Uplink Performance of Time-Reversal MRC in Massive MIMO Systems Subject to Phase Noise

Antonios Pitarokoilis, Saif Khan Mohammed and Erik G. Larsson

Abstract

Multi-user multiple-input multiple-output (MU-MIMO) cellular systems with an excess of base station (BS) antennas (Massive MIMO) offer unprecedented multiplexing gains and radiated energy efficiency. Oscillator phase noise is introduced in the transmitter and receiver radio frequency chains and severely degrades the performance of communication systems. We study the effect of oscillator phase noise in frequency-selective Massive MIMO systems with imperfect channel state information (CSI). In particular, we consider two distinct operation modes, namely when the phase noise processes at the BS antennas are identical (synchronous operation) and when they are independent (non-synchronous operation). We analyze a linear and low-complexity time-reversal maximum-ratio combining (TR-MRC) reception strategy. For both operation modes we derive a lower bound on the sum-capacity and we compare the performance of the two modes. Based on the derived achievable sum-rate, we show that with the proposed receive processing an $O(\sqrt{M})$ array gain is achievable. Due to the phase noise drift the estimated effective channel becomes progressively outdated. Therefore, phase noise effectively limits the length of the interval used for data transmission and the number of scheduled users. The derived achievable rates provide insights into the optimum choice of the data interval length and the number of scheduled users.

Index Terms
I. INTRODUCTION

Multiple-input multiple-output (MIMO) technology has been shown to offer significant gains in the total throughput of wireless networks [2]. The large number of degrees of freedom enable many users to share the same time-frequency resources, paving the way for multi-user MIMO (MU-MIMO) systems [3]. Related to the increase in demand for higher data rates, is the increase in the power consumption of the base stations (BS). Therefore, it is required that future cellular systems provide increased spectral efficiency as well as improved energy efficiency.

MU-MIMO systems with an excess of BS antennas, termed as Massive MIMO, large MIMO or Large Scale Antenna Systems (LSAS), have recently attracted significant interest [4]–[6]. They promise a significant increase in the total cell throughput by means of simple signal processing. At the same time the radiated power can be scaled down with the number of BS antennas, $M$. More specifically, in [7] the authors show that in a MU-MIMO uplink with linear receivers and imperfect channel state information (CSI), by increasing the number of BS antennas from 1 to $M$, one can reduce the total transmit power by $O(\sqrt{M})$ while maintaining a fixed per-user information rate. The crucial assumption in Massive MIMO systems is that the squared Euclidean norm of the channel vector of each user grows as $O(M)$, whereas the inner products between channel vectors of different users grow at a lesser rate. This assumption can be justified in the MU-MIMO setting since the users are typically separated by many wavelengths, which implies that their channel vectors become asymptotically (in the number of BS antennas) orthogonal. Extensive measurements have confirmed the validity of this assumption [5].

Phase noise is inevitable in communication systems due to imperfections in the circuitry of the local oscillators that are used for the conversion of the baseband signal to passband and vice versa. To be specific, phase noise is the instantaneous drift of the phase of the carrier wave and results in a widening of the power spectral density of the generated waveform. Due to phase noise the channel is time-varying and accurate channel estimation becomes a challenge. Further, phase noise causes a partial loss of coherency between the acquired channel estimate and the true channel gains during data transmission. This can result in severe degradation of the system performance, especially when coherent signaling is used.
In MIMO an array power gain is obtained by coherently combining signals received by several antennas, using estimated channel responses. Since phase noise distorts the received data, it is crucial to examine its effect on the performance. Significant research work is available on phase noise. However, most of it is concerned with single-user single-antenna multi-carrier transmission, since multi-carrier transmission is more sensitive to phase noise compared to single-carrier transmission [8]. In [9] a method to calculate the bit-error-rate (BER) of a single-user orthogonal frequency division multiplexing (OFDM) system impaired with phase noise is provided. The authors in [10] study the signal-to-interference-and-noise-ratio (SINR) degradation in OFDM and propose a method to mitigate the effect of phase noise. In [11] a method to characterize phase noise in OFDM systems is developed and an algorithm to compensate for the degradation is described. Finally, in [12] the authors propose a method to jointly estimate the channel coefficients and the phase noise in a single-user MIMO system and, based on this method, they propose a phase noise mitigation algorithm.

From an information-theoretic point of view, the calculation of capacity of phase noise channels is particularly challenging. To the best of our knowledge, the exact capacity of typical phase noise-impaired channels under realistic models is not known. The behavior of the capacity of such channels is only known asymptotically for some cases in the high signal-to-noise-ratio (SNR) regime [13]. In [14] the authors derive a non-asymptotic upper bound on the capacity of a single-user deterministic MIMO channel impaired with Wiener phase noise, which is tight in the high-SNR regime. To the authors’ knowledge, we present the first analysis of the effect of phase noise in a multi-user multi-antenna scenario with imperfect channel state information where single-carrier transmission is used.

Specifically, in this work, we consider a single-cell frequency-selective MU-MIMO uplink, where a number of non-cooperative users transmit independent data streams to a base station with a large number of antennas. Since the channel is assumed to be unknown, CSI is acquired via uplink training. There are phase noise sources both at the transmitters and at the receiver. We consider and compare two distinct cases. In the first case, which is termed *synchronous operation* mode, the phase noise processes at the BS antennas are identical. In the second case, which is termed *non-synchronous operation* mode, the phase noise processes at the BS antennas are independent. These two operation modes correspond to the cases of a common phase reference versus independent phase references, respectively. A time-reversal maximum-ratio combining
(TR-MRC) strategy is proposed and achievable sum-rates are derived for both operation modes.

Based on the derived expressions of the achievable sum-rates, we show that for a fixed desired per-user information rate, by doubling the number of BS antennas, the total transmit power can be reduced by a factor of $\sqrt{2}$. We observe that the use of independent phase noise sources can yield higher sum-rate performance and we support this possibly counter-intuitive result by a simple toy example for which the exact capacity is calculated. Further, the achievable rate expressions reveal a fundamental trade-off between the length of the time interval spent on data transmission and the sum-rate performance. The rate expressions also provide valuable insight into the optimum number of scheduled users.

II. System Model

We consider a frequency-selective MU-MIMO uplink channel with $M$ BS antennas and $K$ single-antenna users. The channel between the $k$-th user and the $m$-th BS antenna is modeled as a finite impulse response (FIR) filter with $L$ equally spaced channel taps. The $l$-th channel tap is given by $g_{m,k,l} \triangleq \sqrt{d_{k,l}h_{m,k,l}}$, where $h_{m,k,l}$ and $d_{k,l}$ model the fast and slow time-varying components, respectively. In this paper we assume a block fading model where $h_{m,k,l}$ is fixed during the transmission of a block of $KL + N_D$ symbols and varies independently from one block to another. $N_D$ denotes the number of channel uses utilized for data transmission (see Fig. 1). The parameters $d_{k,l} \geq 0$, $l = 0, \ldots, L - 1$ model the power delay profile (PDP) of the frequency-selective channel for the $k$-th user. Since $\{d_{k,l}\}$ vary slowly with time and spatial location, we assume them to be fixed for the entire communication and independent of $m$. We further assume $h_{m,k,l}$ to be i.i.d. $\mathcal{CN}(0,1)$ distributed. Further, the PDP for every user is normalized such that the average received power is the same irrespectively of the length of the channel impulse response. Therefore, it holds that

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

for $1 \leq k \leq K$. Finally, we assume exact knowledge of the channel statistics at the BS, but not of the particular channel realizations.
A. Phase Noise Model

Phase noise is introduced at the transmitter during up-conversion, when the baseband signal is multiplied with the carrier generated by the local oscillator. The phase of the generated carrier drifts randomly, resulting in a phase distortion of the transmitted signal. A similar phenomenon also happens at the receiver side during down-conversion of the bandpass signal to baseband. In the following, \( \theta_k, \ k = 1, \ldots, K \) denotes the phase noise process at the \( k \)-th single-antenna user. Since the users are not cooperating, the transmitter phase noise processes are assumed to be mutually independent. On the other hand, at the receiver side two distinct operation modes are considered. We term these operation modes as synchronous and non-synchronous operation depending on whether the phase noise processes at the BS antennas are identical or independent.

For the synchronous case, all BS antennas are subject to the same phase noise process and \( \phi \) denotes this common phase noise process at each BS antenna. This models the scenario of a centralized BS with a single oscillator feeding the down-conversion module in each receiver. For the case of non-synchronous operation, \( \phi_m, \ m = 1, \ldots, M \) denotes the phase noise process at the \( m \)-th BS antenna. This models a completely distributed scenario where each BS antenna uses a distinct oscillator for down-conversion. We further assume that the phase noise processes \( \theta_k, \ k = 1, \ldots, K \) and \( \phi \) (or \( \phi_m, \ m = 1, \ldots, M \)) for the case of synchronous (or non-synchronous) operation mode are mutually independent.

In this study each phase noise process is modeled as an independent Wiener process, which is a well-established model \cite{11}, \cite{15}. Therefore, the discrete-time phase noise process at the \( k \)-th user at time \( i \) is given by:

\[
\theta_k[i] = \theta_k[i - 1] + w_k^i[i],
\]

where \( w_k^i[i] \sim \mathcal{N}(0, \sigma_\theta^2) \) are independent identically distributed zero-mean Gaussian increments with variance \( \sigma_\theta^2 \triangleq 4\pi^2 f_c^2 c T_s^2 \), \( f_c \) is the carrier frequency, \( T_s \) is the symbol interval and \( c \) is a constant that depends on the oscillator. Depending on the operation mode, the phase noise processes \( \phi[i] \) and \( \phi_m[i] \) at the \( M \) BS antennas are defined in a manner similar to (2), where the increments have variance \( \sigma_\phi^2 \). We assume throughout that \( \sigma_\phi^2 = \sigma_\theta^2 \).

\footnote{The discrete-time phase noise model is used since we will be working with the discrete-time complex baseband representation of the transmit and receive signals.}
B. Received Signal

Let $x_k[i]$ be the symbol transmitted from the $k$-th user at time $i$. The received signal at the $m$-th BS antenna element at time $i$ is then given by, for the synchronous operation

$$y_m[i] = \sqrt{P} \sum_{k=1}^{K} \sum_{l=0}^{L-1} e^{-j\phi[i]} g_{m,k,l} e^{j\theta_k[i-l]} x_k[i-l] + n_m[i].$$

(3)

and for the non-synchronous operation

$$y_m[i] = \sqrt{P} \sum_{k=1}^{K} \sum_{l=0}^{L-1} e^{-j\phi_m[i]} g_{m,k,l} e^{j\theta_k[i-l]} x_k[i-l] + n_m[i],$$

(4)

where $n_m[i] \sim \mathcal{CN}(0,\sigma^2)$ is the additive white Gaussian noise (AWGN) at the $m$-th receiver at time $i$. Each user transmits a stream of i.i.d. $\mathcal{CN}(0,1)$ information symbols (i.e., $x_k[i] \sim \mathcal{CN}(0,1)$), that are independent of the information symbols of the other users. $P$ denotes the average uplink transmitted power from each user.

III. Transmission Scheme and Receive Processing

Motivated by the need for low-complexity channel estimation and detection algorithms, we propose the following block-based uplink transmission scheme. In the proposed scheme, a transmission block of $KL+N_D$ channel uses consists of $KL$ channel uses dedicated to uplink channel training followed by $N_D$ channel uses for data transmission.

A. Channel Estimation

For coherent demodulation, the BS needs to estimate the uplink channel. This is facilitated through the transmission of uplink pilot symbols during the training phase of each transmission block. The users transmit uplink training signals sequentially in time, i.e., at any given time only one user is transmitting uplink training signals and all other users are silent. To be precise, the $k$-th user sends an impulse of amplitude $\sqrt{P_p KL}$ at the $(k-1)L$-th channel use and is idle for the remaining portion of the training phase. Here, $P_p$ is the average power transmitted by a user during the training phase. We choose the proposed training sequence since it allows for a very simple channel estimation scheme at the BS and since it facilitates our derivation of achievable rates. However, our results are expected to be approximately valid also for other training schemes. Therefore, using (3) and (4), respectively, the signal received at the $m$-th BS
receiver at time $(k - 1)L + l$, $l = 0, \ldots, L - 1$, $k = 1, \ldots, K$ is given by, for synchronous operation

$$y_m[(k - 1)L + l] = \sqrt{P_p KL} g_{m,k,l} e^{-j\phi_m[(k - 1)L + l]} e^{j\theta_k[(k - 1)L]} + n_m[(k - 1)L + l], \quad (5)$$

and for non-synchronous operation

$$y_m[(k - 1)L + l] = \sqrt{P_p KL} g_{m,k,l} e^{-j\phi_m[(k - 1)L + l]} e^{j\theta_k[(k - 1)L]} + n_m[(k - 1)L + l]. \quad (6)$$

The corresponding channel estimates are then given by, for synchronous operation

$$\hat{g}_{m,k,l} = \frac{1}{\sqrt{P_p KL}} y_m[(k - 1)L + l]$$

$$= g_{m,k,l} e^{-j\phi_m[(k - 1)L + l]} e^{j\theta_k[(k - 1)L]} + \frac{1}{\sqrt{P_p KL}} n_m[(k - 1)L + l], \quad (7)$$

and for non-synchronous operation

$$\hat{g}_{m,k,l} = \frac{1}{\sqrt{P_p KL}} y_m[(k - 1)L + l]$$

$$= g_{m,k,l} e^{-j\phi_m[(k - 1)L + l]} e^{j\theta_k[(k - 1)L]} + \frac{1}{\sqrt{P_p KL}} n_m[(k - 1)L + l]. \quad (8)$$

As expected, the channel estimate is distorted by the AWGN and by the phase noise at the user and at the BS.
B. Maximum Ratio Combining

Using (3) and (4), the received signal during the data phase is given by, for synchronous operation

\[ y_m[i] = \sqrt{P_D} \sum_{k=1}^{K} \sum_{l=0}^{L-1} e^{-j\phi_m[i]} g_{m,k,l} e^{j\theta_k[i-l]} x_k[i-l] + n_m[i], \]  

(9)

and for non-synchronous operation

\[ y_m[i] = \sqrt{P_D} \sum_{k=1}^{K} \sum_{l=0}^{L-1} e^{-j\phi_m[i]} g_{m,k,l} e^{j\theta_k[i-l]} x_k[i-l] + n_m[i], \]  

(10)

where \( i = KL, \ldots, N_D + KL - 1 \) and \( P_D \) is the per-user average transmit power constraint during the data phase. Motivated by the need for low-complexity detection, we consider the TR-MRC receiver at the BS. The TR-MRC receiver reverses the order of the received symbols, \( y_m[i] \), in the time-domain and convolves them with the complex conjugate of the estimated channel impulse response. Therefore, the detected symbol, \( \hat{x}_k[i] \), is given by

\[ \hat{x}_k[i] = \sum_{l=0}^{L-1} \sum_{m=1}^{M} \hat{g}_{m,k,l}^* y_m[i + l]. \]  

(11)

For the synchronous operation \( \hat{g}_{m,k,l} \) is given by (7) and \( y_m[i] \) is given by (9), and for the non-synchronous operation \( \hat{g}_{m,k,l} \) is given by (8) and \( y_m[i] \) is given by (10).

IV. Achievable sum-rate

We use the information sum-rate as the performance metric for quantifying the effects of phase noise. To this end, using (7) and (9) for the synchronous operation and using (8) and (10) for the non-synchronous operation, (11) can be written as

\[ \hat{x}_k[i] = A_k[i] x_k[i] + ISI_k[i] + MUI_k[i] + AN_k[i], \]  

(12)

where it holds for the non-synchronous operation that

\[ A_k[i] \triangleq \sqrt{P_D} \sum_{m=1}^{M} \sum_{l=0}^{L-1} |g_{m,k,l}|^2 e^{-j(\phi_m[i+l] - \phi_m[(k-1)L+l])} e^{j(\theta_k[i] - \theta_k[(k-1)L])} \]  

(13)
where an effective noise term. This results in the following equivalent expression

\[
\text{ISI}_k[i] \triangleq \sqrt{P_D} \sum_{m=1}^{M} \sum_{l=0}^{L-1} \sum_{p=0}^{L-1} g_{m,k,l}^* g_{m,k,p} e^{-j(\phi_m[i+l]-\phi_m[(k-1)L+l])} e^{j(\theta_k[i+l-p]-\theta_k[(k-1)L])} x_k[i+l-p]
\]

(14)

\[
\text{MUI}_k[i] \triangleq \sqrt{P_D} \sum_{m=1}^{M} \sum_{q=1}^{K} \sum_{l=0}^{L-1} \sum_{p=0}^{L-1} g_{m,k,l}^* g_{m,q,p} e^{-j(\phi_m[i+l]-\phi_m[(k-1)L+l])} e^{j(\theta_q[i+l-p]-\theta_k[(k-1)L])} x_q[i+l-p]
\]

(15)

\[
\text{AN}_k[i] \triangleq \sqrt{\frac{P_D}{P_p KL}} \sum_{m=1}^{M} \sum_{q=1}^{K} \sum_{l=0}^{L-1} \sum_{p=0}^{L-1} g_{m,q,p} e^{-j(\phi_m[i+l]-\theta_q[i+l-p])} n_m[(k-1)L+l] x_q[i+l-p]
\]

\[
+ \sum_{m=1}^{M} \sum_{l=0}^{L-1} g_{m,k,l}^* n_m[i+l]
\]

(16)

In (12), \(A_k[i]x_k[i]\) is the desired signal term for the \(k\)-th user, \(\text{ISI}_k[i]\) stands for the intersymbol interference for user \(k\) at time \(i\), caused by the information symbols of the \(k\)-th user transmitted at other time instances, \(\text{MUI}_k[i]\) denotes the multi-user interference due to the information symbols of the other users and finally \(\text{AN}_k[i]\) is an aggregate noise term that incorporates the effects of the channel estimation and the receiver AWGN noise, \(n_m[i]\). The expressions for the terms in (12) for the synchronous operation are obtained from (13)-(16) by substituting \(\phi_1[i] \equiv \ldots \equiv \phi_M[i] \equiv \phi[i]\).

In the following, we derive an achievable information rate for the \(k\)-th user. Similar capacity bounding techniques have been used earlier in e.g. [16, 17]. In (12), we add and subtract the term \(\mathbb{E}[A_k[i]] x_k[i]\), where the expectation is taken over the channel gains, \(g_{m,k,l}\), and the phase noise processes, \(\theta_k\), \(\phi\) for the synchronous operation and \(\theta_k\), \(\phi_m\) for the non-synchronous operation. We relegate the variation around this term, i.e., \(\text{IF}_k[i] \triangleq (A_k[i] - \mathbb{E}[A_k[i]]) x_k[i]\), to an effective noise term. This results in the following equivalent expression

\[
\hat{x}_k[i] = \mathbb{E}[A_k[i]] x_k[i] + \text{EN}_k[i],
\]

(17)

where

\[
\text{EN}_k[i] \triangleq \text{IF}_k[i] + \text{ISI}_k[i] + \text{MUI}_k[i] + \text{AN}_k[i],
\]

(18)
is the effective additive noise term. In (17) the detected symbol, $\hat{x}_k[i]$, is a sum of two uncorrelated terms (i.e., $\mathbb{E} \left[ (\mathbb{E}[A_k[i]]x_k[i]) (\text{EN}_k[i])^* \right] = 0$). The importance of the equivalent representation in (17) is that the scaling factor $\mathbb{E}[A_k[i]]x_k[i]$ of the desired information symbol is a constant, which is known at the BS since the BS has knowledge of the channel statistics. The exact probability distribution of $\text{EN}_k[i]$ is difficult to compute. However, its variance can be easily calculated given that the channel statistics is known at the BS. Therefore, (17) describes an effective single user single-input single-output (SISO) additive noise channel, where the noise is zero mean, has known variance and is uncorrelated with the desired signal term. From the expressions for $A_k[i]$ and $\text{EN}_k[i]$ in (13) and (18) the mean value of $A_k[i]$ and the variance of $\text{EN}_k[i]$ is given by two theorems that follow.

**Theorem 1.** The mean value of $A_k[i]$ in both operation modes is given by

$$\mathbb{E}[A_k[i]] = \sqrt{P_D}M e^{-\frac{\sigma^2}{2} (i - (k-1)L)}. \quad (19)$$

**Proof:** We prove the statement for the non-synchronous operation. The proof for the synchronous operation is nearly identical. From (13), we have

$$\mathbb{E}[A_k[i]] = \mathbb{E} \left[ \sqrt{P_D} \sum_{m=1}^{M} \sum_{l=0}^{L-1} \left| g_{m,k,l} \right|^2 e^{-j(\phi_m[i+l]-\phi_m[(k-1)L+l])} e^{-j(\theta_k[(k-1)L]-\theta_k[i])} \right]$$

$$= (a) \sqrt{P_D} \sum_{m=1}^{M} \sum_{l=0}^{L-1} \mathbb{E} \left[ \left| g_{m,k,l} \right|^2 \right] \mathbb{E} \left[ e^{-j(\phi_m[i+l]-\phi_m[(k-1)L+l])} \right] \mathbb{E} \left[ e^{-j(\theta_k[(k-1)L]-\theta_k[i])} \right]$$

$$= (b) \sqrt{P_D} \sum_{m=1}^{M} \sum_{l=0}^{L-1} d_{k,l} e^{-\frac{\sigma^2}{2} (i - (k-1)L)} e^{-\frac{\sigma^2}{2} (i - (k-1)L)} \quad \overset{(c)}{=} \sqrt{P_D} M e^{-\frac{\sigma^2}{2} (i - (k-1)L)}.$$

In (a) we have used the fact that the channel realizations, $g_{m,k,l}$, the phase noise at the BS, $\phi_m$, and the phase noise at the $k$-th user, $\theta_k$, are mutually independent random processes. The equality (b) is a consequence of the Wiener phase noise model. That is, after a time interval, $\Delta t = i - (k-1)L$, the phase drift of an oscillator is a zero mean Gaussian random variable with variance that is proportional to $\Delta t$,

$$U_{\phi_m} \overset{\Delta}{=} \phi_m[i + l] - \phi_m[(k - 1)L + l] \sim \mathcal{N} (0, \sigma^2_{\phi}(i - (k - 1)L)),$$

$$U_{\theta_k} \overset{\Delta}{=} \theta_k[i] - \theta_k[(k - 1)L] \sim \mathcal{N} (0, \sigma^2_{\theta}(i - (k - 1)L)).$$
Henceforth $\mathbb{E} \left[ e^{-jU_{\phi_m}} \right] = \varphi_{\phi_m}(-1) = e^{-\frac{\sigma^2}{2}(i-(k-1)L)}$ and $\mathbb{E} \left[ e^{jU_{\theta_k}} \right] = \varphi_{\theta_k}(1) = e^{-\frac{\sigma^2}{2}(i-(k-1)L)}$, where $\varphi_{\phi_m}(\cdot)$ and $\varphi_{\theta_k}(\cdot)$ are the characteristic functions of $U_{\phi_m}$ and $U_{\theta_k}$, respectively. The equality (c) follows from (1).

In the following we provide an intuitive explanation of the result in Theorem 1. Observe that in the non-synchronous mode, the proposed estimate of the channel gains to the $m$-th BS antenna in (8) depends on the phase noise process of the $m$-th BS antenna element but is independent of the phase noise process of any other BS antenna element. Additionally, the received signal $y_m[i]$ at the $m$-th BS antenna element, which contains also the symbol to be detected, $x_k[i]$, depends only on the phase noise process at this antenna element but not on $\phi_{m'}[i]$, $m' \neq m$. Since only $y_m[i + l]$, $l = 0, 1, \ldots, L - 1$, $m = 1, \ldots, M$ depend on $x_k[i]$, for detecting $x_k[i]$ we only need to consider the received samples at the time instances $i, i + 1, \ldots, i + L - 1$. Further, for any $l \in \{0, 1, \ldots, L - 1\}$, the useful signal component in $y_m[i + l]$ is $\sqrt{P_D}g_{m,k,l}x_k[i]e^{j\theta_k[i]}e^{-j\phi_m[i+l]}$ (see (10)). For coherent combining of the channel gains at time $i + l$ we would ideally have liked to compute $g_{m,k,l}^*y_m[i + l]$ and summed it over $m = 1, \ldots, M$. However, since the channel gain estimate $\hat{g}_{m,k,l}$ depends only on $g_{m,k,l}$ and is independent of any other channel gain $\hat{g}_{m',k',l'}$ ($m' \neq m$, $k' \neq k$, $l' \neq l$), it is intuitive to compute $\hat{g}_{m,k,l}^*y_m[i + l]$ and sum it over $m = 1, \ldots, M$. The term $\hat{g}_{m,k,l}^*y_m[i + l]$ contains the useful signal term $\sqrt{P_D}|g_{m,k,l}|^2e^{j(\theta_k[i] - \theta_k[(k-1)L])}e^{-j(\phi_m[i+l] - \phi_m[(k-1)L+l])}x_k[i]$. The factor $\eta_{m,k,l,i} = \sqrt{P_D}|g_{m,k,l}|^2e^{j(\theta_k[i] - \theta_k[(k-1)L])}e^{-j(\phi_m[i+l] - \phi_m[(k-1)L+l])}$ is a random variable having a non-zero mean value of $\sqrt{P_D}d_k[e^{-\sigma^2\delta^2(i-(k-1)L)}]$. Note that the mean value depends on the time difference between the time instance of the received sample $y_m[i + l]$ (i.e., $i + l$) and the time when the channel gain $g_{m,k,l}$ was estimated (i.e., $(k-1)L + l$). This mean value is also independent of $m$, i.e., it is the same for all BS antennas due to the identical phase noise process statistics. Hence this mean value is the same for both modes of operation.

Also note that summing the mean value of $\eta_{m,k,l,i}$ over $m = 1, \ldots, M$, $l = 0, \ldots, L - 1$ gives

$$\sum_{m=1}^{M} \sum_{l=0}^{L-1} \mathbb{E} [\eta_{m,k,l,i}] = \mathbb{E} \left[ \sum_{m=1}^{M} \sum_{l=0}^{L-1} \eta_{m,k,l,i} \right] = \mathbb{E} [A_k[i]] = \sqrt{P_D}Me^{-\sigma^2\delta^2(i-(k-1)L)}.$$

The factor $M$ signifies the combining gain in a coherent receiver (i.e., when $\sigma_\phi = 0$). The factor $e^{-\sigma^2\delta^2(i-(k-1)L)}$ signifies the loss in effective amplitude gain due to the non-coherency between the received data samples and the estimated channel gains. Note that this non-coherency arises due to
the fact that the channel gains for the $k$-th user are estimated at $t = (k-1)L + l$, $l = 0, \ldots, L-1$ and the samples for detecting $x_k[i]$ are received at $t = i + l$, $l = 0, \ldots, L-1$, i.e., a time difference of $i - (k-1)L$. The oscillator phase drifts in this time period result in a partial non-coherency. Hence it is clear that the larger this time difference is the smaller the effective amplitude gain (in fact, the effective amplitude $M e^{-\sigma^2(i-(k-1)L)}$ decreases exponentially with increasing time difference $i - (k-1)L$).

**Theorem 2.** The variance $\text{Var}(E_{N_k}[i]) \triangleq \mathbb{E} [ | EN_k[i] - \mathbb{E}[EN_k[i]] |^2 ]$ satisfies, for synchronous operation

$$\text{Var}(E_{N_k}[i]) \leq P_D M^2 (1 - \kappa_k[i]) + P_D MK + \sigma^2 M \left(1 + \frac{P_D}{K P_p} + \frac{\sigma^2}{K P_p}\right), \quad (20)$$

and for non-synchronous operation

$$\text{Var}(E_{N_k}[i]) \leq P_D M (1 - \kappa_k[i]) (1 + M \kappa_k[i]) + P_D MK + \sigma^2 M \left(1 + \frac{P_D}{K P_p} + \frac{\sigma^2}{K P_p}\right), \quad (21)$$

where $\kappa_k[i] \triangleq e^{-\sigma^2(i-(k-1)L)}$.

**Proof:** See Appendix. \(\blacksquare\)

The last term on the right-hand-side of the expressions (20) and (21) is the contribution of the additive noise term $AN_k[i]$. For the sake of clarity of the result we provide the following remark.

**Remark 1.** The contribution of the additive noise term $AN_k[i]$ in the variance expressions (20) and (21). $\mathbb{E} [|AN_k[i]|^2] = \sigma^2 M \left(1 + \frac{P_D}{K P_p} + \frac{\sigma^2}{K P_p}\right)$ is a sum of three terms. The term $\sigma^2 M$ corresponds to the AWGN during the data phase as it is filtered in (11). The term $\sigma^2 M \frac{P_D}{K P_p}$ corresponds to the cross-correlation between the channel estimation error term in (7) and (8) and the received symbols in (9) and (10), respectively. Finally, the last term $\sigma^2 M \frac{\sigma^2}{K P_p}$ corresponds to the channel estimation error.

In the following we provide a coding strategy that justifies the achievable rates we are interested in deriving. From Theorems 1 and 2 it is obvious that $\mathbb{E}[A_k[i]]$ and the upper bound on $\text{Var}(E_{N_k}[i])$ depend on $i$ and are different for different $i = KL, \ldots, KL + N_D - 1$. Further, for a given $i$, across multiple transmission blocks, the terms $\mathbb{E}[A_k[i]]$ and $\text{Var}(E_{N_k}[i])$ are the same and the realizations of $E_{N_k}[i]$ are i.i.d. Hence, for each $i$, we have an additive noise
SISO channel. This motivates us to consider $N_D$ channel codes for each user, one for each $i = KL, \ldots, KL + N_D - 1$. At the $k$-th transmitter (user), the symbols of the $i$-th channel code ($x_k[i]$) are transmitted only during the $i$-th channel use of each transmission block. Similarly, at the BS, for the $k$-th user, the $i$-th received and processed symbols (i.e., $\hat{x}_k[i]$) across different transmission blocks are jointly decoded. Essentially, this implies that, at the BS we have $N_D$ parallel channel decoders for each user. We propose the above scheme of $N_D$ parallel channel codes for each user only to derive a lower bound on the achievable information rate. In practice, due to reasons of complexity, channel coding/decoding would not only be performed across different transmission blocks, but also across consecutive channel uses within each transmission block.

Given the previously described coding strategy, we are now interested in computing a lower bound on the reliable rate of communication for each of the $N_D$ channel codes. Since the data symbols $x_k[i]$ are Gaussian, for each $i = KL, \ldots, KL + N_D - 1$ a lower bound on the information rate for the effective channel in (17) can be computed by considering the worst case (in terms of mutual information) uncorrelated additive noise. With Gaussian information symbols, it is known that the worst case uncorrelated noise is Gaussian with the same variance as that of $E_n[k][i]$ [16]. Consequently, a lower bound on $I(\hat{x}_k[i]; x_k[i])$ (i.e., the mutual information rate for the $i$-th channel code for user $k$) is given by Theorem 3.

**Theorem 3.** The achievable rates for the $i$-th channel code for the $k$-th user is given by, for synchronous operation

$$I(\hat{x}_k[i]; x_k[i]) > R_k^s[i] \Delta \log_2 \left( 1 + \frac{P_D \kappa_k^2[i]}{\sigma^2 M (1 - \kappa_k^2[i])} + \frac{P_D}{\sigma^2 K} + 1 + \frac{P_D}{T_p} + \frac{\sigma^2}{K T_p} \right),$$

and for non-synchronous operation

$$I(\hat{x}_k[i]; x_k[i]) > R_k^{ns}[i] \Delta \log_2 \left( 1 + \frac{P_D M \kappa_k^2[i]}{\sigma^2 (1 - \kappa_k[i])(1 + M \kappa_k[i])} + \frac{P_D}{\sigma^2 K} + 1 + \frac{P_D}{T_p} + \frac{\sigma^2}{K T_p} \right).$$

**Corollary 1.** The rates achieved by the proposed time-reversed maximum-ratio combining (TR-
**MRC** receiver given by (22) and (23) in Theorem 3 are invariant of the PDP of the frequency-selective channel.

**Corollary 2.** The proposed TR-MRC receiver exhibits better performance in the case of non-synchronous operation.

**Proof of Corollary 2.** Looking at the first term in the denominators of (22)-(23) (all other terms are equal) we see that

$$\frac{\frac{P_D M^2 (1 - \kappa_k[i])}{\sigma^2}}{\frac{P_D M (1 - \kappa_k[i])(1 + M\kappa_k[i])}{\sigma^2}} = \frac{M - 1}{1 + M\kappa_k[i]} + 1 > 1 \Rightarrow R^s_k[i] < R^n[i].$$

Corollary 2 conveys a seemingly counter-intuitive result. However, this is not the first time that such a result is reported. In [18, Section III.A] the authors study the effect of phase noise in single-user beamforming with perfect transmitter channel knowledge. The performance measure they use is the error vector magnitude (EVM) and they show that EVM is smallest in the desired direction when uncorrelated phase noise sources are used. In the following, we provide a simple example to illustrate that Corollary 2 does not seem to be an artifact of the bounding technique used to derive the lower bounds on the information rate in Theorem 3 but is likely to be the result of a more fundamental phenomenon.

For the purpose of illustration we consider a very simple channel with only phase noise and no AWGN (essentially a limiting result when the variance of AWGN vanishes, i.e., at very high signal-to-noise-ratio) see Fig. 2. Here $X \in \{\pm 1\}$, $\Pr\{X = +1\} = p$, $\Pr\{X = -1\} = 1 - p$ be the input to the channel. The input $X$ is rotated by $\varphi_1$ and $\varphi_2$ to form $Y_1$ and $Y_2$, respectively. Let the random variables $\varphi_1$, $\varphi_2$, that model the effect of memoryless, stationary and ergodic phase noise processes, have the following probability mass functions.

![Fig. 2. System model for the example.](image-url)
(p.m.f): \( \varphi_i \in \{ -\frac{\pi}{2}, 0, \frac{\pi}{2} \} \), \( \Pr\{ \varphi_i = -\frac{\pi}{2} \} = \Pr\{ \varphi_i = 0 \} = \Pr\{ \varphi_i = \frac{\pi}{2} \} = \frac{1}{3}, \ i = 1, 2 \). The output of this discrete memoryless channel (DMC) is given by

\[
Y = \frac{1}{2} \left( e^{j\varphi_1} + e^{j\varphi_2} \right) X.
\]

(24)

We now consider two cases, firstly when the two phase noise processes are synchronous (i.e., \( \varphi_1 \equiv \varphi_2 \)) and secondly when they are non-synchronous and independent of each other. For the case of synchronous phase noise processes, \( \varphi_1 \equiv \varphi_2 \) and the system reduces to \( Y = e^{j\varphi_1} X \). It is straightforward to see that \( Y \) takes values in \( Y_s = \{ +1, +j, -1, -j \} \). The input p.m.f. induces the following p.m.f on the output symbols;

\[
\Pr\{ Y = +1 \} = \frac{p}{3}, \Pr\{ Y = -1 \} = \frac{(1 - p)}{3}, \Pr\{ Y = \pm j \} = \frac{1}{3}.
\]

The capacity of this channel can be calculated as follows

\[
C_s = \max_p I(X; Y) = \max_p H(Y) - H(Y|X) = \max_p \log_2 3 + \frac{1}{3} H_2(p) - \log_2 3
\]

\[= \max_p \frac{1}{3} H_2(p) = \frac{1}{3} \text{ bits,}\]

where \( H_2(p) \) is the binary entropy function. The capacity is achieved for equiprobable input symbols (i.e., \( p = 1/2 \)).

For the case of non-synchronous phase noise processes, where \( \varphi_1 \) and \( \varphi_2 \) are independent of each other, the system model is given by (24) and the output variable takes values in \( Y_{ns} = \{ +1, \frac{1}{2}(1+j), \frac{1}{2}(1-j), j, 0, -j, -\frac{1}{2}(1+j), -\frac{1}{2}(1-j), -1 \} \). The p.m.f. of the output induced by the input distribution is given by \( \Pr\{ Y = +1 \} = \frac{p}{9}, \Pr\{ Y = (1 \pm j)/2 \} = \frac{2p}{9}, \Pr\{ Y = \pm j \} = \frac{1}{9}, \Pr\{ Y = 0 \} = \frac{2}{9}, \Pr\{ Y = -(1 \pm j)/2 \} = \frac{2(1 - p)}{9}, \text{ and } \Pr\{ Y = -1 \} = \frac{(1 - p)}{9} \). We can compute that \( H(Y) = \frac{5}{9} H_2(p) + \log_2 9 - 6/9 \) and \( H(Y|X = \pm 1) = \log_2 9 - 6/9 \). Then, the capacity is given by

\[
C_{ns} = \max_p I(X; Y) = \max_p H(Y) - H(Y|X) = \max_p \frac{5}{9} H_2(p) + \log_2 9 - 6/9 - \log_2 9 + 6/9
\]

\[= \max_p \frac{5}{9} H_2(p) = \frac{5}{9} \text{ bits.}\]

The capacity is achieved for equiprobable input symbols (i.e., \( p = 1/2 \)). Since \( C_s < C_{ns} \), it is concluded that the capacity of the channel in Fig. 2 with independent phase noise processes is
larger than the capacity of the same channel when the phase noise processes are fully correlated.

Since no data transmission happens during the training phase, the overall effective information rate achievable by the $k$-th user is given by, for synchronous operation

$$R_k^s = \frac{1}{KL + N_D} \sum_{i=KL}^{KL+N_D-1} R_k[i].$$  \hfill (25)

and for non-synchronous operation

$$R_k^{ns} = \frac{1}{KL + N_D} \sum_{i=KL}^{KL+N_D-1} R_k^{ns}[i].$$  \hfill (26)

The achievable sum-rate is therefore given by, for synchronous operation

$$R_s = \sum_{k=1}^{K} R_k^s = \frac{1}{KL + N_D} \sum_{k=1}^{K} \sum_{i=KL}^{KL+N_D-1} R_k^s[i].$$  \hfill (27)

and for non-synchronous operation

$$R_{ns} = \sum_{k=1}^{K} R_k^{ns} = \frac{1}{KL + N_D} \sum_{k=1}^{K} \sum_{i=KL}^{KL+N_D-1} R_k^{ns}[i].$$  \hfill (28)

We conclude this section by formalizing an intuitively obvious observation.

**Remark 2.** Phase noise degrades the sum-rate performance in both operation modes.

**Proof:** The sum-rate for the no-phase-noise case can be derived from (22), (25) and (27) (or (23), (26) and (28)) by setting $c = 0$ (i.e., $\sigma_\phi^2 = 0$) and is given by

$$R = \frac{K N_D}{N_D + KL} \log_2 \left( 1 + \frac{P_D}{\sigma^2} M + \frac{P_D}{\sigma^2} K + 1 + \frac{P_D}{\sigma^2} \frac{M}{P_p} K + 1 + \frac{P_D}{\sigma^2} \frac{M}{KP_p} \right).$$  \hfill (29)

For the synchronous case, $\frac{P_D}{\sigma^2} M > \frac{P_D}{\sigma^2} M \kappa_k^2[i]$ and $\frac{P_D}{\sigma^2} M (1 - \kappa_k^2[i]) + \frac{P_D}{\sigma^2} K + 1 + \frac{P_D}{\sigma^2} \frac{M}{P_p} K + 1 + \frac{P_D}{\sigma^2} \frac{M}{KP_p} > \frac{P_D}{\sigma^2} K + 1 + \frac{P_D}{\sigma^2} \frac{M}{P_p} + \frac{\sigma^2}{KP_p}$, which implies that $R_s < R$ (see (22)). Similarly we can argue that $R_{ns} < R$.

In the following $\beta = \frac{P_p}{P_D} > 0$ denotes the ratio between the per-user average transmit power during the training phase and that during transmission phase.

---

3With this example we do not argue that capacity always increases if we use independent phase noise sources. We rather argue that there are cases where the use of independent phase noise sources can be beneficial.
V. ASYMPTOTIC RESULTS

In this section we present the main asymptotic results that follow from the derived achievable sum-rates. We start by a remark on the effect of phase noise in the low-SNR regime.

Remark 3. In the low SNR regime, the performance loss due to phase noise is not significant.

Proof: The sum-rate of the system when phase noise is present is given by (27) and (28) for synchronous and non-synchronous operation, respectively. From (22) and (23) it is clear that in the low SNR regime, i.e., when \( P_D/\sigma^2 \ll 1 \), the dominating factor in the denominator of the argument of the \( \log_2 \) function is, in both operation modes, the term \( \frac{\sigma^2}{K\beta P_D} \). From (29) (after the substitution \( P_p = \beta P_D \)) it is clear that the term \( \frac{\sigma^2}{K\beta P_D} \) is also the dominating term in the denominator of the achievable rate expression in the no-phase-noise case. Therefore, the performance loss of both operation modes compared to the no-phase-noise scenario is small.

We proceed with a result on the sum-rate performance in the high-SNR regime.

Proposition 1. Saturation in the high-SNR regime. In the presence of phase noise the effective information rate of the \( k \)-th user saturates for \( \frac{P_D}{\sigma^2} \to \infty \) to the values, for synchronous operation

\[
R_s^k \to \frac{1}{N_D + KL} \sum_{i=KL}^{N_D+KL-1} \log_2 \left( 1 + \frac{M\kappa^2_k[i]}{M(1 - \kappa^2_k[i]) + K} \right), \tag{30}
\]

and for non-synchronous operation

\[
R_{ns}^k \to \frac{1}{N_D + KL} \sum_{i=KL}^{N_D+KL-1} \log_2 \left( 1 + \frac{M\kappa^2_k[i]}{(1 - \kappa_k[i])(1 + M\kappa_k[i]) + K} \right). \tag{31}
\]

Proof: The result follows immediately from (22), (23) and the definitions of \( R_s^k \) and \( R_{ns}^k \) in (25) and (26).

In the high-SNR regime, MRC is known to be suboptimal since intersymbol interference and multi-user interference dominate the effective noise term. Therefore saturation in the high-SNR regime is observed also in the no-phase-noise case due to the MRC reception strategy.

A particularly desirable property of massive MIMO systems is the array power gain that they offer, facilitating the design of highly power-efficient communication systems. Previous work has shown that an \( O(M) \) array power gain can be achieved when perfect CSI is available [19]. The authors in [7] prove that with simple linear processing and imperfect CSI, which is acquired via uplink training, an \( O(\sqrt{M}) \) array power gain can be achieved. In the present work we extend this
result to the phase noise-impaired single-carrier massive MIMO uplink with TR-MRC receive processing.

**Proposition 2.** An $O(\sqrt{M})$ array gain is achievable for the frequency-selective MU-MIMO uplink in the presence of phase noise and imperfect channel estimation, i.e., for a fixed number of users $K$, with a sufficiently large antenna array at the BS, the average transmitted power $P_D$ can be reduced by roughly 1.5dB for every doubling in the number of BS antennas while maintaining a constant positive information rate for each user.

**Proof:** We start by proving the proposition for the synchronous case. Let $P_D = \frac{E_u}{M^\alpha}$, where $E_u$ is fixed. Based on the derived achievable rates in Theorem 3, we compute the maximum possible exponent, $\alpha > 0$, such that a fixed, non-zero rate for the $i$-th code of user $k$ can be achieved, while the transmit power of each user is scaled as $1/M^\alpha$ with increasing $M$. From (22) we have

\[
R_{s}^{s}[i] = \log_2 \left( 1 + \frac{E_u}{\sigma^2 M^\alpha} M \kappa_k^2[i] \right) \quad \text{for } M \rightarrow \infty \quad \text{and} \quad \lim_{M \rightarrow \infty} R_{s}^{s}[i] > 0 \text{ if } \alpha - 1 \leq 0 \text{ and } 2\alpha - 1 \leq 0 \Rightarrow \alpha \leq 1/2. \tag{32}
\]

As $M \rightarrow \infty$ we have $\lim_{M \rightarrow \infty} R_{s}^{s}[i] > 0$ if $\alpha - 1 \leq 0$ and $2\alpha - 1 \leq 0 \Rightarrow \alpha \leq 1/2$. For $\alpha = 1/2$ the rate $R_{s}^{s}$ converges to the value (as $M \rightarrow \infty$)

\[
R_{s}^{s} \rightarrow \frac{1}{N_D + KL} \sum_{i=KL}^{N_D+KL-1} \log_2 \left( 1 + \frac{E_u}{\sigma^2 (1 - \kappa_k^2[i])} + \frac{\sigma^2}{\kappa^2 E_u} \right). \tag{33}
\]

Similarly, it can be proved that the array gain for the non-synchronous operation is $O(\sqrt{M})$ and the rate approaches (as $M \rightarrow \infty$) the value

\[
R_{ns}^{s} \rightarrow \frac{1}{N_D + KL} \sum_{i=KL}^{N_D+KL-1} \log_2 \left( 1 + \frac{E_u}{\sigma^2 (1 - \kappa_k^2[i])} + \frac{\sigma^2}{\kappa^2 E_u} \right). \tag{33}
\]

We conclude this section with an observation of the dependence of the derived achievable sum-rates on the length of the channel impulse response, $L$.

**Proposition 3.** The achievable sum-rates (27) and (28) are monotonically decreasing in $L$.

**Proof:** We prove the proposition for the synchronous case. The non-synchronous case
follows similarly. Note that the rate \( R_s[k] \) can be expressed as
\[
R_s[k] = -\log_2 \left( 1 - \frac{A}{B} e^{-2\sigma_\phi^2 i - (k-1)L} \right),
\]
for \( i = KL, KL+1, \ldots, N_D + KL - 1 \), where \( A \triangleq \frac{P_D}{\sigma_\phi^2} M \) and \( B \triangleq \frac{P_D}{\sigma_\phi^2} (M + K) + 1 + \frac{1}{\beta} + \frac{\sigma_\phi^2}{\beta P_D} > A \). Note that the starting index of \( i \), (i.e., \( KL \)) also increases with increasing \( L \). So we cannot treat \( i \) as constant. Instead, letting \( t \triangleq i - KL \) we have
\[
R_s[k][t] = -\log_2 \left( 1 - \frac{A}{B} e^{-2\sigma_\phi^2 t e^{-2\sigma_\phi^2 ((K-k+1)L)}} \right),
\]
t = 0, 1, \ldots, N_D - 1. Note that the range of \( t \), (i.e., \( 0, 1, \ldots, N_D - 1 \)) is now independent of \( L \). It is now clear that \( R_s[k][t] \) reduces with increasing \( L \) for all \( t \in \{0, 1, \ldots, N_D - 1\} \).

VI. Numerical Examples

Throughout this section, the plots used to illustrate the main results assume that \( T_s = 0.1\mu s \), \( f_c = 2 \text{ GHz} \) and \( c = 4.7 \times 10^{-18} \text{(rad Hz)}^{-1} \), unless otherwise stated. The selected parameters correspond to typical values of a wideband wireless communication system, such as a WLAN IEEE 802.11 [20]. The selected \( c \) corresponds to an oscillator with phase noise \(-133 \text{ dBc/Hz}\) at a frequency offset \( f_m = 20 \text{ MHz} \) from the carrier. This is an inexpensive oscillator of rather poor quality since one can find oscillators with phase noise performance as low as \(-157 \text{ dBc/Hz}\) at 20 MHz but at much greater cost. In typical cellular systems the delay spread is of the order of microseconds. We select \( L = 20 \), which corresponds to 2\( \mu s \) of delay spread for the selected symbol rate. Since the derived achievable sum-rates are independent of the power delay profile, we do not specify one. Also, the constant of proportionality between \( P_D \) and \( P_p \) is fixed to \( \beta = 1 \), i.e., \( P_p = P_D \). In Fig. 3 the sum-rate performance of the system, as given by (27) and (28), is plotted as a function of \( \frac{P_D}{\sigma_\phi^2} \) for \( N_D = [100 1000] \) with \( M = 100, K = 10 \). The sum-rate achieved without phase noise (29) is plotted for the sake of comparison. We observe that at low SNR, the loss in sum-rate performance is insignificant. This observation supports the result in Remark 3.

A significant desirable property of massive MIMO systems is the array power gain that they offer, facilitating the design of highly power-efficient communication systems [4], [7], [19]. Proposition 2 extends this result to the case of single-carrier frequency-selective Massive MU-
MIMO systems impaired with phase noise. The above observation is further supported through Fig. 4, where the minimum per-user $\frac{P_D}{\sigma^2}$ required to achieve a fixed per-user information rate of $r = 1$ bpcu is plotted as a function of the number of BS antennas for $N_D = 1000$ and $K = 10$ for two choices of the oscillator constant $c$, namely, $4.7 \times 10^{-18}$ and $2.35 \times 10^{-17}$ (rad Hz)$^{-1}$. The plot for the phase-noise-free case is also given for the sake of comparison.
From Fig. 4 we are motivated to study the gap in required $\frac{P_d}{\sigma^2}$ between the phase-noise-impaired cases and the no-phase-noise operation. In Table I we present numerical results on this gap. Each row corresponds to a different oscillator constant $c$, namely, $9.4 \times 10^{-19}$, $4.7 \times 10^{-18}$ and $2.35 \times 10^{-17} (\text{rad Hz})^{-1}$. In order to give a more intuitive measure of the disturbance introduced by phase noise, we list the vertical $\frac{P_d}{\sigma^2}$ gap as a function of the standard deviation of the phase noise drifts at a time difference of $N_D$ channel uses (i.e., the time difference between the end of the training phase and the end of the data phase). This result is shown in Table I. As expected, the performance gap is minimal for small phase noise drift and increases as the standard deviation of the phase noise drift increases.

| Gap in required $\frac{P_d}{\sigma^2}$ [dB] | Synchronous | Non-Synchronous |
|------------------------------------------|-------------|----------------|
| $\sigma_\phi \sqrt{N_D}$ (degrees)     | M=500       | M=2500         | M=500          | M=2500         |
| 6.98°                                    | 0.1095      | 0.0950         | 0.0766         | 0.0705         |
| 15.61°                                   | 0.5647      | 0.5071         | 0.3926         | 0.3485         |
| 34.91°                                   | 3.9727      | 3.3913         | 2.1076         | 1.8546         |

It is also interesting to study the gap in required $\frac{P_d}{\sigma^2}$ as a function of the desired per-user information rate. For this purpose we provide Table II. There, we tabulate the gap in required $\frac{P_d}{\sigma^2}$ in dB for various values of the per-user desired information rate for the synchronous and non-synchronous mode, for $N_D = 1000$ channel uses, oscillator constant $c = 4.7 \times 10^{-18} (\text{rad Hz})^{-1}$, $K = 10$ users and $M = 500$ BS antennas. In the low spectral efficiency regime this gap is minimal. However, as the desired per-user information rate increases the gap increases at a faster rate. When the desired per-user information rate increases from 2 bpcu to 2.5 bpcu, which corresponds to 25% increase, the gap in dB in the case of non-synchronous operation doubles, whereas in the synchronous operation mode the vertical gap increases more than two times. This happens because the desired per-user rate is close to the high-SNR saturation rate for the case of synchronous receivers. As a result, a large increase in the transmit power is required in order

\[4\] With the selected parameters, the high-SNR saturation value for the synchronous operation is 2.8 bpcu per user.
to achieve the desired information rate.

**TABLE II**

| Gap in required $\frac{P}{\sigma^2}$ due to phase noise for $N_D = 1000$, $c = 4.7 \times 10^{-18} \text{(rad Hz)}^{-1}$, $K = 10$ users and $M = 500$ BS antennas for various values of the desired per-user information rate in bits per channel use [BPCU]. |
|---|---|---|
| Per-user rate | Synchronous | Non-Synchronous |
| 0.25 | 0.2594 | 0.2337 |
| 0.5 | 0.3355 | 0.2746 |
| 1 | 0.5647 | 0.3912 |
| 2 | 1.8591 | 0.9547 |
| 2.5 | 4.3465 | 1.7142 |

In Proposition 2, the rates converge to their corresponding value when the power is scaled by $O(\sqrt{M})$. However, it is also important to consider the case of $\alpha$ where $0 \leq \alpha < 1/2$. In the no-phase noise case, the sum-rate is lower bounded by

$$R = \frac{K N_D}{N_D + KL} \log_2 \left( 1 + \frac{E_u}{\sigma^2} K \beta M^{1-2\alpha} \right).$$

Since $1 - 2\alpha > 0$, $R$ becomes unbounded. However, in the phase-noise-impaired system the rates remain bounded to

$$R_{ks}^{s} \to \frac{1}{N_D + KL} \sum_{i=KL}^{N_D+KL-1} \log_2 \left( 1 + \frac{\kappa_k^2[i]}{(1 - \kappa_k[i])} \right) \quad (34)$$

for the synchronous operation and to

$$R_{ns}^{ns} \to \frac{1}{N_D + KL} \sum_{i=KL}^{N_D+KL-1} \log_2 \left( 1 + \frac{\kappa_k[i]}{(1 - \kappa_k[i])} \right) \quad (35)$$

for the non-synchronous operation. Note that the above expressions are independent of $M$ and $\frac{E_u}{\sigma^2}$. This implies that for a phase-noise-impaired system the achievable rates are bounded. Therefore, with power scaling there is a rate region that cannot be achieved in a phase noise impaired system even in the case when we are allowed to deploy infinitely many BS antennas or spend infinite power $E_u$. This result agrees with earlier studies of sum-rate performance of MIMO with transceiver impairments [21]. In Fig. 5, equations (34) and (35) are plotted for two choices of the oscillator constant.
For fixed $M$, $K$ and $L$ there is a fundamental trade-off between the length of the data interval, $N_D$, and the achievable sum-rate performance. A fraction $\frac{KL}{N_D+KL}$ of each coherence interval is spent on training. Since a fixed time interval of $KL$ channel uses is required for channel estimation, a small data interval, $N_D$, leads to underutilization of the available resources, yielding a low sum-rate performance. As $N_D$ increases, more resources are utilized for the data transmission, increasing the sum-rate performance. However, as it can be seen from (22) and (23), $R^s_k[i] < R^s_k[i-1]$ and $R^{ns}_k[i] < R^{ns}_k[i-1]$, which implies that the gain of increasing the data interval diminishes with increasing $N_D$. In fact, the individual rates $R^s_k[i]$ and $R^{ns}_k[i]$ approach 0 as $i \to \infty$. This phenomenon occurs because with large $N_D$, the phase noise drift in the oscillators is so large such that there is a total loss of coherency between the received symbols during the data phase and the estimated channel at the beginning of the transmission block. In Fig. 6 the sum-rate performance is plotted as a function of $N_D$ for various values of the parameter $c$. In the no-phase-noise case the optimal value of $N_D$ is infinity. However, there is a clear trade-off between the sum-rate and the length of the data interval in the phase-noise-impaired operation modes. As expected, the larger the variance of the phase noise processes the smaller the optimal $N_D$.

Further insight can be obtained by considering the optimum number of scheduled users. In
Fig. 6. Sum-rate performance as a function of $N_D$, with fixed $\frac{P_D}{\sigma^2} = 10$ dB, $M = 100$ BS antennas, $K = 10$ users and $L = 20$ taps for various values of the oscillator constant, $c$.

practice, the coherence interval is finite and therefore the training overhead upper-bounds the optimum number of scheduled users. Now, consider the case where the coherence interval is arbitrarily long. Then for the no-phase noise case, the optimal $N_D$ is unbounded. In that case one can increase the number of users, thereby achieving an increase in the sum-rate performance. In the presence of phase noise increasing the number of scheduled users, $K$, not only increases the length of the training overhead, but it also increases the phase drift between the estimated channel coefficients and the actual realizations of the effective channel impulse responses during the data interval. That is, by increasing the number of users, $K$, the partial loss of coherency between the estimated channel coefficients and the actual effective channels during data transmission is also increased. As a result, with increasing $K$ the increase in the achievable sum-rate during the data interval may eventually become insignificant to compensate for the reduction in sum-rate due to this partial loss of coherency. In Fig. 7 for every $K$ the maximum achievable sum-rate performance is found by maximizing with respect to $N_D$ and, subsequently, this maximum sum-rate performance is plotted as a function of $K$ for $\frac{P_D}{\sigma^2} = 10$ dB, $M = 80$ BS antennas, $c = 7.4 \times 10^{-17}$ (rad Hz)$^{-1}$ and $L = 20$ taps for the no phase noise case, the synchronous operation mode and the non-synchronous operation mode. It is clear that the sum-rate performance is not monotonically increasing in the phase-noise-impaired cases as it is in the no phase noise case.
This implies that in practice the optimum number of scheduled users is not only upper-bounded by the length of the coherence interval, but it is also upper-bounded as a consequence of the phase noise.

We conclude the discussion with an observation of the dependence of the derived achievable sum-rates on the length of the channel impulse response $L$. In Proposition 3 we state that
the achievable sum-rates (27) and (28) for both operation modes are monotonically decreasing with increasing $L$. We support this statement by Fig. 8. In Fig. 8 for every $L$ we calculate the maximum achievable sum-rate greedily over the choices of $N_D$ and then we plot it as a function of $L$ for fixed $\frac{P_D}{\sigma^2} = 0$ dB, $M = 100$ BS antennas, $K = 10$ users for two different values of the oscillator constant, $c$. The maximum sum-rate with respect to $N_D$ in the no-phase noise case is invariant of $L$, as can be seen from (29), where the maximum is achieved for $N_D \rightarrow \infty$. The no-phase noise curve is also shown in Fig. 8 for comparison. Further, it is clear that the sum-rate performance loss with increasing $L$ is more severe when the quality of the oscillators becomes worse.

VII. Conclusions

Phase noise is an inevitable hardware impairment in communication systems. We studied the effect of phase noise on the sum-rate performance of single-carrier transmission in a MU-MIMO uplink with an excess of BS antennas. Two distinct operation modes in terms of the phase noise processes at the BS antennas are considered, namely, synchronous and non-synchronous operation. Since the knowledge of the exact channel realizations is not available, CSI is acquired via uplink training. The BS uses TR-MRC receive processing to detect the information symbols. An analytical expression for the achievable sum-rate is rigorously derived for both operation modes. Based on the derived achievable sum-rates, we observe that it can be beneficial to use independent instead of fully synchronous phase noise sources. It is also shown that there is a minimal loss in performance in the low SNR regime due to phase noise. Further, the proposed receive processing achieves an $O(\sqrt{M})$ array power gain, extending earlier results where phase noise was not considered. Finally, due to the progressive phase noise drift at the oscillators, there is a fundamental trade-off between the length of the time interval used for data transmission and the sum-rate performance.

Appendix

In this appendix we state the proof of Theorem 2. For both operation modes, we have

$$\text{Var}(\text{EN}_k[i]) \overset{\Delta}{=} \mathbb{E} \left[ |\text{EN}_k[i] - \mathbb{E}[\text{EN}_k[i]]|^2 \right]$$

$$= \text{Var}(\text{IF}_k[i]) + \text{Var}(\text{ISI}_k[i]) + \text{Var}(\text{MUI}_k[i]) + \text{Var}(\text{AN}_k[i])$$
since the terms in $E_{k}^{i}$ are mutually uncorrelated. We start by computing the terms $\text{Var} (\text{ISI}_{k}^{i})$, $\text{Var} (\text{MUI}_{k}^{i})$, $\text{Var} (\text{AN}_{k}^{i})$ for the non-synchronous case, which are the same for both operation modes and conclude with the term $\text{Var} (\text{IF}_{k}^{i})$, the calculation of which is different depending on the operation mode. First we compute the variance of the ISI term.

$$\mathbb{E}[|\text{ISI}_{k}^{i}|^2] = \mathbb{E}[|\sqrt{P_{D}} \sum_{m=1}^{M} \sum_{l=0}^{L-1} \sum_{q=0}^{L-1} g_{m,k,l}^{*} g_{m,k,q} e^{-j(\phi_{m}[i+l]-\phi_{m}[(k-1)L+l]+\theta_{k}[(k-1)L]-\theta_{k}[i+l-q])} \cdot x_{k}[i+l-q]|^2] = P_{D} \sum_{m=1}^{M} \sum_{l=0}^{L-1} \sum_{q=0}^{L-1} \sum_{q' \neq l} \mathbb{E}[g_{m,k,l}^{*} g_{m,q,p} g_{m',k,p'} g_{m',k,l'}] \cdot \mathbb{E}\left[e^{-j(\phi_{m}[i+l]-\phi_{m}[(k-1)L+l]+\phi_{m}[(k-1)L]+\theta_{k}[(k-1)L]})]\right] \cdot \mathbb{E}\left[e^{j(\theta_{k}[i+l-p]-\theta_{k}[(k-1)L]-\theta_{k}[i+l'-p']-\theta_{k}[(k-1)L])}\right] \cdot \mathbb{E}\left[x_{k}[i+l-p] x_{k}^{*}[i+l'-p']\right] = P_{D} \sum_{m=1}^{M} \sum_{l=0}^{L-1} \sum_{q=0}^{L-1} \sum_{q' \neq l} d_{k,l} d_{k,q} = P_{D} M \left(1 - \sum_{l=0}^{L-1} d_{k,l}^2\right),$$

where we have used the fact that the channel coefficients, the phase noise processes and the data symbols are mutually independent. The last step follows from the normalization of the PDP (see (1)). We will make use of these facts in all the following derivations as well. We proceed with the calculation of the multi-user interference.

$$\mathbb{E}[|\text{MUI}_{k}^{i}|^2] = \mathbb{E}[|\sqrt{P_{D}} \sum_{m=1}^{M} \sum_{q=1}^{K} \sum_{l=0}^{L-1} \sum_{p=0}^{L-1} g_{m,k,l}^{*} g_{m,q,p} e^{-j(\phi_{m}[i+l]-\phi_{m}[(k-1)L+l]+\theta_{k}[(k-1)L]-\theta_{q}[i+l-p])} \cdot x_{q}[i+l-p]|^2] = P_{D} \sum_{m=1}^{M} \sum_{m'\neq m} \sum_{q=1}^{K} \sum_{q'\neq q} \sum_{l=0}^{L-1} \sum_{l'\neq l} \sum_{p=0}^{L-1} \sum_{p'\neq p} \mathbb{E}[g_{m,k,l}^{*} g_{m,q,p} g_{m',k,p'} g_{m',k,l'}] \cdot \mathbb{E}\left[e^{-j(\phi_{m}[i+l]-\phi_{m}[(k-1)L+l]+\phi_{m}[(k-1)L]+\theta_{q}[(k-1)L]})]\right] \cdot \mathbb{E}\left[e^{j(\theta_{q}[i+l-p]-\theta_{q}[(k-1)L]-\theta_{q'}[(k-1)L]+\theta_{k}[(k-1)L])}\right] \cdot \mathbb{E}\left[x_{q}[i+l-p] x_{q}^{*}[i+l'-p']\right] = P_{D} \sum_{m=1}^{M} \sum_{q=1}^{K} \sum_{l=0}^{L-1} \sum_{q' \neq q} d_{q,l} d_{q,p} = P_{D} M (K-1),$$

We conclude the first part of the proof with the calculation of the variance of the additive noise term.

$$\mathbb{E}[|\text{AN}_{k}^{i}|^2] = \mathbb{E}[|\sqrt{\frac{P_{D}}{P_{p} K L}} \sum_{m=1}^{M} \sum_{q=1}^{K} \sum_{l=0}^{L-1} \sum_{p=0}^{L-1} g_{m,q,p} e^{-j(\phi_{m}[i+l]-\theta_{q}[i+l-p])} n_{m}[(k-1)L+l] x_{q}[i+l-p]|^2]$$
We proceed by calculating the variance of the term \( E \). Based on the result of Theorem 1 it is sufficient to calculate the mode. We start with the synchronous operation.

\[
\mathbb{E}[g_{m,q,p}e^{-j(\phi_m[i+l]-\theta_q[i+l-p])}n_m[(k-1)L+l]x_q[i+l-p]] \\
\mathbb{E}[(g_{m',q',p'}e^{-j(\phi_{m'}[i+l']-\theta_{q'}[i+l'-p'])}n_{m'}[(k-1)L+l']x_{q'}[i+l'-p'])^*] + \sigma^2 \sum_{m=1}^{M} \sum_{l=0}^{L-1} \mathbb{E}[|\hat{g}_{m,k,l}|^2]
\]

\[
= \frac{P_D \sigma^2}{P_p KL} \sum_{m=1}^{M} \sum_{q=1}^{K} \sum_{l=0}^{L-1} \sum_{a=1-L}^{L-a} d_{q,l-a} + \sigma^2 \sum_{m=1}^{M} \sum_{l=0}^{L-1} \left( \frac{\sigma^2}{P_p KL} + \mathbb{E}[|g_{m,k,l}|^2] \right)
\]

\[
= \sigma^2 M \left( \frac{P_D}{P_p} + \frac{\sigma^2}{P_p K} + 1 \right)
\]

We proceed by calculating the variance of the term \( \mathbb{I}_k[i] \). It holds

\[
\text{Var}(\mathbb{I}_k[i]) = \mathbb{E} \left[ |(A_k[i] - \mathbb{E}[A_k[i]])x_k[i]|^2 \right] = \mathbb{E} \left[ |A_k[i]|^2 \right] - \left( \mathbb{E}[A_k[i]] \right)^2.
\]

Based on the result of Theorem 1 it is sufficient to calculate \( \mathbb{E} \left[ |A_k[i]|^2 \right] \) for each operation mode. We start with the synchronous operation.

\[
\mathbb{E} \left[ |A_k[i]|^2 \right] = P_D \sum_{m=1}^{M} \sum_{l=0}^{L-1} \mathbb{E}[|g_{m,k,l}|^4] + P_D \sum_{m=1}^{M} \sum_{l=0}^{L-1} \mathbb{E}[|g_{m,k,l}|^2] \mathbb{E}[|g_{m,k,l'}|^2]
\]

\[
\cdot \mathbb{E}[e^{-j(\phi[i+l]-\phi[i+l']-\phi[(k-1)L+l]+\phi[(k-1)L+l'])}]
\]

\[
+ P_D \sum_{m=1}^{M} \sum_{m'=1}^{M} \sum_{l=0}^{L-1} \sum_{l'=0}^{L-1} \mathbb{E}[|g_{m,k,l}|^2] \mathbb{E}[|g_{m',k,l'}|^2]
\]

\[
\cdot \mathbb{E}[e^{-j(\phi[i+l]-\phi[i+l']-\phi[(k-1)L+l]+\phi[(k-1)L+l'])}] = P_D M \sum_{l=0}^{L-1} 2d_{k,l}^2
\]

\[
+ P_D M \sum_{l=0}^{L-1} \sum_{l'=0}^{L-1} d_{k,l}d_{k,l'}e^{-\sigma_0^2|l-l'|} + P_D M (M - 1) \sum_{l=0}^{L-1} \sum_{l'=0}^{L-1} d_{k,l}d_{k,l'}e^{-\sigma_0^2|l-l'|}
\]

\[
= P_D M \sum_{l=0}^{L-1} d_{k,l}^2 + P_D M^2 \sum_{l=0}^{L-1} \sum_{l'=0}^{L-1} d_{k,l}d_{k,l'}e^{-\sigma_0^2|l-l'|}
\]
where we have used the fact that \(e^{-\sigma_o^2|i-l'|} \leq 1\). The last step is a consequence of the PDP normalization (see (1)). We conclude with the calculation of the term \(\mathbb{E}[|A_k[i]|^2]\) for the non-synchronous mode.

\[
\mathbb{E}[|A_k[i]|^2] = P_D \sum_{m=1}^{M} \sum_{l=0}^{L-1} \mathbb{E} [|g_{m,k,l}|^4] + P_D \sum_{m=1}^{M} \sum_{l=0}^{L-1} \sum_{l' \neq l} \mathbb{E} [|g_{m,k,l}|^2]\mathbb{E} [|g_{m,k,l'}|^2]
+ \mathbb{E} [e^{-j(\phi_m[i+l] - \phi_m[i+l'] - \phi_m[(k-1)L+l] + \phi_m[(k-1)L+l'])}]
+ P_D \sum_{m=1}^{M} \sum_{m' = 1}^{M} \sum_{l=0}^{L-1} \sum_{l' = 0}^{L-1} \mathbb{E} [|g_{m,k,l}|^2]\mathbb{E} [|g_{m',k,l'}|^2]
+ \mathbb{E} [e^{-j(\phi_m[i+l] - \phi_m[i+l'] - \phi_m[(k-1)L+l] + \phi_m[(k-1)L+l'])}]
= P_D \sum_{l=0}^{L-1} 2d_{k,l}^2
+ P_D \sum_{l=0}^{L-1} \sum_{l' \neq l} d_{k,l}d_{k,l'}e^{-\sigma_o^2|i-l'|} + P_D (M-1) \sum_{l=0}^{L-1} \sum_{l' \neq l} d_{k,l}d_{k,l'}e^{-\sigma_o^2(i-(k-1)L)}
= P_D \sum_{l=0}^{L-1} d_{k,l}^2 + P_D (M-1)e^{-\sigma_o^2(i-(k-1)L)}\]

The same arguments for upper-bounding the term in the synchronous operation mode can be used also for the non-synchronous operation.

**REFERENCES**

[1] A. Pitarokoilis, S. K. Mohammed, and E. G. Larsson, “Effect of oscillator phase noise on the uplink performance of large MU-MIMO systems,” in *50th Allerton Conference on Communication Control and Computing*, pp. 1190–1197, Oct. 2012.

[2] G. Foschini and M. Gans, “On limits of wireless communications in a fading environment when using multiple antennas,” *Wireless Personal Communications*, vol. 6, pp. 311–335, Mar. 1998.

[3] D. Gesbert, M. Kountouris, R. W. Heath Jr., C.-B. Chae, and T. Sälzer, “Shifting the MIMO Paradigm,” *IEEE Signal Processing Magazine*, vol. 24, pp. 36–46, September 2007.
[4] T. L. Marzetta, “Noncooperative cellular wireless with unlimited numbers of base station antennas,” *IEEE Transactions on Wireless Communications*, vol. 9, pp. 3590–3600, Nov. 2010.

[5] F. Rusek, D. Persson, B. K. Lau, E. Larsson, T. Marzetta, O. Edfors, and F. Tufvesson, “Scaling Up MIMO: Opportunities and Challenges with Very Large Arrays,” *IEEE Signal Processing Magazine*, vol. 30, pp. 40–60, Jan. 2013.

[6] E. G. Larsson, F. Tufvesson, O. Edfors, and T. L. Marzetta, “Massive MIMO for Next Generation Wireless Systems,” *submitted to IEEE Communications Magazine, arXiv: 1304.6690*, Apr. 2013.

[7] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, “Energy and spectral efficiency of very large multiuser MIMO systems,” *IEEE Transactions on Communications, (accepted)*, vol. arXiv:1112.3810, 2013.

[8] T. Pollet, M. Van Bladel, and M. Moeneclaey, “BER sensitivity of OFDM systems to carrier frequency offset and wiener phase noise,” *IEEE Transactions on Communications*, vol. 43, pp. 191–193, Feb/Mar/Apr 1995.

[9] L. Tomba, “On the effect of Wiener phase noise in OFDM systems,” *IEEE Transactions on Communications*, vol. 46, pp. 580–583, May 1998.

[10] S. Wu and Y. Bar-Ness, “OFDM systems in the presence of phase noise: consequences and solutions,” *IEEE Transactions on Communications*, vol. 52, pp. 1988–1996, Nov. 2004.

[11] D. Petrovic, W. Rave, and G. Fettweis, “Effects of phase noise on OFDM systems with and without PLL: Characterization and compensation,” *IEEE Transactions on Communications*, vol. 55, pp. 1607–1616, Aug. 2007.

[12] H. Mehrpouyan, A. Nasir, S. Blostein, T. Eriksson, G. Karagiannidis, and T. Svensson, “Joint Estimation of Channel and Oscillator Phase Noise in MIMO Systems,” *IEEE Transactions on Signal Processing*, vol. 60, pp. 4790–4807, Sep. 2012.

[13] A. Lapidoth, “On phase noise channels at high SNR,” in *Proceedings of the 2002 IEEE Information Theory Workshop*, pp. 1–4, Oct. 2002.

[14] G. Durisi, A. Tarable, C. Camarda, and G. Montorsi, “On the capacity of MIMO Wiener phase-noise channels,” in *Proc. Inf. Theory Applicat. Workshop (ITA), San Diego, CA, U.S.A.*, Feb. 2013.

[15] A. Demir, A. Mehrotra, and J. Roychowdhury, “Phase noise in oscillators: a unifying theory and numerical methods for characterization,” *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 47, pp. 655–674, May 2000.

[16] B. Hassibi and B. Hochwald, “How much training is needed in multiple-antenna wireless links?,” *IEEE Transactions on Information Theory*, vol. 49, pp. 951–963, Apr. 2003.

[17] T. L. Marzetta, “How much training is required for multiuser MIMO?,” in *Fortieth Asilomar Conference on Signals, Systems and Computers, 2006. ACSSC ’06.*, pp. 359–363, November 2006.

[18] T. Höhne and V. Ranki, “Phase noise in beamforming,” *IEEE Transactions on Wireless Communications*, vol. 9, pp. 3682–3689, Dec. 2010.

[19] A. Pitarokoilis, S. K. Mohammed, and E. G. Larsson, “On the optimality of single-carrier transmission in large-scale antenna systems,” *IEEE Wireless Communications Letters*, vol. 1, pp. 276–279, August 2012.

[20] D. Petrovic, W. Rave, and G. Fettweis, “Common phase error due to phase noise in OFDM-estimation and suppression,” in *15th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), 2004*, vol. 3, pp. 1901–1905 Vol.3, Sept. 2004.

[21] E. Björnson, P. Zetterberg, M. Bengtsson, and B. Ottersten, “Capacity limits and multiplexing gains of MIMO channels with transceiver impairments,” *IEEE Communications Letters*, vol. 17, pp. 91–94, Jan. 2013.