On Wiener Phase Noise Channels at High Signal-to-Noise Ratio

Hassan Ghozlan
Department of Electrical Engineering
University of Southern California
Los Angeles, CA 90089 USA
ghozlan@usc.edu

Gerhard Kramer
Institute for Communications Engineering
Technische Universität München
80333 Munich, Germany
gerhard.kramer@tum.de

Abstract—**Consider a waveform channel where the transmitted signal is corrupted by Wiener phase noise and additive white Gaussian noise (AWGN). A discrete-time channel model that takes into account the effect of filtering on the phase noise is developed. The model is based on a multi-sample receiver which, at high Signal-to-Noise Ratio (SNR), achieves a rate that grows logarithmically with the SNR if the number of samples per symbol grows with the square-root of the SNR. Moreover, the pre-log factor is at least 1/2 in this case.**

I. INTRODUCTION

Phase noise is an impairment that often arises in coherent communication systems. Different models are adopted for the phase noise process depending on the application. In [1], Katz and Shamai studied a discrete-time model of a phase noise channel (partially coherent channel) in which the phase noise is independent and identically distributed (i.i.d.) with a Tikhonov distribution. This model is reasonable for the residual phase error of a phase-tracking scheme, such as a Phase-Locked Loop (PLL). In [2], the authors investigate white (Gaussian) phase noise for which they observed a "spectral loss" phenomenon. The white phase noise approximates the nonlinear effect of cross-phase modulation (XPM) in a Wavelength-Division Multiplexing (WDM) optical communication system. Lapidoth studied in [3] a *discrete-time* phase noise channel

$$Y_k = X_k e^{j\Theta_k} + N_k \qquad (1)$$

at high SNR, where $\{Y_k\}$ is the output, $\{X_k\}$ is the input, $\{\Theta_k\}$ is the phase noise process and $\{N_k\}$ is the additive noise. He considered both memoryless phase noise and phase noise with memory. He showed that the capacity grows *logarithmically* with the SNR with a pre-log factor 1/2, where the pre-log is due to amplitude modulation only. The phase modulation contributes a bounded number of bits only.

In this paper, we study a communication system in which the transmitted *waveform* is corrupted by Wiener phase noise and AWGN. The model is

$$r(t) = x(t) \, \exp(j\theta(t)) + n(t), \text{ for } t \in \mathbb{R} \qquad (2)$$

where $x(t)$ and $r(t)$ are the transmitted and received signals, respectively, while $n(t)$ and $\theta(t)$ are the additive and phase noise, respectively. A detailed description of the model is given in Sec. II. One application for such a channel model is optical communication under linear propagation, in which the laser phase noise is a continuous-time Wiener process (see [4] and references therein). Since the sampling of a continuous-time Wiener process yields a discrete-time Wiener process (Gaussian random walk), it is tempting to use the model (1) with $\{\Theta\}$ as a discrete-time Wiener process, but this ignores the effect of *filtering* prior to sampling. It was pointed out in [4] that "even coherent systems relying on amplitude modulation (phase noise is obviously a problem in systems employing phase modulation) will suffer some degradation due to the presence of phase noise". This is because the filtering converts phase fluctuations to amplitude variations. It is worth mentioning that filtering is necessary before sampling to limit the variance of the noise samples.

The model (1) thus does not fit the channel (2) and it is not obvious whether a pre-log 1/2 is achievable. The model that takes the effect of (matched) filtering into account is

$$Y_k = X_k H_k + N_k \qquad (3)$$

where $\{H_k\}$ is a fading process. The model (3) falls in the class of non-coherent fading channels, i.e., the transmitter and receiver have knowledge of the distribution of the fading process $\{H_k\}$, but have no knowledge of its realization. For such channels, Lapidoth and Moser showed in [5] that, at high SNR, the capacity grows *double-logarithmically* with the SNR, when the process $\{H_k\}$ is stationary, ergodic, and *regular*.

Rather than using a matched filter and sampling its output at the symbol rate, we use a multi-sample receiver, i.e., a filter whose output is sampled many times per symbol. We show that this receiver achieves a rate that grows *logarithmically* with the SNR if the number of samples per symbol grows with the square-root of the SNR. Furthermore, we show that a pre-log of 1/2 is achievable through amplitude modulation. In this paper, we study only rectangular pulses but we believe that the results hold qualitatively for other pulses.

The paper is organized as follows. The continuous-time model is described in Sec. II and the discretization is described in Sec. III. We derive a lower bound on the capacity in Sec. IV and discuss our result in Sec. V. Finally, we conclude the paper with Sec. VI.

II. CONTINUOUS-TIME MODEL

We use the following notation: $j = \sqrt{-1}$, $*$ denotes the complex conjugate, δ_D is the Dirac delta function, $\lceil \cdot \rceil$ is the ceiling operator, $\Re[\cdot]$ is the real part of a complex number, $\log(\cdot)$ is the natural logarithm and we use X_1^k to denote the k-tuple $(X_1, X_2, \ldots, X_k)$. Suppose the transmit-waveform is $x(t)$ and the receiver observes

$$r(t) = x(t) \, \exp(j\theta(t)) + n(t) \tag{4}$$

where $n(t)$ is a realization of a white circularly-symmetric complex Gaussian process $N(t)$ with

$$\mathbb{E}\left[N(t)\right] = 0$$
$$\mathbb{E}\left[N(t_1)N^*(t_2)\right] = \sigma_N^2 \, \delta_D(t_2 - t_1). \tag{5}$$

The phase $\theta(t)$ is a realization of a Wiener process $\Theta(t)$:

$$\Theta(t) = \Theta(0) + \int_0^t W(\tau)d\tau \tag{6}$$

where $\Theta(0)$ is uniform on $[-\pi, \pi)$ and $W(t)$ is a real Gaussian process with

$$\mathbb{E}\left[W(t)\right] = 0 \tag{7}$$
$$\mathbb{E}\left[W(t_1)W(t_2)\right] = 2\pi\beta \, \delta_D(t_2 - t_1). \tag{8}$$

The processes $N(t)$ and $\Theta(t)$ are independent of each other and independent of the input as well. $N_0 = 2\sigma_N^2$ is the single-sided power spectral density of the additive noise. The parameter β is called the full-width at half-maximum (FWHM), because the power spectral density of $e^{j\Theta(t)}$ has a Lorentzian shape, for which β is the full-width at half the maximum. The transmitted waveforms must satisfy the power constraint

$$\mathbb{E}\left[\frac{1}{T}\int_0^T |X(t)|^2 dt\right] \leq \mathcal{P} \tag{9}$$

where T is the transmission interval.

III. DISCRETE-TIME MODEL

Let $(x_1, x_2, \ldots, x_n)$ be the codeword sent by the transmitter. Suppose the transmitter uses a unit-energy rectangular pulse, i.e., the waveform sent by the transmitter is

$$x(t) = \sum_{m=1}^n x_m \, g(t - (m-1)T_{\text{symbol}}) \tag{10}$$

where T_{symbol} is the symbol interval and

$$g(t) \equiv \begin{cases} \sqrt{1/T_{\text{symbol}}}, & 0 \leq t < T_{\text{symbol}}, \\ 0, & \text{otherwise}. \end{cases} \tag{11}$$

Let L be the number of samples per symbol ($L \geq 1$) and define the sample interval Δ as

$$\Delta = \frac{T_{\text{symbol}}}{L}. \tag{12}$$

The received waveform $r(t)$ is filtered using an integrator over a sample interval to give the output signal

$$y(t) = \int_{t-\Delta}^t r(\tau) \, d\tau. \tag{13}$$

where $y(t)$ is a realization of $Y(t)$. The output $Y(t)$ is sampled every Δ seconds which yields the discrete-time model:

$$Y_k = X_{\lceil k/L \rceil}\Delta \, e^{j\Theta_k} \, F_k + N_k \tag{14}$$

for $k = 1, \ldots, nL$, where $Y_k \equiv Y(k\Delta)$, $\Theta_k \equiv \Theta((k-1)\Delta)$,

$$F_k \equiv \frac{1}{\Delta} \int_{(k-1)\Delta}^{k\Delta} e^{j(\Theta(\tau) - \Theta_k)} \, d\tau \tag{15}$$

and

$$N_k \equiv \int_{(k-1)\Delta}^{k\Delta} N(\tau) \, d\tau. \tag{16}$$

The process $\{N_k\}$ is an i.i.d. circularly-symmetric complex Gaussian process with mean 0 and $\mathbb{E}[|N_k|^2] = \sigma_N^2\Delta$ while the process $\{\Theta_k\}$ is the discrete-time Wiener process:

$$\Theta_k = \Theta_{k-1} + W_k \tag{17}$$

where Θ_1 is uniform on $[-\pi, \pi)$ and $\{W_k\}$ is an i.i.d. real Gaussian process with mean 0 and $\mathbb{E}[|W_k|^2] = 2\pi\beta\Delta$. The process $\{F_k\}$ is an i.i.d. process. Moreover, $\{F_k\}$ and $\{W_k\}$ are independent of $\{N_k\}$ but not independent of each other.

Equations (9) – (11) imply the power constraint

$$\frac{1}{n}\sum_{m=1}^n \mathbb{E}[|X_m|^2] \leq P = \mathcal{P}T_{\text{symbol}}. \tag{18}$$

IV. LOWER BOUND

For the kth input symbol X_k we have L outputs, so it is convenient to group the L samples per symbol in one vector and define $\mathbf{Y}_k \equiv (Y_{(k-1)L+1}, Y_{(k-1)L+2}, \ldots, Y_{(k-1)L+L})$. We further define $X_A \equiv |X|$ and $X_\Phi \equiv \angle X$. We decompose the mutual information using the chain rule into two parts:

$$I(X_1^n; \mathbf{Y}_1^n) = I(X_{A,1}^n; \mathbf{Y}_1^n) + I(X_{\Phi,1}^n; \mathbf{Y}_1^n | X_{A,1}^n). \tag{19}$$

The first term represents the contribution of the amplitude modulation while the second term represents the contribution of the phase modulation. We focus on the amplitude contribution and use $I(X_{\Phi,1}^n; \mathbf{Y}_1^n | X_{A,1}^n) \geq 0$ to obtain the lower bound

$$I(X_1^n; \mathbf{Y}_1^n) \geq I(X_{A,1}^n; \mathbf{Y}_1^n). \tag{20}$$

Suppose that $X_{A,1}^n$ is i.i.d. Hence, we have

$$\begin{aligned} I(X_{A,1}^n; \mathbf{Y}_1^n) &\overset{(a)}{=} \sum_{k=1}^n I(X_{A,k}; \mathbf{Y}_1^n | X_{A,1}^{k-1}) \\ &\overset{(b)}{=} \sum_{k=1}^n H(X_{A,k}) - H(X_{A,k} | \mathbf{Y}_1^n \, X_{A,1}^{k-1}) \\ &\overset{(c)}{\geq} \sum_{k=1}^n I(X_{A,k}; \mathbf{Y}_k) \\ &\overset{(d)}{\geq} \sum_{k=1}^n I(X_{A,k}; V_k) \end{aligned} \tag{21}$$

where

$$V_k = \sum_{\ell=1}^{L} |Y_{(k-1)L+\ell}|^2. \qquad (22)$$

Step (a) follows from the chain rule of mutual information, (b) follows from the independence of $X_{A,1}, X_{A,2}, \ldots, X_{A,n}$, (c) holds because conditioning does not increase entropy, and (d) follows from the data processing inequality. Since $X_{A,1}^n$ is identically distributed, then V_1^n is also identically distributed and we have, for $k \geq 2$,

$$I(X_{A,k}; V_k) = I(X_{A,1}; V_1). \qquad (23)$$

In the rest of this section, we consider only one symbol $(k = 1)$ and drop the time index. Moreover, we assume that $T_{\text{symbol}} = 1$ for simplicity. By combining (22) and (14), we have

$$V = \sum_{\ell=1}^{L} \left(X_A^2 \Delta^2 |F_\ell|^2 + 2 X_A \Delta \Re[e^{j\Phi_X} e^{j\Theta_\ell} F_\ell N_\ell^*] + |N_\ell|^2 \right)$$
$$= X_A^2 \Delta G + 2 X_A \Delta Z_1 + Z_0 \qquad (24)$$

where G, Z_1 and Z_0 are defined as

$$G \equiv \frac{1}{L} \sum_{\ell=1}^{L} |F_\ell|^2 \qquad (25)$$

$$Z_1 \equiv \sum_{\ell=1}^{L} \Re[e^{j\Phi_X} e^{j\Theta_\ell} F_\ell N_\ell^*] \qquad (26)$$

$$Z_0 \equiv \sum_{\ell=1}^{L} |N_\ell|^2. \qquad (27)$$

The second-order statistics of Z_1 and Z_0 are

$$\begin{aligned}
\mathbb{E}[Z_1] &= 0 & \text{Var}[Z_1] &= \mathbb{E}[G]\sigma_N^2/2 \\
\mathbb{E}[Z_0] &= \sigma_N^2 & \text{Var}[Z_0] &= \sigma_N^4 \Delta \\
\mathbb{E}[Z_1(Z_0 - \mathbb{E}[Z_0])] &= 0.
\end{aligned} \qquad (28)$$

By using the Auxiliary-Channel Lower Bound Theorem in [6, Sec. VI], we have

$$I(X_A; V) \geq \mathbb{E}[-\log Q_V(V)] + \mathbb{E}[\log Q_{V|X_A}(V|X_A)] \qquad (29)$$

where $Q_{V|X_A}(v|x_A)$ is an arbitrary auxiliary channel and

$$Q_V(v) = \int P_{X_A}(x_A) Q_{V|X_A}(v|x_A) dx_A \qquad (30)$$

where $P_{X_A}(\cdot)$ is the *true* input distribution, i.e., $Q_V(\cdot)$ is the output distribution obtained by connecting the true input source to the auxiliary channel. $\mathbb{E}[\cdot]$ is the expectation according to the *true* distribution. We choose the auxiliary channel

$$Q_{V|X_A}(v|x_A) = \frac{1}{\sqrt{4\pi x_A^2 \Delta^2 \sigma_N^2}} \exp\left(-\frac{(v - x_A^2 \Delta - \sigma_N^2)^2}{4 x_A^2 \Delta^2 \sigma_N^2} \right). \qquad (31)$$

It follows that

$$\mathbb{E}[-\log(Q_{V|X_A}(V|X_A))] = \mathbb{E}\left[\frac{(V - X_A^2\Delta - \sigma_N^2)^2}{4 X_A^2 \Delta^2 \sigma_N^2} \right]$$
$$+ \log \Delta + \frac{1}{2}\log(4\pi\sigma_N^2) + \frac{1}{2}\mathbb{E}[\log(X_A^2)]. \qquad (32)$$

By using (24), we have

$$\begin{aligned}
&\left(V - X_A^2 \Delta - \sigma_N^2 \right)^2 \\
&= \left(X_A^2 \Delta(G-1) + 2X_A\Delta Z_1 + (Z_0 - \sigma_N^2) \right)^2 \\
&= X_A^4 \Delta^2 (G-1)^2 + 4 X_A^2 \Delta^2 Z_1^2 + (Z_0 - \sigma_N^2)^2 \\
&\quad + 4 X_A^3 \Delta^2 (G-1) Z_1 + 2 X_A^2 \Delta (G-1)(Z_0 - \sigma_N^2) \\
&\quad + 4 X_A \Delta Z_1 (Z_0 - \sigma_N^2)
\end{aligned} \qquad (33)$$

and hence, using the second-order statistics (28), we have

$$\mathbb{E}\left[\frac{(V - X_A^2\Delta - \sigma_N^2)^2}{4 X_A^2 \Delta^2 \sigma_N^2} \right]$$
$$= \frac{1}{4\sigma_N^2} P \, \mathbb{E}\left[(G-1)^2\right] + \frac{1}{2}\mathbb{E}[G] + \frac{\sigma_N^2}{4\Delta}\mathbb{E}\left[\frac{1}{X_A^2}\right] \qquad (34)$$

where we also used

$$\mathbb{E}[(G-1)Z_1] = 0. \qquad (35)$$

Substituting (34) into (32) and using $\mathbb{E}[G] \leq 1$ yield

$$\mathbb{E}[-\log(Q_{V|X_A}(V|X_A))]$$
$$\leq \log \Delta + \frac{1}{2}\log(4\pi\sigma_N^2) + \frac{1}{2}\mathbb{E}[\log(X_A^2)]$$
$$+ \frac{P}{4\sigma_N^2}\mathbb{E}\left[(G-1)^2\right] + \frac{1}{2} + \frac{\sigma_N^2}{4\Delta}\mathbb{E}\left[\frac{1}{X_A^2}\right]. \qquad (36)$$

It is convenient to define $X_P \equiv X_A^2$. We choose the input distribution

$$P_{X_P}(x_P) = \begin{cases} \frac{1}{\lambda}\exp\left(-\frac{x_P - P_{\min}}{\lambda}\right), & x_P \geq P_{\min} \\ 0, & \text{otherwise} \end{cases} \qquad (37)$$

where $0 < P_{\min} < P$ and $\lambda = P - P_{\min}$, so that

$$\mathbb{E}[X_P] = \mathbb{E}[X_A^2] = P. \qquad (38)$$

It follows from (30) and (37) that

$$Q_V(v) = \int_{P_{\min}}^{\infty} \frac{1}{\lambda}\exp\left(-\frac{x_P - P_{\min}}{\lambda}\right) Q_{V|X_P}(v|x_P) \, dx_P$$
$$\leq \exp(P_{\min}/\lambda) \, F_V(v) \qquad (39)$$

where

$$Q_{V|X_P}(v|x_P) = Q_{V|X_A}(v|\sqrt{x_P}) \qquad (40)$$

and

$$F_V(v) \equiv \int_0^{\infty} \frac{1}{\lambda}\exp\left(-\frac{x_P}{\lambda}\right) Q_{V|X_P}(v|x_P) dx_P. \qquad (41)$$

The inequality (39) follows from the non-negativity of the integrand. By combining (31), (40), (41) and making the change of variables $x = x_P \Delta$, we have

$$F_V(v)$$
$$= \int_0^{\infty} \frac{e^{-x/(\lambda\Delta)}}{\lambda\Delta} \frac{1}{\sqrt{4\pi x \Delta \sigma_N^2}} \exp\left(-\frac{(v - x - \sigma_N^2)^2}{4 x \Delta \sigma_N^2} \right) dx$$
$$= \frac{1}{\sqrt{\lambda\Delta(\lambda\Delta + 4\Delta\sigma_N^2)}} \times$$
$$\exp\left(\frac{2}{4\Delta\sigma_N^2}\left[v - \sigma_N^2 - |v - \sigma_N^2|\sqrt{1 + \frac{4\Delta\sigma_N^2}{\lambda\Delta}} \right] \right) \qquad (42)$$

where we used equation (140) in Appendix A of [7]:

$$\int_0^\infty \frac{1}{a}\exp\left(-\frac{x}{a}\right)\frac{1}{\sqrt{\pi bx}}\exp\left(-\frac{(u-x)^2}{bx}\right)dx$$
$$=\frac{1}{\sqrt{a(a+b)}}\exp\left(\frac{2}{b}\left[u-|u|\sqrt{1+\frac{b}{a}}\right]\right). \quad (43)$$

Therefore, we have

$$\mathbb{E}[-\log(F_V(V))]$$
$$=\frac{1}{2}\log(\Delta^2(\lambda^2+4\lambda\sigma_N^2))$$
$$-\frac{1}{2\Delta\sigma_N^2}\left[\mathbb{E}[V-\sigma_N^2]-\mathbb{E}[|V-\sigma_N^2|]\sqrt{1+\frac{4\sigma_N^2}{\lambda}}\right]$$
$$\overset{(a)}{\geq}\log(\Delta\lambda)+\frac{1}{2\sigma_N^2\Delta}\mathbb{E}[V-\sigma_N^2]\left[\sqrt{1+\frac{4\sigma_N^2}{\lambda}}-1\right]$$
$$\overset{(b)}{\geq}\log(\Delta\lambda) \quad (44)$$

where (a) holds because the logarithmic function is monotonic and $\mathbb{E}[|\cdot|]\geq\mathbb{E}[\cdot]$, and (b) holds because

$$\mathbb{E}[V-\sigma_N^2]$$
$$=\mathbb{E}[X_A^2]\Delta\mathbb{E}[G]+2\mathbb{E}[X_A]\Delta\mathbb{E}[Z_1]+\mathbb{E}[Z_0]-\sigma_N^2$$
$$=P\Delta\mathbb{E}[G]\geq 0. \quad (45)$$

The monotonicity of the logarithmic function and (39) yield

$$\mathbb{E}[-\log(Q_V(V))]\geq\mathbb{E}\left[-\log\left(e^{P_{\min}/\lambda}F_V(V)\right)\right]$$
$$\geq\log\Delta+\log\lambda-\frac{P_{\min}}{\lambda} \quad (46)$$

where the last inequality follows from (44). It follows from (29), (36) and (46) that

$$I(X_A;V)\geq\log\lambda-\frac{P_{\min}}{\lambda}-\frac{1}{2}\log(4\pi\sigma_N^2)-\frac{1}{2}\mathbb{E}[\log(X_A^2)]$$
$$-\frac{P}{4\sigma_N^2}\mathbb{E}\left[(G-1)^2\right]-\frac{1}{2}-\frac{\sigma_N^2}{4\Delta}\mathbb{E}\left[\frac{1}{X_A^2}\right]. \quad (47)$$

If $P_{\min}=P/2$, then $\lambda=P-P_{\min}=P/2$ and we have

$$\mathbb{E}\left[\frac{1}{X_P}\right]\leq\frac{1}{P_{\min}}=\frac{2}{P} \quad (48)$$

and

$$\mathbb{E}\left[\log\left(X_P\right)\right]=\int_\lambda^\infty\frac{1}{\lambda}e^{-(x-\lambda)/\lambda}\log(x)dx$$
$$\overset{(a)}{=}\log\lambda+\int_1^\infty e^{-(u-1)}\log(u)du$$
$$\overset{(b)}{\leq}\log\lambda+1 \quad (49)$$

where (a) follows by the change of variables $u=x/\lambda$, and (b) holds because $\log(u)\leq u-1$ for all $u>0$. Substituting into (47), we obtain

$$I(X_A;V)-\frac{1}{2}\log\mathsf{SNR}\geq -2-\frac{1}{2}\log(8\pi)-\frac{1}{2\mathsf{SNR}\Delta}$$
$$-\frac{1}{4}\mathsf{SNR}\,\mathbb{E}\left[(G-1)^2\right] \quad (50)$$

where $\mathsf{SNR}=P/\sigma_N^2$. Suppose L grows with SNR such that

$$L=\left\lceil\beta\sqrt{\mathsf{SNR}}\right\rceil. \quad (51)$$

Since $\Delta=1/L$, then we have

$$\lim_{\mathsf{SNR}\to\infty}\mathsf{SNR}\Delta=\infty\quad\text{and}\quad\lim_{\mathsf{SNR}\to\infty}\mathsf{SNR}\Delta^2=\frac{1}{\beta^2} \quad (52)$$

which implies

$$\lim_{\mathsf{SNR}\to\infty}I(X_A;V)-\frac{1}{2}\log\mathsf{SNR}\geq -2-\frac{1}{2}\log(8\pi)-\frac{\pi^2}{36} \quad (53)$$

because (see Appendix)

$$\lim_{\Delta\to 0}\frac{\mathbb{E}[(G-1)^2]}{\Delta^2}=\frac{(\pi\beta)^2}{9}. \quad (54)$$

By combining (20), (21), (23) and (53), we have

$$\lim_{\mathsf{SNR}\to\infty}\frac{1}{n}I(X_1^n;\mathbf{Y}_1^n)-\frac{1}{2}\log\mathsf{SNR}\geq -2-\frac{1}{2}\log(8\pi)-\frac{\pi^2}{36}. \quad (55)$$

This shows that the information rate grows logarithmically at high SNR with a pre-log factor of 1/2.

V. DISCUSSION

There is a wide literature on the design of receivers for the channel model (1) with a discrete-time Wiener phase noise, e.g., see [8], [9], [10] and references therein. One may want to make use of these designs, which raises the following question: "when is it justified to approximate the non-coherent fading model (3) with the discrete-time phase noise model (1)?" Our result suggests that this approximation may be justified when the phase variation is small over one symbol interval (i.e., when the phase noise linewidth is small compared to the symbol rate) *and* also the SNR is low to moderate. It must be noted that the SNR at which the high-SNR asymptotics start to manifest themselves depends on the application.

We remark that the authors of [11] treated on-off keying transmission in the presence of Wiener phase noise by using a double-filtering receiver, which is composed of an intermediate frequency (IF) filter, followed by an envelope detector (square-law device) and then a post-detection filter. They showed that by optimizing the IF receiver bandwidth the double-filtering receiver outperforms the single-filtering (matched filter) receiver. Furthermore, they showed via computer simulation that the optimum IF bandwidth increases with the SNR. This is similar to our result in the sense that we require the number of samples per symbol to increase with the SNR in order to achieve a rate that grows logarithmically with the SNR.

Finally, we remark that we have not computed the contribution of phase modulation to the information rate. We believe that using the multi-sample receiver it is possible to achieve an overall pre-log that is larger than 1/2. This matter is currently under investigation.

VI. Conclusion

We studied a communication system impaired by Wiener phase noise and AWGN. A discrete-time channel model based on filtering and oversampling is considered. The model accounts for the filtering effects on the phase noise. It is shown that at high SNR the multi-sample receiver achieves rates that grow logarithmically with at least a 1/2 pre-log factor if the number of samples per symbol grows with the square-root of the SNR.

Acknowledgment

Appendix

We discuss the limit in (54). We express $\mathbb{E}[(G-1)^2]$ as

$$
\begin{aligned}
\mathbb{E}[(G-1)^2] &= \mathsf{Var}(G) + (\mathbb{E}[G]-1)^2 \\
&= \frac{1}{L}\mathsf{Var}(|F_1|^2) + \left(\mathbb{E}[|F_1|^2]-1\right)^2 \quad (56)
\end{aligned}
$$

where the last equality follows from the definition of G in (25) and because $\{F_k\}$ is i.i.d.

Next, we outline the steps for computing $\mathbb{E}[|F_1|^4]$ and $\mathbb{E}[|F_1|^2]$. Let M be a positive integer, $\mathbf{c} = (c_1,\ldots,c_M)^T$ be a constant vector, $\mathbf{t} = (t_1,\ldots,t_M)^T$ be a non-negative real vector and $\mathbf{\Theta}(\mathbf{t}) = (\Theta(t_1)-\Theta(0),\ldots,\Theta(t_M)-\Theta(0))^T$ where $\Theta(t)$ is defined in (6). We have

$$
\begin{aligned}
&\mathbb{E}\left[\frac{1}{\Delta^M}\int\cdots\int_0^\Delta \exp(j\mathbf{c}^T\mathbf{\Theta}(\mathbf{t}))d\mathbf{t}\right] \\
&\overset{(a)}{=} \frac{1}{\Delta^M}\int\cdots\int_0^\Delta \mathbb{E}\left[\exp(j\mathbf{c}^T\mathbf{\Theta}(\mathbf{t}))\right]d\mathbf{t} \\
&\overset{(b)}{=} \frac{1}{\Delta^M}\int\cdots\int_0^\Delta \exp\left(-\frac{1}{2}\mathbf{c}^T\Sigma(\mathbf{t})\mathbf{c}\right)d\mathbf{t} \\
&\overset{(c)}{=} \int\cdots\int_0^1 \exp\left(-\frac{\Delta}{2}\mathbf{c}^T\Sigma(\mathbf{t})\mathbf{c}\right)d\mathbf{u} \quad (57)
\end{aligned}
$$

where $d\mathbf{t} = dt_M\ldots dt_1$ and $\Sigma(\mathbf{t})$ is the covariance matrix of $\mathbf{\Theta}(\mathbf{t})$ whose entries are given by

$$
\Sigma_{ij}(\mathbf{t}) = 2\pi\beta\min\{t_i,t_j\}, \text{ for } i,j=1,\ldots,M. \quad (58)
$$

Step (a) follows from the linearity of expectation, (b) follows by using the characteristic function of a Gaussian random vector, and (c) follows from the transformation of variables $\mathbf{t} = \mathbf{u}\,\Delta$. We define

$$
a = e^{-\pi\beta\Delta} \quad (59)
$$

and use $M = 2$ and $\mathbf{c} = (-1,1)^T$ in (57) to compute

$$
\mathbb{E}[|F_1|^2] = 2\frac{a-1-\log a}{(\log a)^2}. \quad (60)
$$

We also have, using $M = 4$ and $\mathbf{c} = (-1,1,-1,1)^T$ in (57),

$$
\mathbb{E}[|F_1|^4] \quad (61)
$$
$$
= \frac{783 - 784a + a^4 + 540\log a + 240a\log a + 144(\log a)^2}{18(\log a)^4}.
$$

Computing the integrals is tedious but straightforward. Finally, it follows from (56), and (59) – (61) that

$$
\lim_{\Delta\to 0}\frac{\mathbb{E}[(G-1)^2]}{\Delta^2} = (\pi\beta)^2 \lim_{a\to 1}\frac{\mathbb{E}[(G-1)^2]}{(\log a)^2} = \frac{(\pi\beta)^2}{9}. \quad (62)
$$

References

[1] M. Katz and S. Shamai. On the capacity-achieving distribution of the discrete-time noncoherent and partially coherent AWGN channels. *IEEE Trans. Inf. Theory*, 50(10):2257 – 2270, Oct. 2004.

[2] B. Goebel, R. 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*, 57(9):5720 –5736, Sep. 2011.

[3] A. Lapidoth. Capacity bounds via duality: A phase noise example. In *Proc. 2nd Asian-Euro. Workshop on Inf. Theory*, pages 58 – 61, 2002.

[4] G.J. Foschini and G. Vannucci. Characterizing filtered light waves corrupted by phase noise. *IEEE Trans. Inf. Theory*, 34(6):1437 –1448, Nov. 1988.

[5] A. Lapidoth and S.M. Moser. Capacity bounds via duality with applications to multiple-antenna systems on flat-fading channels. *IEEE Trans. Inf. Theory*, 49(10):2426 – 2467, Oct. 2003.

[6] D.M. Arnold, H.-A. Loeliger, P.O. Vontobel, A. Kavcic, and Wei Zeng. Simulation-based computation of information rates for channels with memory. *IEEE Trans. Inf. Theory*, 52(8):3498 –3508, Aug. 2006.

[7] S.M. Moser. Capacity results of an optical intensity channel with input-dependent gaussian noise. *IEEE Trans. Inf. Theory*, 58(1):207 –223, Jan. 2012.

[8] A. Barbieri, G. Colavolpe, and G. Caire. Joint iterative detection and decoding in the presence of phase noise and frequency offset. *IEEE Trans. Commun.*, 55(1):171 –179, Jan. 2007.

[9] A. Spalvieri and L. Barletta. Pilot-aided carrier recovery in the presence of phase noise. *IEEE Trans. Commun.*, 59(7):1966 –1974, July 2011.

[10] A. Barbieri and G. Colavolpe. On the information rate and repeat-accumulate code design for phase noise channels. *IEEE Trans. Commun.*, 59(12):3223 –3228, Dec. 2011.

[11] G.J. Foschini, L.J. Greenstein, and G. Vannucci. Noncoherent detection of coherent lightwave signals corrupted by phase noise. *IEEE Trans. Commun.*, 36(3):306 –314, Mar. 1988.