Capacity of Discrete-Time Wiener Phase Noise Channels to Within a Constant Gap

Luca Barletta
Politecnico di Milano, Italy
luca.barletta@polimi.it

Stefano Rini
National Chiao Tung University, Taiwan
stefano@nctu.edu.tw

***Abstract*—The capacity of the discrete-time channel affected by both additive Gaussian noise and Wiener phase noise is studied. Novel inner and outer bounds are presented, which differ of at most 6.65 bits per channel use for all channel parameters. The capacity of this model can be subdivided in three regimes: (i) for large values of the frequency noise variance, the channel behaves similarly to a channel with circularly uniform iid phase noise; (ii) when the frequency noise variance is small, the effects of the additive noise dominate over those of the phase noise, while (iii) for intermediate values of the frequency noise variance, the transmission rate over the phase modulation channel has to be reduced due to the presence of phase noise.**

I. Introduction

In the discrete-time Wiener phase noise (WPN) channel, the channel input is affected by both additive white Gaussian noise (AWGN) and multiplicative Wiener phase noise. The Wiener phase process can be used to model a number of random phenomena: from imperfections in the oscillator circuits at the transceivers, to slow fading effects in wireless environments or oscillations in the laser frequency in optical communications. Despite its relevance in many practical scenarios, the capacity of the WPN channel remains an open problem as the presence of memory in the phase noise process makes the analysis challenging. In this paper we provide the first approximate characterization of capacity for the discrete-time WPN channel for all channel parameters and provide an input distribution which performs close to optimal.

State of the Art: Channel models encompassing both AWGN and multiplicative phase noise have been considered in the literature under different assumptions on the distribution of the phase noise process. The channel model in which the phase noise is randomly selected at the beginning of transmission and is kept fixed through the transmission block-length is referred to as *block-memoryless phase noise channel*. In [1] the authors prove that the capacity-achieving input distribution for this model exhibits a circular symmetry and that the distribution of the amplitude of the input is discrete with an infinite number of mass points. In [2], the author presents capacity outer and inner bounds that capture the first two terms of the asymptotic expansion of capacity as the signal-to-noise ratio (SNR) goes to infinity.

When the phase noise process is composed of iid circularly uniform samples, the channel is referred to as *non-coherent phase noise channel*. The capacity of this model is first studied in [3] where it is shown that the capacity-achieving distribution is not Gaussian. The authors of [4] improve on the results of [3] by showing that the capacity-achieving distribution, similarly to the block-memoryless phase noise channel, is discrete and possesses an infinite number of mass points.

The WPN channel encompasses the block memoryless channel and the non-coherent channel as the two limiting cases in which the variance of the innovation process tends to zero and infinity, respectively. This model was fist studied in [5] where the high SNR capacity is derived using duality arguments. The authors of [6] propose a numerical method of evaluating tight information rate bounds for this model. In [7] the authors derive analytical approximations to capacity which are shown to be tight through numerical evaluations.

Capacity bounds for more complex models taking into account oversampling at the receiver and/or effect of imperfect matched filtering are proposed in [8]–[11]: here the bounds are valid only at high SNR.

Contributions: In this paper we derive the capacity of the discrete-time WPN channel to within 6.65 bits per channel use (bpcu) for all channel parameters, namely SNR and frequency noise variance. This result is shown by separately considering three regimes of the frequency noise variance: *small*, *intermediate*, and *high* variance. The practical insights into these regimes are as follows. In the small frequency noise variance regime, the effects of the additive noise dominates over those of the phase noise: for this reason the channel behaves essentially as an AWGN channel. In the high frequency noise variance regime, the instantaneous phase variations dominate over the memory of the process and the channel resembles a non-coherent phase noise channel. In the intermediate frequency noise variance regime, part of the transmission rate over the phase modulation channel has to be sacrificed due to the presence of the phase noise.

Organization: The channel model and known results in literature are presented in Sec. II. Some preliminary results, useful for obtaining tight capacity bounds, are detailed in Sec. III. Outer bounds are derived in Sec. IV, while the capacity to within a constant gap is shown in Sec. V. Conclusions are drawn in Sec. VI.

II. Channel Model and Known Results

The discrete-time Wiener phase noise (WPN) channel is described by the input-output relationship

$$Y_i = X_i e^{j\Theta_i} + W_i, \quad i = 1, \dots, N \tag{1}$$

where $j = \sqrt{-1}$, $W_i \sim \mathcal{CN}(0,2)$, *i.e.* W_i is a circularly-symmetric complex Gaussian random variable (RV) with zero mean and variance 2. The channel input X_i is subject to an average power constraint over the transmission block-length N, that is $\sum_{i=1}^{N} \mathbb{E}[|X_i|^2] \leq NP$, while $\{\Theta_i,\ i \in \mathbb{N}\}$ is a Wiener process defined by the recursive equation

$$\Theta_0 = 0, \quad \Theta_{i+1} = \Theta_i + \Delta_i, \quad i \geq 0, \tag{2}$$

where $\Delta_i \sim \mathcal{N}(0, \sigma_\Delta^2)$, and σ_Δ^2 is the frequency noise variance. Standard definitions of code, achievable rate and capacity are assumed in the following.

Known Results: As noted in [5], conditioned on $|X_i| = |x_i|$, the modulus square of the output has a non-central chi-square distribution with non-centrality parameter $|x_i|^2$ and two degrees of freedom, *i.e.*

$$|Y_i|^2 \,|\, |X_i| = |x_i| \sim \chi_2^2(|x_i|^2). \tag{3}$$

Moreover, in the capacity-achieving distribution, the input has iid circularly symmetric phase $\angle X_i,\ i \in [1, N]$.

Lemma II.1. Ergodic phase noise capacity [5, Sec. VI]. *The capacity of the model in* (1) *when* $\{\Theta_i,\ i \in \mathbb{N}\}$ *is any stationary ergodic process with finite entropy rate is obtained as*

$$\mathcal{C} = \frac{1}{2}\log\left(1 + \frac{P}{2}\right) + \log(2\pi) - h(\{\Theta_k\}) + o(1), \tag{4}$$

where $h(\{\Theta_k\})$ *is the entropy rate of the phase noise process and* $o(1)$ *vanishes as* $P \to \infty$.

The result in Lem. II.1 establishes the high SNR capacity of the phase noise channel for a large class of phase processes but does not apply to the WPN channel, as Wiener process is not stationary.

Specializing a result of [10] to model (1) we obtain the following lemma.

Lemma II.2. WPN channel capacity inner bound [10, Sec. III]. *The capacity of the model in* (1)-(2) *is lower-bounded as*

$$\mathcal{C} \geq \sup_{b \geq 0, \nu > 0} \left\{ \frac{1}{2}\log\left(\frac{P - 2b + 2\nu}{\pi^2 e \nu \rho}\right) - \left(\frac{2 + b^{-1}}{\nu} + \frac{3b}{P - 2b}\right)\log(e) \right\}, \tag{5}$$

where $\rho = 1 - (1 - \mathbb{E}[|X_i|^{-2}])^2 e^{-\sigma_\Delta^2/2}$.

The inner bound in (5) is obtained by letting the input be a truncated exponential distribution, that is

$$p_{|X_i|^2}(x) = \frac{1}{P/2 - b}\exp\left(-\frac{x - b}{P/2 - b}\right), \qquad x \geq b. \tag{6}$$

III. Preliminary Results

In this section we present two theorems that will aid, in the following sections, the development of tight inner and outer bounds to capacity. The first theorem bounds the entropy of a wrapped Gaussian RV while the second the entropy of a chi and chi-square RVs.

Theorem III.1. Wrapped Gaussian entropy. *The entropy of the circularly wrapped Gaussian distribution* Δ *of variance* σ_Δ^2 *is lower-bounded as*

$$h(\Delta) \geq \begin{cases} \log(2\pi) - 2\frac{e^{-\sigma_\Delta^2}}{1 - e^{-\sigma_\Delta^2}}\log(e) & \sigma_\Delta^2 > 2\pi/e \\ \frac{1}{2}\log(2\pi e \sigma_\Delta^2) + g(\sigma_\Delta^2)\log(e) & \sigma_\Delta^2 \leq 2\pi/e \end{cases} \tag{7}$$

where $g(\sigma_\Delta^2)$ *is obtained as*

$$g(\sigma_\Delta^2) = \frac{1}{2}\mathrm{erf}\left(\frac{\pi}{\sqrt{2\sigma_\Delta^2}}\right) - \frac{e^{-\frac{\pi^2}{2\sigma_\Delta^2}}}{\sqrt{2\pi\sigma_\Delta^2}}\left(\pi + \frac{4(\pi + \frac{\sigma_\Delta^2}{\pi})}{1 - e^{-\frac{\pi^2}{\sigma_\Delta^2}}}\right) - \frac{1}{2}.$$

Proof: The bound involving the term $g(\sigma_\Delta^2)$ is derived in [11]: the following derivation is an alternative bound to

$$h(\Delta) \geq \frac{1}{2}\log(2\pi e \sigma_\Delta^2) + g(\sigma_\Delta^2)\log(e). \tag{8}$$

The pdf of a zero-mean wrapped Gaussian can be written, using Jacobi's triple product as

$$p_\Delta(x) = \frac{1}{2\pi}\prod_{n=1}^{\infty}(1 - q^n)(1 + q^{n-1/2}e^{+ix})(1 + q^{n-1/2}e^{-ix}) \tag{9}$$

where $q = e^{-\sigma_\Delta^2}$ so that

$$h(\Delta) = -\log\left(\frac{\phi(q)}{2\pi}\right) + 2\sum_{j=1}^{\infty}\frac{(-1)^j}{j}\frac{q^{j(j+1)/2}}{1 - q^j}\log(e) \tag{10}$$

where $\phi(x)$ is the Euler function. Note that $\phi(q)$ in (10) is less than one by definition, so that $-\log(\phi(q)) \geq 0$. Moreover, the function $\kappa(j, q)$ defined as

$$\kappa(j, q) = \frac{1}{j}\frac{q^{j(j+1)/2}}{1 - q^j} \tag{11}$$

is decreasing in j when $q \in (0, 1)$ so that

$$\sum_{j=1}^{\infty}\frac{(-1)^j}{j}\frac{q^{j(j+1)/2}}{1 - q^j} \geq -\frac{q}{1 - q}. \tag{12}$$

The upper bound in (14) follows by noting that

$$h(\Delta) = h(\angle e^{jZ}) \leq h(Z). \tag{13}$$

■

The bound in (7) essentially states that the entropy of the wrapped Gaussian $\Delta = \angle e^{jZ}$ for $Z \sim \mathcal{N}(0, \sigma_\Delta^2)$ is well approximated as

$$h(\Delta) \approx \min\left\{\frac{1}{2}\log(2\pi e \sigma_\Delta^2), \log(2\pi)\right\}, \tag{14}$$

that is, the minimum between the entropy of the Gaussian RV Z and a uniformly distributed RV with support $[0, 2\pi]$. As shown in Fig. 1, the approximation in (14) is rather tight: this figure plots the entropy of the uniform RV over $[0, 2\pi]$, the wrapped Gaussian RV and the Gaussian RV for different

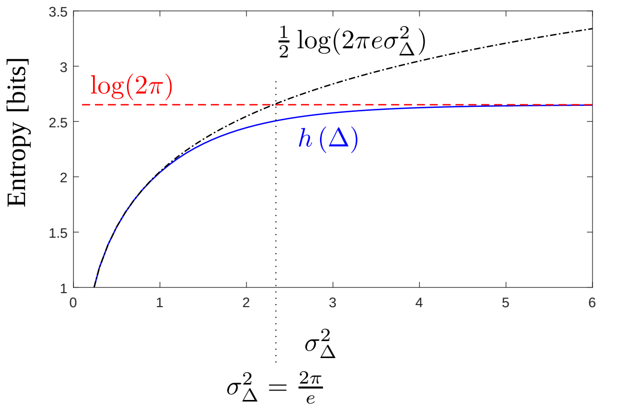

Fig. 1: The entropy of a wrapped Gaussian RV (blue solid line), a Gaussian RV (black dash-dotted line), and a uniform RV in $[0, 2\pi]$ (red dashed line) as in (14) for $\sigma_\Delta^2 \in [0, 6]$.

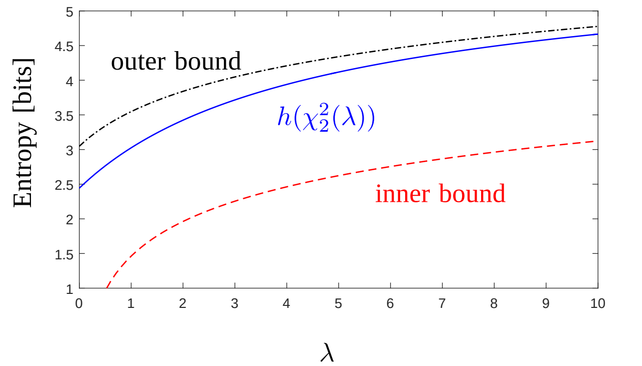

Fig. 2: The entropy of a chi-square distribution (blue solid line), the inner bound (red dashed line) and the outer bound (black dash-dotted line) in (15) for the non-centrality parameter $\lambda \in [0, 10]$.

values of σ_Δ^2.

Theorem III.2. Non-central χ_2^2 and χ_2 entropy. *The entropy of a non-central chi-square distribution with two degrees of freedom and non-central parameter λ is bounded as*

$$\frac{1}{2}\log(8\pi e\lambda) - \log(3) \le h(\chi_2^2(\lambda)) \le \frac{1}{2}\log(8\pi e(1+\lambda)). \tag{15}$$

Similarly, the entropy of a chi distribution with non-central parameter λ is lower-bounded as

$$h(\chi_2(\lambda)) \ge \frac{1}{2}\log(8\pi e) - \log(6) + \frac{1}{2}\mathrm{Ei}(-\lambda/2)\log(e). \tag{16}$$

Proof: The pdf of $T \sim \chi_2^2(\lambda)$ is

$$p_T(t) = \frac{1}{2}e^{-\frac{t+\lambda}{2}} I_0(\sqrt{\lambda t}) \tag{17}$$

where

$$\mathbb{E}[T] = 2 + \lambda \tag{18a}$$
$$\mathbb{V}\mathrm{ar}[T] = 4(1+\lambda). \tag{18b}$$

Using the bound for $I_0(x)$ of Corollary A.2, we have

$$h(T) = \mathbb{E}[-\log p_T(T)] \tag{19}$$
$$\ge \mathbb{E}\left[3\log(2) + \frac{1}{2}\left(\sqrt{T} - \sqrt{\lambda}\right)^2 \log(e) + \frac{1}{4}\log(\lambda T) - \log(\sqrt{\pi}+1)\right]. \tag{20}$$

Note that $\frac{1}{2}(\sqrt{T} - \sqrt{\lambda})^2$ is convex in T for $T \ge 0$, so that we have

$$\mathbb{E}\left[\frac{1}{2}\left(\sqrt{T} - \sqrt{\lambda}\right)^2\right] \ge \frac{1}{2}\left(\sqrt{\mathbb{E}[T]} - \sqrt{\lambda}\right)^2 = \frac{1}{2}\left(\sqrt{\lambda+2} - \sqrt{\lambda}\right)^2 \ge 0. \tag{21}$$

Furthermore, we have

$$\mathbb{E}[\log T] = \log(\lambda) - \mathrm{Ei}(-\lambda/2)\log(e) \ge \log(\lambda), \tag{22}$$

where $\mathrm{Ei}(\cdot)$ is the exponential integral function, so that the entropy bound reads as

$$h(T) \ge \mathbb{E}\left[\frac{1}{2}\log(64) + \frac{1}{4}\log(\lambda^2) \; - \log(\sqrt{\pi}+1)\right] \tag{23}$$
$$\ge \frac{1}{2}\log\left(\frac{64}{(\sqrt{\pi}+1)^2}\lambda\right)$$
$$\ge \frac{1}{2}\log(8\pi e\lambda) - \frac{1}{2}\log(3). \tag{24}$$

An upper bound on $h(T)$ can be obtained through the "Gaussian maximizes entropy" property as

$$h(T) \le \frac{1}{2}\log(2\pi e\mathbb{V}\mathrm{ar}[T])$$
$$= \frac{1}{2}\log(8\pi e(1+\lambda)). \tag{25}$$

For the entropy of a chi distribution, by change of variable, we can write

$$h(\sqrt{T}) = h(T) - \mathbb{E}[\log\sqrt{T}] - \log(2)$$
$$\overset{(15)}{\ge} \frac{1}{2}\log(8\pi e\lambda) - \log(6) - \frac{1}{2}\mathbb{E}[\log T]$$
$$\overset{(22)}{=} \frac{1}{2}\log(8\pi e) - \log(6) + \frac{1}{2}\mathrm{Ei}(-\lambda/2)\log(e). \tag{26}$$

■

The bounds (15) are plotted in Fig. 2 as a function of the non-centrality parameter λ.

IV. Outer Bounds

This section introduces novel outer bounds to the capacity of phase noise channel which connect this model to the non-coherent phase noise channel and the memoryless phase noise channel.

In the non-coherent phase noise channel, the phase noise process $\{\Theta_i, \;\; i \in \mathbb{N}\}$ is an iid sequence of uniformly distributed RVs in $[0, 2\pi]$: for this reason reliable communication can take place only over the amplitude modulation channel $p_{|Y|\,|\,|X|}$. The rate attainable over this channel is bounded in the next theorem.

Theorem IV.1. Outer bound on capacity over non-coherent AWGN channel. *The capacity of the non-coherent phase noise channel is upper-bounded as*

$$\sup_{p_X:\,\mathbb{E}|X|^2\le P} I\left(|X|\,;|X+Z|\right) \le \frac{1}{2}\log(2\pi e(P+2)), \quad (27)$$

where $Z \sim \mathcal{CN}(0,2)$.

Proof: Using the Gaussian maximum entropy bound we bound the positive entropy term as $h(|X+Z|) \le 1/2\log(2\pi e(P+2))$. For the conditional entropy we write

$$h(|X+Z|\,|\,|X|) \ge \max_{x\ge 0} h(|X+Z|\,|\,|X|, |X|>x)$$
$$\ge \max_{x\ge 0}\int_x^\infty h(|t+Z|)\,\mathrm{d}F_{|X|}(t)$$
$$\ge \max_{x\ge 0}\mathbb{P}(|X|>x)\left(\frac{1}{2}\log\left(\frac{8\pi e}{6^2}\right)+\frac{1}{2}\mathrm{Ei}\left(-\frac{x^2}{2}\right)\log(e)\right),$$

where in the last step we used the bound in Thm. III.2 and the (increasing) monotonicity of the bound in (16). Since

$$\max_{x\ge 0} f(x) = \max\left\{\max_{0\le x<r} f(x), \max_{x\ge r} f(x)\right\}, \quad (28)$$

and by choosing r such that $\frac{1}{2}\log(8\pi e) - \log(6) + \frac{1}{2}\mathrm{Ei}(-r^2/2)\log(e) = 0$ (there exists only one value of r with this property, namely $r \approx 0.937$), we conclude that $h(|X+Z|\,|\,|X|)$ must be positive, since the maximization for $0 \le x < r$ gives a negative number, while the maximization for $x \ge r$ gives a positive number. ■

The next outer bound is a refinement of the result in Lem. II.1 to yield an outer bound to the capacity of the model in (1)-(2) valid at finite SNRs. The bound is derived by revealing the past phase realization to the receiver, which results in a memoryless phase noise channel.

Theorem IV.2. Memoryless phase noise channel outer bound. *The capacity of the WPN channel in* (1)-(2) *can be upper-bounded as*

$$\mathcal{C} \le \min\left\{\frac{1}{2}\log(2\pi e(P+2)) + \log(2\pi) - h(\Delta),\right.$$
$$\left.\log\left(1+P/2\right)\right\}, \quad (29)$$

where $h(\Delta)$ *is the entropy of a wrapped Gaussian with variance* σ_Δ^2.

Proof: A trivial capacity outer bound is $\mathcal{C} \le \log(1+P/2)$ and is obtained by providing the phase noise sequence to the receiver. Another capacity outer bound is obtained as follows:

$$I(X^N;Y^N) \le \sum_k I(X^N,\Theta_{k-1};Y_k|Y^{k-1})$$
$$= \sum_k \left(I(X^N;Y_k|\Theta_{k-1}) + I(\Theta_{k-1};Y_k|Y^{k-1})\right)$$
$$= \sum_k \left(I(X_k;Y_k|\Theta_{k-1}) + I(\Theta_{k-1};Y_k|Y^{k-1})\right).$$

Since the additive noise W is circularly symmetric, a uniformly distributed input phase $\underline{/X_k}$ in $[0,2\pi)$ is capacity-achieving. This also implies that

$$I(\Theta_{k-1};Y_k|Y^{k-1})$$
$$= I(\Theta_{k-1};|X_k|e^{j(\Theta_{k-1}+\Delta_{k-1}+\angle X_k)}+W_k|Y^{k-1}) = 0,$$

given that $\underline{/X_k}$ is independent of Θ_{k-1}. Using the polar decomposition $X = |X|e^{j\angle X}$, and dropping the time index for convenience of notation, write

$$I\left(X_k;Y_k\,|\,\Theta_{k-1}\right) = I\left(X\,;Xe^{j\Delta}+W\right) \quad (30a)$$
$$= I\left(|X|\,;|X+W|\right) + I\left(\underline{/X};Xe^{i\Delta}+W\,\middle|\,|X|\right) \quad (30b)$$
$$\le I\left(|X|\,;|X+W|\right) + I\left(\underline{/X}\,;\underline{/X}\oplus\Delta\right) \quad (30c)$$
$$\le \frac{1}{2}\log(2\pi e(P+2)) + \log(2\pi) - h(\Delta), \quad (30d)$$

where (30b) follows by circular symmetry of W, (30c) by revealing W to the receiver, and (30d) from Thm. IV.1 and the circular symmetry of X. The symbol $\oplus$ in (30c) denotes the addition modulo 2π. ■

V. Main Result

Theorem V.1. Capacity to within a constant gap. *The capacity of the WPN channel in* (1)-(2) *is upper-bounded as*

$$\mathcal{C} \le \frac{1}{2}\log(1+P/2)$$
$$+\begin{cases} \frac{1}{2}\log(4\pi e) + 2\frac{e^{-\frac{2\pi}{e}}}{1-e^{-\frac{2\pi}{e}}}\log(e) & \sigma_\Delta^2 > \frac{2\pi}{e} \\ \frac{1}{2}\log\left(\frac{2}{\sigma_\Delta^2}\right) + \log(2\pi) + \log^2(e) & P^{-1} \le \sigma_\Delta^2 \le \frac{2\pi}{e} \\ \frac{1}{2}\log(1+P/2) & P^{-1} > \sigma_\Delta^2 \end{cases} \quad (31)$$

and the exact capacity is to within $\mathcal{G}$ bpcu *from the outer bound in* (31), *where*

$$\mathcal{G} \le \begin{cases} 4 & \sigma_\Delta^2 > \frac{2\pi}{e} \\ 6.65 & P^{-1} \le \sigma_\Delta^2 \le \frac{2\pi}{e} \\ 1.21 & P^{-1} > \sigma_\Delta^2 \end{cases} \quad (32)$$

Proof: The achievability proof relies on a simple transmission scheme employing iid complex Gaussian inputs while the converse proof is developed from the outer bound in Thm. IV.2.

Achievability: As in [10, Eq. (18)], the WPN channel capacity can be lower-bounded as

$$\mathcal{C} \ge \frac{1}{N}\sum_{k=1}^N I\left(|X_k|;Y_1^N\,\middle|\,X_1^{k-1}\right) + I\left(\underline{/X_k};Y_1^N\,\middle|\,X_1^{k-1},|X_k|\right)$$
$$\ge I\left(|X_1|^2\,;|Y_1|^2\right) + I\left(\underline{/X_1};Y_0^1\,\middle|\,X_0,|X_1|\right). \quad (33)$$

Next we lower-bound the term $I_{||} = I\left(|X_1|^2\,;|Y_1|^2\right)$ and $I_\angle = I\left(\underline{/X_1};Y_0^1\,\middle|\,X_0,|X_1|\right)$, which we refer to as the amplitude and phase modulation channel, respectively.

• **Amplitude modulation channel:** Let the input distribution be $X \sim \mathcal{CN}(0,P)$. Recognizing that $|Y|^2 = (1+P/2)K$ where $K \sim \chi_2^2$, we obtain

$$I\left(|X|^2\,;|Y|^2\right) = \log\left(e(2+P)\right) - h(|Y|^2\,|\,|X|^2). \quad (34)$$

Using the outer bound of Thm. III.2, we bound the term $h(|Y|^2\,|\,|X|^2)$ as

$$h(|Y|^2\,|\,|X|^2) \leq \frac{1}{2}\mathbb{E}\log(8\pi e(|X|^2+1)) \leq \frac{1}{2}\log(8\pi e(1+P)). \tag{35}$$

Finally, combining (34) and (35) yields

$$I_{||} \geq \frac{1}{2}\log\left(1+\frac{P}{2}\right) - \frac{1}{2}\log\left(\frac{2\pi}{e}\right) + \frac{1}{2}\log\left(\frac{1+P/2}{1+P}\right). \tag{36}$$

• **Phase modulation channel:** The second term in the RHS of (33) can be lower-bounded as follows:

$$\begin{aligned} I_\angle &= I\left(\angle X_1; Y_0^1 \,\middle|\, X_0, |X_1|\right) \\ &= I\left(\angle X_1; X_0+W_0 \,|\, X_0, |X_1|\right) \\ &\quad + I\left(\angle X_1; X_1 e^{j\Delta_0}+W_1 \,\middle|\, X_0, |X_1|, X_0+W_0\right) \\ &= I\left(\angle X_1; X_1 e^{j\Delta_0}+W_1 \,\middle|\, |X_1|\right) \\ &\geq I\left(\angle X_1; \angle X_1 \oplus \Delta_0 \oplus N \,|\, |X_1|\right) \\ &= \log(2\pi) - h(\Delta_0 \oplus N |\, |X_1|), \end{aligned} \tag{37}$$

where (37) follows by considering just the phase of Y_1, and $N = \angle |X_1| + W_1$. Since $\Delta_0 \oplus N$ is defined over the support $[0, 2\pi]$, we can apply the maximum entropy theorem to upper-bound the conditional entropy term $h(\Delta_0 \oplus N |\, |X_1|)$ with the entropy of a wrapped Gaussian RV with variance $\sigma_\Delta^2 + 1/|X_1|^2$:

$$h(\Delta_0 \oplus N |\, |X_1|) \tag{38a}$$

$$\leq \frac{1}{2}\mathbb{E}\left[\log\left(2\pi e\left(\sigma_\Delta^2 + \frac{1}{|X_1|^2}\right)\right)\right] \tag{38b}$$

$$= \frac{1}{2}\mathbb{E}\left[\log\left(2\pi e\left(\sigma_\Delta^2|X_1|^2+1\right)\right)\right] - \frac{1}{2}\mathbb{E}\left[\log|X_1|^2\right]$$

$$\leq \frac{1}{2}\log\left(2\pi e\left(\frac{\sigma_\Delta^2 P+1}{P}\right)\right) + \frac{\gamma}{2}\log(e), \tag{38c}$$

where (38b) follows by Thm. III.1, and (38c) from Jensen's inequality for the first term and from the fact that $\mathbb{E}\log|X_1|^2 = \log(Pe^{-\gamma})$, where γ is the Euler-Mascheroni constant.

Putting together the contributions of amplitude and phase modulation in (36) and (38) respectively, we obtain the inner bound

$$\mathcal{C} \geq \frac{1}{2}\log\left(1+\frac{P}{2}\right) + \frac{1}{2}\log\left(\frac{1+P/2}{1+P}\frac{Pe^{-\gamma}}{\sigma_\Delta^2 P+1}\right). \tag{39}$$

Note that, for any fixed σ_Δ^2, the pre-log at large P given by (39) is $1/2$. Also, note that if $\sigma_\Delta^2 \leq 1/P$, then the pre-log at large P is 1.

Converse and gap from capacity: The outer bound in Th. IV.2 can be sub-divided in the three regimes of small, intermediate, and high frequency noise variance.

• **High frequency noise variance:** $\sigma_\Delta^2 > 2\pi/e$**:** The outer bound from Th. IV.2 together with the condition $\sigma_\Delta^2 > 2\pi/e$ yields the outer bound

$$\mathcal{C} \leq \begin{cases} \frac{1}{2}\log(1+P/2) & P \leq 10 \\ \frac{1}{2}\log(2\pi e(P+2)) + 2\frac{e^{-\frac{2\pi}{e}}}{1-e^{-\frac{2\pi}{e}}}\log(e) & P > 10 \end{cases} \tag{40}$$

By comparing the outer bound in (40) with the inner bound in (39) we obtain the capacity gap

$$\mathcal{G} \leq \begin{cases} \frac{1}{2}\log(1+P) + \frac{1}{2}\log\left(\frac{2\pi}{e}\right) & P \leq 10 \\ \frac{1}{2}\log\left(\frac{1+P}{2+P}\right) + 2\frac{e^{-\frac{2\pi}{e}}}{1-e^{-\frac{2\pi}{e}}}\log(e) + \log(4\pi) & P > 10 \end{cases} \tag{41}$$

which is smaller than 4 bpcu for any $P > 0$ and $\sigma_\Delta^2 > 2\pi/e$.

• **Small frequency noise variance:** $\sigma_\Delta^2 < P^{-1}$**:** In this regime we consider the trivial outer bound $\mathcal{C} \leq \log(1+P/2)$. When $P < 2$, capacity is necessarily less than $1\,\mathrm{bpcu}$. When $P \geq 2$, the gap between the trivial outer bound and the inner bound in (39) is at most $1.21\,\mathrm{bpcu}$.

• **Intermediate frequency noise variance:** $\frac{1}{P} \leq \sigma_\Delta^2 \leq \frac{2\pi}{e}$**:** The outer bound in Th. IV.2 in this regime can be rewritten as

$$\begin{aligned} \mathcal{C} &\leq \frac{1}{2}\log\left(\frac{P+2}{\sigma_\Delta^2}\right) + \log(2\pi) - g(\sigma_\Delta^2)\log(e) \\ &\leq \frac{1}{2}\log\left(\frac{P+2}{\sigma_\Delta^2}\right) + \log(2\pi) + \log^2(e), \end{aligned} \tag{42}$$

where in the last step we note that $g(\sigma_\Delta^2)$ is decreasing in σ_Δ^2 so that $g(\sigma_\Delta^2) \geq -\log(e)$. We next compare the inner bound in (39) with the outer bound in (42) and obtain

$$\begin{aligned} \mathcal{G} &= \left(\frac{\gamma}{2}+\log(e)\right)\log(e) + \frac{1}{2}\log\left((2\pi)^2\frac{1+P}{2+P}\frac{1+\sigma_\Delta^2 P}{\sigma_\Delta^2 P}\right) \\ &\leq \left(\frac{\gamma}{2}+\log(e)\right)\log(e) + \log(2\pi) + \frac{3}{2}\log(2) \leq 6.65, \end{aligned}$$

where, in the last passage, we have considered the worst case with $\sigma_\Delta^2 = P^{-1}$ and $P \to \infty$. ■

Note that the inner bound in Th. V.1 relies on iid complex Gaussian inputs, thus showing that this input distribution performs sufficiently close to capacity.

Some further insights on the result in Th. V.1 emerge by simplifying the outer bound in (31) as in the following lemma.

Lemma V.2. Larger gap, simpler expression. *The capacity of the WPN channel in in* (1)-(2) *is upper-bounded as*

$$\mathcal{C} \leq \begin{cases} \frac{1}{2}\log(1+P/2)+4 & \sigma_\Delta^2 > \frac{2\pi}{e} \\ \frac{1}{2}\log(1+P/2) - \frac{1}{2}\log\sigma_\Delta^2 + 5.5 & P^{-1} \leq \sigma_\Delta^2 \leq \frac{2\pi}{e} \\ \log(1+P/2) & P^{-1} > \sigma_\Delta^2 \end{cases} \tag{43}$$

and the exact capacity is to within a gap of 7 bpcu *from the outer bound in* (43).

The three regimes in (43) can be intuitively explained as follows. The inequality $\sigma_\Delta^2 \gtrless 2\pi/e$ in (43) arises from the result in Th. III.1 and whether the phase noise entropy is better approximated by using a uniform RV over $[0, 2\pi]$ or a Gaussian RV. When the frequency noise variance is high, the

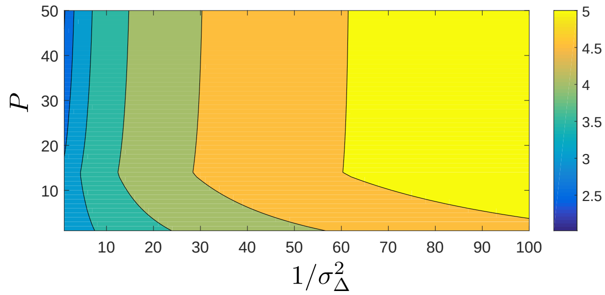

Fig. 3: A contour plot of the exact gap between inner and outer bound in the proof of Th. V.1 for $P \in [0, 50]$ and $1/\sigma_\Delta^2 \in [1, 100]$.

channel behaves similarly to a channel with uniform phase noise, in which transmission takes place only in the amplitude modulation channel. When the frequency noise variance is small, instead, the effect of the multiplicative noise times the channel input has variance $P\sigma^2 \le 1$ which is smaller than the variance of the additive noise. The inequality $P^{-1} \le \sigma_\Delta^2$, as in Th. IV.2, intuitively arises from the rate attainable on the phase modulation channel: the largest rate attainable on this channel is close to $1/2\log(P+1)$, when the frequency noise variance is relatively small. For a higher frequency noise variance, the rate of the phase modulation channel is instead close to $1/2\log(\sigma_\Delta^{-2})$.

Although the largest gap between inner and outer bound is bounded by $6.65 \ \mathrm{bpcu}$, the difference between inner and outer bounds for any parameter regime can be easily evaluated from the proof of Th. V.1. A plot of the exact value of $\mathcal{G}$ in (32) is presented in Fig. 3.

VI. Conclusions

We have derived outer and inner bounds of the discrete-time Wiener phase noise channel and have shown that they differ of at most $6.65 \ \mathrm{bpcu}$ at any SNR and frequency noise variance. Both bounds have rather simple expressions and suggest that all channels can be roughly divided in three parameter regimes: high, intermediate, and small frequency noise variance. Moreover, the analysis of the inner bound shows that a complex Gaussian input distribution performs rather close to capacity.

Appendix

Lemma A.1. Upper bound on modified Bessel function *The zero-th order modified Bessel function of the first kind can be upper-bounded as*

$$I_0(x) \le \frac{e^x}{4\sqrt{x}}\left(\mathrm{erf}(\sqrt{x})\sqrt{\pi} + \frac{2e^{-x}(1-e^{-x})}{\sqrt{x}}\right) \tag{44}$$

for any $x \ge 0$, where

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}\mathrm{d}t. \tag{45}$$

Proof: By definition

$$\begin{aligned} I_0(x) &= \frac{1}{\pi}\int_0^{\pi} e^{x\cos(\theta)}d\theta \\ &= \underbrace{\frac{1}{\pi}\int_0^{\pi/2} e^{x\cos(\theta)}d\theta}_{\mathcal{I}_1} + \underbrace{\frac{1}{\pi}\int_{\pi/2}^{\pi} e^{x\cos(\theta)}d\theta}_{\mathcal{I}_2}. \end{aligned} \tag{46}$$

Using the infinite product formula for the cosine function and a linear lower bound for the cosine, we obtain the bound

$$e^{x\cos(\theta)} \le \begin{cases} e^{x(1-\frac{4}{\pi^2}\theta^2)} & 0 \le \theta \le \pi/2 \\ e^{x(1-\frac{2}{\pi}\theta)} & \pi/2 \le \theta \le \pi \end{cases} \tag{47}$$

so that

$$\begin{aligned} \mathcal{I}_1 &\le \frac{1}{\pi}\int_0^{\pi/2} e^{x\left(1-\frac{4}{\pi^2}\theta^2\right)}d\theta \\ &= \frac{\sqrt{\pi}}{4}\frac{e^x}{\sqrt{x}}\mathrm{erf}(\sqrt{x}) \end{aligned} \tag{48}$$

while

$$\begin{aligned} \mathcal{I}_2 &\le \frac{1}{\pi}\int_{\pi/2}^{\pi} e^{x\left(1-\frac{2}{\pi}\theta\right)}d\theta \\ &= \frac{1-e^{-x}}{2x}. \end{aligned} \tag{49}$$

Combining (48) and (49) we obtain the bound

$$I_0(x) \le \frac{e^x}{4\sqrt{x}}\left(\mathrm{erf}(\sqrt{x})\sqrt{\pi} + \frac{2e^{-x}(1-e^{-x})}{\sqrt{x}}\right) \tag{50}$$

■

The result in Lemma A.1 can be further weakened to obtain an expression which can be more easily manipulated analytically.

Corollary A.2. *An upper bound to $I_0(x)$ is*

$$I_0(x) \le \frac{e^x}{\sqrt{x}}\frac{\sqrt{\pi}+1}{4}. \tag{51}$$

Proof: From Lemma A.1 we have

$$\begin{aligned} I_0(x) &\le \frac{e^x}{4\sqrt{x}}\left(\mathrm{erf}(\sqrt{x})\sqrt{\pi} + \frac{2e^{-x}(1-e^{-x})}{\sqrt{x}}\right) \\ &\le \frac{e^x}{4\sqrt{x}}\left(\sqrt{\pi}+1\right) \end{aligned} \tag{52}$$

where the last equality follows by $\mathrm{erf}(x) \le 1$ and the fact that $2e^{-x}(1-e^{-x}) < \sqrt{x}$. ■

Note that, although simpler, the expression in (51) is not tight as $x \to 0$, where it actually has an asymptote. The expression in (44), although cumbersome, is tight as $x \to 0$.

References

[1] R. Nuriyev and A. Anastasopoulos, “Capacity and coding for the block-independent noncoherent awgn channel,” *IEEE Transactions on Information Theory*, vol. 51, no. 3, pp. 866–883, March 2005.

[2] G. Durisi, “On the capacity of the block-memoryless phase-noise channel,” *IEEE Communications Letters*, vol. 16, no. 8, pp. 1157–1160, 2012.

[3] G. Colavolpe and R. Raheli, “The capacity of the noncoherent channel,” *European transactions on telecommunications*, vol. 12, no. 4, pp. 289–296, 2001.

[4] M. Katz and S. Shamai, “On the capacity-achieving distribution of the discrete-time noncoherent and partially coherent AWGN channels,” *IEEE Transactions on Information Theory*, vol. 50, no. 10, pp. 2257–2270, 2004.

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

[6] 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.

[7] M. R. Khanzadi, R. Krishnan, J. Söder, and T. Eriksson, “On the capacity of the Wiener phase-noise channel: Bounds and capacity achieving distributions,” *IEEE Trans. Commun.*, vol. 63, no. 11, pp. 4174–4184, 2015.

[8] H. Ghozlan and G. Kramer, “On Wiener phase noise channels at high signal-to-noise ratio,” in *IEEE Int. Symp. Inf. Theory (ISIT)*, 2013, pp. 2279–2283.

[9] ——, “Phase modulation for discrete-time Wiener phase noise channels with oversampling at high SNR,” in *Proc. IEEE Int. Symp. Inf. Theory*, June 2014, pp. 1554–1557.

[10] L. Barletta and G. Kramer, “Lower bound on the capacity of continuous-time Wiener phase noise channels,” in *Proc. IEEE Int. Symp. Inf. Theory*, June 2015, pp. 2326–2330.

[11] ——, “Upper bound on the capacity of discrete-time Wiener phase noise channels,” in *Proc. IEEE Inf. Theory Workshop*, April 2015, pp. 1–5.