



CHALMERS

Chalmers Publication Library

Capacity of SIMO and MISO Phase-Noise Channels with Common/Separate Oscillators

This document has been downloaded from Chalmers Publication Library (CPL). It is the author's version of a work that was accepted for publication in:

IEEE Transactions on Communications (ISSN: 0090-6778)

Citation for the published paper:

Khazadi, M. ; Durisi, G. ; Eriksson, T. (2015) "Capacity of SIMO and MISO Phase-Noise Channels with Common/Separate Oscillators". IEEE Transactions on Communications

Downloaded from: <http://publications.lib.chalmers.se/publication/213041>

Notice: Changes introduced as a result of publishing processes such as copy-editing and formatting may not be reflected in this document. For a definitive version of this work, please refer to the published source. Please note that access to the published version might require a subscription.

Chalmers Publication Library (CPL) offers the possibility of retrieving research publications produced at Chalmers University of Technology. It covers all types of publications: articles, dissertations, licentiate theses, masters theses, conference papers, reports etc. Since 2006 it is the official tool for Chalmers official publication statistics. To ensure that Chalmers research results are disseminated as widely as possible, an Open Access Policy has been adopted. The CPL service is administrated and maintained by Chalmers Library.

(article starts on next page)

Capacity of SIMO and MISO Phase-Noise Channels with Common/Separate Oscillators

M. Reza Khanzadi, *Student Member, IEEE*, Giuseppe Durisi, *Senior Member, IEEE*, Thomas Eriksson

Abstract—In multiple antenna systems, phase noise due to instabilities of the radio-frequency (RF) oscillators, acts differently depending on whether the RF circuitries connected to each antenna are driven by separate (independent) local oscillators (SLO) or by a common local oscillator (CLO). In this paper, we investigate the high-SNR capacity of single-input multiple-output (SIMO) and multiple-output single-input (MISO) phase-noise channels for both the CLO and the SLO configurations.

Our results show that the first-order term in the high-SNR capacity expansion is the same for all scenarios (SIMO/MISO and SLO/CLO), and equal to $0.5 \ln(\rho)$, where ρ stands for the SNR. On the contrary, the second-order term, which we refer to as phase-noise number, turns out to be scenario-dependent. For the SIMO case, the SLO configuration provides a diversity gain, resulting in a larger phase-noise number than for the CLO configuration. For the case of Wiener phase noise, a diversity gain of at least $0.5 \ln(M)$ can be achieved, where M is the number of receive antennas. For the MISO, the CLO configuration yields a higher phase-noise number than the SLO configuration. This is because with the CLO configuration one can obtain a coherent-combining gain through maximum ratio transmission (a.k.a. conjugate beamforming). This gain is unattainable with the SLO configuration.

Index Terms—Phase noise, channel capacity, multiple antennas, distributed oscillators, Wiener process.

I. INTRODUCTION

Phase noise due to phase and frequency instability in the local radio-frequency (RF) oscillators used in wireless communication links results in imperfect synchronization between transmitters and receivers, which degrades the system throughput [1]–[4], especially when high-order modulation schemes are used to support high spectral efficiency.

A fundamental way to assess the impact of phase noise on the throughput of wireless links is to determine the corresponding Shannon capacity. Unfortunately, a closed-form expression for the capacity of wireless channels impaired by phase noise is not available (although it is known that the capacity-achieving distribution has discrete amplitude and uniform independent phase when the phase-noise process is stationary and memoryless, with uniform marginal distribution over $[0, 2\pi)$ [5]). Nevertheless, both asymptotic capacity characterizations for large signal-to-noise ratio (SNR) and nonasymptotic capacity bounds

are available in the literature. Specifically, Lapidot [1] characterized the first two terms in the high-SNR expansion of the capacity of a single-input single-output (SISO) stationary phase-noise channel. Focusing on memoryless phase-noise channels, Katz and Shamai [5] provided upper and lower bounds on the capacity that are tight at high SNR. The results in [1], [5] have been generalized to block-memoryless phase-noise channels in [6], [7]. Numerical methods for the calculation of the information rates achievable with specific modulation formats have been proposed in, e.g., [8]–[10].

In multiple-antenna systems, phase noise acts differently depending on whether the RF circuitries connected to each antenna are driven by *separate* (independent) local oscillators (SLO) or by a *common* local oscillator (CLO). Although the CLO configuration is intuitively more appealing because it results in a single phase-noise process to be tracked, the SLO configuration is unavoidable when the spacing between antennas needed to exploit the available spatial degrees of freedom, and, hence, achieve multiplexing or diversity gains, is large [11], [12]. This occurs for example in multiple-antenna line-of-sight microwave backhaul links operating in the 20–40 GHz frequency band, where the spacing between antennas required to exploit the available spatial degrees of freedom can be as large as few meters [4]. In large-antenna-array systems [13]–[15], cost and packaging considerations may also make the SLO configuration attractive.

For the CLO configuration, a high-SNR capacity expansion together with finite-SNR capacity upper and lower bounds have been recently reported in [16], [4]. For both the CLO and the SLO configurations, the multiplexing gain was partly characterized in [17]. In [18], [15], [19], lower bounds on the sum-rate capacity for the case when multiple single-antenna users communicate with a base station equipped with a large antenna array (uplink channel) have been developed for both CLO and SLO. These bounds suggest that the SLO configuration yields a higher sum-rate capacity than the CLO configuration. However, it is unclear whether these lower bounds are tight.

Contributions: We consider the scenario where a multiple-antenna base station communicates with a single-antenna user over an AWGN channel impaired by phase noise and study the first two terms in the high-SNR capacity expansion, for both the uplink (SIMO) and the downlink (MISO) channel, and for both CLO and SLO. We characterize the first term and provide bounds on the second term that are tight for some phase-noise models of practical interest. Our findings are as follows. The first-order term in the high-SNR capacity expansion turns out to be the same in all four scenarios, and equal to $0.5 \ln(\rho)$, where ρ stands for the SNR. In contrast, the second-order term, which we denote as *phase-noise number* (terminology borrowed

This work was partly supported by the Swedish Foundation for Strategic Research under grant SM13-0028.

The material of this paper was presented in part at the 2015 IEEE International Conference on Communications, London, U.K.

M.R. Khanzadi is with the Department of Signals and Systems and the Department of Microtechnology and Nanoscience, Chalmers University of Technology, Gothenburg, Sweden (e-mail: khanzadi@chalmers.se).

G. Durisi and T. Eriksson are with the Department of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden (e-mail: {durisi,thomase}@chalmers.se).

from the fading literature [20]), takes different values in the four cases. For the uplink channel, the SLO phase-noise number is larger than the CLO one. Intuitively, this holds because the SLO configuration provides a diversity gain. For the specific case of Wiener phase noise [21], we show that a diversity gain of at least $0.5 \ln(M)$, where M is the number of receive antennas, can be achieved. This result provides a theoretical justification of the observation reported in [18], [15], [19] that SLO yields a higher sum-rate capacity than CLO for the uplink channel.

For the downlink channel, the ordering is reversed: the CLO configuration results in a higher phase-noise number than the SLO configuration. Coarsely speaking, this holds because CLO allows for maximum-ratio transmission (a.k.a. conjugate beamforming), which yields a coherent-combing gain, whereas this gain is lost in the SLO case. For the case of Wiener phase noise, we determine numerically the extent to which the quality of the local oscillators in the SLO configuration must be improved to overcome the loss of coherent-combing gain.

Our results are derived under the assumption that the continuous-time phase noise process remains constant over the duration of the symbol time. This assumption allows us to obtain a discrete-time equivalent channel model by sampling at Nyquist rate. As shown recently in [22]–[24], by dropping this assumption one may obtain drastically different high-SNR behaviors. In the Wiener phase-noise case, for example, the first-order term in the high-SNR capacity expression was shown in [22] to be at least as large as $0.5 \ln(\rho)$. However, it is unclear whether this lower bound is tight.

Notation: Boldface letters such as \mathbf{a} and \mathbf{A} denote vectors and matrices, respectively. The operator $\text{diag}(\cdot)$, applied to a vector \mathbf{a} , generates a square diagonal matrix having the elements of \mathbf{a} on its main diagonal. With $\mathcal{N}(0, \sigma^2)$ and $\mathcal{CN}(0, \sigma^2)$, we denote the probability distribution of a real Gaussian random variable and of a circularly symmetric complex Gaussian random variable with zero mean and variance σ^2 . Furthermore, $\mathcal{U}[0, 2\pi)$ stands for the uniform distribution over the interval $[0, 2\pi)$, and $\text{Gamma}(\alpha, \beta)$ stands for the Gamma distribution with parameters α and β ; specifically, if s is $\text{Gamma}(\alpha, \beta)$ -distributed, its probability density function (pdf) $q_s(s)$ is

$$q_s(s) = \frac{s^{\alpha-1} e^{-s/\beta}}{\beta^\alpha \Gamma(\alpha)}, \quad s \geq 0 \quad (1)$$

where $\Gamma(\cdot)$ denotes the Gamma function. Throughout the paper, all sums between angles (both random and deterministic) are performed modulus 2π , although this is not always explicitly mentioned so as to keep the notation compact. For a given discrete-time vector-valued random process $\{\boldsymbol{\theta}_k\}$, we denote the sequence $\{\boldsymbol{\theta}_m, \dots, \boldsymbol{\theta}_n\}$, $m < n$ as $\boldsymbol{\theta}_m^n$. When $m = 1$, we omit the subscript. For two functions $f(\cdot)$ and $g(\cdot)$, the notation $f(x) = o(g(x))$, $x \rightarrow \infty$, means that $\lim_{x \rightarrow \infty} |f(x)/g(x)| = 0$. For a given complex vector \mathbf{b} , we denote by $\angle \mathbf{b}$ the vector that contains the phase of the elements of \mathbf{b} . Finally, $\ln(\cdot)$ denotes the natural logarithm.

II. REVIEW OF THE SISO CASE

We start by reviewing the results obtained in [1] for the SISO case. The analysis of the uplink scenario in Section III and of the downlink scenario in Section IV will rely on these results.

Consider the discrete-time SISO phase-noise channel

$$y_k = e^{j\theta_k} h x_k + w_k, \quad k = 1, \dots, n. \quad (2)$$

Here, x_k denotes the input symbol at discrete-time instant k . The constant h is the path-loss coefficient, which is assumed deterministic, time-invariant, and known to the transmitter and the receiver; $\{w_k\}$ are the additive Gaussian noise samples, drawn independently from a $\mathcal{CN}(0, 2)$ distribution.¹ Finally, the phase-noise process $\{\theta_k\}$ is assumed stationary, ergodic, independent of $\{w_k\}$, and with finite differential-entropy rate²

$$h(\{\theta_k\}) > -\infty. \quad (3)$$

Under these assumptions, the capacity of the SISO phase-noise channel (2) is given by

$$C(\rho) = \lim_{n \rightarrow \infty} \frac{1}{n} \sup I(y^n; x^n) \quad (4)$$

where the supremum is over all probability distributions on $x^n = (x_1, \dots, x_n)$ that satisfy the average-power constraint

$$\frac{1}{n} \sum_{k=1}^n \mathbb{E}[|x_k|^2] \leq 2\rho. \quad (5)$$

Here, ρ can be thought of as the SNR (recall that we set the noise variance to 2; hence, the SNR is equal to half the signal power 2ρ). A closed-form expression for the capacity of the phase-noise channel is not available. Lapidoth [1] proved the following asymptotic characterization of $C(\rho)$.

Theorem 1 ([1]): The capacity of the SISO phase-noise channel (2) is given by

$$C(\rho) = \eta \ln(\rho) + \chi + o(1), \quad \rho \rightarrow \infty \quad (6)$$

where $\eta = 1/2$ and

$$\chi = (1/2) \ln(|h|^2/2) + \ln(2\pi) - h(\{\theta_k\}). \quad (7)$$

The factor $\eta = 1/2$ in (6) is the so-called capacity *prelog*, defined as the asymptotic ratio between capacity and the logarithm of SNR as SNR grows to infinity: $\eta = \lim_{\rho \rightarrow \infty} C(\rho)/\ln(\rho)$. The capacity prelog can be interpreted as the fraction of complex dimensions available for communications in the limiting regime of high signal power, or equivalently vanishing noise variance [25]. For the phase-noise channel (2), only the amplitude $|x_k|$ of the transmitted signal x_k can be perfectly recovered in the absence of additive noise, whereas the phase $\angle x_k$ is lost. Hence, the fraction of complex dimensions available for communication is $\eta = 1/2$.

We denote the second term in the high-SNR expansion (6) of $C(\rho)$ as the *phase-noise number* χ

$$\chi = \lim_{\rho \rightarrow \infty} \{C(\rho) - \eta \ln(\rho)\}. \quad (8)$$

We can see from (7) that the phase-noise number depends only on the statistics of the phase-noise process and on the path-loss coefficient h . It is worth mentioning that the approximation

¹As we shall see in, e.g., Appendix I-A, normalizing the noise variance to 2 will turn out convenient.

²Note that the differential-entropy rate of the complex random process $\{e^{j\theta_k}\}$ is equal to $-\infty$. This means that the results obtained in [20] in the context of fading channel are not applicable to (2).

$C(\rho) \approx \eta \ln(\rho) + \chi$, although based on a high-SNR capacity expansion, is often accurate already at low SNR values [5], [7], [4]. Next, we provide closed-form expressions for χ for the phase-noise models commonly used in the wireless literature.

Noncoherent System: Consider the case where the phase-noise process $\{\theta_k\}$ is stationary and memoryless with uniform marginal distribution over $[0, 2\pi)$. This scenario models accurately a noncoherent communication system where the phase of x_k is not used to transmit information (see [5]). The phase-noise number for this case can be readily obtained from (7) by using that $h(\{\theta_k\}) = \ln(2\pi)$.

Partially Coherent System: When a phase tracker such as a phase-locked loop (PLL) is employed at the receiver, the output signal after phase tracking is impaired only by the residual phase error. Systems employing phase trackers are sometimes referred to as partially coherent [5]. It is often accurate to assume that the residual phase-error process $\{\theta_k\}$ is stationary and memoryless. Under this assumption, the phase-noise number for the partially-coherent case simplifies to

$$\chi = (1/2) \ln(|h|^2/2) + \ln(2\pi) - h(\theta) \quad (9)$$

where θ is the random variable modeling the residual phase error. When a PLL is used, the statistics of θ are accurately described by a Tikhonov distribution

$$f_\theta(\theta) = \frac{e^{\lambda \cos \theta}}{2\pi I_0(\lambda)} \quad (10)$$

where $1/\lambda$ is the variance of θ , which depends on the oscillator quality and also on the parameters of the PLL [26]. In this case,

$$h(\theta) = \ln(2\pi I_0(\lambda)) - \lambda I_1(\lambda)/I_0(\lambda) \quad (11)$$

where $I_0(\cdot)$ and $I_1(\cdot)$ stand for the modified Bessel functions of first kind and order 0 and 1, respectively.

The Wiener Process: The case of phase-noise process with memory is relevant when a free-running oscillator is used or when the phase tracker is not able to completely remove the memory of the phase-noise process [21], [27]. The samples $\{\theta_k\}$ of a free-running oscillator are typically modeled using a Wiener process [21], [28], according to which

$$\theta_{k+1} = (\theta_k + \Delta_k) \bmod (2\pi) \quad (12)$$

where $\{\Delta_k\}$ are Gaussian random samples, independently drawn from a $\mathcal{N}(0, \sigma_\Delta^2)$ distribution. Hence, the sequence $\{\theta_k\}$ is a Markov process, i.e.,

$$f_{\theta_k | \theta_{k-1}, \dots, \theta_0} = f_{\theta_k | \theta_{k-1}} = f_\Delta \quad (13)$$

where the wrapped Gaussian distribution

$$f_\Delta(\delta) = \sum_{l=-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma_\Delta^2}} \exp\left(-\frac{(\delta - 2\pi l)^2}{2\sigma_\Delta^2}\right), \quad \delta \in [0, 2\pi) \quad (14)$$

is the pdf of the innovation Δ modulus 2π . Under the assumption that the initial phase-noise sample θ_0 is uniformly distributed over $[0, 2\pi)$, the process $\{\theta_k\}$ is stationary. Hence, its differential-entropy rate is given by the differential entropy of the innovation process

$$h(\{\theta_k\}) = h(\Delta). \quad (15)$$

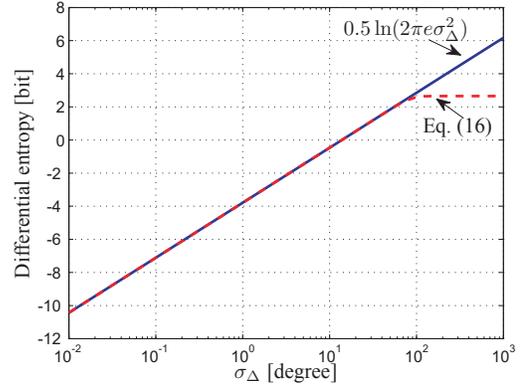


Fig. 1. Differential entropy of a wrapped and an unwrapped Gaussian random variable as a function of its standard deviation.

The phase-noise number of the Wiener phase-noise channel can be readily obtained by using that the differential entropy of the wrapped Gaussian random variable Δ is given by [29]

$$h(\Delta) = -\ln\left(\frac{\varphi(e^{-\sigma_\Delta^2})}{2\pi}\right) + 2 \sum_{n=1}^{\infty} \frac{(-1)^n e^{-\sigma_\Delta^2(n^2+n)/2}}{n(1 - e^{-n\sigma_\Delta^2})} \quad (16)$$

where

$$\varphi(e^{-\sigma_\Delta^2}) = \prod_{l=1}^{\infty} (1 - e^{-l\sigma_\Delta^2}). \quad (17)$$

As shown in Fig. 1, $h(\Delta)$ can be well-approximated by the differential entropy of an unwrapped $\mathcal{N}(0, \sigma_\Delta^2)$ random variable

$$h(\Delta) \approx (1/2) \ln(2\pi e \sigma_\Delta^2) \quad (18)$$

whenever the standard deviation σ_Δ is below 55° . The oscillators commonly used in wireless transceivers result in a phase-noise standard variation that is well below 55° [30, Fig. 2].

III. UPLINK CHANNEL

Building on the results reviewed in Section II, we next analyze the uplink channel of a wireless communication system where a single-antenna terminal communicates with a base station equipped with M antennas over an AWGN channel impaired by phase noise. This yields the following $1 \times M$ single-input multiple-output (SIMO) phase-noise channel:

$$\mathbf{y}_k = \Theta_k \mathbf{h} x_k + \mathbf{w}_k, \quad k = 1, \dots, n. \quad (19)$$

Here, the matrix $\Theta_k = \text{diag}([e^{j\theta_{1,k}}, \dots, e^{j\theta_{M,k}}])$ contains the phase-noise samples. We assume that, for each $m = 1, \dots, M$, the phase-noise process $\{\theta_{m,k}\}$ is stationary, ergodic, independent of the additive-noise process $\{\mathbf{w}_k\}$, and has finite differential-entropy rate. Note that we do not necessarily assume that the phase-noise processes $\{\theta_{m,k}\}$, $m = 1, \dots, M$ are independent. It will turn out convenient to define also the phase-noise vector-valued process $\{\boldsymbol{\theta}_k\}$ where $\boldsymbol{\theta}_k = [\theta_{1,k}, \dots, \theta_{M,k}]^T$. The vector $\mathbf{h} = [h_1, \dots, h_M]^T$ contains the path-loss coefficients, which, similarly to the SISO case, are assumed to be deterministic, time-invariant, and known to the transmitter and the receiver. Finally, the vector $\mathbf{w}_k =$

$[w_{1,k}, \dots, w_{M,k}]^T$ contains the AWGN samples, which are drawn independently from a $\mathcal{CN}(0, 2)$ distribution. Similarly to (4), the capacity of the SIMO phase-noise channel (19) is

$$C(\rho) = \lim_{n \rightarrow \infty} \frac{1}{n} \sup I(\mathbf{y}^n; x^n) \quad (20)$$

where the supremum is over all probability distributions on x^n that satisfy the average-power constraint (5).

A. Uplink, Common Local Oscillator (UL-CLO)

In the CLO configuration, we have that $\theta_{1,k} = \dots = \theta_{M,k} = \theta_k$ for all k . Hence, the input-output relation (19) simplifies to

$$\mathbf{y}_k = e^{j\theta_k} \mathbf{h}x_k + \mathbf{w}_k. \quad (21)$$

By projecting \mathbf{y}_k on $\mathbf{h}/\|\mathbf{h}\|$, i.e., by performing coherent/maximal-ratio combining, we obtain a sufficient statistics for the detection of x_k from \mathbf{y}_k . Through this projection, the SIMO phase-noise channel (21) is transformed into an equivalent SISO phase-noise channel with channel gain $\|\mathbf{h}\|$. Therefore, using Theorem 1, we conclude that the prelog for the UL-CLO case is $\eta_{\text{ul-clo}} = 1/2$ and that the phase-noise number is

$$\chi_{\text{ul-clo}} = (1/2) \ln(\|\mathbf{h}\|^2/2) + \ln(2\pi) - h(\{\theta_k\}). \quad (22)$$

B. Uplink, Separate Local Oscillators (UL-SLO)

In the SLO case, the M phase-noise processes $\{\theta_{m,k}\}$, are independent and identically distributed (i.i.d.) across the receive antennas. Hence, coherent combining does not yield a sufficient statistics. In Theorem 2 below, we provide a characterization of the high-SNR capacity of $C(\rho)$, which holds irrespectively of the dependency between the M phase-noise processes $\{\theta_{m,k}\}$, $m = 1, \dots, M$.

Theorem 2: The prelog of the SIMO phase-noise channel (19) is given by $\eta_{\text{ul}} = 1/2$. Furthermore, the phase-noise number is bounded by

$$\chi_{\text{ul}} \geq (1/2) \ln(\|\mathbf{h}\|^2/2) + \ln(2\pi) - h(\phi_0 | \boldsymbol{\theta}_0 + \phi_0, \boldsymbol{\theta}_{-\infty}^{-1}) \quad (23a)$$

$$\chi_{\text{ul}} \leq (1/2) \ln(\|\mathbf{h}\|^2/2) + \ln(2\pi) - h(\phi_0 | \boldsymbol{\theta}_0 + \phi_0) + I(\boldsymbol{\theta}_0; \boldsymbol{\theta}_{-\infty}^{-1}) \quad (23b)$$

where $\{\phi_k\}$ is a stationary memoryless process, with marginal distribution uniform over $[0, 2\pi)$.

Proof: See Appendix I. ■

Remark 1: A more accurate characterization of the phase-noise number may be obtained by adapting to the case of phase noise the tools developed in [31] for the analysis of stationary SIMO fading channels at high SNR. We leave this refinement for future work.

Remark 2: The upper and lower bounds in (23) match when the phase noise processes are memoryless. Indeed, under this assumption,

$$\chi_{\text{ul}} = (1/2) \ln(\|\mathbf{h}\|^2/2) + \ln(2\pi) - h(\phi_0 | \boldsymbol{\theta}_0 + \phi_0). \quad (24)$$

Remark 3: For the CLO case where $\theta_{1,k} = \dots = \theta_{M,k} = \theta_k$ for all k , the bounds in (23) match and reduce to (22). Indeed, for the lower bound we have that

$$\chi_{\text{ul}} \geq \frac{1}{2} \ln\left(\frac{\|\mathbf{h}\|^2}{2}\right) + \ln(2\pi) - h(\phi_0 | \boldsymbol{\theta}_0 + \phi_0, \boldsymbol{\theta}_{-\infty}^{-1}) \quad (25)$$

$$= \frac{1}{2} \ln\left(\frac{\|\mathbf{h}\|^2}{2}\right) + \ln(2\pi) - h(\phi_0 | \theta_0 + \phi_0, \theta_{-\infty}^{-1}) \quad (26)$$

$$= \frac{1}{2} \ln\left(\frac{\|\mathbf{h}\|^2}{2}\right) + I(\phi_0; \theta_0 + \phi_0 | \theta_{-\infty}^{-1}) \quad (27)$$

$$= \frac{1}{2} \ln\left(\frac{\|\mathbf{h}\|^2}{2}\right) + h(\theta_0 + \phi_0 | \theta_{-\infty}^{-1}) - h(\theta_0 + \phi_0 | \theta_{-\infty}^{-1}, \phi_0) \quad (28)$$

$$= \frac{1}{2} \ln\left(\frac{\|\mathbf{h}\|^2}{2}\right) + \ln(2\pi) - h(\theta_0 | \theta_{-\infty}^{-1}) = \chi_{\text{ul-clo}}. \quad (29)$$

Here, in (29) we used that $(\theta_0 + \phi_0) \sim \mathcal{U}[0, 2\pi)$, which holds because $\phi_0 \sim \mathcal{U}[0, 2\pi)$. For the upper bound, we have

$$\chi_{\text{ul}} \leq \frac{1}{2} \ln\left(\frac{\|\mathbf{h}\|^2}{2}\right) + \ln(2\pi) - h(\phi_0 | \boldsymbol{\theta}_0 + \phi_0) + I(\boldsymbol{\theta}_0; \boldsymbol{\theta}_{-\infty}^{-1}) \quad (30)$$

$$= \frac{1}{2} \ln\left(\frac{\|\mathbf{h}\|^2}{2}\right) + \ln(2\pi) - h(\phi_0 | \theta_0 + \phi_0) + I(\theta_0; \theta_{-\infty}^{-1}) \quad (31)$$

$$= \frac{1}{2} \ln\left(\frac{\|\mathbf{h}\|^2}{2}\right) + I(\phi_0; \theta_0 + \phi_0) + I(\theta_0; \theta_{-\infty}^{-1}) \quad (32)$$

$$= \frac{1}{2} \ln\left(\frac{\|\mathbf{h}\|^2}{2}\right) + h(\theta_0 + \phi_0) - h(\theta_0 + \phi_0 | \phi_0) + h(\theta_0) - h(\theta_0 | \theta_{-\infty}^{-1}) \quad (33)$$

$$= \frac{1}{2} \ln\left(\frac{\|\mathbf{h}\|^2}{2}\right) + \ln(2\pi) - h(\theta_0 | \theta_{-\infty}^{-1}) = \chi_{\text{ul-clo}}. \quad (34)$$

C. Discussion

The fact that $\eta_{\text{ul-clo}} = \eta_{\text{ul-slo}} = \eta_{\text{siso}}$ comes perhaps as no surprise because adding multiple antennas at the receiver only (SIMO channel) does not yield spacial multiplexing gains. We next compare the phase-noise number of the CLO and the SLO configurations. We see from (22) and (23) that the term $0.5 \ln(\|\mathbf{h}\|^2/2)$ appears in the phase-noise number of both the CLO and SLO configuration. As already pointed out, in the CLO case this term comes from coherently combining the signals received at the M antennas. Coherent combining is possible because, in the CLO case, the received signals at the different antennas are phase-shifted by the same random quantity. In the SLO case, however, coherent combining is not possible because the received signals at the different antennas are subject to independent random phase shifts. It turns out (see Appendix I-A and Appendix I-B) that a coherent-combining gain can be harvested regardless by separately decoding the amplitude and the phase of the transmitted signal, and by adding the square of the received signals when decoding the amplitude.

The CLO and SLO phase-noise numbers coincide in the non-coherent case (stationary, memoryless phase noise, with uniform marginal distribution over $[0, 2\pi)$):

$$\chi_{\text{ul-clo}} = \chi_{\text{ul-slo}} = (1/2) \ln(\|\mathbf{h}\|^2/2). \quad (35)$$

For phase-noise processes with memory, the CLO configuration results in a smaller phase-noise number than the SLO configuration. Indeed by rewriting the second and the third term on the RHS of (22) as follows

$$\ln(2\pi) - h(\{\theta_k\}) = I(\theta_0 + \phi_0; \phi_0 | \theta_{-\infty}^{-1}) \quad (36)$$

$$= \ln(2\pi) - h(\phi_0 | \phi_0 + \theta_0, \theta_{-\infty}^{-1}) \quad (37)$$

where $\phi_0 \sim \mathcal{U}(0, 2\pi]$ is independent of $\{\theta_k\}$, we see that the differential entropy on the RHS of (37) is larger than the differential entropy in the SLO phase-noise lower bound in (23a). To shed further light on the difference between the CLO and the SLO configuration, we now consider the special case of Wiener phase noise. For the CLO configuration, by substituting (18) in (15), and then (15) in (22) we obtain

$$\chi_{\text{ul-clo}} \approx (1/2) \ln(\|\mathbf{h}\|^2/2) + \ln(2\pi) - (1/2) \ln(2\pi e \sigma_{\Delta}^2). \quad (38)$$

For the SLO configuration, we manipulate the lower-bound in (23a) as follows:

$$\chi_{\text{ul-slo}} \geq \frac{1}{2} \ln\left(\frac{\|\mathbf{h}\|^2}{2}\right) + \ln(2\pi) - h(\phi_0 | \theta_0 + \phi_0, \theta_{-\infty}^{-1}) \quad (39)$$

$$= \frac{1}{2} \ln\left(\frac{\|\mathbf{h}\|^2}{2}\right) + \ln(2\pi) - h(\phi_0 | \{\phi_0 + \Delta_{m,-1}\}_{m=1}^M) \quad (40)$$

$$= (1/2) \ln(\|\mathbf{h}\|^2/2) + \ln(2\pi) - h\left(\phi_0 | \{\phi_0 + \Delta_{m,-1}\}_{m=1}^M, \phi_0 + \frac{1}{M} \sum_{m=1}^M \Delta_{m,-1}\right) \quad (41)$$

$$\geq (1/2) \ln(\|\mathbf{h}\|^2/2) + \ln(2\pi) - h\left(\phi_0 | \phi_0 + \frac{1}{M} \sum_{m=1}^M \Delta_{m,-1}\right) \quad (42)$$

$$= \frac{1}{2} \ln\left(\frac{\|\mathbf{h}\|^2}{2}\right) + \ln(2\pi) - h\left(\frac{1}{M} \sum_{m=1}^M \Delta_{m,-1}\right) \quad (43)$$

$$\approx (1/2) \ln(\|\mathbf{h}\|^2/2) + \ln(2\pi) - (1/2) \ln(2\pi e \sigma_{\Delta}^2/M). \quad (44)$$

Here, (40) follows by (15) and by the Markov property of the Wiener process; (41) holds because ϕ_0 and $\phi_0 + \sum_{m=1}^M \Delta_{m,-1}/M$ are conditionally independent given $\{\phi_0 + \Delta_{m,-1}\}_{m=1}^M$; in (42) we used again that conditioning does not increase differential entropy; finally, (44) follows from (18).

By comparing (38) and (44) we see that for the Wiener phase-noise case, $\chi_{\text{ul-slo}} \geq \chi_{\text{ul-clo}}$. This gain can be explained as follows: in the SLO case, we have M independent noisy observations of the phase of the transmitted signal. These independent noisy observations can be used to improve the estimation of the transmitted phase. Specifically, a diversity gain at least equal to $0.5 \ln(M)$ can be achieved by using separate oscillators instead of a common oscillator. Equivalently, in order to obtain equal phase-noise numbers in the CLO and SLO configurations, the phase-noise variance σ_{Δ}^2 in the CLO case must be at least M times lower than in the SLO case. For large-antenna arrays, we expect the throughput gains resulting from the SLO configuration to occur only at very high SNR. Indeed, whereas in the CLO case the phase-noise tracker can leverage on the (large) antenna array gain, which yields a fast convergence to the high-SNR asymptotics, this is not the case in the SLO configuration, where each phase-noise process needs to be tracked separately,

without relying on any array gain. It is perhaps also worth mentioning that the SLO gains cannot be achieved in the CLO case simply by independently phase-shifting the signal received at each antenna. In fact, this strategy does not even achieve the CLO phase-noise number (22).

A configuration that is perhaps more relevant from a practical point of view is the one where the M phase-noise processes $\{\theta_{m,k}\}$, $m = 1, \dots, M$ result from the sum of the phase-noise contribution $\theta_k^{(\text{tx})}$ at the transmitter and of M independent phase-noise contributions $\{\theta_{m,k}^{(\text{rx})}\}$, $m = 1, \dots, M$ at the receivers. Assuming that both $\theta_k^{(\text{tx})}$ and $\{\theta_{m,k}^{(\text{rx})}\}$ evolve according to independent Wiener processes with iid innovations $\Delta_k^{(\text{tx})} \sim \mathcal{N}(0, \sigma_{\Delta, \text{tx}}^2)$ and $\Delta_{m,k}^{(\text{rx})} \sim \mathcal{N}(0, \sigma_{\Delta, \text{rx}}^2)$, we obtain

$$\chi_{\text{ul-slo}} \geq (1/2) \ln(\|\mathbf{h}\|^2/2) + \ln(2\pi) - h\left(\Delta_{-1}^{(\text{tx})} + \frac{1}{M} \sum_{m=1}^M \Delta_{m,-1}^{(\text{rx})}\right) \quad (45)$$

$$\approx (1/2) \ln(\|\mathbf{h}\|^2/2) + \ln(2\pi) - (1/2) \ln(2\pi e (\sigma_{\Delta, \text{tx}}^2 + \sigma_{\Delta, \text{rx}}^2/M)). \quad (46)$$

Here, (46) follows by proceeding as in (39)–(44). The case where a single oscillator is used at the receiver can be obtained from (46) by setting $M = 1$. Also for this setup, we see that using independent oscillators at the receiver is advantageous, although the gain is smaller than what suggested by (44).

IV. DOWNLINK CHANNEL

We next analyze the downlink channel, i.e., the scenario where a base station equipped with M antennas communicates with a single-antenna terminal. This yields the following $M \times 1$ multiple-input single-output (MISO) phase-noise channel

$$y_k = \mathbf{h}^T \Theta_k \mathbf{x}_k + w_k. \quad (47)$$

Here, the phase-noise process $\{\Theta_k\}$ and the path-loss vector \mathbf{h} are defined as in Section III; $\mathbf{x}_k = [x_{1,k}, \dots, x_{M,k}]^T$, where $x_{m,k}$ denotes the symbol transmitted from antenna m at time instant k ; finally, $\{w_k\}$ is the additive noise process, with samples drawn independently from a $\mathcal{CN}(0, 2)$ distribution. Similarly to (4) and (20), the capacity of the MISO phase-noise channel (47) is

$$C(\rho) = \lim_{n \rightarrow \infty} \frac{1}{n} \sup I(y^n; \mathbf{x}^n) \quad (48)$$

where the supremum is over all probability distributions on $\mathbf{x}^n = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ that satisfy the average-power constraint

$$\frac{1}{n} \sum_{k=1}^n \mathbb{E}[\|\mathbf{x}_k\|^2] \leq 2\rho. \quad (49)$$

A. Downlink, Common Local Oscillator (DL-CLO)

In the CLO case, we have that $\theta_{1,k} = \dots = \theta_{M,k} = \theta_k$ for all k . Hence, the input-output relation (47) simplifies to

$$y_k = e^{j\theta_k} \mathbf{h}^T \mathbf{x}_k + w_k. \quad (50)$$

Maximum ratio transmission, i.e., setting $\mathbf{x}_k = s_k \mathbf{h}^* / \|\mathbf{h}\|$, with $\{s_k\}$ chosen so that (49) holds is capacity achieving. Indeed, set $s_k = \mathbf{h}^T \mathbf{x}_k / \|\mathbf{h}\|$. Then

$$I(y_k; \mathbf{x}_k) = I(y_k; \mathbf{x}_k, s_k) - I(y_k; s_k | \mathbf{x}_k) \quad (51)$$

$$= I(y_k; s_k) + \underbrace{I(y_k; \mathbf{x}_k | s_k)}_{=0} - I(y_k; s_k | \mathbf{x}_k) \quad (52)$$

$$\leq I(y_k; s_k). \quad (53)$$

Here, the first equality follows from the chain rule for mutual information, the second equality follows because y_k and \mathbf{x}_k are conditionally independent given s_k (see (50)), and in the last step we used that mutual information is nonnegative. Note now that the upper bound is tight whenever $I(y_k; s_k | \mathbf{x}_k) = 0$. This is achieved by choosing \mathbf{x}_k so that the transformation $\mathbf{x}_k \mapsto s_k$ is invertible. This implies that, in order to achieve capacity, one should set $\mathbf{x}_k = s_k \mathbf{h}^* / \|\mathbf{h}\|$.

With conjugate beamforming, the MISO channel is transformed into a SISO channel. Hence, Theorem 1 allows us to conclude that, for the DL-CLO case, $\eta_{\text{dl-clo}} = 1/2$ and

$$\chi_{\text{dl-clo}} = (1/2) \ln(\|\mathbf{h}\|^2/2) + \ln(2\pi) - h(\{\theta_k\}). \quad (54)$$

B. Downlink, Separate Local Oscillator (DL-SLO)

In Theorem 3 below, we characterize the prelog and provide bounds on the phase-noise number of the MISO phase-noise channel (47). Afterwards, we shall discuss specific phase-noise models for which the bounds on the phase-noise number are tight. Note that Theorem 3 holds irrespectively of the dependency between the M phase-noise processes $\{\theta_{m,k}\}$.

Theorem 3: The prelog of the MISO phase-noise channel (47) is given by $\eta_{\text{dl}} = 1/2$. Furthermore, the phase-noise number is bounded by

$$\chi_{\text{dl}} \geq \ln(2\pi) + \max_{m=1, \dots, M} \left\{ \frac{1}{2} \ln \left(\frac{|h_m|^2}{2} \right) - h(\{\theta_{m,k}\}) \right\} \quad (55a)$$

$$\chi_{\text{dl}} \leq \ln(2\pi) + \sup_{\|\hat{\mathbf{x}}\|=1} \left\{ \frac{1}{2} \ln \left(\frac{1}{2} \mathbb{E} \left[|\mathbf{h}^T \Theta_0 \hat{\mathbf{x}}|^2 \right] \right) \right\} - \inf_{\|\hat{\mathbf{x}}\|=1} \left\{ h(\underline{\mathbf{h}^T \Theta_0 \hat{\mathbf{x}} | \theta_{-\infty}^{-1}}) \right\} \quad (55b)$$

where $\hat{\mathbf{x}}$ is a unit-norm vector in \mathbb{C}^M .

Proof: See Appendix II. ■

Remark 4: The lower bound on the phase-noise number in (55a) is achieved by antenna selection, i.e., by activating only the transmit antenna that leads to the largest SISO phase-noise number. The other $M - 1$ transmit antennas are switched off.

Remark 5: The bounds on the phase-noise number reported in (55) may be tightened using the tools developed in [32], [33] in the context of MIMO fading channels. This tightening is left for future work.

Remark 6: In the CLO case where $\theta_{1,k} = \dots = \theta_{1,k} = \theta_k$ for all k , the upper bound in (55b) is tight. Indeed,

$$\chi_{\text{dl}} \leq \ln(2\pi) + \sup_{\|\hat{\mathbf{x}}\|=1} \left\{ \frac{1}{2} \ln \left(\frac{1}{2} \mathbb{E} \left[|\mathbf{h}^T \Theta_0 \hat{\mathbf{x}}|^2 \right] \right) \right\} - \inf_{\|\hat{\mathbf{x}}\|=1} \left\{ h(\underline{\mathbf{h}^T \Theta_0 \hat{\mathbf{x}} | \theta_{-\infty}^{-1}}) \right\} \quad (56)$$

$$= \ln(2\pi) + \sup_{\|\hat{\mathbf{x}}\|=1} \left\{ \frac{1}{2} \ln \left(\frac{1}{2} |\mathbf{h}^T \hat{\mathbf{x}}|^2 \right) \right\} - \inf_{\|\hat{\mathbf{x}}\|=1} \left\{ h(\underline{\mathbf{h}^T \hat{\mathbf{x}} + \theta_0 | \theta_{-\infty}^{-1}}) \right\} \quad (57)$$

$$= \ln(2\pi) + (1/2) \ln(\|\mathbf{h}\|^2/2) - h(\{\theta_k\}) = \chi_{\text{dl-clo}}. \quad (58)$$

Note that the lower bound in (55a) is not tight because antenna selection is not optimal for the CLO case. The upper and lower bounds in (55) match when the phase-noise processes are independent across antennas, and have uniform marginal distributions over $[0, 2\pi)$. This occurs in noncoherent systems and for the Wiener model. We formalize this result below.

Theorem 4: The phase-noise number of the MISO phase-noise channel (47) under the additional assumptions that the M phase-noise processes $\{\theta_{m,k}\}$, $m = 1, \dots, M$ i) are independent and identically distributed (i.i.d.) across antennas (SLO configuration), ii) have uniform marginal distributions over $[0, 2\pi)$, is given by

$$\chi_{\text{dl-slo}} = \frac{1}{2} \max_{m=1, \dots, M} \ln \left(\frac{|h_m|^2}{2} \right) + \ln(2\pi) - h(\{\theta_k\}) \quad (59)$$

where $h(\{\theta_k\})$ is the differential-entropy rate of one of the i.i.d. phase-noise processes.

Proof: See Appendix III. ■

We next compare $\chi_{\text{dl-clo}}$ and $\chi_{\text{dl-slo}}$. For the noncoherent case, we have that

$$\chi_{\text{dl-clo}} = (1/2) \ln(\|\mathbf{h}\|^2/2) \quad (60)$$

$$\chi_{\text{dl-slo}} = (1/2) \max_{m=1, \dots, M} \ln(|h_m|^2/2). \quad (61)$$

For the Wiener case, by substituting (18) in (15) and then (15) in (54) and in (59), we obtain

$$\chi_{\text{dl-clo}} \approx (1/2) \ln(\|\mathbf{h}\|^2/2) + \ln(2\pi) - (1/2) \ln(2\pi e \sigma_{\Delta}^2) \quad (62)$$

$$\chi_{\text{dl-slo}} \approx (1/2) \max_{m=1, \dots, M} \ln(|h_m|^2/2) + \ln(2\pi) - (1/2) \ln(2\pi e \sigma_{\Delta}^2). \quad (63)$$

In both the noncoherent and the Wiener case, we see that the SLO configuration results in no coherent-combining gain: $\|\mathbf{h}\|^2$ is replaced by $\max_{m=1, \dots, M} |h_m|^2$. The resulting throughput loss is most pronounced when the entries of \mathbf{h} have all the same magnitude. To shed further light on this loss, we depart from the model we considered so far, where the $\{h_m\}$, $m = 1, \dots, M$ are deterministic, and move to a quasi-static fading model [34, p. 2631], [35, Sec. 5.4.1], where the $\{h_m\}$ are independently drawn from a $\mathcal{CN}(0, 1)$ distribution and stay constant over the duration of a codeword. We also assume that the $\{h_m\}$ are perfectly known to the transmitter and the receiver. In this scenario, $0.5 \ln(\rho) + \chi$, where χ is now a function of the instantaneous

channel gains, is the rate supported by the channel in the high-SNR regime, for a given channel realization.

In Fig. 2a, we plot the cumulative distribution function of $0.5 \ln(\rho) + \chi$, which is a high-SNR approximation of the outage capacity. We consider the case of Wiener phase noise with standard deviation $\sigma_\Delta = 6^\circ$ and set $M = 20$ and $\rho = 20$ dB. For a given outage probability, the rate supported in the SLO case is smaller than that in the CLO case. For example, for a target outage probability of $\varepsilon = 0.1$, the rate supported in the SLO case is 1.36 bit/channel use lower than that in the CLO case. To achieve the same rate at $\varepsilon = 0.1$, the standard deviation σ_Δ of the phase-noise process in the SLO case must be set to 2.34° .

Fig. 2b shows the difference between the outage capacity in the CLO and the SLO cases, for a target outage probability $\varepsilon = 0.1$, as a function of the number of antennas. As expected, the gap increases as the number of antennas get large. Note that the gap does not depend on the variance of the phase-noise processes (see (54) and (59))

V. CONCLUSIONS

We studied the capacity of multiple-antenna systems affected by phase noise. Specifically, we analyzed the first two terms in the high-SNR expansion of the capacity of both the uplink and the downlink channel of a system where wireless communications occur between a base station equipped with M antennas and a single-antenna user. Our analysis covers two different configurations: the case when the RF circuitries connected to each antenna at the base station are driven by separate local oscillators, and the case when a common oscillator drives all the antennas.

For all four cases (uplink/downlink, common/separate oscillators) the first term in the high-SNR capacity expansion is equal to $0.5 \ln(\rho)$, whereas the second term, which we denote as phase noise number, turns out to take different values depending on which case is considered. For the uplink channel, the phase noise number is larger when separate oscillators are used. For the specific case of Wiener phase noise, a gain of at least $0.5 \ln(M)$ can be achieved. This gain, which is due to diversity, implies that to achieve the same throughput in the high-SNR regime, the oscillator used in the common oscillator configuration must be at least M times better than any of the oscillators used in the separate configuration.

In contrast, the phase noise number of the downlink channel is larger when a common oscillator drives all the antennas. This is due to the fact that conjugate beamforming, which provides a coherent-combining gain for the common oscillator configuration, does not achieve the phase-noise number when separate oscillators are used. The capacity achieving-strategy for the separate oscillator configuration turns out to be antenna selection, i.e., activating only the transmit antenna that yields the largest SISO high-SNR capacity, and switching off all other antennas.

APPENDIX I PROOF OF THEOREM 2

As the proof consists of several steps, we organized it in three subsections. In Appendix I-A, we prove Theorem 2 for the special case of stationary memoryless phase-noise processes

$\{\theta_{m,k}\}$, $m = 1, \dots, M$, with uniform marginal distribution over $[0, 2\pi)$ (noncoherent system). Building on this result, in Appendix I-B we generalize the proof to the case of arbitrary stationary memoryless phase-noise processes (partially coherent system). For these first two cases, our bounds are tight, and the phase noise number is characterized in closed form. Finally, we tackle the case of phase-noise processes with memory in Section I-C. Before we proceed further, we state in Lemma 5 below a property of the capacity-achieving input distribution, which will be used throughout this section.

Lemma 5: The process $\{x_k\}$ that achieves the capacity of the channel (19) can be assumed circularly symmetric, i.e., $\{\underline{x}_k\}$ can be taken to be a stationary memoryless process with uniform marginal distribution over $[0, 2\pi)$ and independent of $\{|x_k|\}$.

Proof: The proof follows along the same lines as the proof of [33, Prop. 7]. ■

A. Noncoherent System

We focus on the case where the M phase-noise processes $\{\theta_{m,k}\}$, $m = 1, \dots, M$ are stationary memoryless with uniform marginal distribution over $[0, 2\pi)$. Stationarity and lack of memory imply that the time index k in (19) can be dropped and that the capacity expression in (20) simplifies to

$$C(\rho) = \sup I(\mathbf{y}; x) \quad (64)$$

where

$$\mathbf{y} = \Theta \mathbf{h}x + \mathbf{w} \quad (65)$$

and where the supremum in (64) is over all probability distributions on x that satisfy

$$\mathbb{E}[|x|^2] \leq 2\rho. \quad (66)$$

We show next that the prelog and phase-noise number of the SIMO channel (65) are

$$\eta = 1/2, \quad \chi = (1/2) \ln(\|\mathbf{h}\|^2/2). \quad (67)$$

To prove (67), we establish a lower bound and an upper bound on $C(\rho)$ that match up to a $o(1)$ term as $\rho \rightarrow \infty$.

Lower bound: To lower-bound $C(\rho)$, we evaluate the mutual information on the right-hand side (RHS) of (64) for a specific probability distribution, namely, we choose \underline{x} to be uniformly distributed over $[0, 2\pi)$, and we set $|x|^2 = 4\rho s$ where the random variable s is independent of \underline{x} and is Gamma(1/2, 1)-distributed, i.e., its pdf $f_S(s)$ is given by

$$f_S(s) = e^{-s}/\sqrt{\pi s}, \quad s \geq 0. \quad (68)$$

Note that, with this choice, (66) holds with equality because $\mathbb{E}[s] = 1/2$. Let $t = \|\mathbf{y}\|^2$. The mutual information in (64) can be lower-bounded as follows

$$I(\mathbf{y}; x) \geq I(t; x) = h(t) - h(t|x). \quad (69)$$

Here, (69) follows from the data-processing inequality [36, Th. 2.8.1]. Observe now that, given x , the random variable t

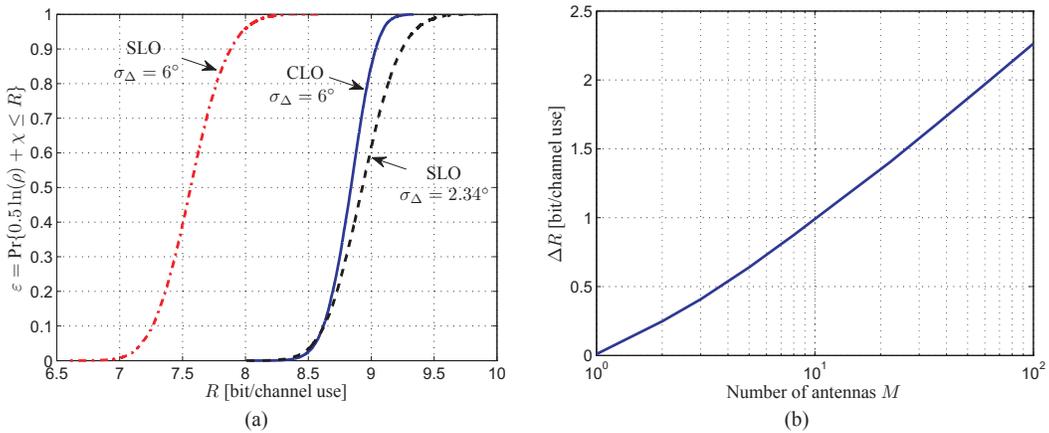


Fig. 2. (a) High-SNR approximation of the outage probability for the CLO and the SLO configurations. A Wiener phase-noise model is considered. Furthermore, $M = 20$, and $\rho = 20$ dB. (b) Difference ΔR between the outage capacity of the CLO and the SLO configurations for a target outage probability of $\varepsilon = 0.1$, as a function of the number of antennas.

follows a noncentral chi-squared distribution³ with $2M$ degrees of freedom and noncentrality parameter $\|\mathbf{h}x\|^2$, i.e.,

$$t \sim \left\| \|\mathbf{h}x\| + w_1 \right\|^2 + \sum_{m=2}^M |w_m|^2 \quad (70)$$

where \sim denotes equality in distribution and $\{w_m\}$, $m = 1, \dots, M$ are i.i.d. $\mathcal{CN}(0, 2)$ -distributed.

Let ϕ be uniformly distributed over $[0, 2\pi)$ and independent of x and of the $\{w_m\}$. Using (70), we lower-bound the first term in (69) as follows:

$$h(t) = h\left(\left\| \|\mathbf{h}x\| + w_1 \right\|^2 + \sum_{m=2}^M |w_m|^2\right) \quad (71)$$

$$\geq h\left(\left\| \|\mathbf{h}x\| + w_1 \right\|^2\right) \quad (72)$$

$$= h(e^{j\phi} \|\mathbf{h}x\| + w_1) - \ln(\pi) \quad (73)$$

$$\geq h(e^{j\phi} \|\mathbf{h}x\|) - \ln(\pi) \quad (74)$$

$$= h(\|\mathbf{h}x\|^2) \quad (75)$$

$$= h(s) + \ln(4\rho) + \ln(\|\mathbf{h}\|^2) \quad (76)$$

$$= \frac{1}{2}\psi\left(\frac{1}{2}\right) + \frac{1}{2} + \frac{1}{2}\ln(\pi) + \ln(4\rho) + \ln(\|\mathbf{h}\|^2). \quad (77)$$

To obtain (73), we used first that $h(w) = h(|w|^2) + \ln(\pi)$ for every circularly symmetric random variable w [20, Eq. (320)], and then that w_1 is circularly symmetric, which implies that $e^{j\phi}w_1 \sim w_1$; (75) follows again from [20, Eq. (320)]; in (76) we used that $h(aq) = h(q) + \ln a$ for every real-valued random variable q and every positive constant a , and also that $|x|^2 = 4\rho s$; finally, (77) holds because the differential entropy of the Gamma-distributed random variable s is [1, Eq. (19)]

$$h(s) = (1/2)\psi(1/2) + 1/2 + (1/2)\ln(\pi) \quad (78)$$

³Normalizing the noise variance to 2 (see Section II) is crucial to obtain a chi-squared distribution. The more common unitary normalization for the noise variance would result in an additional constant factor, which is tedious to track.

with $\psi(\cdot)$ denoting Euler's digamma function. We next upper-bound the second term of (69) by proceeding as in [1]:

$$\begin{aligned} h(t|x) &\leq \frac{1}{2} \mathbb{E}[\ln(8\pi e(M + 4\rho s \|\mathbf{h}\|^2))] \quad (79) \\ &= \frac{1}{2} \ln(32\pi e \|\mathbf{h}\|^2 \rho) + \frac{1}{2} \psi\left(\frac{1}{2}\right) + o(1), \quad \rho \rightarrow \infty. \quad (80) \end{aligned}$$

In (79) we used that the conditional variance of t given $|x|^2 = 4\rho s$ is $4(M + 4\rho s \|\mathbf{h}\|^2)$ and that the Gaussian distribution maximizes differential entropy under a variance constraint; (80) follows because $\mathbb{E}[\ln s] = \psi(1/2)$ [1, Eq. (18)]. Substituting (77) and (80) into (69) and then (69) into (64), we obtain

$$C(\rho) \geq \frac{1}{2} \ln(\rho) + \frac{1}{2} \ln\left(\frac{\|\mathbf{h}\|^2}{2}\right) + o(1), \quad \rho \rightarrow \infty. \quad (81)$$

Upper Bound: Since the RHS of (81) coincides with the asymptotic capacity expansion for the UL-CLO case (set $h(\{\theta_k\}) = \ln(2\pi)$ in (22)), to establish (67) it is sufficient to show that the capacity in the UL-CLO case is no smaller than that in the UL-SLO case. Let ϕ be uniformly distributed over $[0, 2\pi)$ and independent of all other random variables in (65). Furthermore, let

$$\tilde{\mathbf{y}} = e^{j\phi} \mathbf{h}x + \mathbf{w} \quad (82)$$

where $\tilde{\mathbf{y}} = [\tilde{y}_1, \dots, \tilde{y}_M]^T$ and all the other quantities are defined as in (65). Recall also that, by definition, $\Theta = \text{diag}\{[e^{j\theta_1}, \dots, e^{j\theta_M}]\}$. We upper-bound the mutual information on the RHS of (64) by proceeding as follows:

$$I(\mathbf{y}; x) \leq I(\mathbf{y}, \{\theta_m - \phi\}_{m=1}^M; x) \quad (83)$$

$$= I(\mathbf{y}; x | \{\theta_m - \phi\}_{m=1}^M) \quad (84)$$

$$= I(\tilde{\mathbf{y}}; x). \quad (85)$$

Here, (84) follows because x and $\{\theta_m - \phi\}_{m=1}^M$ are independent. Since (82) coincides with the input-output relation for the UL-CLO case, we conclude that

$$C(\rho) \leq \frac{1}{2} \ln(\rho) + \frac{1}{2} \ln\left(\frac{\|\mathbf{h}\|^2}{2}\right) + o(1), \quad \rho \rightarrow \infty. \quad (86)$$

The upper bound (86) matches the lower bound (81) up to a $o(1)$ term. This implies (67).

B. Partially Coherent System

We next analyze the partially coherent case where the phase-noise processes are stationary memoryless with arbitrary marginal probability distribution (we do not require the processes to be independent across antennas), and prove that prelog and phase-noise numbers are $\eta = 1/2$ and

$$\chi = (1/2) \ln(\|\mathbf{h}\|^2/2) + \ln(2\pi) - h(\phi | \phi + \boldsymbol{\theta}) \quad (87)$$

where ϕ is uniformly distributed over $[0, 2\pi)$ and where, by definition, $\boldsymbol{\theta} = [\theta_1, \dots, \theta_M]^T$. Similarly to Appendix I-A, we establish this result by deriving an upper bound and a lower bound on $C(\rho) = \sup I(\mathbf{y}; x)$ that match up to a $o(1)$ term.

Lower Bound: We choose the same input distribution as in the noncoherent case, i.e., x is circularly symmetric with $|x|^2 = 4\rho s$ where $s \sim \text{Gamma}(1/2, 1)$. Let $\phi = \underline{x}$ and $r = |x|$. We lower-bound the mutual information in (64) as follows

$$I(\mathbf{y}; x) = I(\mathbf{y}; r) + I(\mathbf{y}; \phi | r) \quad (88)$$

$$\geq I(\|\mathbf{y}\|^2; r) + I(\mathbf{y}; \phi | r). \quad (89)$$

Here, (89) holds because of the data processing inequality. To evaluate the first term on the RHS of (89), we use (81) and obtain

$$I(\|\mathbf{y}\|^2; r) \geq \frac{1}{2} \ln(\rho) + \frac{1}{2} \ln\left(\frac{\|\mathbf{h}\|^2}{2}\right) + o(1), \quad \rho \rightarrow \infty. \quad (90)$$

We lower-bound the second term on the RHS of (89) as

$$I(\mathbf{y}; \phi | r) = I\left(\{y_m\}_{m=1}^M, \{\underline{y}_m\}_{m=1}^M; \phi | r\right) \quad (91)$$

$$\geq I\left(\{\underline{y}_m\}_{m=1}^M; \phi | r\right) \quad (92)$$

$$= I\left(\{\phi + \theta_m + \underline{r} |h_m| + w_m\}_{m=1}^M; \phi | r\right) \quad (93)$$

$$= \ln(2\pi) - h\left(\phi \left| \left\{ \phi + \theta_m + \underline{r} |h_m| + w_m \right\}_{m=1}^M, r\right.\right). \quad (94)$$

Fix an arbitrary $\xi_0 > 0$. We upper-bound the second term on the RHS of (94) as follows:

$$h\left(\phi \left| \left\{ \phi + \theta_m + \underline{r} |h_m| + w_m \right\}_{m=1}^M, r\right.\right) \quad (95)$$

$$= \int_0^\infty f_r(a) h\left(\phi \left| \left\{ \phi + \theta_m + \underline{a} |h_m| + w_m \right\}_{m=1}^M\right.\right) da \quad (96)$$

$$= \int_0^{\xi_0} f_r(a) h\left(\phi \left| \left\{ \phi + \theta_m + \underline{a} |h_m| + w_m \right\}_{m=1}^M\right.\right) da$$

$$+ \int_{\xi_0}^\infty f_r(a) h\left(\phi \left| \left\{ \phi + \theta_m + \underline{a} |h_m| + w_m \right\}_{m=1}^M\right.\right) da \quad (97)$$

$$\leq \Pr\{0 \leq r \leq \xi_0\}$$

$$\times \max_{0 \leq a \leq \xi_0} h\left(\phi \left| \left\{ \phi + \theta_m + \underline{a} |h_m| + w_m \right\}_{m=1}^M\right.\right)$$

$$+ \Pr\{r \geq \xi_0\}$$

$$\times \max_{a \geq \xi_0} h\left(\phi \left| \left\{ \phi + \theta_m + \underline{a} |h_m| + w_m \right\}_{m=1}^M\right.\right) \quad (98)$$

$$= \Pr\left\{0 \leq s \leq \frac{\xi_0^2}{4\rho}\right\} h\left(\phi \left| \left\{ \phi + \theta_m + \underline{w}_m \right\}_{m=1}^M\right.\right)$$

$$+ \Pr\left\{s \geq \frac{\xi_0^2}{4\rho}\right\} h\left(\phi \left| \left\{ \phi + \theta_m + \underline{\xi_0} |h_m| + w_m \right\}_{m=1}^M\right.\right) \quad (99)$$

$$= h\left(\phi \left| \left\{ \phi + \theta_m + \underline{\xi_0} |h_m| + w_m \right\}_{m=1}^M\right.\right) + o(1). \quad (100)$$

Here, in (99) we used that $r = \sqrt{4\rho s}$; (100) holds because for every fixed ξ_0

$$\lim_{\rho \rightarrow \infty} \Pr\{0 \leq s \leq \xi_0^2/4\rho\} = 0. \quad (101)$$

The differential entropy in (100) can be made arbitrarily close to $h(\phi | \phi + \boldsymbol{\theta})$ by choosing ξ_0 sufficiently large. By substituting (100) in (94), (90) and (94) into (89), and by letting ξ_0 tend to infinity, we conclude that

$$C(\rho) \geq (1/2) \ln(\rho) + (1/2) \ln(\|\mathbf{h}\|^2/2) + \ln(2\pi) - h(\phi | \phi + \boldsymbol{\theta}) + o(1), \quad \rho \rightarrow \infty. \quad (102)$$

Upper bound: Let $r = |x|$; let also $\phi = \underline{x}$, which—without loss of generality—we shall assume uniformly distributed over $[0, 2\pi)$ and independent of r (see Lemma 5). Using chain rule, we decompose the mutual information $I(\mathbf{y}; x)$ as

$$I(\mathbf{y}; x) = I(\mathbf{y}; r) + I(\mathbf{y}; \phi | r). \quad (103)$$

We next upper-bound both terms on the RHS of (103). The first term can be bounded as follows

$$I(\mathbf{y}; r) \leq I(\mathbf{y}; \Theta; r) \quad (104)$$

$$= I(\mathbf{y}; r | \Theta) \quad (105)$$

$$= I(\mathbf{h}r e^{j\phi} + \mathbf{w}; r) \quad (106)$$

$$= (1/2) \ln(\rho) + (1/2) \ln(\|\mathbf{h}\|^2/2) + o(1). \quad (107)$$

Here, in (105) we used that x and $\boldsymbol{\theta}$ are independent and (107) follows because the mutual information on the RHS of (106) coincides with that in the noncoherent UL-CLO case.⁴

Next, we upper-bound the second term on the RHS of (103):

$$I(\mathbf{y}; \phi | r) = I(\mathbf{y}, r; \phi) \quad (108)$$

$$\leq I\left(\mathbf{y}, r, \{e^{j(\theta_m + \phi)}\}_{m=1}^M; \phi\right) \quad (109)$$

$$= I\left(\{e^{j(\theta_m + \phi)}\}_{m=1}^M; \phi\right) \quad (110)$$

$$= \ln(2\pi) - h(\phi | \phi + \boldsymbol{\theta}). \quad (111)$$

Here, in (108) we used that ϕ and r are independent and (110) follows because ϕ and the pair (\mathbf{y}, r) are conditionally independent given $\{e^{j(\theta_m + \phi)}\}_{m=1}^M$. Substituting (107) and (111) into (103), we obtain

$$C(\rho) \leq (1/2) \ln(\rho) + (1/2) \ln(\|\mathbf{h}\|^2/2) + \ln(2\pi) - h(\phi | \phi + \boldsymbol{\theta}) + o(1), \quad \rho \rightarrow \infty. \quad (112)$$

The upper bound (112) matches the lower bound (102) up to a $o(1)$ term. This implies (87).

C. Phase Noise with Memory

We establish (23) by proceeding similarly as in Appendix I-A and I-B.

⁴Note that ϕ in (106) plays the role of uniform phase noise, although it is the phase of the transmitted signal. This is because the mutual information in (106) is between the channel output and the amplitude of the transmitted signal.

Lower Bound: Fix an arbitrary $\xi_0 > 0$ and some positive integer γ . We evaluate the mutual information on the RHS of (20) for an i.i.d. input process $\{x_k\}$ having the same marginal distribution as in the noncoherent case, i.e., uniform phase and amplitude distributed as in (68). Using the chain rule for mutual information and the nonnegativity of mutual information, we obtain

$$I(x^n; \mathbf{y}^n) = \sum_{k=1}^n I(x_k; \mathbf{y}^n | x^{k-1}) \quad (113)$$

$$\geq \sum_{k=\gamma+1}^{n-\gamma} I(x_k; \mathbf{y}^k | x^{k-1}). \quad (114)$$

For every $k \in [\gamma+1, n-\gamma]$, we lower-bound $I(x_k; \mathbf{y}^k | x^{k-1})$ as follows:

$$I(x_k; \mathbf{y}^n | x^{k-1}) = I(x_k; x^{k-1}, \mathbf{y}^k) \quad (115)$$

$$\geq I(x_k; x_{k-\gamma}^{k-1}, \mathbf{y}_{k-\gamma}^{k-1}, \mathbf{y}_k) \quad (116)$$

$$= I(x_k; x_{k-\gamma}^{k-1}, \mathbf{y}_{k-\gamma}^{k-1}, \boldsymbol{\theta}_{k-\gamma}^{k-1}, \mathbf{y}_k) - \underbrace{I(x_k; \boldsymbol{\theta}_{k-\gamma}^{k-1} | x_{k-\gamma}^{k-1}, \mathbf{y}_{k-\gamma}^{k-1}, \mathbf{y}_k)}_{\leq \epsilon(\rho, \xi_0, \gamma)} \quad (117)$$

$$\geq I(x_k; x_{k-\gamma}^{k-1}, \mathbf{y}_{k-\gamma}^{k-1}, \mathbf{y}_k, \boldsymbol{\theta}_{k-\gamma}^{k-1}) - \epsilon(\rho, \xi_0, \gamma) \quad (118)$$

$$= I(x_k; \mathbf{y}_k, \boldsymbol{\theta}_{k-\gamma}^{k-1}) - \epsilon(\rho, \xi_0, \gamma) \quad (119)$$

$$= I(x_{\gamma+1}; \mathbf{y}_{\gamma+1} | \boldsymbol{\theta}_1^\gamma) - \epsilon(\rho, \xi_0, \gamma). \quad (120)$$

Here, (115) follows because the $\{x_k\}$ are i.i.d.; in (117) we upper-bounded the second mutual information by a function, which we denote by $\epsilon(\rho, \xi_0, \gamma)$, that depends only on ρ, ξ_0 , and γ and that satisfies (see Appendix IV)

$$\lim_{\xi_0 \rightarrow \infty} \lim_{\rho \rightarrow \infty} \epsilon(\rho, \xi_0, \gamma) = 0 \quad (121)$$

for all γ ; (119) follows because x_k and the pair $(\mathbf{y}_{k-\gamma}^{k-1}, x_{k-\gamma}^{k-1})$ are conditionally independent given $(\mathbf{y}_k, \boldsymbol{\theta}_{k-\gamma}^{k-1})$; finally, in (120) we used stationarity and that $x_{\gamma+1}$ and $\boldsymbol{\theta}_1^\gamma$ are independent.

Substituting (120) into (114) and then (114) into (20),

$$C(\rho) \geq I(x_{\gamma+1}; \mathbf{y}_{\gamma+1} | \boldsymbol{\theta}_1^\gamma) - \epsilon(\rho, \xi_0, \gamma). \quad (122)$$

Apart from the side information $(\boldsymbol{\theta}_1^\gamma)$, the mutual information on the RHS of (122) coincides with that of the memoryless partially coherent channel we analyzed in Appendix I-B. Proceeding as in Appendix I-B, and letting ξ_0 tend to infinity, we obtain

$$I(x_{\gamma+1}; \mathbf{y}_{\gamma+1} | \boldsymbol{\theta}_1^\gamma) \geq \frac{1}{2} \ln(\rho) + \frac{1}{2} \ln\left(\frac{\|\mathbf{h}\|^2}{2}\right) + \ln(2\pi) - h(\phi_{\gamma+1} | \boldsymbol{\theta}_{\gamma+1} + \phi_{\gamma+1}, \boldsymbol{\theta}_1^\gamma) + o(1), \quad \rho \rightarrow \infty. \quad (123)$$

Substituting (123) into (122), using stationarity, and letting γ tend to infinity, we finally obtain the following capacity lower bound

$$C(\rho) \geq (1/2) \ln(\rho) + (1/2) \ln(\|\mathbf{h}\|^2/2) + \ln(2\pi) - h(\phi_0 | \boldsymbol{\theta}_0 + \phi_0, \boldsymbol{\theta}_{-\infty}^{-1}) + o(1). \quad (124)$$

Upper bound: We use the chain rule for mutual information on the RHS of (20) and obtain

$$I(x^n, \mathbf{y}^n) = \sum_{k=1}^n I(x^n; \mathbf{y}_k | \mathbf{y}^{k-1}). \quad (125)$$

We next proceed as in [31, Eq. (77)] and upper-bound each term on the RHS of (125) as

$$I(x^n; \mathbf{y}_k | \mathbf{y}^{k-1}) \leq I(x^n, \mathbf{y}^{k-1}; \mathbf{y}_k) \quad (126)$$

$$= I(x_k, x^{k-1}, \mathbf{y}^{k-1}; \mathbf{y}_k) \quad (127)$$

$$\leq I(x_k, x^{k-1}, \mathbf{y}^{k-1}, \boldsymbol{\theta}^{k-1}; \mathbf{y}_k) \quad (128)$$

$$= I(x_k, \boldsymbol{\theta}^{k-1}; \mathbf{y}_k) \quad (129)$$

$$= I(x_k; \mathbf{y}_k) + I(\boldsymbol{\theta}^{k-1}; \mathbf{y}_k | x_k) \quad (130)$$

$$\leq I(x_k; \mathbf{y}_k) + I(\boldsymbol{\theta}^{k-1}; \mathbf{y}_k, x_k, \boldsymbol{\theta}_k) \quad (131)$$

$$= I(x_k; \mathbf{y}_k) + I(\boldsymbol{\theta}^{k-1}; \boldsymbol{\theta}_k) \quad (132)$$

$$\leq I(x_0; \mathbf{y}_0) + I(\boldsymbol{\theta}_0; \boldsymbol{\theta}_{-\infty}^{-1}). \quad (133)$$

Here, (129) holds because \mathbf{y}_k and $(\mathbf{y}^{k-1}, x^{k-1})$ are conditionally independent given $(x_k, \boldsymbol{\theta}^{k-1})$; (132) holds because $\boldsymbol{\theta}^{k-1}$ and (x_k, \mathbf{y}_k) are conditionally independent given $\boldsymbol{\theta}_k$; finally, (133) follows because of the stationarity of the phase-noise processes and the nonnegativity of mutual information. Substituting (133) into (125), then (125) into (20), we obtain

$$C(\rho) \leq \sup\{I(x_0; \mathbf{y}_0)\} + I(\boldsymbol{\theta}_0; \boldsymbol{\theta}_{-\infty}^{-1}) \quad (134)$$

$$\leq (1/2) \ln(\rho) + (1/2) \ln(\|\mathbf{h}\|^2/2) + \ln(2\pi)$$

$$- h(\phi | \phi + \boldsymbol{\theta}_0) + I(\boldsymbol{\theta}_0; \boldsymbol{\theta}_{-\infty}^{-1}) + o(1) \quad (135)$$

where the last step follows from (112).

APPENDIX II PROOF OF THEOREM 3

As a first step, we adapt Lemma 5, which describes the structure of the capacity-achieving distribution for the SIMO phase-noise channel (19) to the MISO phase-noise channel (47). Let $\{\phi_k\}$ be a stationary memoryless process with uniform marginal distribution over $[0, 2\pi)$. We say that a vector process $\{\mathbf{x}_k\}$ is circularly symmetric if $\{\mathbf{x}_k e^{j\phi_k}\} \sim \{\mathbf{x}_k\}$.

Lemma 6: The input process that achieves the capacity of the channel (47) can be assumed to be circularly symmetric.

Proof: The proof follows along the same lines as the proof of [33, Prop. 7]. ■

By Lemma 6, we can express $\{\mathbf{x}_k\}$ as $\{\mathbf{x}_k = \|\mathbf{x}_k\| \hat{\mathbf{x}}_k e^{j\phi_k}\}$ where $\|\hat{\mathbf{x}}_k\| = 1$ for all k and $\{\phi_k\}$ is a stationary memoryless process, independent of $\{\|\mathbf{x}_k\|, \hat{\mathbf{x}}_k\}$ and with uniform marginal distribution over $[0, 2\pi)$.

Lower Bound: See Remark 4.

Upper Bound: We use chain rule for mutual information on the RHS of (48) and obtain

$$I(\mathbf{x}^n; \mathbf{y}^n) = \sum_{k=1}^n I(\mathbf{x}^n; \mathbf{y}_k | \mathbf{y}^{k-1}). \quad (136)$$

Proceeding similarly to (126)–(133), we next upper-bound each term on the RHS of (136):

$$I(\mathbf{x}^n; y_k | y^{k-1}) \leq I(\mathbf{x}^n, y^{k-1}; y_k) \quad (137)$$

$$= I(\mathbf{x}^k, y^{k-1}; y_k) \quad (138)$$

$$\leq I(\mathbf{x}_k, \mathbf{x}^{k-1}, y^{k-1}, \boldsymbol{\theta}^{k-1}; y_k) \quad (139)$$

$$= I(\mathbf{x}_k, \boldsymbol{\theta}^{k-1}; y_k) \quad (140)$$

$$= I(\mathbf{x}_k; y_k) + I(\boldsymbol{\theta}^{k-1}; y_k | \mathbf{x}_k) \quad (141)$$

$$\leq I(\mathbf{x}_k; y_k) + I(\boldsymbol{\theta}^{k-1}; y_k, \sqrt{\mathbf{h}^T \Theta_k \hat{\mathbf{x}}_k} | \mathbf{x}_k) \quad (142)$$

$$= I(\mathbf{x}_k; y_k) + I(\boldsymbol{\theta}^{k-1}; \sqrt{\mathbf{h}^T \Theta_k \hat{\mathbf{x}}_k} \| \mathbf{x}_k \|, \hat{\mathbf{x}}_k) \quad (143)$$

$$\leq I(\mathbf{x}_0; y_0) + I(\boldsymbol{\theta}_{-\infty}^{-1}; \sqrt{\mathbf{h}^T \Theta_0 \hat{\mathbf{x}}_0} \| \mathbf{x}_0 \|, \hat{\mathbf{x}}_0). \quad (144)$$

Here, (138) follows because \mathbf{x}_{k+1}^n and y_k are conditionally independent given $(\mathbf{x}^{k-1}, y^{k-1})$; (140) holds because the pair $(\mathbf{x}^{k-1}, y^{k-1})$ and y_k are conditionally independent given $\boldsymbol{\theta}^{k-1}$ and \mathbf{x}_k ; similarly, (143) holds because $\boldsymbol{\theta}^{k-1}$ and y_k are conditionally independent given $\sqrt{\mathbf{h}^T \Theta_k \hat{\mathbf{x}}_k}$ and \mathbf{x}_k and because of Lemma 6; finally, in (144) we used that the phase-noise processes are stationary and that mutual information is nonnegative.

The first term on the RHS of (144), which corresponds to the mutual information achievable on a partially coherent MISO phase-noise channel, can be further upper-bounded as follows:

$$I(\mathbf{x}_0; y_0) = I(\| \mathbf{x}_0 \|, \hat{\mathbf{x}}_0; y_0) + I(\phi_0; y_0 | \| \mathbf{x}_0 \|, \hat{\mathbf{x}}_0) \quad (145)$$

$$\leq I(\| \mathbf{x}_0 \|, \hat{\mathbf{x}}_0; y_0) + I(\phi_0; y_0, \sqrt{\mathbf{h}^T \Theta_0 \mathbf{x}_0} \| \| \mathbf{x}_0 \|, \hat{\mathbf{x}}_0) \quad (146)$$

$$= I(\| \mathbf{x}_0 \|, \hat{\mathbf{x}}_0; y_0) + I(\phi_0; \sqrt{\mathbf{h}^T \Theta_0 \mathbf{x}_0} \| \| \mathbf{x}_0 \|, \hat{\mathbf{x}}_0) \quad (147)$$

$$= I(\| \mathbf{x}_0 \|, \hat{\mathbf{x}}_0; y_0) + I(\phi_0; \phi_0 + \sqrt{\mathbf{h}^T \Theta_0 \hat{\mathbf{x}}_0} \| \| \mathbf{x}_0 \|, \hat{\mathbf{x}}_0). \quad (148)$$

Here, (147) follows because ϕ_0 and y_0 are conditionally independent given $(\sqrt{\mathbf{h}^T \Theta_0 \hat{\mathbf{x}}_0}, \| \mathbf{x}_0 \|, \hat{\mathbf{x}}_0)$. Substituting (148) into (144) and using that ϕ_0 is uniformly distributed over $[0, 2\pi)$ and independent of $(\hat{\mathbf{x}}_0, \| \mathbf{x}_0 \|)$, we obtain

$$\begin{aligned} \frac{1}{n} I(\mathbf{x}^n; y^n) &\leq I(\| \mathbf{x}_0 \|, \hat{\mathbf{x}}_0; y_0) + \ln(2\pi) \\ &\quad - h(\sqrt{\mathbf{h}^T \Theta_0 \hat{\mathbf{x}}_0} \| \| \mathbf{x}_0 \|, \hat{\mathbf{x}}_0, \boldsymbol{\theta}_{-\infty}^{-1}) \end{aligned} \quad (149)$$

$$\begin{aligned} &\leq \sup_{Q_{\| \mathbf{x}_0 \|, \hat{\mathbf{x}}_0}} \{ I(\| \mathbf{x}_0 \|, \hat{\mathbf{x}}_0; y_0) \} + \ln(2\pi) \\ &\quad - \inf_{\| \hat{\mathbf{x}} \| = 1} h(\sqrt{\mathbf{h}^T \Theta_0 \hat{\mathbf{x}}} | \boldsymbol{\theta}_{-\infty}^{-1}). \end{aligned} \quad (150)$$

Here, the supremum is over all joint probability distributions $Q_{\| \mathbf{x}_0 \|, \hat{\mathbf{x}}_0}$ on $(\| \mathbf{x}_0 \|, \hat{\mathbf{x}}_0)$ that satisfy $\mathbb{E}[\| \mathbf{x}_0 \|^2] \leq 2\rho$, and the infimum is over all deterministic unit-norm vectors $\hat{\mathbf{x}}$ in \mathbb{C}^M .

To conclude the proof, we next characterize the first term on the RHS of (150). Let

$$C_0(\rho) = \sup_{Q_{\| \mathbf{x}_0 \|, \hat{\mathbf{x}}_0}} I(\| \mathbf{x}_0 \|, \hat{\mathbf{x}}_0; y_0). \quad (151)$$

Then,

$$\lim_{\rho \rightarrow \infty} C_0(\rho) / \ln(\rho) = 1/2. \quad (152)$$

Indeed, choosing $\hat{\mathbf{x}}_0$ to be the m th element of the canonical basis for \mathbb{C}^M (this corresponds to switching off all but the m th transmit antenna), the phase-noise channel

$$y_0 = e^{j\phi_0} \| \mathbf{x}_0 \| \mathbf{h}^T \Theta_0 \hat{\mathbf{x}}_0 + w_0 \quad (153)$$

reduces to the noncoherent SISO phase noise channel $y_0 = e^{j(\phi_0 + \theta_{m,0})} h_m \| \mathbf{x}_0 \| + w_0$, for which the prelog is $1/2$ (see Theorem 1). Conversely,

$$C_0(\rho) \leq \sup_{Q_{\| \mathbf{x}_0 \|, \hat{\mathbf{x}}_0}} I(\| \mathbf{x}_0 \|, \hat{\mathbf{x}}_0; y_0, \Theta_0) \quad (154)$$

$$= \sup_{Q_{\| \mathbf{x}_0 \|, \hat{\mathbf{x}}_0}} I(\| \mathbf{x}_0 \|, \hat{\mathbf{x}}_0; y_0 | \Theta_0) \quad (155)$$

$$= \sup_{Q_{\| \mathbf{x}_0 \|}} I(\| \mathbf{x}_0 \|; e^{j\phi_0} \| \mathbf{h} \| \| \mathbf{x}_0 \| + w_0) \quad (156)$$

$$= (1/2) \ln(\rho) + (1/2) \ln(\| \mathbf{h} \|^2 / 2) + o(1). \quad (157)$$

Here, (156) follows because setting $\hat{\mathbf{x}}_0 = \Theta_0^H \mathbf{h}^* / \| \mathbf{h} \|^2$ with probability 1 (w.p.1) achieves the supremum in (155); (157) follows from Theorem 1.

Fix now an arbitrary $\xi_0 > 0$ and let $C_{\xi_0}(\rho)$ be the capacity of the channel (153) when the input is subject to the additional constraint that $\| \mathbf{x}_0 \| \geq \xi_0$ w.p.1. It follows from (152), from [31, Thm. 8] and from [20, Thm. 4.12] that

$$C_0(\rho) = C_{\xi_0}(\rho) + o(1). \quad (158)$$

This fact is often referred to as ‘‘escaping-to-infinity’’ property of the capacity achieving distribution [20, Def. 4.11].

Take an arbitrary input distribution that satisfies $\mathbb{E}[\| \mathbf{x}_0 \|^2] \leq 2\rho$ and $\| \mathbf{x}_0 \|^2 \geq \xi_0$ w.p.1. Furthermore, let $q(\cdot)$ be the pdf of a Gamma($1/2, 2\mathbb{E}[|y_0|^2]$)-distributed random variable:

$$q(r) = e^{-r} / (2\mathbb{E}[|y_0|^2]) / \sqrt{2\pi r \mathbb{E}[|y_0|^2]}. \quad (159)$$

Here, the expectation is computed with respect to the probability distribution induced on $|y_0|^2$ by the chosen input distribution and by the channel transition probability. Then

$$I(\| \mathbf{x}_0 \|, \hat{\mathbf{x}}_0; y_0) = I(\| \mathbf{x}_0 \|, \hat{\mathbf{x}}_0; |y_0|^2) \quad (160)$$

$$\leq -\mathbb{E}[\ln(q(|y_0|^2))] - h(|y_0|^2 | \| \mathbf{x}_0 \|^2, \hat{\mathbf{x}}_0). \quad (161)$$

Here, (160) holds because $|y_0|^2$ is a sufficient statistics for the detection of $(\hat{\mathbf{x}}, \| \mathbf{x} \|)$ from y (recall that ϕ_0 in (153) is uniformly distributed over $[0, 2\pi)$), and (161) follows from [20, Thm. 5.1]. Substituting (159) into (161), we obtain

$$\begin{aligned} I(\| \mathbf{x}_0 \|, \hat{\mathbf{x}}_0; y_0) &\leq 1/2 + (1/2) \ln(2\pi) + (1/2) \ln(\mathbb{E}[|y_0|^2]) \\ &\quad + (1/2) \mathbb{E}[\ln(|y_0|^2)] - h(|y_0|^2 | \| \mathbf{x}_0 \|^2, \hat{\mathbf{x}}_0). \end{aligned} \quad (162)$$

Note now that

$$\mathbb{E}[|y_0|^2] \leq 2\rho \mathbb{E}[| \mathbf{h}^T \Theta_0 \hat{\mathbf{x}}_0 |^2] + 2 \quad (163)$$

$$\leq 2\rho \sup_{\hat{\mathbf{x}}} \mathbb{E}[| \mathbf{h}^T \Theta_0 \hat{\mathbf{x}} |^2] + 2 \quad (164)$$

where the supremum is over all deterministic unit-norm vectors $\hat{\mathbf{x}}$ in \mathbb{C}^M . Furthermore,

$$(1/2) \mathbb{E}[\ln(|y_0|^2)] - h(|y_0|^2 | \| \mathbf{x}_0 \|^2, \hat{\mathbf{x}}_0) \quad (165)$$

$$\leq (1/2) \mathbb{E}[\ln(|y_0|^2)] - h(|y_0|^2 | \| \mathbf{x}_0 \|^2, \hat{\mathbf{x}}_0, \Theta_0) \quad (166)$$

$$\begin{aligned} &\leq \sup_{\hat{\mathbf{x}}} \sup_{\xi \geq \xi_0} \left\{ (1/2) \mathbb{E}[\ln(|\xi \mathbf{h}^T \Theta_0 \hat{\mathbf{x}} + w|^2)] \right. \\ &\quad \left. - h(|\xi \mathbf{h}^T \Theta_0 \hat{\mathbf{x}} + w|^2 | \Theta_0) \right\}. \end{aligned} \quad (167)$$

Since for every fixed $\hat{\mathbf{x}}$, [1, Eq. (9)]

$$\begin{aligned} \lim_{\xi \rightarrow \infty} \frac{1}{2} \mathbb{E} \left[\ln \left(|\xi \mathbf{h}^T \Theta_0 \hat{\mathbf{x}} + w|^2 \right) \right] - h \left(|\xi \mathbf{h}^T \Theta_0 \hat{\mathbf{x}} + w|^2 \mid \Theta_0 \right) \\ = -(1/2) \ln(8\pi e) \end{aligned} \quad (168)$$

we can make (167) arbitrarily close to $-0.5 \ln(8\pi e)$ by choosing ξ_0 sufficiently large.

Substituting (164) and (167) into (162), then using (168) and (158), and finally letting ξ_0 tend to infinity, we obtain

$$C_0(\rho) \leq \frac{1}{2} \ln(\rho) + \sup_{\hat{\mathbf{x}}} \frac{1}{2} \ln \left(\mathbb{E} \left[|\mathbf{h}^T \Theta \hat{\mathbf{x}}|^2 \right] / 2 \right) + o(1). \quad (169)$$

Substituting (169) into (150), we obtain the desired result.

APPENDIX III PROOF OF THEOREM 4

We start by evaluating the second term on the RHS of (55b).

Let $\hat{\mathbf{x}} = [\hat{x}_1, \dots, \hat{x}_M]^T$. Then

$$\sup_{\|\hat{\mathbf{x}}\|=1} \mathbb{E} \left[|\mathbf{h}^T \Theta_0 \hat{\mathbf{x}}|^2 \right] = \sup_{\|\hat{\mathbf{x}}\|=1} \mathbb{E} \left[\left| \sum_{m=1}^M e^{j\theta_{m,0}} h_m \hat{x}_m \right|^2 \right] \quad (170)$$

$$= \sup_{\|\hat{\mathbf{x}}\|=1} \sum_{m=1}^M |h_m|^2 |\hat{x}_m|^2 \quad (171)$$

$$= \max_{m=1, \dots, M} |h_m|^2. \quad (172)$$

In (171) we used that the random variables $\{\theta_{m,0}\}$, $m = 1, \dots, M$ are independent and uniformly distributed over $[0, 2\pi)$, which implies that $\mathbb{E}[e^{j\theta_m}] = 0$, $m = 1, \dots, M$.

We next bound the third term on the RHS of (55b) as follows:

$$\begin{aligned} \inf_{\|\hat{\mathbf{x}}\|=1} \left\{ h \left(\sqrt{\mathbf{h}^T \Theta_0 \hat{\mathbf{x}}} \mid \boldsymbol{\theta}_{-\infty}^{-1} \right) \right\} \\ \geq \inf_{\|\hat{\mathbf{x}}\|=1} h \left(\sqrt{\mathbf{h}^T \Theta_0 \hat{\mathbf{x}}_0} \mid \boldsymbol{\theta}_{-\infty}^{-1}, \theta_{2,0}, \dots, \theta_{M,0} \right) \end{aligned} \quad (173)$$

$$= h(\theta_{1,0} \mid \boldsymbol{\theta}_{-\infty}^{-1}) \quad (174)$$

$$= h(\{\theta_k\}). \quad (175)$$

Here in (174) we used that the phase-noise processes are independent, and in (175) that they are identically distributed. Substituting (172) and (175) in (55b), we obtain the following upper bound on the phase-noise number

$$\chi_{\text{dl-slo}} \leq \frac{1}{2} \ln \left(\max_{m=1, \dots, M} |h_m|^2 / 2 \right) + \ln(2\pi) - h(\{\theta_k\}). \quad (176)$$

This bound is tight. Indeed consider the transmission scheme where only the transmit antenna that experiences the largest channel gain $\max_{m=1, \dots, M} |h_m|^2$ is switched on, whereas all other antennas are switched off. By Theorem 1, the phase-noise number achievable with this transmission scheme coincides with the RHS of (176). This concludes the proof of Theorem 4.

APPENDIX IV

We shall prove that

$$I \left(x_k; \boldsymbol{\theta}_{k-\gamma}^{k-1} \mid x_{k-\gamma}^{k-1}, \mathbf{y}_{k-\gamma}^{k-1}, \mathbf{y}_k, \mathbf{y}_{k+1}^{k+\gamma} \right) \leq \epsilon(\rho, \xi_0, \gamma) \quad (177)$$

where $\epsilon(\rho, \xi_0, \gamma)$ is a function that depend only on ρ , ξ_0 , and γ and that satisfies (121) for all γ .

We proceed as follows:

$$\begin{aligned} I \left(x_k; \boldsymbol{\theta}_{k-\gamma}^{k-1} \mid x_{k-\gamma}^{k-1}, \mathbf{y}_{k-\gamma}^{k-1}, \mathbf{y}_k, \mathbf{y}_{k+1}^{k+\gamma} \right) \\ = h \left(\boldsymbol{\theta}_{k-\gamma}^{k-1} \mid x_{k-\gamma}^{k-1}, \mathbf{y}_{k-\gamma}^{k-1}, \mathbf{y}_k, \mathbf{y}_{k+1}^{k+\gamma} \right) \\ - h \left(\boldsymbol{\theta}_{k-\gamma}^{k-1} \mid x_k, x_{k-\gamma}^{k-1}, \mathbf{y}_{k-\gamma}^{k-1}, \mathbf{y}_k, \mathbf{y}_{k+1}^{k+\gamma} \right) \end{aligned} \quad (178)$$

$$\begin{aligned} \leq h \left(\boldsymbol{\theta}_{k-\gamma}^{k-1} \mid x_{k-\gamma}^{k-1}, \mathbf{y}_{k-\gamma}^{k-1} \right) \\ - h \left(\boldsymbol{\theta}_{k-\gamma}^{k-1} \mid x_k, x_{k-\gamma}^{k-1}, \mathbf{y}_{k-\gamma}^{k-1}, \mathbf{y}_k, \mathbf{y}_{k+1}^{k+\gamma}, x_{k+1}^{k+\gamma}, \boldsymbol{\theta}_k^{k+\gamma} \right) \end{aligned} \quad (179)$$

$$\begin{aligned} = h \left(\boldsymbol{\theta}_{k-\gamma}^{k-1} \mid x_{k-\gamma}^{k-1}, \mathbf{y}_{k-\gamma}^{k-1} \right) \\ - h \left(\boldsymbol{\theta}_{k-\gamma}^{k-1} \mid x_{k-\gamma}^{k-1}, \mathbf{y}_{k-\gamma}^{k-1}, \boldsymbol{\theta}_k^{k+\gamma} \right) \end{aligned} \quad (180)$$

$$= I \left(\boldsymbol{\theta}_{k-\gamma}^{k-1}; \boldsymbol{\theta}_k^{k+\gamma} \mid x_{k-\gamma}^{k-1}, \mathbf{y}_{k-\gamma}^{k-1} \right) \quad (181)$$

$$\begin{aligned} = h \left(\boldsymbol{\theta}_k^{k+\gamma} \mid x_{k-\gamma}^{k-1}, \mathbf{y}_{k-\gamma}^{k-1} \right) \\ - h \left(\boldsymbol{\theta}_k^{k+\gamma} \mid x_{k-\gamma}^{k-1}, \mathbf{y}_{k-\gamma}^{k-1}, \boldsymbol{\theta}_{k-\gamma}^{k-1} \right) \end{aligned} \quad (182)$$

$$\begin{aligned} = h \left(\boldsymbol{\theta}_k^{k+\gamma} \mid x_{k-\gamma}^{k-1}, \left\{ \sqrt{\mathbf{h}_l^T \mathbf{x}_l + \mathbf{w}_l} \right\}_{l=k-\gamma}^{k-1} \right) \\ - h \left(\boldsymbol{\theta}_k^{k+\gamma} \mid \boldsymbol{\theta}_{k-\gamma}^{k-1} \right) \end{aligned} \quad (183)$$

$$\begin{aligned} \leq h \left(\boldsymbol{\theta}_k^{k+\gamma} \mid \left\{ \sqrt{\mathbf{h}_l^T \boldsymbol{\xi}_0 + \mathbf{w}_l} \right\}_{l=k-\gamma}^{k-1} \right) \\ - h \left(\boldsymbol{\theta}_k^{k+\gamma} \mid \boldsymbol{\theta}_{k-\gamma}^{k-1} \right) + o(1), \quad \rho \rightarrow \infty. \end{aligned} \quad (184)$$

Here, (184) follows by a multivariate extension of the steps taken in (96)–(100). The RHS of (184) depends only on γ and ξ_0 and can be made arbitrarily close to zero by choosing ξ_0 sufficiently large.

REFERENCES

- [1] A. Lapidith, "On phase noise channels at high SNR," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Bangalore, India, Oct. 2002, pp. 1–4.
- [2] G. Colavolpe, A. Barbieri, and G. Caire, "Algorithms for iterative decoding in the presence of strong phase noise," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 9, pp. 1748–1757, Sep. 2005.
- [3] 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 Trans. Signal Process.*, vol. 60, no. 9, pp. 4790–4807, Sep. 2012.
- [4] G. Durisi, A. Tarable, C. Camarda, R. Devassy, and G. Montorsi, "Capacity bounds for MIMO microwave backhaul links affected by phase noise," *IEEE Trans. Commun.*, vol. 62, no. 3, pp. 920–929, Mar. 2014.
- [5] M. Katz and S. Shamai (Shitz), "On the capacity-achieving distribution of the discrete-time noncoherent and partially coherent AWGN channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2257–2270, Oct. 2004.
- [6] R. Nuriyev and A. Anastasopoulos, "Capacity and coding for the block-independent noncoherent AWGN channel," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 866–883, Mar. 2005.
- [7] G. Durisi, "On the capacity of the block-memoryless phase-noise channel," *IEEE Commun. Lett.*, vol. 16, no. 8, pp. 1157–1160, Aug. 2012.
- [8] J. Dauwels and H.-A. Loeliger, "Computation of information rates by particle methods," *IEEE Trans. Commun.*, vol. 54, no. 1, pp. 406–409, Jan. 2008.
- [9] L. Barletta, M. Magarini, and A. Spalvieri, "The information rate transferred through the discrete-time Wiener's phase noise channel," *J. Lightw. Technol.*, vol. 30, no. 10, pp. 1480–1486, May 2012.
- [10] L. Barletta, M. Magarini, S. Pecorino, and A. Spalvieri, "Upper and lower bounds to the information rate transferred through first-order Markov channels with free-running continuous state," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 3834–3844, Jul. 2014.
- [11] D. Gesbert, H. Bölcskei, D. Gore, and A. Paulraj, "Outdoor MIMO wireless channels: models and performance prediction," *IEEE Trans. Commun.*, vol. 50, no. 12, pp. 1926–1934, Dec. 2002.

- [12] D. Chizhik, G. J. Foschini, M. J. Gans, and R. A. Valenzuela, "Keyholes, correlations, and capacities of multielement transmit and receive antennas," *IEEE Trans. Wireless Commun.*, vol. 1, no. 2, pp. 361–368, Apr. 2002.
- [13] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.
- [14] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, pp. 40–60, Jan. 2013.
- [15] E. Björnson, J. Hoydis, M. Kountouris, and M. Debbah, "Massive MIMO systems with non-ideal hardware: Energy efficiency, estimation, and capacity limits," *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 7112–7139, Nov. 2014.
- [16] 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, pp. 1–7.
- [17] G. Durisi, A. Tarable, and T. Koch, "On the multiplexing gain of MIMO microwave backhaul links affected by phase noise," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Budapest, Hungary, Jun. 2013, pp. 3209–3214.
- [18] A. Pitarokoilis, S. K. Mohammed, and E. G. Larsson, "Uplink performance of time-reversal MRC in massive MIMO system subject to phase noise," *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, pp. 711–723, Feb. 2015.
- [19] E. Björnson, M. Matthaiou, and M. Debbah, "Circuit-aware design of energy-efficient massive MIMO systems," in *Int. Symp. Commun., Cont., Signal Process. (ISCCSP)*, Athens, Greece, May 2014.
- [20] A. Lapidath and S. M. Moser, "Capacity bounds via duality with applications to multiple-antenna systems on flat-fading channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2426–2467, Oct. 2003.
- [21] A. Demir, A. Mehrotra, and J. Roychowdhury, "Phase noise in oscillators: A unifying theory and numerical methods for characterization," *IEEE Trans. Circuits Syst. I*, vol. 47, no. 5, pp. 655–674, May 2005.
- [22] H. Ghazlan and G. Kramer, "On Wiener phase noise channels at high signal-to-noise ratio," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Istanbul, Turkey, Jul. 2013, pp. 2279–2283.
- [23] —, "Phase modulation for discrete-time Wiener phase noise channels with oversampling at high SNR," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Honolulu, HI, Jun. 2014, pp. 1554–1557.
- [24] L. Barletta and G. Kramer, "On continuous-time white phase noise channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Honolulu, HI, Jun. 2014, pp. 2426–2429.
- [25] G. Durisi and H. Bölcskei, "High-SNR capacity of wireless communication channels in the noncoherent setting: A primer," *Int. J. Electron. Commun. (AEÜ)*, vol. 65, no. 8, pp. 707–712, Aug. 2011, invited paper.
- [26] A. Viterbi, "Phase-locked loop dynamics in the presence of noise by Fokker-Planck techniques," *Proc. IEEE*, vol. 51, no. 12, pp. 1737–1753, Dec. 1963.
- [27] M.R. Khanzadi, D. Kuylenstierna, A. Panahi, T. Eriksson, and H. Zirath, "Calculation of the performance of communication systems from measured oscillator phase noise," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 61, no. 5, pp. 1553–1565, May 2014.
- [28] G. Colavolpe, "Communications over phase-noise channels: A tutorial review," in *Advanced Satellite Multimedia Systems Conference (ASMS) and 12th Signal Processing for Space Communications Workshop (SPSC)*, Sep. 2012, pp. 316–327.
- [29] K. V. Mardia and P. E. Jupp, *Directional statistics*. Chichester, U.K.: John Wiley & Sons, 2009.
- [30] M. R. Khanzadi, R. Krishnan, D. Kuylenstierna, and T. Eriksson, "Oscillator phase noise and small-scale channel fading in higher frequency bands," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Austin, TX, U.S.A., Dec. 2014, to appear.
- [31] A. Lapidath and S. M. Moser, "The fading number of single-input multiple-output fading channels with memory," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 437–453, Feb. 2006.
- [32] S. M. Moser, "On the fading number of multiple-input single-output fading channels with memory," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Seattle, WA, Jul. 2006, pp. 1056–1060.
- [33] —, "The fading number of multiple-input multiple-output fading channels with memory," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2716–2755, Jun. 2009.
- [34] E. Biglieri, J. G. Proakis, and S. Shamai (Shitz), "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.
- [35] D. N. C. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
- [36] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York, NY, U.S.A.: Wiley, 2006.



M. Reza Khanzadi (S'10) received his M.Sc. degree in communication engineering from Chalmers University of Technology, Gothenburg, Sweden, in 2010. He is currently a Ph.D. candidate at the Department of Signals and Systems in collaboration with the Department of Microtechnology and Nanoscience of the same university. From October to December 2014, he was a Research Visitor in the University of Southern California, Los Angeles, CA. Bayesian inference, statistical signal processing, and information theory are his current research interests. His Ph.D. project is mainly focused on radio frequency oscillator modeling, oscillator phase noise estimation/compensation, and determining the effect of oscillator phase noise on the performance of communication systems. He has been the recipient of the S2 Pedagogical Prize 2012 from Department of Signals and Systems, as well as a 2013 and 2014 Ericsson's Research Foundation grants, 2014 Chalmers Foundation grant, and 2014 Gothenburg Royal Society of Arts and Sciences grant.



Giuseppe Durisi (S'02, M'06, SM'12) received the Laurea degree summa cum laude and the Doctor degree both from Politecnico di Torino, Italy, in 2001 and 2006, respectively. From 2002 to 2006, he was with Istituto Superiore Mario Boella, Torino, Italy. From 2006 to 2010 he was a postdoctoral researcher at ETH Zurich, Zurich, Switzerland. Since 2010, he has been with Chalmers University of Technology, Gothenburg, Sweden, where is now associate professor. He is also guest researcher at Ericsson, Sweden.

Dr. Durisi is a senior member of the IEEE. He is the recipient of the 2013 IEEE ComSoc Best Young Researcher Award for the Europe, Middle East, and Africa Region, and is co-author of a paper that won a "student paper award" at the 2012 International Symposium on Information Theory, and of a paper that won the 2013 IEEE Sweden VT-COM-IT joint chapter best student conference paper award. From 2011 to 2014 he served as publications editor for the IEEE Transactions on Information Theory. His research interests are in the areas of communication and information theory, and compressed sensing.



Thomas Eriksson Thomas Eriksson received the Ph.D. degree in Information Theory in 1996, from Chalmers University of Technology, Gothenburg, Sweden. From 1990 to 1996, he was at Chalmers. In 1997 and 1998, he was at AT&T Labs - Research in Murray Hill, NJ, USA, and in 1998 and 1999 he was at Ericsson Radio Systems AB, Kista, Sweden. Since 1999, he has been at Chalmers University, where he is a professor in communication systems. Further, he was a guest professor at Yonsei University, S. Korea, in 2003–2004. He is currently vice head of the department of Signals and Systems at Chalmers, with responsibility for undergraduate and master education. His research interests include communication, data compression, and modeling and compensation of non-ideal hardware components (e.g. amplifiers, oscillators, modulators in communication transmitters and receivers).