}|^2] = 1 - \frac{L_r}{N_r}\sigma^2_{MSE,bs}\frac{\pi^2}{12} \quad E[|\boldsymbol{\Psi}_{R,i,j}|^2] = \frac{L_r}{N_r}\sigma^2_{MSE,bs}\frac{\pi^2}{12} \tag{34}$$

and the same for $\boldsymbol{\Psi}_R$ with L_t and N_t.

A comparison with the simulation results, in terms of the SINR degradation, is shown in Fig. 15 of Section VIII, showing the good precision of the theoretical expression (34).

VIII. Numerical Results and Discussion

A. Equivalence between singular values and channel coefficients

The equivalence between the coefficients of the beam-space representation (3) and the singular values is confirmed by Fig. 4, where the rates obtained with the actual SVD and the coefficients α_p of the beam-space channel representation are compared. Clearly we can see two effects:

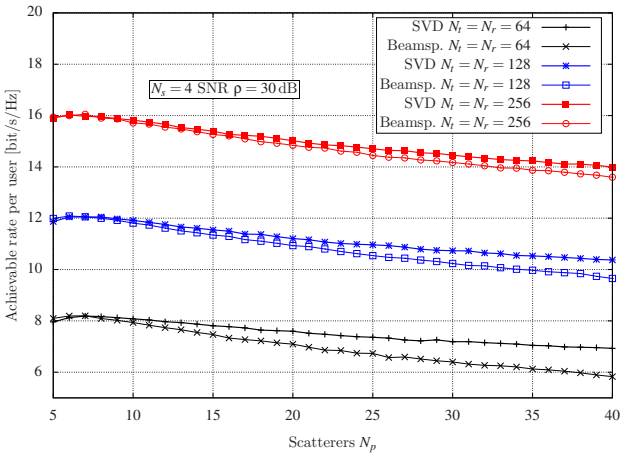


Fig. 4. Comparison between the rates with $\Lambda_i = |\alpha_i|^2|$ and with the actual channel singular values, for $N_s = 4$ and different values $N_t = N_r$.

- When the number of antennas reaches the typical values of massive MIMO systems, the difference between the rates greatly reduces.
- Only when the number of scatterers increases, which is not the case of sparse mmWave channels, a certain difference appears between the rates.

B. SINR statistics

First we show that the statistics of the SINR (11) are almost perfectly described by a Gaussian distribution. This is in accordance with similar results on the statistics of the dominant

eigenvalue of a MIMO channel [45]. This is quantified by the Kolmogorov-Smirnov (KS) test and the Kullback-Leibler (KL) divergence between the SINR PDF and a Gaussian PDF with the same mean and variance, which are presented in Table I for some representative values of N_s and N_p.

	KS test	KL divergence
Log-Rayleigh $N_s = 4$ $N_p = 10$	$7.4 \cdot 10^{-2}$	$1.3 \cdot 10^{-1}$
Gaussian $N_s = 4$ $N_p = 10$	$5.7 \cdot 10^{-3}$	$2.4 \cdot 10^{-3}$
Gaussian $N_s = 4$ $N_p = 20$	$2.1 \cdot 10^{-2}$	$7.9 \cdot 10^{-3}$
Gaussian $N_s = 8$ $N_p = 20$	$2.2 \cdot 10^{-2}$	$1.0 \cdot 10^{-2}$

TABLE I
VALUES OF KS TEST AND KL DIVERGENCE BETWEEN THE SINR PDF AND LOG-RAYLEIGH OR GAUSSIAN PDF.

Given the accuracy of the Gaussian assumption for the SINR, we compare in Fig. 5 the analytical and simulation values of the mean and standard deviation as a function of the number of scatterers N_p of the channel. We can note

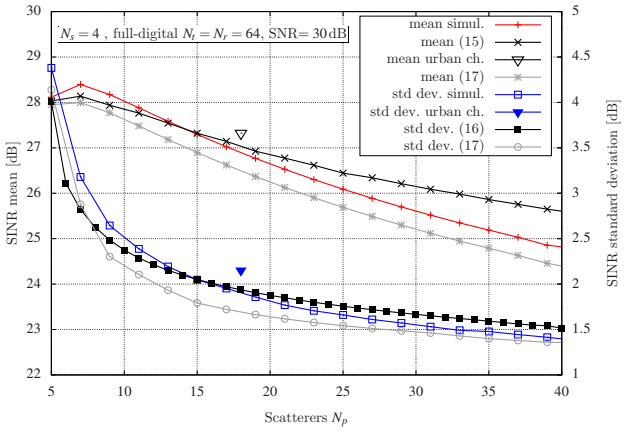


Fig. 5. SINR mean and standard deviation for the full-digital case as a function of the number of scatterers N_p, for $N_t = N_r = 64$ and $N_t = N_r = 128$.

that the analytical model of (15)-(16) provides a very good approximation, which increases when the number of antennas is large, as the case addressed of massive MIMO. This is expected since the orthogonality between the steering vectors is higher for increasing values of N_t, N_r and the singular values are well approximated by the squared gains α_p. A slightly better result is obtained by resorting to the integration of the joint PDF (41). The decrease of the standard deviation with N_p is expected, since the largest N_s values tend to be closer to each other increasing the size of the set. The decrease of the mean value is due to the power normalization of the channel, with coefficients α_p having variance $1/N_p$. In the figure we present also the comparison with the values obtained by a realistic channel model [26] implemented by a software simulator [27]. It can be seen that the theoretical values are a very good approximation, by considering the number of scatterers N_p of this model equal to the number of singular values of the channel, i.e. the channel rank. If we increase the number of streams N_s, more singular values contribute to the SINR. In Fig. 6 we plot the SINR mean and standard deviation as a function of the number of streams N_s, for $N_t = N_r = 64$. The

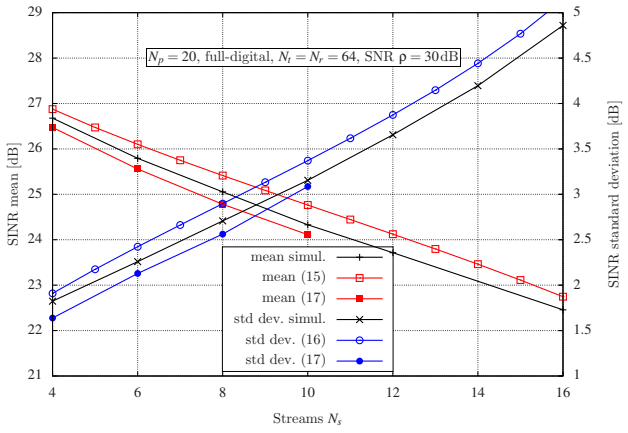


Fig. 6. SINR mean and standard deviation for the full-digital case as a function of the number of streams N_s, for $N_p = 20$ scatterers, $N_t = N_r = 64$, and SNR $\rho = 30$dB.

evaluation of multiple integrals of the analytical PDF (17) becomes lengthy for large N_s and it is limited to $N_s = 10$ in the figure. However the analytical expressions (15)-(16) can be evaluated for any values of N_s and give values of the SINR moments which differ only fractions of dB from the actual values. Note that considering more scatterers increases the variability of the SINR, hence the variance, but also reduces the mean value due to the channel gain normalization.

For the hybrid analog-digital beamforming similar results are presented in Fig. 7, which shows the mean SINR value for different numbers of RF chains with $N_t = N_r = 64$ antennas. In this case the results of the full-digital scheme and the lower bounds (LB) for the hybrid scheme for $N_p \geq (N_t - L_t + N_s)$ are compared with the values obtained by simulation. We can

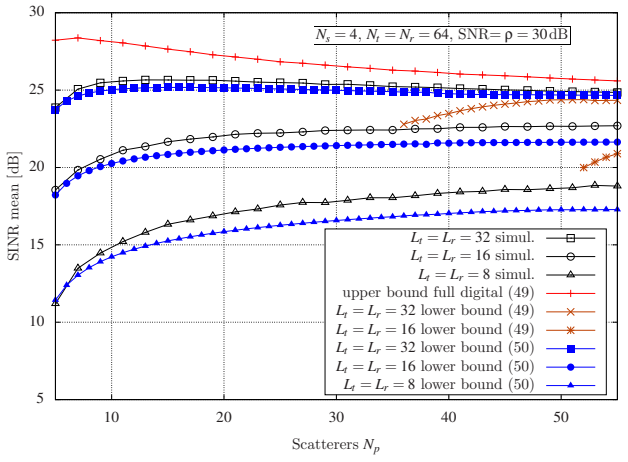


Fig. 7. SINR mean value as a function of the number of scatterers N_p, for $N_s = 4$, $L_t = L_r = 8,\ 16,\ 32$, $N_t = N_r = 64$ antennas, and SNR $\rho = 30$dB.

note that the bound (18) related to the generalized Cauchy interlacing theorem is useful only in the case of a large number of scatterers, which is not very suited to typical mmWave channels. On the other hand, the bound (19) based on the array pattern provides a very good approximation to the actual mean in the whole range of N_p,. This latter bound becomes slightly looser when increasing the number of scatterers N_p. Anyway, the validity of the approximation is good for all the values of

N_p, in particular for realistic values of N_p in mmWave, which are around 10–20 [24], [26], [27]. The difference between the full-digital case and the hybrid analog-digital beamforming becomes smaller if the number of RF chains increases from $L_t = L_r = 8$ to 16 and 32, where the hybrid analog-digital scheme gets nearer to the full-digital one. In any case a clear idea of the trend is obtained by the lower bound, with errors smaller than 1 dB. In Fig. 8 the values of the SINR standard deviation are presented for the same system setup, as a function of the number of scatterers N_p. We should note

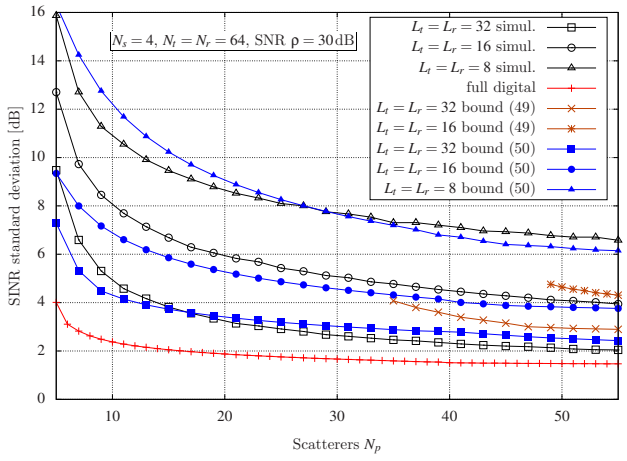


Fig. 8. SINR standard deviation as a function of the number of scatterers N_p, for $N_s = 4$, $L_t = L_r = 8, 16, 32$, $N_t = N_r = 64$ antennas, and SNR $\rho = 30$dB.

first of all that a lower and upper bound on the singular values does not reflect automatically on a central moment such as the standard deviation, but actually on the first and second non-central moments. In any case, also for the standard deviation the considerations on the suitability of bound (18) to a mmWave channel apply. On the other hand, the bound (19) provides a good match to the simulated values, within fractions of dB. In this figure we present the performance of both the hybrid analog-digital and the full-digital scheme. The latter can be approached also by more sophisticated hybrid analog-digital beamforming techniques, for which the SINR formulas (15)-(16) provide a good match.

C. Effect of the phase noise

Also in the presence of phase noise the PDF of the SINR fits almost perfectly a Gaussian PDF, for both the full-digital and the hybrid case with DFT-based matrix, as shown by the values of the Kolmogorov-Smirnov (KS) test and the Kullback-Leibler divergence (KL) presented in Table II. Due to the good

PN σ_θ^2	Full-dig. KS	Full-dig. KL	Hyb. KS	Hyb. KL
0	$2.1 \cdot 10^{-2}$	$7.9 \cdot 10^{-3}$	$2.3 \cdot 10^{-2}$	$3.0 \cdot 10^{-2}$
0.01	$2.8 \cdot 10^{-2}$	$2.2 \cdot 10^{-2}$	$2.5 \cdot 10^{-2}$	$5.3 \cdot 10^{-2}$
0.05	$3.0 \cdot 10^{-2}$	$2.4 \cdot 10^{-2}$	$2.4 \cdot 10^{-2}$	$5.4 \cdot 10^{-2}$
0.1	$3.0 \cdot 10^{-2}$	$2.5 \cdot 10^{-2}$	$2.4 \cdot 10^{-2}$	$5.4 \cdot 10^{-2}$

TABLE II
VALUES OF THE KS TEST AND KL DIVERGENCE BETWEEN THE SINR PDF AND A GAUSSIAN PDF FOR $N_s = 4$, $N_p = 10$.

match between the PDF and its Gaussian approximation, one can characterize the SINR by its mean and variance.

In order to compare different channel models, in Figs. 9–10 we present the achievable rates per stream obtained: by simulation, by the semi-analytical approach with the exact SVD matrices of Method 1 and by the approximate Method 2 with (29)–(30). Fig. 9 shows the full-digital case and Fig. 10 the hybrid scheme. Together with the channel model of (3) with N_p scatterers, in Fig. 9 we present also a case in which the scatterers are grouped into clusters (10 clusters) with a model like [26], [27]. It can be seen that the theoretical analysis

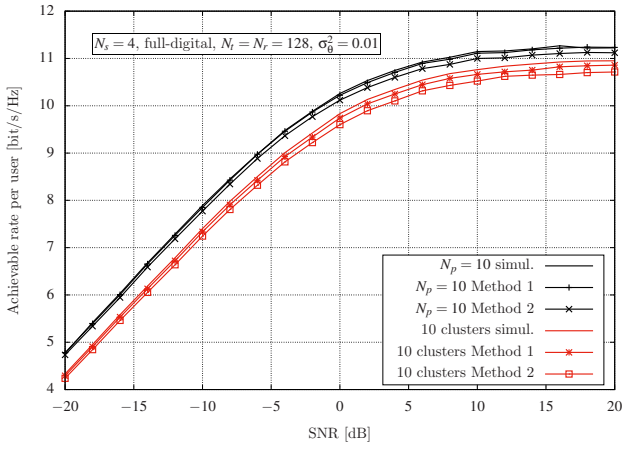


Fig. 9. Achievable rate per stream, with PN variance $\sigma_\theta^2 = 0.01$, $N_p = 10$ scatterers and $N_t = N_r = 128$: comparison of analytical and simulation results.

suits well different channel models. Moreover, the difference between the beam-space model (3) and more sophisticated models, with scatterers grouped in clusters, is reduced.

In the following, since we proved the accuracy of both the singular value characterization and the accuracy of the PN effects, Method 3 will be used. In Fig. 10 the results are shown for the hybrid scheme with $L_t = L_r = 8$ RF chains and $N_t = N_r = 64$ antennas. The same channel models as in Fig. 9 are used, together with a further model, in which we consider single scatterers (not grouped into clusters), but with a minimum separation between the angles of different scatterers, set to 10 degrees in the case of Fig. 9. With separated scatterers, the equivalence between the beam-space model (3) and a matrix spectral decomposition holds almost perfectly. In any case, for the full-digital case theoretical and simulated values are almost overlapped, while for the hybrid case a very small difference is noticeable only at high SNR.

In order to validate the theoretical analysis of Section VI, in Fig. 11 we present the achievable rate per stream for a hybrid system with $L_t = L_r = 16$ RF chains and $N_t = N_r = 64$ antennas, for a channel with $N_p = 20$ scatterers. It can be clearly seen that the analysis gives a very good approximation especially in the range of PN of practical interest. The accuracy tends to become looser for a very high PN variance, which is not practical, since the degradation would be too severe, as it can be seen in the following results.

A first comparison between the effects of PN on the full-digital and the hybrid schemes is shown in Fig. 12. The results are obtained by the analytical approach of Method 3. Note that

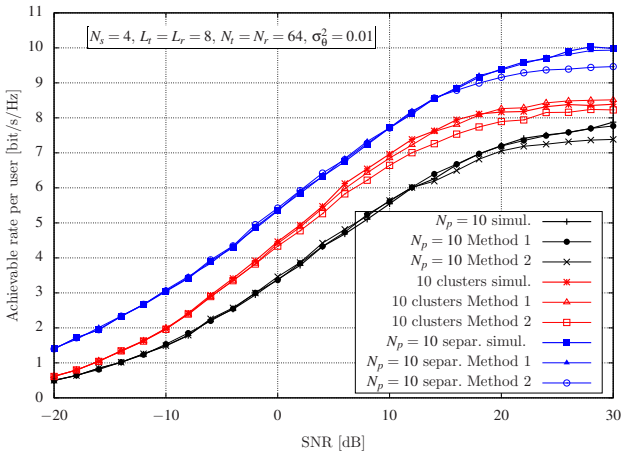


Fig. 10. Achievable rate per stream: comparison of analytical and simulation results, with PN variance $\sigma_\theta^2 = 0.01$, $L_t = L_r = 8$ and $N_t = N_r = 64$.

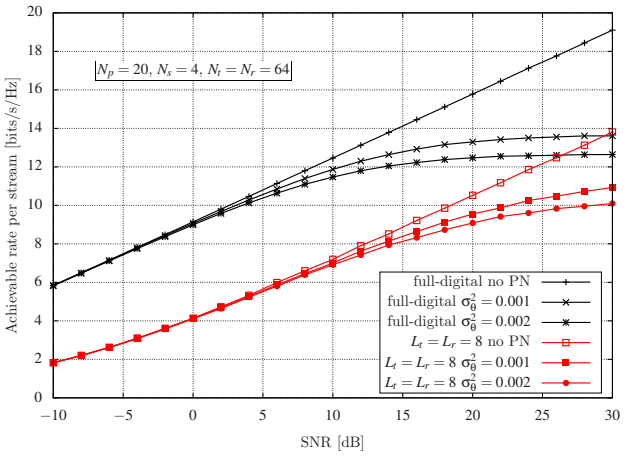


Fig. 12. Achievable rate per stream with different values of σ_θ^2, $N_p = 20$ scatterers, $N_t = N_r = 64$, and $L_t = L_r = 8$ RF chains for the hybrid scheme.

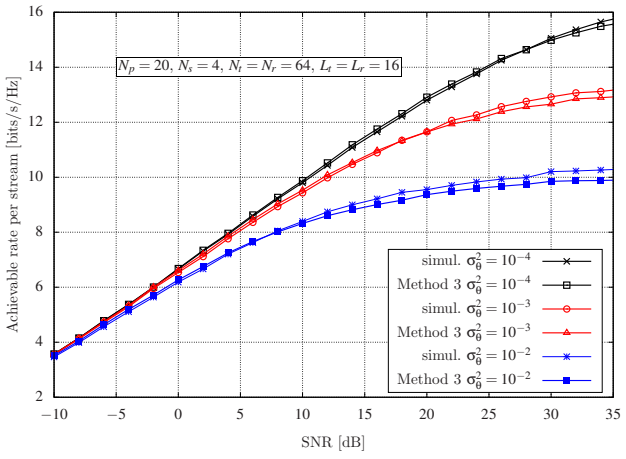


Fig. 11. Achievable rate per stream with a channel with $N_p = 20$ scatterers, for different values of σ_θ^2, $L_t = L_r = 16$, and $N_t = N_r = 64$.

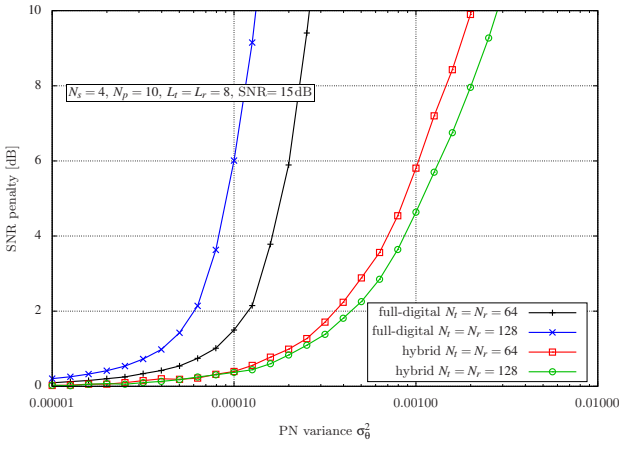


Fig. 13. SNR penalty at the reference SNR of 15 dB, for a channel with $N_p = 10$ scatterers as a function of the PN variance σ_θ^2, $L_t = L_r = 8$ RF chains and $N_t = N_r = 64$ or $N_t = N_r = 128$ antennas.

the large gap between the full-digital approach and the hybrid solution is due to the limited number of RF chains compared to the number of antennas, having a fixed DFT analog matrix. However it can be noticed a smaller sensitivity to the PN experienced by the hybrid scheme. In fact, if we consider the sensitivity to the PN, the SNR penalty is presented in Fig. 13, at the reference SNR of 15 dB, for $N_t = N_r = 64, 128$ and $L_t = L_r = 8$, again by the analytical approach of Method 3. We note that the full-digital scheme is much more sensitive to the effects of PN than hybrid schemes. A greater number of antennas gives a worse degradation in the full-digital scheme, due to the increased number of independent oscillators, while for hybrid analog-digital schemes the sensitivity to PN is determined by the number of RF chains.

D. Sensitivity to the channel estimation error

In Fig. 14 the sensitivity to the channel estimation error is presented for an additive estimation error on each matrix element characterized by a mean squared error (MSE) relative to the channel matrix value. Also in this case the degradation is measured by the *SNR penalty*, defined as the increase of SNR necessary to reach the same achievable rate as in the absence of estimation errors. It is clear that the sensitivity of the full-digital scheme to the channel estimation error in this case is lower than for hybrid schemes. On the other hand, if the estimation is performed directly on the beam-space, the sensitivity to the estimation error is much bigger and it is plotted in Fig. 15, showing an opposite effect, with much greater sensitivity of the full-digital scheme with respect to the hybrid scheme. By combining the effects of PN and of the channel estimation error, we compare the rate of the full-digital and the hybrid schemes in Fig. 16, using simulation results, since the joint analysis in not tractable, for a channel with $N_p = 10$ scatterers, SNR of 15 dB, for $N_t = N_r = 64$ and $L_t = L_r = 8$. We can see that when the effects of the channel estimation error and of the PN are combined, the difference between the achievable rate of the full-digital approach and the DFT-based hybrid scheme greatly reduces, even with a limited number of RF chains with respect to the number of antennas. From the results considering jointly both PN and the estimation error, we can say that the overall penalty is well approximated by the sum of the two penalties, suggesting that the two effects can be treated separately.

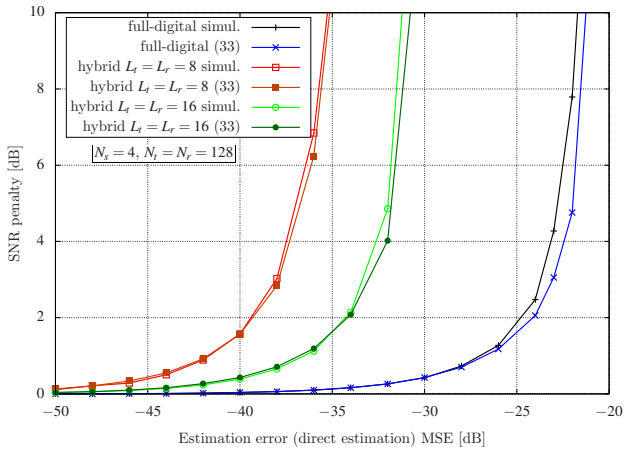


Fig. 14. SNR penalty at SNR= 15 dB, for a channel with $N_p = 10$ scatterers as a function of the estimation error MSE (with direct estimation of the channel matrix), $N_t = N_r = 128$ antennas and $L_t = L_r = 8$ or $L_t = L_r = 16$ RF chains.

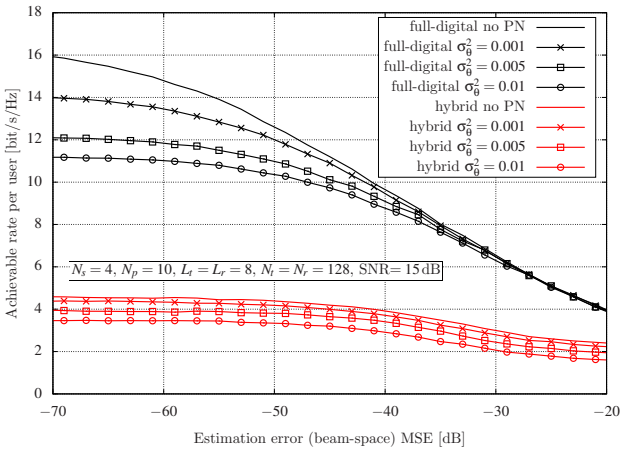


Fig. 16. Achievable rate per user at SNR= 15 dB as a function of the estimation error MSE (in the beam-space), for a channel with $N_p = 10$ scatterers, different levels of PN, $N_t = N_r = 128$, and $L_t = L_r = 8$.

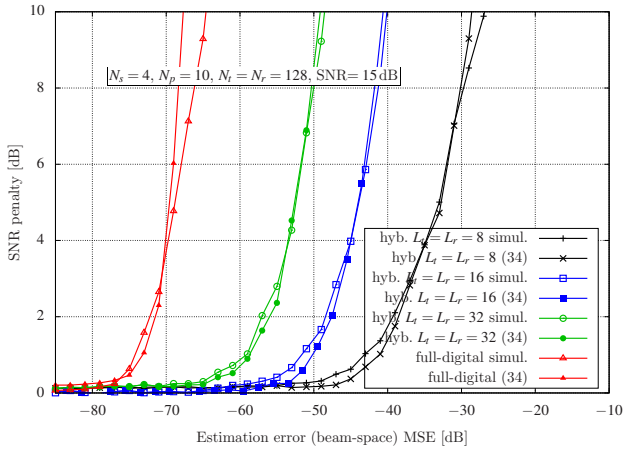


Fig. 15. SNR penalty at SNR= 15 dB, for a channel with $N_p = 10$ scatterers as a function of the estimation error MSE (in the beam-space), $N_t = N_r = 128$ antennas and different numbers of RF chains.

E. Discussion

A first observation concerns the mmWave channel: the singular values determining the SINR can be considered equal to the channel coefficients of the beam-space representation. The validity of this assumption increases with the number of antennas, becoming very suitable to massive MIMO systems. The sparsity of the mmWave channel can be captured by the beam-space representation. For practical mmWave channels the number of scatterers is limited to about 10–20, with a consequent limit in the number of parallel streams that can be spatially multiplexed. Moreover, the SINR in a log-scale is well modeled by a Gaussian distribution: this holds both in the full-digital case and in hybrid digital-analog schemes. Therefore, a mean-variance characterization of the SINR provides a full statistical description.

Impairments such as the PN and the channel estimation error introduce an inter-stream interference, due to precoding and combinig matrices not matched to the channel SVD. For the effect of PN, the choice of the number of RF chains determines a trade-off between the reduction of achievable rate with respect to the full-digital case and the PN degradation. Although it may be unfeasible in massive MIMO, when the number of RF chains approaches half the number of antennas, $L_t \approx N_t/2$, the rate obtained is almost the same as in a full-digital scheme. Moreover, the sensitivity to PN is much higher for the full-digital scheme than for hybrid schemes (Fig. 13), with an interference term due to PN almost proportional to the number of antennas in the full-digital scheme and to the number of RF chains in hybrid schemes. Concerning the channel estimation error, these schemes have a different sensitivity, depending on the channel estimation method. With a pilot-based estimation obtaining directly the channel matrix, the full-digital approach proves to be more robust, and the interference decreases with the number of antennas, as in (33). If the sparsity of the mmWave channel is exploited by the estimation method, identifying the beam-space parameters, hybrid schemes are more robust, with an interference in the SINR which depends on L_t/N_t. For practical values of PN ($\sigma_\theta^2 \leq 10^{-2}$) and channel estimation MSE (≤ -60 dB), the penalty due to the estimation error is more severe than the one due to PN. However, with a beam-space channel estimation, the combination of the two effects can reduce the rate of full-digital systems to values comparable with hybrid schemes.

IX. Conclusions

We derived the SINR statistics in massive MIMO for the mmWave channel, for full-digital and hybrid analog-digital SVD-based schemes. The full-digital case can be taken as reference also for hybrid schemes with the joint optimization of the analog and digital parts. The hybrid scheme presented here is sub-optimum, but with a simpler design, since it does not require the knowledge of the full channel matrix but only of the reduced equivalent channel at the RF chains level. The main conclusions are: i) The analytical expression of the SINR mean and variance and the joint PDF of the largest singular values have been derived. With the Gaussian model of the SINR, this allows the theoretical evaluation of the achievable rate. ii) In the presence of PN we obtained the analytical expression of the interference due to the imperfect SVD, giving the SINR degradation and the achievable rate.

iii) In the presence of channel estimation errors we derived the expression of the SINR degradation as a function of the estimation MSE, both for a uniform estimation error on the channel matrix and for an error in the beam-space parameters. A practical design solution for hybrid analog-digital schemes is the use of an analog matrix of fixed phase-shifters, with columns taken from the DFT matrix. This scheme is more robust to PN and channel estimation errors, narrowing the gap in terms of achievable rate with the full-digital approach, if $N_t \leq 8L_t$. For more antennas the rate reduction is still notable.

Appendix A
Derivation of the SINR mean and variance

Consider N_p iid random variables with PDF $f(a)$ and CDF $F(a)$. The i-th ordered (in decreasing order) random variable has PDF given by [40]

$$f_i(a_i) = \frac{N_p!}{(N_p-i)!(i-1)!} F(a_i)^{N_p-i} [1-F(a_i)]^{i-1} f(a_i) \quad (35)$$

From (3), $\Lambda_p = |\alpha_p|^2$ are exponentially distributed, $f_{\Lambda_p}(a) = \beta e^{-\beta a}$, with $\beta = N_p$. The SINR on the i-th stream, corresponding to the i-th largest Λ_p, is $\mathrm{SINR}_{i,dB} = 10\log_{10}\rho + \frac{10}{\ln(10)} y_i$, where the random variable $y_i = \ln \Lambda_i$, for $i = 1,\ldots,N_s$ has PDF

$$f_{y_i}(a_i) = \frac{N_p! \left[1-e^{-\beta e^{a_i}}\right]^{N_p-i}}{(N_p-i)!\,(i-1)!} e^{-\beta(i)e^{a_i}} \beta e^{a_i} \quad (36)$$

By using the following integrals [46]

$$\int_{-\infty}^{+\infty} x e^{-\alpha e^x + x} dx = \frac{1}{\alpha}\left[\left(\ln\frac{1}{\alpha} - \gamma\right)\right] \quad (37)$$

$$\int_{-\infty}^{+\infty} x^2 e^{-\alpha e^x + x} dx = \frac{1}{\alpha}\left[\frac{\pi^2}{6} + \left(\gamma - \ln\frac{1}{\alpha}\right)^2\right], \quad (38)$$

where γ is the Euler-Mascheroni constant, the mean and the variance of y_i turn out to be

$$m_{y_i} = E[y_i] = \frac{N_p!}{(i-1)!\,(N_p-i)!} \sum_{n=0}^{N_p-i} \binom{N_p-i}{n} \frac{(-1)^n}{n+1} \times \left[\ln\frac{1}{\beta(n+1)} - \gamma\right] \quad (39)$$

$$M_{y_i} = E[y_i^2] = \frac{N_p!}{(i-1)!\,(N_p-i)!} \sum_{n=0}^{N_p-i} \binom{N_p-i}{n} \frac{(-1)^n}{i+n+1} \times \left[\frac{\pi^2}{6} + \left(\ln\frac{1}{\beta(n+1)} - \gamma\right)^2\right]. \quad (40)$$

A more detailed analysis is obtained by considering the joint statistics of the N_s largest singular values among N_p. From the general results [40], the joint PDF of the random variables y_i is

$$f_{y_1,\ldots,y_{N_s}}(a_1,\ldots,a_{N_s}) = \frac{N_p! \left[1-e^{-\beta e^{a_{N_s}}}\right]^{(N_p-N_s)}}{(N_p-N_s)!} \prod_{i=1}^{N_s} \beta e^{a_i - \beta e^{a_i}}, \quad (41)$$

for $a_1 \geq a_2 \geq \ldots \geq a_{N_s}$. Then the expected value and all the moments can be obtained by integration of the PDF (41). Note however that the numerical evaluation of the multiple integrals with the PDF (41) becomes lengthy. The approximation of separating the integrals for each variable leads to (15), (16).

Appendix B
Perturbation analysis

Assume that the SVD of $\boldsymbol{H}$ is $\boldsymbol{H} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{V}^H$. A perturbation $\boldsymbol{H}_e = \boldsymbol{H} + \delta\boldsymbol{H}$, can be decomposed as $\boldsymbol{H}_e = \boldsymbol{U}_e\boldsymbol{\Lambda}_e\boldsymbol{V}_e^H$ with

$$\boldsymbol{\Lambda}_e = \boldsymbol{\Lambda} + \delta\boldsymbol{\Lambda} \qquad \boldsymbol{U}_e = \boldsymbol{U} + \delta\boldsymbol{U} \qquad \boldsymbol{V}_e = \boldsymbol{V} + \delta\boldsymbol{V} \quad (42)$$

The theory of [43] gives the matrices $\boldsymbol{U}_e$, $\boldsymbol{\Lambda}_e$ and $\boldsymbol{V}_e$, in a general framework, up to the second order. Considering the first-order perturbation and $\delta\boldsymbol{H}$ with iid elements with zero-mean and variance σ^2_{MSE} (as in Section VII-A), we have

$$\delta\boldsymbol{\Lambda} = \boldsymbol{U}^H\delta\boldsymbol{H}\boldsymbol{V} \quad \delta\boldsymbol{U} = \delta\boldsymbol{H}\boldsymbol{V}\boldsymbol{\Lambda}^{-1} \quad \delta\boldsymbol{V} = \delta\boldsymbol{H}\boldsymbol{U}\boldsymbol{\Lambda}^{-1}. \quad (43)$$

The error comes from the application of the matrices derived from $\boldsymbol{H}_e$ to the actual channel $\boldsymbol{H}$. We have

$$\boldsymbol{H}_{bb} = \boldsymbol{U}_e^H\boldsymbol{H}\boldsymbol{V}_e = (\boldsymbol{U}+\delta\boldsymbol{U})^H\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{V}^H(\boldsymbol{V}+\delta\boldsymbol{V}). \quad (44)$$

Using (43) and considering only the first-order perturbations, with respect to the ideal case $\boldsymbol{H}_{bb,e}$ we have an error $\delta\boldsymbol{H}_{bb,e}$

$$\boldsymbol{H}_{bb} = \boldsymbol{H}_{bb,e} + \delta\boldsymbol{H}_{bb,e} = \boldsymbol{H}_{bb,e} + 3\boldsymbol{U}^H\delta\boldsymbol{H}\boldsymbol{V}. \quad (45)$$

By the approximation of considering the columns of $\boldsymbol{U}$ and $\boldsymbol{V}$ as Gaussian vectors with iid zero-mean elements with variance $1/L_r$ and $1/L_t$, the interference term (33) is

$$\sum_{j\neq i}^{N_s} E\left[|\boldsymbol{H}_{bb}(i,j)|^2\right] = \frac{3(Ns-1)}{L_rL_t}\sigma^2_{\mathrm{MSE}}. \quad (46)$$

Appendix C
Effect of errors in the beam-space

We assume that the channel coefficients, α_p and the angles $\vartheta_{T,p}$, $\vartheta_{R,p}$ of (3) are affected by errors $\delta\alpha_p$, $\delta\vartheta_{T,p}$, $\delta\vartheta_{R,p}$, modeled by iid zero-mean Gaussian random variables with variance $\sigma^2_{MSE,bs}$. It can be easily proven that the effect of the errors $\delta\alpha_p$ are negligible with respect to the errors $\delta\vartheta_{T,p}$, $\delta\vartheta_{R,p}$ on the angles, due to the high antenna directivity. For $\vartheta_{T,p}$ and $\vartheta_{R,p}$ uniformly distributed in $(-\pi/2,\pi/2)$, the variance of $\delta\vartheta_{T,p}$, $\delta\vartheta_{R,p}$ is $\sigma^2_{MSE,bs}\pi^2/12$. The derivation of the interference terms in $\boldsymbol{H}_{bb}$ is similar to the PN effect of Section VI. The difference is that now the columns of the precoding and combining matrices correspond to steering vectors to the angles $\vartheta_{T,p}+\delta\vartheta_{T,p}$ and $\vartheta_{R,p}+\delta\vartheta_{R,p}$. With reference to Fig. 3, the elements of the product of the combining and the SVD matrices are now the inner products (5) for $\delta\vartheta_{R,i}$, scaled by $\sqrt{L_r/N_r}$. The Taylor expansion of (5), for small MSE, gives

$$E\left[|\boldsymbol{\Psi}_{R,i,i}|^2\right] = 1 - \frac{L_r}{N_r}\sigma^2_{MSE,bs}\frac{\pi^2}{12} \quad E\left[|\boldsymbol{\Psi}_{R,i,j}|^2\right] = \frac{L_r}{N_r}\sigma^2_{MSE,bs}\frac{\pi^2}{12} \quad (47)$$

Then the same steps as in Section VI lead to (34).

References

[1] T. E. Bogale and L. B. Le, "Massive MIMO and mmWave for 5G wireless HetNet: potential benefits and challenges," *IEEE Veh. Technol. Mag.*, vol. 11, no. 1, pp. 64–75, March 2016.

[2] J. Zhu, W. Xu and N. Wang, "Secure massive MIMO systems with limited RF chains," *IEEE Trans. Veh. Technol.*, vol. 66, no. 6, pp. 5455–5460, June 2017.

[3] R. W. Heath Jr., N. Gonzalez-Prelcic, S. Rangan, W. Roh, and A. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems", *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, April 2016.

[4] M. Sanchez-Fernandez, A. Tulino, E. Rajo-Iglesias, J. Llorca, and A. Garcia Armada, "Blended antenna wearables for an unconstrained mobile experience," *IEEE Commun. Mag.*, pp. 2–10, Feb. 2017.

[5] D. Soldani, P. Airas, T. Hoglund, H. Rasanen, and D. Debrecht, "5G to the Home," *IEEE 85th Veh. Technol. Conf. (VTC Spring)*, Sydney, 2017.

[6] Y. Wu, C. K. Wen, D. W. K. Ng, R. Schober, and A. Lozano, "Low-complexity MIMO precoding with discrete signals and statistical CSI," *IEEE Int. Conf. on Commun. (ICC)*, Kuala Lumpur, 2016.

[7] W. B. Sun, Q. Y. Yu, W. X. Meng, and C. Li, "Random beamforming for multiuser multiplexing in downlink correlated Rician channel," *IEEE Int. Conf. on Commun. (ICC)*, Kuala Lumpur, 2016.

[8] S. Park, J. Park, A. Yazdan, and R. W. Heath, "Exploiting spatial channel covariance for hybrid precoding in massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 65, no. 14, pp. 3818–3832, July 2017.

[9] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressed channel sensing: a new approach to estimating sparse multipath channels," *Proc. IEEE*, vol. 98, no. 6, pp. 1058–1076, June 2010.

[10] J. Brady, N. Behdad, and A. Sayeed, "Beamspace MIMO for millimeter-wave communications: System architecture, modeling, analysis, and measurements," *IEEE Trans. Antennas Propag.*, vol. 61, no. 7, pp. 3814–3827, July 2013.

[11] H. Noh, Y. Kim, J. Lee, and C. Lee, "Codebook design of generalized space shift keying for FDD massive MIMO systems in spatially correlated channels," *IEEE Trans. Veh. Technol.*, vol. 64, no. 2, pp. 513-523, Feb. 2015.

[12] X. Gao, L. Dai, S. Han, C. I, and X. Wang, "Reliable beamspace channel estimation for millimeter-wave massive MIMO systems with lens antenna array," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 6010–6021, Sept. 2017.

[13] K. Venugopal, M. C. Valenti and R. W. Heath, "Device-to-device millimeter wave communications: interference, coverage, rate, and finite topologies," *IEEE Trans. Wireless Commun.*, vol. 15, no. 9, pp. 6175–6188, Sept. 2016.

[14] V. Petrov, M. Komarov, D. Moltchanov, J. M. Jornet and Y. Koucheryavy, "Interference and SINR in millimeter wave and Terahertz communication systems with blocking and directional antennas," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1791–1808, March 2017.

[15] A. K. Papazafeiropoulos, "Impact of general channel aging conditions on the downlink performance of massive MIMO," *IEEE Trans. Veh. Technol.*, vol. 66, no. 2, pp. 1428–1442, Feb. 2017.

[16] Y. Zhang, D. Wang, X. Xia, and X. You, "Downlink performance of hybrid precoding in massive MIMO systems subject to phase noise," *9th Int. Conf. on Wireless Commun. and Signal Process. (WCSP)*, Nanjing, 2017.

[17] A. Alkhateeb, Y. Nam, M. S. Rahman, J. Zhang, and R. W. Heath, "Initial beam association in millimeter wave cellular systems: analysis and design insights," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2807–2821, May 2017.

[18] I. A. Hemadeh, K. Satyanarayana, M. El-Hajjar, and L. Hanzo, "Millimeter-Wave Communications: Physical Channel Models, Design Considerations, Antenna Constructions, and Link-Budget," *IEEE Commun. Surveys & Tutorials*, vol. 20, no. 2, pp. 870–913, Secondquarter 2018.

[19] S. Sun, T. S. Rappaport, and M. Shaft, "Hybrid beamforming for 5G millimeter-wave multi-cell networks," *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, Honolulu, HI, 2018, pp. 589–596.

[20] R. Corvaja, A. García Armada, "Effect of multipath and antenna diversity in MIMO-OFDM systems with imperfect channel estimation and phase noise compensation," *Physical Communication*, Vol. 1, Issue 4, 2008, pp. 288–297.

[21] R. Corvaja, A. García Armada, M. A. Vázquez, and A. Pérez-Neira, "Design of precoding and combining in hybrid analog-digital massive MIMO with phase noise,", *IEEE 25th European Signal Processing Conference (EUSIPCO)*, 2017.

[22] S. Park, A. Alkhateeb, and R. W. Heath, "Dynamic subarrays for hybrid precoding in wideband mmWave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2907–2920, May 2017.

[23] V. V. Ratnam, A. F. Molisch, O. Y. Bursalioglu, and H. C. Papadopoulos, "Hybrid beamforming with selection for multiuser massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 66, no. 15, pp. 4105–4120, Aug. 2018.

[24] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE J. Sel. Areas Commun.*, vol.32, no. 6, pp. 1164–1179, June 2014.

[25] A. Maltsev, A. Pudeyev, A. Lomayev, and I. Bolotin, "Channel modeling in the next generation mmWave Wi-Fi: IEEE 802.11ay standard," *European Wireless 2016*, Oulu, Finland, 2016.

[26] T. S. Rappaport, G. R. MacCartney, Jr., S. Sun, H. Yan, and S. Deng, "Small-scale, local area, and transitional millimeter wave propagation for 5G communications," *IEEE Trans. Antennas Propag.*, vol. 65, no. 12, pp. 6474–6490, Dec. 2017.

[27] NYUSIM: The open source 5G channel model simulator software, available at *http://wireless.engineering.nyu.edu/5g-millimeter-wave-channel-modeling-software*, July 2016.

[28] M. R. Khanzadi, G. Durisi, and T. Eriksson, "Capacity of SIMO and MISO Phase-Noise Channels With Common/Separate Oscillators," *IEEE Trans. Commun.*, vol. 63, no. 9, pp. 3218–3231, Sept. 2015.

[29] R. Krishnan *et al.*, "Linear Massive MIMO Precoders in the Presence of Phase NoiseA Large-Scale Analysis," in *IEEE Trans. Veh. Technol.*, vol. 65, no. 5, pp. 3057–3071, May 2016.

[30] R. Corvaja and A. García Armada, "Phase Noise Degradation in Massive MIMO Downlink With Zero-Forcing and Maximum Ratio Transmission Precoding," *IEEE Trans. Veh. Technol.*, vol. 65, no. 10, Oct. 2016.

[31] X. Gao, L. Dai, and A. M. Sayeed, "Low RF-Complexity technologies to enable millimeter-Wave MIMO with large antenna array for 5G wireless communications," *IEEE Commun. Mag.*, vol. 56, no. 4, pp. 211–217, April 2018.

[32] B. Rivet, L. Girin, and C. Jutten, "Log-Rayleigh distribution: A simple and efficient statistical representation of log-spectral coefficients," *IEEE Trans. Speech Audio Process.*, vol. 15, no. 3, pp. 796–802, March 2007.

[33] H. Yang and T. L. Marzetta, "Performance of conjugate and Zero-Forcing beamforming in large-scale antenna systems", *IEEE J. Sel. Areas in Commun.*, vol. 31, no. 2, pp. 172–179, Feb. 2013.

[34] T. L. Marzetta et al., *Fundamentals of massive MIMO*, Cambridge University Press, Nov. 2016.

[35] M. -. Alouini and M. K. Simon, "An MGF-based performance analysis of generalized selection combining over Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 401–415, March 2000.

[36] A.M. Tulino, S. Verdù, "Random matrix theory and wireless communications", in *Foundations and Trends in Communications and Information Theory*, vol. 1, no. 1, pp. 1–182, 2004.

[37] A. Zanella, M. Chiani, and M. Z. Win, "On the marginal distribution of the eigenvalues of Wishart matrices," *IEEE Trans. Commun.*, vol. 57, no. 4, pp. 1050–1060, April 2009.

[38] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO Systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, March 2014.

[39] A. Dax, "From Eigenvalues to Singular Values: A Review," *Advances in Pure Mathematics*, Vol. 3 No. 9B, 2013, pp. 8–24.

[40] M. Ahsanullah, V. B. Nevzorov, and M. Shakil, "Distributions of order statistics," in *An introduction to order statistics*. Atlantis Press, 2013.

[41] M. T. Ivrlac and J. A. Nossek, "Transmit and receive array gain of uniform linear arrays of isotrops," 2009 *IEEE Sarnoff Symposium*, Princeton, NJ, 2009, pp. 1–6.

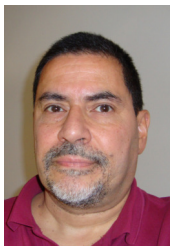
[42] Taesang Yoo and A. Goldsmith, "Capacity and power allocation for fading MIMO channels with channel estimation error," *IEEE Trans. Inform. Theory*, vol. 52, no. 5, pp. 2203–2214, May 2006.

[43] Zhengyuan Xu, "Perturbation analysis for subspace decomposition with applications in subspace-based algorithms," *IEEE Trans. Signal Process.*, vol. 50, no. 11, pp. 2820–2830, Nov. 2002.

[44] R. Janaswamy, "Angle and time of arrival statistics for the Gaussian scatter density model," *IEEE Trans. Wireless Commun.*, vol. 1, no. 3, pp. 488–497, July 2002.

[45] C. Martin and B. Ottersten, "Asymptotic eigenvalue distributions and capacity for MIMO channels under correlated fading," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1350–1359, July 2004.

[46] Wolfram Alpha LLC. 2018. Wolfram|Alpha. www.wolframalpha.com (access December, 2018).



Roberto Corvaja Roberto Corvaja (M98 SM17) received the Ph.D. degree in Electronic Engineering and Telecommunications at the University of Padova (Italy) in September 1994. Since 1994 he is with the Department of Information Engineering of the University of Padova as assistant professor in Telecommunications. He was visiting researcher at Hewlett-Packard Labs in Bristol, UK, and visiting professor at the University Carlos III of Madrid, Spain. His research activity covered several aspects of digital communications, from coherent optical communications systems, to wireless communications and MIMO-OFDM wireless systems. In the last years his research has included massive MIMO systems with hybrid schemes. Recently his research has included also quantum communications and he took part to several projects on free-space classical and quantum optical communications. He is co-author of two communications books and of several journal and conference papers. He served as TPC member for IEEE conferences and as reviewer for different IEEE journals. He is currently associate Editor of the IEEE Open Journal of the Communications Society.

Ana García Armada Ana García Armada (S96 A98 M00 SM08) received the Ph.D. degree in electrical engineering from the Polytechnical University of Madrid in February 1998. She is currently a Professor at University Carlos III of Madrid, Spain. She is leading the Communications Research Group at this university. She has been visiting scholar at Stanford University, Bell Labs and University of Southampton. She has participated (and coordinated most of them) in more than 30 national and 10 international research projects as well as 20 contracts with the industry, all of them related to wireless communications. She is the co-author of eight book chapters on wireless communications and signal processing. She has published around 150 papers in international journals and conference proceedings and she holds four patents. She has contributed to international standards organizations, such as ITU and ETSI, and is member of the expert group of the European 5G PPP and member of the advisory committee 5JAC of the ESA as expert appointed by Spain on 5G. She has served on the editorial boards of Physical Communication (2008–2017), IET Communications (2014–2017). She has been serving on the editorial board of IEEE Communications Letters since 2016 (Editor until Feb 2019, Senior Editor from Mar 2019, Exemplary Editor Award 2017 and 2018) and IEEE Transactions on Communications since 2019. She has served on the TPC of more than 40 conferences and she has been member of the organizing committee of IEEE Globecom 2019, IEEE Vehicular Technology Conference (VTC) Fall 2018, Spring 2018 and 2019 and IEEE 5G Summit 2017, among others. She will be the General Chair of Globecom 2021. She was the Newsletter Editor of the IEEE ComSoc Signal Processing and Consumer Electronics Committee (2017–2018) and is now the Secretary of this committee (since 2019). She has been the Secretary of the IEEE ComSoc Women in Communications Engineering Standing Committee (2016–2017) and the Chair of this committee (2018–2019). She has received the Young Researchers Excellence Award, the Award to Outstanding achievement in research, teaching and management and the Award to Best Practices in Teaching, all from University Carlos III of Madrid, She was awarded the third place Bell Labs Prize 2014 for shaping the future of information and communications technology. She received the Outstanding service award from the IEEE ComSoc Signal Processing and Communications Electronics (SPCE) technical committee in 2019. Her main interests are multi-carrier and multi-antenna techniques and signal processing applied to wireless communications.