*(Article begins on next page)*

# On the Capacity of the Block-Memoryless Phase-Noise Channel

Giuseppe Durisi, *Senior Member, IEEE*

***Abstract*—Bounds are presented on the capacity of the block-memoryless phase-noise channel. The bounds capture the first two terms in the asymptotic expansion of capacity for SNR going to infinity and turn out to be tight for a large range of SNR values of practical interest. Through these bounds, the capacity dependency on the coherence time of the phase-noise process is determined.**

## I. Introduction

The AWGN channel with phase noise is a widely used model to capture imperfect carrier-phase tracking in wireless communications and certain impairments in fiber-optic communications [1]. The surging data-rate demands in microwave backhaul links, which can be accurately modeled as AWGN channels impaired by phase noise, has recently motivated a renewed interest in characterizing the capacity of phase-noise channels.

When the phase-noise process varies slowly, i.e., its *coherence time* is much larger than the inverse of the signal bandwidth, the phase-noise samples in the discretized channel input-output (I/O) relation are correlated. A simple way to model this correlation is to assume that the phase-noise samples remain constant over a block of $N \geq 1$ samples before changing to an independent realization [2]. The resulting channel model is commonly referred to as *block-memoryless phase-noise* channel. This model is attractive because correlation is captured by a single parameter, i.e., the *coherence time* $N$ of the phase-noise process.

To date, the capacity of the block-memoryless phase-noise channel is not known in closed form. Nuriyev and Anastasopoulos [2] proved that the capacity-achieving input distribution exhibits a circular symmetry and that the resulting input-amplitude distribution is discrete with an infinite number of mass points. They also showed that in the low-SNR regime one can approximate capacity accurately by using only few mass points, whose position can be found by numerically solving a nonconvex optimization problem. In the medium- and high-SNR regimes (SNR above $15\,\mathrm{dB}$), however, this numerical approach is unfeasible due to the large number of mass points needed to approximate capacity accurately.

*Contributions:* In this letter, we derive bounds on the capacity of the block-memoryless phase-noise channel that are tight over a large range of SNR values of practical interest. Specifically, the bounds allow us to identify the first two terms in the asymptotic expansion of capacity for SNR going to infinity, and, hence, to characterize capacity accurately at high SNR.

This work has been partly supported by the Swedish Agency for Innovation Systems (VINNOVA), within the project P36604-1 MAGIC.

G. Durisi is with the Department of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden (e-mail: durisi@chalmers.se).

*Notation:* Uppercase boldface letters denote matrices and lowercase boldface letters designate vectors. The $N\times N$ identity matrix is denoted by $\mathbf{I}_N$; $\Gamma(\cdot)$ stands for the Gamma function and $\psi(\cdot)$ is Euler's digamma function. For two functions $f(x)$ and $g(x)$, the notation $f(x) = \mathcal{O}(g(x))$, $x \to \infty$, means that $\limsup_{x\to\infty}\bigl|f(x)/g(x)\bigr| < \infty$, and $f(x) = o(g(x))$, $x\to\infty$, means that $\lim_{x\to\infty}\bigl|f(x)/g(x)\bigr| = 0$. We denote expectation by $\mathbb{E}[\cdot]$, and use the notation $\mathbb{E}_s[\cdot]$ to stress that expectation is taken with respect to the random variable $s$. With $\mathcal{CN}(\mathbf{0},\mathbf{R})$ we designate the distribution of a circularly-symmetric complex Gaussian random vector with covariance matrix $\mathbf{R}$. We say that a random variable $r$ has Gamma distribution with parameters $\alpha > 0$ and $\beta > 0$, and write $r \sim \mathrm{Gamma}(\alpha,\beta)$, if its probability density function (pdf) $q_r(r)$ is given by

$$q_r(r) = r^{\alpha-1}e^{-r/\beta}/\left(\beta^{\alpha}\Gamma(\alpha)\right), \quad r\geq 0. \tag{1}$$

Finally, $\log(\cdot)$ indicates the natural logarithm.

## II. System Model

We consider a discrete-time AWGN channel impaired by phase noise. The phase-noise process is assumed to stay constant over a block of $N$ samples and to change independently from block to block. Within one block, the channel I/O relation is given by

$$y_k = e^{j\theta}x_k + w_k, \quad k = 1,\dots,N.$$

Here, $\theta$ denotes the phase noise, which is assumed uniformly distributed on $[0,2\pi)$. Stacking the symbols transmitted within a block in a vector $\mathbf{x} = [x_1 \cdots x_N]$ and, similarly, stacking noise and output signals in corresponding vectors $\mathbf{y}$ and $\mathbf{w}$ enables us to write the I/O relation in vector form as follows:

$$\mathbf{y} = e^{j\theta}\mathbf{x} + \mathbf{w}. \tag{2}$$

We assume that $\mathbf{w} \sim \mathcal{CN}(\mathbf{0},\mathbf{I}_N)$ is independent of $\theta$, and that $\mathbf{x}$ is independent of $\theta$ and $\mathbf{w}$. We focus on the scenario where coding is performed over multiple blocks. For this scenario, the relevant performance metric is the channel ergodic capacity, which is given by

$$C(\rho) = \frac{1}{N}\sup_{Q_{\mathbf{x}}} I(\mathbf{x};\mathbf{y}) \tag{3}$$

as a consequence of the block-memoryless assumption. The supremum in (3) is over the set of input probability distributions $Q_{\mathbf{x}}$ that satisfy the average-power constraint

$$\mathbb{E}\left[\|\mathbf{x}\|^2\right] \leq N\rho. \tag{4}$$

Because the noise variance is normalized, $\rho$ is equal to the SNR.

No closed-form expression for $C(\rho)$ is available to date. For the case $N=1$, Lapidoth [3] determined the first two terms in the asymptotic expansion of $C(\rho)$ for $\rho\to\infty$. Specifically, he showed that

$$C(\rho)=\frac{1}{2}\log(\rho)-\frac{1}{2}\log(2)+o(1),\quad \rho\to\infty. \tag{5}$$

Non-asymptotic capacity bounds for $N=1$ are presented in [4]. For the general case $N\geq 1$, the following asymptotic capacity expansion is available [2]:

$$C(\rho)=\left(1-\frac{1}{2N}\right)\log(\rho)+\mathcal{O}(1),\quad \rho\to\infty. \tag{6}$$

Note that the asymptotic capacity expansion in (6) (for the general case $N\geq 1$) is less accurate than the one in (5) (for the special case $N=1$) because in (6) the second term in the expansion of $C(\rho)$ for $\rho\to\infty$ is not determined explicitly.

In this letter, we present non-asymptotic bounds on $C(\rho)$ for the general case $N\geq 1$. The bounds turn out to be tight for a large range of SNR values of practical interest. Furthermore, they allow us to refine (6) and determine the second term in the asymptotic expansion of $C(\rho)$ for $\rho\to\infty$. Specifically, we establish the following result.

*Theorem 1:* The capacity of the channel (2) is given by

$$C(\rho)=\left(1-\frac{1}{2N}\right)\log(\rho)+c_N+o(1),\quad \rho\to\infty \tag{7}$$

where

$$c_N\triangleq\left(1-\frac{1}{2N}\right)\log\left(\frac{2N}{2N-1}\right)+\frac{1}{N}\left[\log\left(\frac{\Gamma(N-1/2)}{\Gamma(N)}\right)-\frac{1}{2}\log(4\pi)\right]. \tag{8}$$

*Proof:* See Section III-D. ■

Note that by setting $N=1$ in (7) and (8), and recalling that $\Gamma(1/2)=\sqrt{\pi}$, one recovers (5).

The rest of this letter is organized as follows: in Section III-A we provide some intuition on the structure of the capacity-achieving input distribution at high SNR. Then, we use this intuition to construct a capacity lower bound (Section III-B) and an upper bound (Section III-C) that agree up to a $o(1)$ term, and, hence, allow us to establish Theorem 1. In Section IV, we present an additional capacity lower bound that, although not asymptotically tight in the sense of (7), yields (together with the upper bound in Section III-C) an accurate capacity characterization for a large range of SNR values of practical interest.

## III. Bounding Capacity at High SNR

### A. Geometric Intuition

We start by providing some geometric intuition that sheds light on the way the capacity bounds used to establish Theorem 1 are constructed. Let the capacity *pre-log* $\chi$ be defined as the asymptotic ratio between capacity and the logarithm of SNR as SNR grows to infinity, i.e.,

$$\chi=\lim_{\rho\to\infty}\big(C(\rho)/\log(\rho)\big).$$

The capacity pre-log can be interpreted as the fraction of signal-space dimensions available for communications [5]. We shall next heuristically determine this fraction for the block-memoryless phase-noise channel in (2). The multiplication of the input vector $\mathbf{x}\in\mathbb{C}^N$ by the phase noise term $e^{j\theta}$ makes one of the $2N$ real parameters characterizing $\mathbf{x}$ not recoverable (from $e^{j\theta}\mathbf{x}$) at the receiver. This means that, even in the absence of the additive noise $\mathbf{w}$, the received signal carries only $2N-1$ real parameters describing $\mathbf{x}$. Hence, the fraction of signal-space dimensions available for communication is $(2N-1)/(2N)=1-1/(2N)$, in agreement with (6) and (7).

On the basis of this observation, we choose the following distribution to evaluate the mutual information on the right-hand side (RHS) of (3) and, hence, obtain a capacity lower bound: we take $\mathbf{x}$ isotropically distributed, to exploit the circular symmetry of the I/O relation (2), and $\|\mathbf{x}\|^2$ distributed as the sum of the square of $2N-1$ independent real Gaussian random variables. This results in a Gamma distribution. To obtain a capacity upper bound that matches the lower bound (up to a $o(1)$ term), we use the *duality approach*, a technique introduced in [6] to characterize the capacity of fading channels under no *a priori* channel knowledge at the transmitter and the receiver. The essence of duality is that it allows one to by-pass the supremization in (3) and obtain tight capacity upper bounds by choosing an appropriate probability distribution on the output $\mathbf{y}$. As distribution we choose the one induced on the noiseless channel output $e^{j\theta}\mathbf{x}$ by the probability distribution on $\mathbf{x}$ used to obtain the lower bound. The approach just outlined generalizes to $N\geq 1$ the proof technique used in [3] for the case $N=1$.

### B. A Lower Bound on Capacity

To obtain a capacity lower bound, we evaluate the mutual information on the RHS of (3) for the probability distribution introduced in Section III-A. Specifically, let $\mathbf{x}=\|\mathbf{x}\|\cdot\mathbf{v}_x$, where $\mathbf{v}_x=\mathbf{x}/\|\mathbf{x}\|$. We take $\mathbf{v}_x$ uniformly distributed on the unit sphere in $\mathbb{C}^N$. Furthermore, we choose $\|\mathbf{x}\|^2=N\rho\, s/(N-1/2)$, where $s\sim\mathrm{Gamma}(N-1/2,1)$ is independent of $\mathbf{v}_x$. As $\mathbb{E}[s]=N-1/2$, the average power constraint (4) is satisfied with equality. Next, we use that, by definition, $I(\mathbf{x};\mathbf{y})=h(\mathbf{y})-h(\mathbf{y}\,|\,\mathbf{x})$ and bound the two differential entropy terms separately. For the first term, we proceed as follows:

$$\begin{aligned}
h(\mathbf{y}) &\overset{(a)}{\geq} h(\mathbf{y}\,|\,\mathbf{w}) \overset{(b)}{=} h(e^{j\theta}\mathbf{x})\\
&\overset{(c)}{=} h(\|\mathbf{x}\|^2)+\log\big(\pi^N/\Gamma(N)\big)+(N-1)\,\mathbb{E}\big[\log\big(\|\mathbf{x}\|^2\big)\big]\\
&\overset{(d)}{=} N\log\left(\frac{2N\rho}{2N-1}\right)+h(s)+\log\left(\frac{\pi^N}{\Gamma(N)}\right)\\
&\qquad+(N-1)\,\mathbb{E}[\log(s)]\\
&\overset{(e)}{=} N\log\left(\frac{2N\rho}{2N-1}\right)+\log\left(\frac{\Gamma(N-1/2)}{\Gamma(N)}\right)+(N-1/2)\\
&\qquad+\log(\pi^N)+\frac{1}{2}\psi(N-1/2).
\end{aligned} \tag{9}$$

Here, (a) follows because conditioning reduces entropy [7, Sec. 8.6], in (b) we used that differential entropy is invariant to translations [7, Thm. 8.6.3] and that $\mathbf{w}$ and $(\mathbf{x},\theta)$ are independent; in (c) we used the change of variable lemma to compute $h(e^{j\theta}\mathbf{x})$ in polar coordinates [6, Lem. 6.17 and Lem. 6.15]; we also exploited that $e^{j\theta}\mathbf{x}$ is isotropically distributed; (d) follows because $\|\mathbf{x}\|^2 = N\rho\, s/(N-1/2)$ and from [7, Eq. (8.66)]; finally, (e) follows because, for $z \sim \text{Gamma}(\alpha, 1)$,

$$\mathbb{E}[\log(z)] = \psi(\alpha)$$
$$h(z) = (1-\alpha)\psi(\alpha) + \alpha + \log\bigl(\Gamma(\alpha)\bigr).$$

We next bound $h(\mathbf{y}\,|\,\mathbf{x})$. Let $\{\tilde{w}_l\}_{l=1}^{N}$ be independent and identically distributed $\mathcal{CN}(0,1)$ random variables. Furthermore, let $\tilde{\mathbf{y}}$ be a $N$-dimensional random vector with entries $\tilde{y}_1 = e^{j\theta}\|\mathbf{x}\| + \tilde{w}_1$ and $\tilde{y}_l = \tilde{w}_l, l = 2,\dots,N$. As $\mathbf{y}$ and $\tilde{\mathbf{y}}$ are related by a unitary transformation, we have that

$$h(\mathbf{y}\,|\,\mathbf{x}) = h(\tilde{\mathbf{y}}\,|\,\mathbf{x}) = \log(\pi e)^{N-1} + h(\tilde{y}_1\,|\,\|\mathbf{x}\|). \tag{10}$$

Because the phase of $\tilde{w}_1$ is uniformly distributed on $[0, 2\pi)$, the random variable $\tilde{y}_1 = e^{j\theta}\|\mathbf{x}\| + \tilde{w}_1$ has the same distribution as $\hat{y}_1 = e^{j\theta}[\|\mathbf{x}\| + \hat{w}_1]$ where $\hat{w}_1 \sim \mathcal{CN}(0,1)$. Let now $\hat{\theta}_1$ denote the phase of $\hat{y}_1$. Then,

$$\begin{aligned}
h(\tilde{y}_1\,|\,\|\mathbf{x}\|) &= h(\hat{y}_1\,|\,\|\mathbf{x}\|)\\
&\overset{(a)}{=} h(\hat{\theta}_1\,|\,|\hat{y}_1|^2, \|\mathbf{x}\|) - \log(2) + h(|\hat{y}_1|^2\,|\,\|\mathbf{x}\|)\\
&\overset{(b)}{=} \log(\pi) + h(|\hat{y}_1|^2\,|\,\|\mathbf{x}\|) \qquad (11)\\
&\overset{(c)}{\le} \log(\pi) + \frac{1}{2}\mathbb{E}_s\left[\log\left[(2\pi e)\left(1 + \frac{4N\rho}{2N-1}s\right)\right]\right]. \qquad (12)
\end{aligned}$$

Here, (a) follows from [6, Lem. 6.16]; in (b) we used that $\theta$ is uniformly distributed on $[0, 2\pi)$ and, hence, $\hat{\theta}_1$ is uniformly distributed on $[0, 2\pi)$ as well, and independent of $|\hat{y}_1|^2$ and $\|\mathbf{x}\|$; finally (c) follows because $|\hat{y}_1|^2$ has variance $1 + 2\|\mathbf{x}\|^2$ given $\|\mathbf{x}\|$, and because the Gaussian distribution maximizes differential entropy under a variance constraint [7, Thm. 8.6.5]. Substituting (12) into (10), subtracting (10) from (9), and then dividing by $N$, we obtain: $C(\rho) \ge L_1(\rho)$, where

$$\begin{aligned}
L_1(\rho) \triangleq \log\left(\frac{2N\rho}{2N-1}\right) + \frac{1}{2N}\Biggl\{\log\left(\frac{\Gamma(N-1/2)}{\Gamma(N)}\right) - \log(2\pi)\\
+ \psi(N-1/2) - \mathbb{E}_s\left[\log\left(1 + \frac{4N\rho}{2N-1}s\right)\right]\Biggr\} \qquad (13)
\end{aligned}$$

with $s \sim \text{Gamma}(N-1/2, 1)$.

### C. An Upper Bound on Capacity

Let $q_{\mathbf{y}}(\mathbf{y})$ denote an arbitrary pdf on $\mathbf{y}$. By duality [6, Thm. 5.1], for every input probability distribution $Q_{\mathbf{x}}$ we have that

$$I(\mathbf{x};\mathbf{y}) \le -\mathbb{E}_{\mathbf{y}}[\log(q_{\mathbf{y}}(\mathbf{y}))] - h(\mathbf{y}\,|\,\mathbf{x}). \tag{14}$$

The expectation on the RHS of (14) is with respect to the probability distribution induced on $\mathbf{y}$ by $Q_{\mathbf{x}}$ through (2). Note also that for every input distribution $Q_{\mathbf{x}}$ satisfying (4)

$$1 - \Bigl[\mathbb{E}\bigl[\|\mathbf{x}\|^2\bigr] + N\Bigr] / \Bigl[N(\rho+1)\Bigr] \ge 0. \tag{15}$$

Fix now $\lambda \ge 0$ and an arbitrary pdf $q_{\mathbf{y}}(\mathbf{y})$ on $\mathbf{y}$. We can upper-bound $C(\rho)$ in (3) using (14) and (15) as follows:

$$\begin{aligned}
C(\rho) \le \frac{1}{N}\sup_{Q_{\mathbf{x}}}\Biggl\{ -\mathbb{E}_{\mathbf{y}}[\log(q_{\mathbf{y}}(\mathbf{y}))] - h(\mathbf{y}\,|\,\mathbf{x})\\
+ \lambda\left(1 - \frac{\mathbb{E}\bigl[\|\mathbf{x}\|^2\bigr] + N}{N(\rho+1)}\right)\Biggr\}. \qquad (16)
\end{aligned}$$

As in (3), the supremum is over the set of $Q_{\mathbf{x}}$ satisfying (4). Let $\mathbf{y} = \sqrt{r}\cdot\mathbf{v}_y$, where $r = \|\mathbf{y}\|^2$ and $\mathbf{v}_y = \mathbf{y}/\|\mathbf{y}\|$. To evaluate the first term on the RHS of (16), we take $q_{\mathbf{y}}(\mathbf{y})$ so that $r \sim \text{Gamma}(\alpha, \beta)$, with $\alpha$ to be optimized later, and $\beta = N(\rho+1)/\alpha$. Furthermore, we take $\mathbf{v}_y$ uniformly distributed on the unit sphere in $\mathbb{C}^N$ and independent of $r$. Let $q_r(r)$ denote the resulting pdf of $r$. By using polar coordinates,

$$\begin{aligned}
&-\mathbb{E}_{\mathbf{y}}[\log(q_{\mathbf{y}}(\mathbf{y}))] = -\mathbb{E}_{\mathbf{y}}\bigl[\log\bigl(q_{\mathbf{y}}(\sqrt{r}\cdot\mathbf{v}_y)\bigr)\bigr]\\
&\overset{(a)}{=} -\mathbb{E}_{\mathbf{y}}[\log(q_r(r))] + \log\bigl(\pi^N/\Gamma(N)\bigr) + (N-1)\,\mathbb{E}_{\mathbf{y}}[\log(r)]\\
&\overset{(b)}{=} (N-\alpha)\,\mathbb{E}_{\mathbf{y}}[\log(r)] + \alpha\frac{\mathbb{E}_{\mathbf{y}}[r]}{N(\rho+1)}\\
&\quad + \log\bigl(\pi^N\Gamma(\alpha)/\Gamma(N)\bigr) + \alpha\log(N(\rho+1)/\alpha). \qquad (17)
\end{aligned}$$

Here, in (a) we used that

$$q_{r,\mathbf{v}_y}(r, \mathbf{v}_y) = q_{\mathbf{y}}\bigl(\sqrt{r}\cdot\mathbf{v}_y\bigr)\cdot r^{N-1}/2$$

as a consequence of the change of variable theorem, and that

$$q_{r,\mathbf{v}_y}(r, \mathbf{v}_y) = q_r(r)\cdot\Gamma(N)/(2\pi^N)$$

by construction; (b) follows from (1) with $\beta = N(\rho+1)/\alpha$. Let

$$d_{\lambda,\alpha} \triangleq \log(\Gamma(\alpha)/\Gamma(N)) + \lambda - N + 1.$$

Substituting (17) and (10) into (16), and then using (11) and that $\mathbb{E}_{\mathbf{y}}[r] = \mathbb{E}\bigl[\|\mathbf{y}\|^2\bigr] = \mathbb{E}\bigl[\|\mathbf{x}\|^2\bigr] + N$, we obtain

$$\begin{aligned}
C(\rho) \le \frac{1}{N}\sup_{Q_{\mathbf{x}}}\Biggl\{\alpha\log\left(\frac{N(\rho+1)}{\alpha}\right) + d_{\lambda,\alpha}\\
+ (N-\alpha)\,\mathbb{E}_{\mathbf{y}}[\log(r)] - h(|\hat{y}_1|^2\,|\,\|\mathbf{x}\|)\\
+ (\alpha-\lambda)\frac{\mathbb{E}\bigl[\|\mathbf{x}\|^2\bigr] + N}{N(\rho+1)}\Biggr\}. \qquad (18)
\end{aligned}$$

To eliminate the supremum over $Q_{\mathbf{x}}$, we next bound the last three terms on the RHS of (18) (the only terms that depend on $Q_{\mathbf{x}}$) as follows. Let

$$\begin{aligned}
g_{\lambda,\alpha}(s,\rho) \triangleq (N-\alpha)\,\mathbb{E}_{\theta,\mathbf{w}}\bigl[\log(r)\,|\,\|\mathbf{x}\| = \sqrt{s}\bigr]\\
- h(|\hat{y}_1|^2\,|\,\|\mathbf{x}\| = \sqrt{s}) + (\alpha-\lambda)\frac{s+N}{N(\rho+1)}.
\end{aligned}$$

Then

$$\begin{aligned}
(N-\alpha)\,\mathbb{E}_{\mathbf{y}}[\log(r)] - h(|\hat{y}_1|^2\,|\,\|\mathbf{x}\|) + (\alpha-\lambda)\frac{\mathbb{E}\bigl[\|\mathbf{x}\|^2\bigr] + N}{N(\rho+1)}\\
\le \max_{s\ge 0}\Bigl\{g_{\lambda,\alpha}(s,\rho)\Bigr\}. \qquad (19)
\end{aligned}$$

Substituting (19) into (18), and minimizing the resulting bound over $\alpha > 0$ and $\lambda \ge 0$, we obtain: $C(\rho) \le U(\rho)$, where

$$\begin{aligned}
U(\rho) = \min_{\alpha>0}\min_{\lambda\ge 0}\frac{1}{N}\Biggl\{\alpha\log\left(\frac{N(\rho+1)}{\alpha}\right)\\
+ d_{\lambda,\alpha} + \max_{s\ge 0}\Bigl\{g_{\lambda,\alpha}(s,\rho)\Bigr\}\Biggr\}. \qquad (20)
\end{aligned}$$

### D. *Proof of Theorem 1*

To prove Theorem 1, we show that the lower bound $L_1(\rho)$ in (13) and a refined version of the upper bound $U(\rho)$ in (20) have the same asymptotic expansion as the one in (7). For $L_1(\rho)$, it is sufficient to note that

$$\mathbb{E}_s\left[\log\left(1+\frac{4N\rho}{2N-1}s\right)\right] = \mathbb{E}[\log(s)] + \log\left(\frac{4N\rho}{2N-1}\right) + o(1), \quad \rho\to\infty \quad (21)$$

with $\mathbb{E}[\log(s)] = \psi(N-1/2)$, and to substitute (21) into (13).

To refine $U(\rho)$, we exploit the fact that the high-SNR behavior of $C(\rho)$ does not change if we constrain the input distribution to be supported outside a sphere of arbitrary radius. This result, known as *escape-to-infinity* property of the capacity-achieving input distribution [6, Def. 4.11], is formalized in the following lemma (see [6], [5] for an intuitive interpretation).

*Lemma 2:* Fix an arbitrary $s_0 > 0$ and let $\mathcal{K} = \{\mathbf{x} \in \mathbb{C}^N : \|\mathbf{x}\|^2 \geq s_0\}$. Denote by $C^{(\mathcal{K})}(\rho)$ the capacity of the channel (2) when the input signal is subject to the average-power constraint (4) and to the additional constraint that $\mathbf{x}\in\mathcal{K}$ almost surely. Then

$$C(\rho) = C^{(\mathcal{K})}(\rho) + o(1), \quad \rho\to\infty$$

with $C(\rho)$ given in (3).

*Proof:* The lemma follows directly from [8, Thm. 8] and [6, Thm. 4.12]. ■

Fix $s_0 > 0$. By performing the same steps leading to (20), but accounting for the additional constraint that $\mathbf{x}\in\mathcal{K}$ almost surely and also setting $\alpha = N-1/2$ (as for the lower bound) and $\lambda=\alpha$, we obtain: $C^{(\mathcal{K})}(\rho) \leq U^{(\mathcal{K})}(\rho)$, where

$$U^{(\mathcal{K})}(\rho) \triangleq \left(1-\frac{1}{2N}\right)\log\left(\frac{2N(\rho+1)}{2N-1}\right) + \frac{1}{N}\left[\log\left(\frac{\Gamma(N-1/2)}{\Gamma(N)}\right) + \frac{1}{2} + \max_{s\geq s_0}\{\tilde{g}(s)\}\right] \quad (22)$$

with $\tilde{g}(s) \triangleq g_{\alpha,\lambda}(s,\rho)\,|_{\alpha=\lambda=N-1/2}$. As $\lim_{s\to\infty}\tilde{g}(s) = -(1/2)\log(4\pi e)$ (see [3, Eq. (9)] and [6, App. X]), we can make (22) to be arbitrarily close to (7) by choosing $s_0$ sufficiently large. This concludes the proof of Theorem 1.

## IV. The Low- and Medium-SNR Regimes

Differently from the upper bound $U(\rho)$, the lower bound $L_1(\rho)$ turns out to be not accurate for small $\rho$ values. Lack of tightness of $L_1(\rho)$ is due to the inequality (a) in (9), which is rather crude at low SNR. To avoid (a), one can take $\mathbf{x}\sim\mathcal{CN}(\mathbf{0},\rho\mathbf{I}_N)$, which yields

$$h(\mathbf{y}) = N\log(1+\rho) + \log(\pi e)^N. \quad (23)$$

Combining (23) with (10) and proceeding as in (12), one gets $C(\rho)\geq L_2(\rho)$, where

$$L_2(\rho) \triangleq \log(1+\rho) - \frac{1}{2N}\left\{\log(2\pi e) + \mathbb{E}_s[\log(1+2\rho s)]\right\} \quad (24)$$

with $s\sim\mathrm{Gamma}(N,1)$. We remark that the Gaussian input distribution yielding (24) was also used in [2] to establish (6). Also note that $L_2(\rho)$ is not asymptotically tight in the sense of (7).

## V. Numerical Results and Conclusions

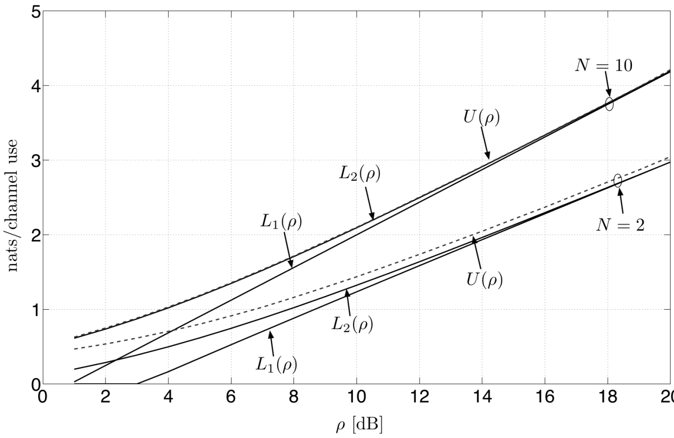


Fig. 1. The capacity lower bounds $L_1(\rho)$ and $L_2(\rho)$ (solid lines), and the capacity upper bound $U(\rho)$ (dashed line) as a function of SNR $\rho$ for $N=2$ and $N=10$.

Fig. 1 shows the capacity lower bounds $L_1(\rho)$ and $L_2(\rho)$ and the upper bound $U(\rho)$ as a function of $\rho$ for $N=2$ and $N=10$. The bounds $L_2(\rho)$ and $U(\rho)$ are surprisingly tight over the entire range of SNR values considered in the figure, and, hence, describe capacity accurately. Although $L_2(\rho)$ is not asymptotically tight in the sense of (7), the asymptotic gap between $L_2(\rho)$ and $C(\rho)$ decays quickly as a function of $N$. For the case $N=10$, this gap is smaller than $7\times10^{-5}$.

*Concluding remarks:* We conclude by observing that the capacity bounds presented in this letter are derived under the assumption of uniform phase noise. An interesting open issue is whether our analysis can be generalized to other phase-noise distributions commonly used in the wireless and fiber-optic communities (e.g., wrapped Gaussian, truncated Gaussian and von Mises/Tykhonov distributions).

## References


[1] B. Goebel, R.-J. Essiambre, G. Kramer, P. J. Winzer, and N. Hanik, "Calculation of mutual information for partially coherent Gaussian channels with applications to fiber optics," *IEEE Trans. Inf. Theory*, vol. 57, no. 9, pp. 5720–5736, Sep. 2011.

[2] 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.

[3] A. Lapidoth, "On phase noise channels at high SNR," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Bangalore, India, Oct. 2002, pp. 1–4.

[4] 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.

[5] 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.

[6] A. Lapidoth 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.

[7] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York, NY, U.S.A.: Wiley, 2006.

[8] A. Lapidoth 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.