Performance of the Massive MIMO Uplink with OFDM and Phase Noise

Antonios Pitarokoilis, Emil Björnson and Erik G. Larsson

Linköping University Post Print

N.B.: When citing this work, cite the original article.

Performance of the Massive MIMO Uplink with OFDM and Phase Noise

Antonios Pitarokoilis, *Student Member, IEEE*, Emil Björnson, *Member, IEEE*, and Erik G. Larsson, *Fellow, IEEE*

Abstract—The performance of multi-user Massive MIMO-OFDM uplink systems in the presence of base station (BS) phase noise impairments is investigated. Closed-form achievable rate expressions are rigorously derived under two different operations, namely the case of a common oscillator (synchronous operation) at the BS and the case of independent oscillators at each BS antenna (non-synchronous operation). It is observed that the non-synchronous operation exhibits superior performance due to the averaging of a portion of the intercarrier interference. Further, radiated power scaling laws are derived, which are identical to the phase-noise-free case.

Index Terms—Phase Noise, Communication Systems, MIMO Systems.

I. INTRODUCTION

The demand for more data traffic in cellular systems has been the driving force for research in the field during the last decades. Base station (BS) densification and allocation of more bandwidth have been pivotal to handling more traffic, but this is expected to change in the future. Already, BSs with multiple antennas have been deployed to increase the performance in terms of energy and spectral efficiency. Recently, *Massive multiple-input multiple-output (MIMO)* systems [1], i.e., systems where the BS is equipped with an unprecedentedly large number of antennas and communicates with a few tens of non-cooperative users over the same time and frequency resources, has become one of the most prominent candidates for the evolution of cellular systems.

The realization of affordable Massive MIMO requires the use of a large number of inexpensive, and potentially low-quality, components. Fortunately, it has been shown that Massive MIMO systems are, in general, robust against hardware impairments. One of the most important hardware impairments is phase noise (PN), which is a random and varying phase rotation of the information signal. PN arises in coherent communication systems due to the various noise sources in the circuits of the local oscillators (LOs) that are used for the modulation of the information signal from the baseband to passband at the transmitter and vice versa at the receiver.

The effect of PN has been extensively studied under various systems models. The effect of PN in Massive MIMO systems has also been recently investigated for single-carrier systems [2], [3]. However, most contemporary wireless systems employ orthogonal frequency division multiplexing (OFDM), i.e., they are multi-carrier. PN destroys the orthogonality of subcarriers in OFDM and, therefore, degrades significantly the system performance [4]. PN-impaired Massive MIMO-OFDM systems

appear to be less well understood. To the authors' knowledge, the only available study of the problem is [5]. In [5], rate expressions for a single-user Massive MIMO-OFDM uplink were given that are valid under specific assumptions and in the asymptotic limit when the number of base station antennas goes to infinity. In the present work, we revisit the PN-impaired Massive MIMO-OFDM uplink, but for a multi-user system with autonomous, spatially multiplexed terminals, and develop rigorous lower bounds on ergodic capacity that are valid for any finite number of antennas.

The contributions of this work can be summarized as follows. First, the frequency-domain system model definition is based on the time-domain description of the channel, which gives a parsimonious description of the problem under study. Further, achievable rates are rigorously derived under the assumptions of perfect CSI at the BS and MRC processing. Closed-form expressions for the proposed achievable rates are provided, which hold for any finite number of BS antennas. Two distinct LO operations are considered, i.e., the synchronous operation and the non-synchronous operation, and compared. Based on the derived closed-form expressions, radiated power scaling laws are derived and compared with previous studies on single-carrier systems.

II. SYSTEM MODEL

A BS equipped with M antenna elements communicates in the uplink with K non-cooperative single-antenna users. The complex baseband equivalent channel impulse response (CIR) of each user to the BS is assumed to be frequency-selective and is modeled as a finite impulse response (FIR) filter with L sample-spaced taps, where T_s is the sampling period. The l-th channel gain of the k-th user to the m-th antenna element is noted as $g_{m,k,l} = \sqrt{d_{k,l}}h_{m,k,l}$. The coefficients $h_{m,k,l} \sim \mathcal{N}_{\mathbb{C}}(0,1)$ are i.i.d. circularly symmetric complex Gaussian random variables and correspond to the small scale fading. The sequence of positive reals $d_{k,0}, \ldots, d_{k,L-1}$ corresponds to the power delay profile (PDP) of the channel from user k to the BS antenna array and it is the same to all the BS antenna elements. The PDPs are assumed to be constant and known and the following condition holds

$$\mathbb{E}\left[\left|\sum_{l=0}^{L-1} g_{m,k,l}\right|^2\right] = \sum_{l=0}^{L-1} d_{k,l} = \beta_k, \qquad (1)$$

so that the total energy contained in each channel impulse response is independent of L. The fast fading coefficients are assumed to remain constant for a time interval T_{coh} and then they change to an independent realization. OFDM transmission
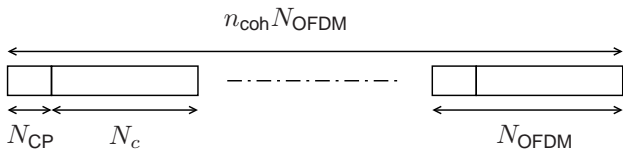
Fig. 1: The frequency-selective channel remains constant for a coherence time of n_{coh} OFDM symbols.

is used, where N_c is the number of subcarriers, N_{CP} is the length of the cyclic prefix (CP) in samples. Hence, an OFDM symbol is $N_{\text{OFDM}} = N_c + N_{\text{CP}}$ samples long. It is assumed that n_{coh} OFDM symbols span the interval $T_{\text{coh}} = n_{\text{coh}} N_{\text{OFDM}} T_s$.

The received signal is impaired by PN at the receiver, $\phi_m[i]$. Free-running LOs are assumed, hence the PN realizations are generated by Wiener processes, i.e., the PN realization at the m-th BS antenna at time i is

$$\phi_m[i] = \phi_m[i-1] + \sigma_\phi w_m[i], \quad (2)$$

where $w_m[i] \sim \mathcal{N}_{\mathbb{R}}(0,1)$ is a sequence of independent and identically distributed real Gaussian random variables. The Wiener model is well-established for free-running oscillators [6] and has been extensively used in the communications literature [7], [8]. The constant σ_ϕ^2 is the variance of the increments of the PN processes, which is a measure of the quality of the LOs ($\sigma_\phi^2 = 0$ for ideal LOs). Two distinct operations are considered, namely, the case where the PN processes are independent among the BS antennas (non-synchronous operation) and the case where the same PN process impairs all the BS antennas (synchronous operation), i.e., $w_1[i] \equiv \cdots \equiv w_M[i]$. The non-synchronous operation corresponds to a distributed deployment, where each BS antenna element is equipped with a different LO and the synchronous operation corresponds to a centralized scenario, where the same LO is used for all the BS antennas.

The received signal at time i at the m-th BS antenna, when the time domain information symbol sequences $x_k[i - L + 1], \ldots, x_k[i]$ are sent, is given by

$$y_m[i] = \sqrt{\rho} \sum_{k=1}^{K} \sum_{l=0}^{L-1} \sqrt{d_{k,l}} h_{m,k,l} e^{-j\phi_m[i]} x_k[i-l] + z_m[i],$$

where $z_m[i] \sim \mathcal{N}_{\mathbb{C}}(0,1)$ additive white Gaussian noise (AWGN). The time-domain samples received at the BS within the n-th OFDM symbol, $1 \le n \le n_{\text{coh}}$, after the CP is removed are given by

$$\boldsymbol{y}(n) = \sqrt{\rho} \sum_{k=1}^{K} \boldsymbol{\Phi}(n) \boldsymbol{G}_k \boldsymbol{x}_k(n) + \boldsymbol{z}(n), \quad (3)$$

where $\boldsymbol{y}(n) \triangleq [\boldsymbol{y}[n,1]^T, \cdots, \boldsymbol{y}[n,N_c]^T]^T \in \mathbb{C}^{MN_c}$,

$$\boldsymbol{y}[n,i] \triangleq \begin{bmatrix} y_1[nN_{\text{OFDM}} - N_c + i] \\ \vdots \\ y_M[nN_{\text{OFDM}} - N_c + i] \end{bmatrix} \in \mathbb{C}^M \quad \text{and}$$

$$\boldsymbol{x}_k(n) \triangleq [x_k[nN_{\text{OFDM}} - N_c + 1], \cdots, x_k[nN_{\text{OFDM}}]]^T \in \mathbb{C}^{N_c}.$$

The matrix $\boldsymbol{G}_k$ is block circulant with each block being a column vector of length M, corresponding to the time-domain CIRs of the user k to the BS. In particular, the (ν_1, ν_2) M-column vector of $\boldsymbol{G}_k$ is given by

$$[\boldsymbol{G}_k]_{\nu_1, \nu_2} = \begin{cases} \boldsymbol{g}_{k, (\nu_1 - \nu_2) \mod N_c}, \\ \quad \text{if } 0 \le (\nu_1 - \nu_2) \mod N_c \le L - 1 \\ \boldsymbol{0}, \text{ otherwise.} \end{cases}$$

Finally, $\boldsymbol{\Phi}(n) = \text{blkdiag}\{\boldsymbol{\Phi}[n,1], \cdots, \boldsymbol{\Phi}[n,N_c]\}$ and $\boldsymbol{\Phi}[n,i] = \text{diag}\{e^{-j\phi_1[nN_{\text{OFDM}} - N_c + i]}, \cdots, e^{-j\phi_M[nN_{\text{OFDM}} - N_c + i]}\}$. The frequency-domain received signal is given by

$$\mathbf{y}(n) = \sqrt{\rho} \sum_{k=1}^{K} \mathbf{G}_k(n) \mathbf{x}_k(n) + \mathbf{z}(n), \quad (4)$$

where $\mathbf{y}(n) \triangleq (\boldsymbol{F}_{N_c} \otimes \boldsymbol{I}_M) \boldsymbol{y}(n)$, $\mathbf{x}_k(n) \triangleq \boldsymbol{F}_{N_c} \boldsymbol{x}_k(n)$, $\mathbf{z}(n) \triangleq (\boldsymbol{F}_{N_c} \otimes \boldsymbol{I}_M) \boldsymbol{z}(n)$ are the frequency-domain counterparts of the corresponding time-domain quantities and $\mathbf{G}_k(n) \triangleq (\boldsymbol{F}_{N_c} \otimes \boldsymbol{I}_M) \boldsymbol{\Phi}(n) \boldsymbol{G}_k \boldsymbol{F}_{N_c}^H$ is the effective frequency-domain channel matrix for user k. By $\boldsymbol{F}_{N_c}$ we denote the unitary DFT matrix of size N_c and $\otimes$ is the Kronecker product. Observe that due to the presence of $\boldsymbol{\Phi}(n)$, the product $\boldsymbol{\Phi}(n) \boldsymbol{G}_k$ is no longer block-circulant and consequently cannot be diagonalized by the DFT. The non-zero elements that appear on the off-diagonal blocks of $\mathbf{G}_k(n)$ are the channel gains of the intercarrier interference.

III. ACHIEVABLE RATES

The physical propagation channel from user k, which we denote[1] as $\mathbf{G}_k = (\boldsymbol{F}_{N_c} \otimes \boldsymbol{I}_M) \boldsymbol{G}_k \boldsymbol{F}_{N_c}^H = \text{blkdiag}\{\mathbf{g}_{k,1}, \ldots, \mathbf{g}_{k,N_c}\}$, is assumed to be known perfectly at the BS. The vector $\mathbf{g}_{k,p} \in \mathbb{C}^{M \times 1}$ is the frequency response of the propagation channel of user k at the p-th subcarrier. With maximum ratio combining (MRC) the detected information vector for user k, $\hat{\mathbf{x}}_k(n)$ is given by

$$\hat{\mathbf{x}}_k(n) = \mathbf{G}_k^H \mathbf{y}(n) = \sqrt{\rho} \boldsymbol{\mathcal{G}}_k(n) \mathbf{x}_k(n) + \mathsf{EN}_k(n), \quad (5)$$

where $\boldsymbol{\mathcal{G}}_k(n) \triangleq \mathbb{E}\left[\mathbf{G}_k^H \mathbf{G}_k(n)\right]$ and

$$\mathsf{EN}_k(n) = \sqrt{\rho} \left(\mathbf{G}_k^H \mathbf{G}_k(n) - \boldsymbol{\mathcal{G}}_k(n)\right) \mathbf{x}_k(n)$$
$$+ \sqrt{\rho} \sum_{p=1, p \ne k}^{K} \mathbf{G}_k^H \mathbf{G}_p(n) \mathbf{x}_p(n) + \mathbf{G}_k^H \mathbf{z}(n). \quad (6)$$

An achievable rate, $R_k(n)$, for the k-th user at the n-th OFDM symbol is given by Proposition 1.

Proposition 1: With $\mathbf{x}_k(n) \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{I}_M)$, an achievable rate for the k-th user at the n-th OFDM symbol with MRC processing when the propagation channel $\mathbf{G}_k$ is available at the receiver is given by

$$R_k(n) = \log_2 \left| \boldsymbol{I}_{N_c} + \rho \boldsymbol{\mathcal{G}}_k^H(n) \mathbf{C}_k^{-1}(n) \boldsymbol{\mathcal{G}}_k(n) \right|, \quad (7)$$

where $\mathbf{C}_k(n)$ is the covariance matrix of $\mathsf{EN}_k(n)$.

Proof: See Appendix. ∎

Remark 1: We note that the derived achievable rate is a function of the OFDM symbol index, n. Observe that the effective channel matrix, $\boldsymbol{\mathcal{G}}_k(n)$, and the effective noise statistics are also functions of n. Hence, the rate in Proposition 1 for the n-th OFDM symbol can be achieved by using a dedicated, for the n-th OFDM symbol, Gaussian codebook

---

[1] We stress the notational difference between the frequency-domain physical propagation channel, $\mathbf{G}_k$, and the effective frequency-domain channel matrix with PN, $\mathbf{G}_k(n)$, at the n-th OFDM symbol for user k.

with codewords that span over multiple coherence intervals of the channel fading. Similar coding strategies have earlier been proposed in [2].

Definition 1: The effective, over the coherence interval, net rate per subcarrier for the k-th user, $\hat{R}_k$, is defined as

$$\hat{R}_k \triangleq \frac{1}{n_{\text{coh}}} \sum_{n=1}^{n_{\text{coh}}} \hat{R}_k(n) = \frac{1}{n_{\text{coh}}} \sum_{n=1}^{n_{\text{coh}}} \frac{1}{N_c} \frac{N_c}{N_{\text{OFDM}}} R_k(n), \quad (8)$$

where $\hat{R}_k(n)$ is the effective net rate per subcarrier at the n-th OFDM symbol. The fraction $\frac{N_c}{N_{\text{OFDM}}}$ accounts for the rate loss due to the CP and $\frac{1}{N_c}$ is the per-subcarrier normalization. In the following we provide closed-form expressions for the achievable rates in Proposition 1 for the two different operations: synchronous and non-synchronous. The expressions are derived by straightforward algebraic manipulations.

Proposition 2: For both operations the effective channel matrix $\boldsymbol{\mathcal{G}}_k(n)$ is given by $\boldsymbol{\mathcal{G}}_k(n) = M N_c \boldsymbol{\Psi}(n)$, where

$$\boldsymbol{\Psi}(n) \triangleq \left(\boldsymbol{F}_{N_c} \bar{\boldsymbol{\Phi}}(n) \boldsymbol{F}_{N_c}^H \right) \circ \left(\boldsymbol{F}_{N_c} \boldsymbol{D}_k \boldsymbol{F}_{N_c}^H \right),$$

$$\bar{\boldsymbol{\Phi}}(n) \triangleq \text{diag} \left\{ e^{-\frac{\sigma_\phi^2}{2}|nN_{\text{OFDM}} - N_c + 1|}, \ldots, e^{-\frac{\sigma_\phi^2}{2}|nN_{\text{OFDM}}|} \right\},$$

$\boldsymbol{D}_k = \text{diag} \{ d_{k,0}, \ldots, d_{k,L-1}, \boldsymbol{0} \}$ and $\circ$ is the element-wise Hadamard product.

Proposition 3: The covariance matrix of the effective noise is given by

$$\mathbf{C}_k^\times(n) = \rho M \mathbf{C}_{1,k} + \rho M(M-1) \mathbf{C}_{2,k}^\times(n)$$
$$- \rho \boldsymbol{\mathcal{G}}_k(n) \boldsymbol{\mathcal{G}}_k^H(n) + M \beta_k \left(\rho \sum_{q \neq k}^K \beta_q + 1 \right) \boldsymbol{I}_{N_c}, \quad (9)$$

where $\times = s$ for synchronous operation and $\times = ns$ for non-synchronous operation,

$$[\mathbf{C}_{1,k}]_{\nu_1,\nu_2} = \boldsymbol{w}_{\nu_1}^H \left(\mathbf{D}_{k,\nu_1}^H \mathbf{D}_{k,\nu_2} \otimes \tilde{\boldsymbol{\Phi}} \right) \boldsymbol{w}_{\nu_2}$$
$$+ \frac{\beta_k \mathrm{d}_{\nu_2 - \nu_1}^*}{N_c} \boldsymbol{w}_{\nu_1}^H \left(\boldsymbol{I}_{N_c} \otimes \tilde{\boldsymbol{\Phi}} \right) \boldsymbol{w}_{\nu_2},$$

$$[\mathbf{C}_{2,k}^s(n)]_{\nu_1,\nu_2} = \boldsymbol{w}_{\nu_1}^H \left(\mathbf{D}_{k,\nu_1}^H \mathbf{D}_{k,\nu_2} \otimes \tilde{\boldsymbol{\Phi}} \right) \boldsymbol{w}_{\nu_2},$$

$$[\mathbf{C}_{2,k}^{ns}(n)]_{\nu_1,\nu_2} = N_c^2 \left[\boldsymbol{F}_{N_c} \bar{\boldsymbol{\Phi}}(n) \boldsymbol{F}_{N_c}^H \mathbf{D}_{k,\nu_1}^H \mathbf{D}_{k,\nu_2} \boldsymbol{F}_{N_c} \bar{\boldsymbol{\Phi}}(n) \boldsymbol{F}_{N_c}^H \right]_{\nu_1,\nu_2},$$

$$\boldsymbol{w}_{\nu_3,\nu_2} \triangleq \left[1 \quad e^{-j\frac{2\pi(\nu_3 - \nu_2)}{N_c}} \quad \cdots \quad e^{-j\frac{2\pi(\nu_3 - \nu_2)(N_c-1)}{N_c}} \right]^T,$$

$$\boldsymbol{w}_{\nu_2} \triangleq \left[\boldsymbol{w}_{1,\nu_2}^T \quad \cdots \quad \boldsymbol{w}_{N_c,\nu_2}^T \right]^T, \quad [\tilde{\boldsymbol{\Phi}}]_{\nu_4,\nu_5} = e^{-\frac{\sigma_\phi^2}{2}|\nu_4 - \nu_5|},$$

$$\mathbf{D}_{k,\nu_1} \triangleq \text{diag} \{ \mathrm{d}_{k,1-\nu_1}, \quad \cdots, \quad \mathrm{d}_{k,N_c-\nu_1} \} \text{ and}$$
$$\mathrm{d}_{k,\nu_2 - \nu_1} \triangleq \left[\boldsymbol{F}_{N_c} \boldsymbol{D}_k \boldsymbol{F}_{N_c}^H \right]_{\nu_1,\nu_2}.$$

Remark 2: The matrix $\mathbf{C}_2^\times(n)$ corresponds to part of the interference caused by the uncertainty of the effective channel gain, $\boldsymbol{\mathcal{G}}_k(n)$, due to PN. The elements of the matrix $\mathbf{C}_2^s(n)$ are not a function of n, however, the elements of the matrix $\mathbf{C}_2^{ns}(n)$ have magnitudes which decay with n and N_{OFDM} as $e^{-\sigma_\phi^2 n N_{\text{OFDM}}}$. This implies that part of the interference in the non-synchronous operation averages out, whereas the same part in the synchronous operation does not. Hence, the achievable rates are higher for the non-synchronous operation than the synchronous. Similar observations were drawn in other studies on single-carrier systems [2], [3].

Corollary 1 can be derived from Propositions 2 and 3 in the absence of PN.

Corollary 1: In the absence of PN the effective channel matrix is given by $\boldsymbol{\mathcal{G}}_k(n) = M\beta_k \boldsymbol{I}_{N_c}$ and the covariance of the effective noise is $\mathbf{C}_k = M\beta_k \left(\rho \sum_{q=1}^K \beta_q + 1 \right) \boldsymbol{I}_{N_c}$. Consequently, the rate for the k-th user at the n-th OFDM symbol is given by

$$R_k^{np}(n) \geq N_c \log_2 \left(1 + \frac{\rho M \beta_k}{\rho \sum_{q=1}^K \beta_q + 1} \right), \quad \forall n. \quad (10)$$

Proposition 4: For a fixed desired achievable rate, it is possible to reduce the total radiated power by 3 dB for every doubling of the number of BS antennas, M, i.e., *an $O(M)$ radiated power gain is achievable in the presence of PN.*

Proof: Let $\rho = E_u/M$. Then, as $M \to \infty$, the rate converges to the limiting value

$$R_k^{\times,\lim}(n) = \log \left| \boldsymbol{I}_{N_c} + E_u \boldsymbol{\Psi}^H(n) \left(\mathbf{C}_k^{\times,\lim}(n) \right)^{-1} \boldsymbol{\Psi}(n) \right|,$$

where $\mathbf{C}_k^{\times,\lim}(n) \triangleq E_u \left(\mathbf{C}_2^\times(n) - N_c^2 \boldsymbol{\Psi}(n) \boldsymbol{\Psi}^H(n) \right) + \beta_k \boldsymbol{I}_{N_c}$. ∎ The scaling law given by Proposition 4 is the same as the one that holds in the PN-free case [9].
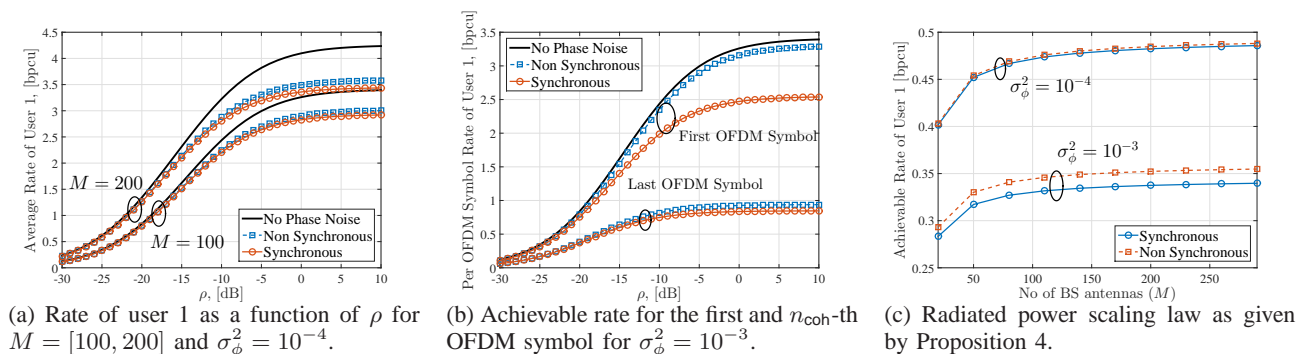
IV. NUMERICAL EXAMPLES

In this section we present some numerical examples that highlight the degradation introduced by PN at Massive MIMO-OFDM systems. In all the examples the following parameters are held constant $N_c = 128$, $N_{\text{CP}} = 16$, $N_{\text{OFDM}} = 144$, $n_{\text{coh}} = 5$, $L = 17$, $K = 10$. The PDPs of all users are assumed to be exponential with $d_{k,l} = \beta_k e^{-\alpha_k l}/\zeta_k$, where $\boldsymbol{\beta} = [\beta_1, \ldots, \beta_K]^T = [0.9940, 0.5852, 0.6289, 0.6984, 0.5370, 0.8420, 0.7012, 0.9914, 0.7011, 0.8103]^T$, $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_K]^T = [1.0557, 0.6844, 0.7120, 0.5773, 1.4138, 1.2067, 1.0578, 0.8134, 0.6662, 1.1225]^T$ and ζ_k is a normalization factor such that (1) holds. The vectors $\boldsymbol{\beta}$, $\boldsymbol{\alpha}$ were generated from uniform distributions $\mathcal{U}[0.5, 1]$ and $\mathcal{U}[0.5, 1.5]$ and kept fixed for the generation of the figures.

In Fig. 2a the effective net rate per subcarrier of user 1, $\hat{R}_k$, given in Definition 1, is shown for $M = [100, 200]$ and $\sigma_\phi^2 = 10^{-4}$. It is clear that the achievable rates increase with M and the performance of non-synchronous operation is superior to the synchronous. This can be attributed to the fact that the part of the interference, $\mathbf{C}_2^{ns}(n)$, due to the uncertainty of the effective channel matrix, $\boldsymbol{\mathcal{G}}_k(n)$, in the non-synchronous operation averages out, as noted in Remark 2. The PN-free bound from (10) is also plotted as a benchmark.

In Fig. 2b the effective net rate per subcarrier of user 1, $\hat{R}_k(n)$, is plotted for the first and the last OFDM symbol of the coherence interval for $\sigma_\phi^2 = 10^{-3}$. The progressive degradation of the performance is attributed to the partial loss of coherency between the physical channel, $\mathbf{G}_k$, and the effective channel, $\mathbf{G}_k(n)$, due to the evolution of the PN processes. It is further observed that the performance difference is more pronounced at the first OFDM symbol rather than the last. In the initial part of the coherence interval and the effective channel has not drifted significantly from the propagation channel and the effective channel gain is relatively

Fig. 2: Numerical Examples

(a) Rate of user 1 as a function of ρ for $M = [100, 200]$ and $\sigma_\phi^2 = 10^{-4}$.

(b) Achievable rate for the first and n_{coh}-th OFDM symbol for $\sigma_\phi^2 = 10^{-3}$.

(c) Radiated power scaling law as given by Proposition 4.

high. In addition, the averaging behavior of a part of the interference in the non-synchronous operation is significant already in the first OFDM symbol. Hence, the difference in achievable rate performance between the synchronous and non-synchronous operation is substantial. On the other hand, at the last OFDM symbol the drift due to PN is significant and the effective channel gain is greatly reduced for both operations. In this case the averaging of part of the interference in the non-synchronous operation does not yield the same gains in performance as in the first OFDM symbol.

In Fig. 2c the effective net rate per subcarrier of user 1 is plotted as a function of M. However, in this case $\rho = E_u/M$, where $E_u = 0$ dB is constant. The rate in all cases saturates at a non-zero limiting value. This establishes the claim in Proposition 4.

V. CONCLUSION

The effect of PN in Massive MIMO-OFDM uplink channels is studied and closed-form achievable rates are rigorously derived for two different operations: synchronous and non-synchronous LOs in the array. It is shown that the non-synchronous operation is superior due to the averaging of the interference due to effective channel uncertainty. A progressive degradation of the achievable rate performance is observed and a radiated power scaling law is provided. The behavior observed here resembles the performance observed in earlier single-carrier studies [2], [3]. The in-depth study of the same system model under PN at the users and with estimated CIRs is of particular interest. However, it increases substantially the complexity and the notation of the problem and, therefore, will be part of the future work based on this initial study.

APPENDIX

We follow the arguments of [10] to justify (7). The mutual information is given by

$$I\left(\mathbf{x}_k(n); \hat{\mathbf{x}}_k(n)\right) = \mathsf{h}\left(\mathbf{x}_k(n)\right) - \mathsf{h}\left(\mathbf{x}_k(n)|\hat{\mathbf{x}}_k(n)\right) \quad (11)$$

The entropy $\mathsf{h}\left(\mathbf{x}_k(n)|\hat{\mathbf{x}}_k(n)\right)$ can be upper bounded by

$$\mathsf{h}\left(\mathbf{x}_k(n)|\hat{\mathbf{x}}_k(n)\right) \leq \mathbb{E}\left[\log\left|\pi e\mathsf{COV}\left(\mathbf{x}_k(n)|\hat{\mathbf{x}}_k(n)\right)\right|\right].$$

By the minimum covariance property we have

$$\mathsf{COV}\left(\mathbf{x}_k(n)|\hat{\mathbf{x}}_k(n)\right) - \mathbb{E}\left[\bar{\mathbf{x}}_k(n)\bar{\mathbf{x}}_k^H(n)\right] \preceq \mathbf{0},$$

where $\bar{\mathbf{x}}_k(n) \triangleq \mathbf{x}_k(n) - \tilde{\mathbf{x}}_k(n)$ and $\tilde{\mathbf{x}}_k(n)$ is any estimate of $\mathbf{x}_k(n)$, which is a function of $\hat{\mathbf{x}}_k(n)$ and $\boldsymbol{\mathcal{G}}_k(n)$. We select $\tilde{\mathbf{x}}_k(n)$ to be the linear MMSE estimate of $\mathbf{x}_k(n)$, i.e.

$$\tilde{\mathbf{x}}_k(n) = \mathbb{E}\left[\mathbf{x}_k(n)\hat{\mathbf{x}}_k^H(n)\right]\left(\mathbb{E}\left[\hat{\mathbf{x}}_k(n)\hat{\mathbf{x}}_k^H(n)\right]\right)^{-1}\hat{\mathbf{x}}_k(n)$$

$$= \sqrt{\rho}\boldsymbol{\mathcal{G}}_k^H(n)\left(\mathbf{C}_k(n) + \sqrt{\rho}\boldsymbol{\mathcal{G}}_k(n)\boldsymbol{\mathcal{G}}_k^H(n)\sqrt{\rho}\right)^{-1}\hat{\mathbf{x}}_k(n).$$

The associated estimation error covariance is given by

$$\mathsf{COV}\left(\mathbf{x}_k(n)|\tilde{\mathbf{x}}_k(n)\right) = \left(\boldsymbol{I}_{N_c} + \rho\boldsymbol{\mathcal{G}}_k^H(n)\mathbf{C}_k^{-1}(n)\boldsymbol{\mathcal{G}}_k(n)\right)^{-1}.$$

Hence, an upper bound on the entropy $\mathsf{h}\left(\mathbf{x}_k(n)|\hat{\mathbf{x}}_k(n)\right)$ is

$$\mathsf{h}\left(\mathbf{x}_k(n)|\hat{\mathbf{x}}_k(n)\right) \leq \log\left|\pi e\mathsf{COV}\left(\mathbf{x}_k(n)|\tilde{\mathbf{x}}_k(n)\right)\right|$$

$$= \log\left|\pi e\left(\boldsymbol{I}_{N_c} + \rho\boldsymbol{\mathcal{G}}_k^H(n)\mathbf{C}_k^{-1}(n)\boldsymbol{\mathcal{G}}_k(n)\right)^{-1}\right|. \quad (12)$$

The result in (7) follows from (11) with $\mathbf{x}_k(n) \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{I}_{N_c})$ and (12).

REFERENCES

[1] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, pp. 3590–3600, Nov. 2010.

[2] A. Pitarokoilis, S. K. Mohammed, and E. G. Larsson, "Uplink performance of time-reversal MRC in massive MIMO systems subject to phase noise," *IEEE Trans. Wireless Commun.*, vol. 14, pp. 711–723, Feb. 2015.

[3] E. Björnson, M. Matthaiou, and M. Debbah, "Massive MIMO with nonideal arbitrary arrays: Hardware scaling laws and circuit-aware design," *IEEE Trans. Wireless Commun.*, vol. 14, pp. 4353–4368, Aug. 2015.

[4] T. Pollet, M. Van Bladel, and M. Moeneclaey, "BER sensitivity of OFDM systems to carrier frequency offset and wiener phase noise," *IEEE Trans. Commun.*, vol. 43, pp. 191–193, Feb/Mar/Apr 1995.

[5] R. Krishnan *et al.*, "On the impact of oscillator phase noise on the uplink performance in a massive MIMO-OFDM system," *CoRR*, vol. abs/1405.0669, 2014.

[6] A. Demir, A. Mehrotra, and J. Roychowdhury, "Phase noise in oscillators: a unifying theory and numerical methods for characterization," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 47, pp. 655–674, May 2000.

[7] L. Tomba, "On the effect of Wiener phase noise in OFDM systems," *IEEE Trans. Commun.*, vol. 46, pp. 580–583, May 1998.

[8] D. Petrovic, W. Rave, and G. Fettweis, "Effects of phase noise on OFDM systems with and without PLL: Characterization and compensation," *IEEE Trans. Commun.*, vol. 55, pp. 1607 –1616, Aug. 2007.

[9] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, pp. 1436–1449, Apr. 2013.

[10] B. Hassibi and B. Hochwald, "How much training is needed in multiple-antenna wireless links?," *IEEE Trans. Inf. Theory*, vol. 49, pp. 951 – 963, Apr. 2003.