

Lower Bound Based on Kalman Carrier Recovery below the Information Rate of Wiener Phase Noise Channel

Luca Barletta, Maurizio Magarini, Arnaldo Spalvieri
Dipartimento di Elettronica e informazione
Politecnico di Milano
Milan, Italy
Email: {barletta, magarini, spalvier}@elet.polimi.it

Abstract—A new lower bound below the information rate transferred through the Additive White Gaussian Noise (AWGN) channel affected by discrete-time multiplicative Wiener's phase noise is proposed in the paper. The proposed lower bound is based on the Kalman approach to data-aided carrier phase recovery, and is less computationally demanding than known methods based on phase quantization and trellis representation of phase memory. Simulation results show that the lower bound is close to the actual channel capacity, especially at low-to-intermediate signal-to-noise ratio.

Keywords— Coherent communication, Phase noise, Wiener phase noise, Channel capacity, Information rate, Kalman filtering, Kalman carrier phase recovery.

I. INTRODUCTION

Multiplicative phase noise is a major source of impairment in coherent communication. In the context of radio transmission, phase noise is introduced by local oscillators used for up-conversion and down-conversion. The impact of phase noise on the performance of OFDM systems is studied, for instance, in [1]. Also, single carrier systems, especially recent systems based on frequency domain equalization as [2], suffer from phase noise and require specific mitigation techniques [3]. With the advent of coherent optics, the role of phase noise is becoming well recognized also in the context of optical transmission, see e.g. [4].

The basic model of phase noise is that of Wiener phase noise, [4], [5], where the power spectral density of the spectral line is a slope of -20 dB/decade. Several methods have been proposed in the literature to combat the detrimental effects of Wiener phase noise. Among these methods we cite the iterative demodulation and decoding techniques of [6]–[8], the insertion of pilot symbols [9]–[11], and soft differential demodulation [12].

Computation of the capacity of the multiplicative Wiener phase noise plus additive white Gaussian noise (AWGN) channel, which is a channel with memory and continuous state, is a challenging problem. A Monte Carlo approach based on phase space quantization and trellis representation of phase memory has been recently proposed in [13]–[15] for computing the constrained channel capacity (i.e. the capacity with a fixed source).

The new result presented in this paper and in the journal paper [16] is a lower bound below the constrained capacity of the AWGN Wiener phase noise channel. The bounding technique proposed here relies upon Monte Carlo simulation of a demodulator aided by the past data, where the transfer function of the causal filter used for phase estimation is worked out by the Kalman approach. Compared to methods based on phase quantization and trellis representation of phase memory, such as [14], the method proposed here is less computationally demanding, since a (Kalman) filter is used in place of a trellis. At low-to-intermediate signal-to-noise ratio, it can be seen that the proposed lower bound is so close to the actual channel capacity that we can claim that the new lower bound gives virtually the actual channel capacity.

The outline of the paper is as follows. In Section II the channel model and the source model are introduced. Section III reports the general method behind the bound, while in Section IV the specific bound is presented. Section V gives simulation results, while in Section VI conclusions are drawn.

II. CHANNEL AND SOURCE MODEL

Let u_i^k indicate the vector $(u_i, u_{i+1}, \dots, u_k)$ with $u_i^k \in \mathcal{U}_i^k$, and let U indicate a stationary and ergodic process, $U = (U_0, U_1, \dots)$, whose generic realization is the sequence (u_0, u_1, \dots) . When \mathcal{U}_i^k is a continuous set, $p(u_i^k)$ is used to indicate the multivariate probability density function, while when \mathcal{U}_i^k is a discrete set $p(u_i^k)$ indicates the multivariate mass probability and $|\mathcal{U}_i^k|$ denotes the number of elements in \mathcal{U}_i^k .

The k -th output of the channel is

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

where j is the imaginary unit, Y is the complex channel output process, X is the channel complex input modulation process, and Φ is the phase noise process which is assumed to be independent of X and W . For concreteness, we assume that the input process X is made of i.i.d. random variables with zero mean and unit variance. Process W is a complex AWGN process with zero mean and variance SNR^{-1} . Process Φ is modelled as a Wiener process:

$$\phi_k = \phi_{k-1} + \gamma v_i, \quad k = 1, 2, \dots \quad (2)$$

where the frequency noise V is a i.i.d. sequence of Gaussian random variables with zero mean and unit variance, γ is a scalar, and ϕ_0 is uniformly distributed in $[-\pi, \pi)$. The phase evolution given in (2) occurs when the power spectral density of the continuous-time complex exponential $e^{j\phi(t)}$, whose samples taken at symbol frequency generate the sequence $e^{j\phi_k}$, is the Lorentzian function

$$\mathcal{L}(f) = \frac{4\gamma^2 T}{\gamma^4 + 16\pi^2 f^2 T^2},$$

where T is the symbol repetition interval and f is the frequency. The parameter γ^2 can be expressed as

$$\gamma^2 = 2\pi B_{FWHM} T,$$

where B_{FWHM} is the full-width half-maximum bandwidth of the spectral line.

Equations (1) and (2) can be casted in the general framework of state-space approach to modelling dynamic systems. Specifically, equation (2) is that of a first-order model where ϕ_k is the state at time k . In this context, the AWGN W is called *measurement noise*, while the frequency noise V is called *process noise*. The interested reader is referred to [17] for a comprehensive book on the subject.

III. THE AUXILIARY PROBABILITY METHOD

The bound that we are going to present is based on the Kullback-Leibler (KL) divergence. The normalized KL divergence between the multivariate mass probability functions $p(u_0^n)$ and $q(u_0^n)$ is

$$\lim_{n \rightarrow \infty} \frac{1}{n} E_p \left\{ \log_2 \frac{p(u_1^n)}{q(u_1^n)} \right\} \geq 0, \quad (3)$$

where $E_p\{\cdot\}$ denotes the expectation over $p(u_0^n)$ and $(n)^{-1}$ is the normalization factor. From the normalized KL divergence one has the following upper bound on the entropy rate of process U :

$$\bar{H}(U) = \lim_{n \rightarrow \infty} \frac{1}{n} E_p \left\{ \log_2 \left(\frac{1}{q(u_1^n)} \right) \right\} \geq H(U). \quad (4)$$

Let us regard the *auxiliary* multivariate mass probability $q(u_1^n)$ as an approximation to $p(u_1^n)$. In this perspective, the KL divergence is a measure of the quality of the fit between $p(u_1^n)$ and $q(u_1^n)$, and the upper bound is equal to the actual entropy rate when the fit is ideal, that is when $q(u_1^n) = p(u_1^n)$.

Assuming that U is ergodic, one can invoke the Shannon-McMillan-Breiman theorem and the chain rule, thus writing for the expectation appearing in (4)

$$\bar{H}(U) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \log_2 \left(\frac{1}{q(u_k | u_0^{k-1})} \right), \quad (5)$$

where u_1^n is generated according to the actual multivariate mass probability $p(u_1^n)$, and the initial condition u_0 is given. The bound can be extended to the conditional entropy rate in a straightforward manner.

IV. LOWER BOUND

Assume discrete input alphabet. The lower bound below the information rate is

$$H(X) - \bar{H}(X|Y) \leq I(X; Y),$$

where the familiar notation is used for the conditional entropy rate and for the mutual information rate. The upper bound

$$\bar{H}(X|Y) \geq H(X|Y)$$

is obtained from (5) as

$$\bar{H}(X|Y) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \log_2 \left(\frac{1}{q(x_k | x_1^{k-1}, y_1^n)} \right). \quad (6)$$

To obtain a tight bound, one has to work out an auxiliary mass probability that closely approximates the actual mass probability. Aiming to simple yet effective methods, we observe that the portion of joint sequence (x_1^{k-1}, y_1^{k-1}) can strongly contribute to a data-aided approximation to the wanted probability, while the portion y_k^n gives a weaker contribution of non-data-aided type. The stronger part of the non-data-aided contribution comes from sample y_k , therefore the auxiliary probability that is hereafter considered is based only on (x_1^{k-1}, y_1^k) , while y_{k+1}^n is ignored. The auxiliary conditional probability proposed here is

$$\begin{aligned} q(x_k | y_1^k, x_1^{k-1}) &= \int_{-\pi}^{\pi} q(\phi_k, x_k | y_1^k, x_1^{k-1}) d\phi_k \\ &= \int_{-\pi}^{\pi} p(x_k | y_k, \phi_k) q(\phi_k | y_1^k, x_1^{k-1}) d\phi_k \\ &\propto \int_{-\pi}^{\pi} q(\phi_k | y_1^k, x_1^{k-1}) p(y_k | x_k, \phi_k) p(x_k) d\phi_k, \end{aligned} \quad (7)$$

where we have exploited the fact that X_k is conditionally independent of (Y_1^{k-1}, X_1^{k-1}) given (Y_k, Φ_k) . The only distribution in (7) that cannot be computed directly from the channel model is

$$\begin{aligned} q(\phi_k | y_1^k, x_1^{k-1}) &\propto q(\phi_k | y_1^{k-1}, x_1^{k-1}) p(y_k | \phi_k) \\ &= q(\phi_k | y_1^{k-1}, x_1^{k-1}) \cdot \sum_{x_k \in \mathcal{X}} p(y_k | \phi_k, x_k) p(x_k), \end{aligned}$$

where the auxiliary probability $q(\phi_k | y_1^{k-1}, x_1^{k-1})$ adopted here is

$$q(\phi_k | y_1^{k-1}, x_1^{k-1}) = g(\hat{\phi}_k, \sigma_k^2; \phi_k),$$

where $g(\eta, \sigma^2; u)$ is a Gaussian distribution with mean η and variance σ^2 over the space spanned by u . The estimate $\hat{\phi}_k$ appearing in the above equation is worked out by a predictive Kalman filter, that is

$$\hat{\phi}_k = E\{\phi_k | y_1^{k-1}, x_1^{k-1}\},$$

and the variance of the estimate is

$$\sigma_k^2 = E\{(\hat{\phi}_k - \phi_k)^2 | y_1^{k-1}, x_1^{k-1}\}.$$

The mean and the variance can be computed in a recursive manner thanks to the update equations of the Kalman filter.

Since the channel model (1) is a nonlinear function of the state, the Kalman filter has to be linearized around the current state estimate [17, Ch. 13.1]. The error that drives the Kalman filter is the one produced by the phase detector of classical data-aided carrier recovery, that is

$$e_k = \Im\{y_k x_k^* e^{-j\hat{\phi}_k}\}, \quad (8)$$

where $\Im\{\cdot\}$ and the superscript $*$ denote the imaginary part and the complex conjugation, respectively. Assuming that the phase error $\phi_k - \hat{\phi}_k$ is small, error (8) can be linearized as

$$e_k \approx \phi_k - \hat{\phi}_k + z_k,$$

where z_k is assumed to be white Gaussian noise with zero mean and time-varying variance $\sigma_{z,k}^2 = (2|x_k|^2 \text{SNR})^{-1}$. The estimate $\hat{\phi}_k$ and error's variance σ_k^2 are computed for $k \geq 0$ according to the iterative equations

$$\hat{\phi}_{k+1} = \hat{\phi}_k + K_k \cdot e_k, \quad (9)$$

$$\sigma_{k+1}^2 = \frac{\sigma_{z,k}^2 \cdot \sigma_k^2}{\sigma_{z,k}^2 + \sigma_k^2} + \gamma^2,$$

where K_k is the Kalman gain at time k :

$$K_k = \frac{\sigma_k^2}{\sigma_k^2 + \sigma_{z,k}^2}.$$

Initial values can be set as $\hat{\phi}_0 = 0$ and $\sigma_0^2 = \infty$. Equations (8) and (9) can be regarded as those of a first-order phase-locked loop (PLL) with time-varying gain, where (8) is the phase detector and (9) describes the loop filter. The analogy between PLL with time-varying gain and Kalman filters has been suggested in [18].

V. SIMULATION RESULTS

In this section, the new lower bound is compared to the actual channel capacity worked out by the computationally demanding trellis-based method of [14]. Specifically, the actual channel capacity has been obtained using a large number of states in the lower bound and in the upper bound of [14]. The number of states is so large that the upper bound and the lower bound become virtually undistinguishable, leading to the actual channel capacity. Fig. 1 reports the results obtained with 4-QAM and with two values of γ . Specifically, $\gamma = 0.125$ is the largest value obtained in the experimental results of [4] and can be regarded as a case of strong phase noise in cases of practical interest. Although less realistic, also the huge $\gamma = 0.5$ is studied, to show the limits of the proposed method. For $\gamma = 0.125$ the lower bound is virtually undistinguishable from the actual capacity in a wide range of information rate, say, below 1.5 bit/2D. This range is the one spanned by codes with rate lower than 0.75, that are codes of large practical interest. For information rate greater than 1.5 bit/2D the bound loses accuracy. This is because we have not exploited the conditioning on y_{k+1}^n , which, at high SNR, could potentially bring a non-negligible contribution to the accuracy of the fit between the auxiliary probability and the actual probability.

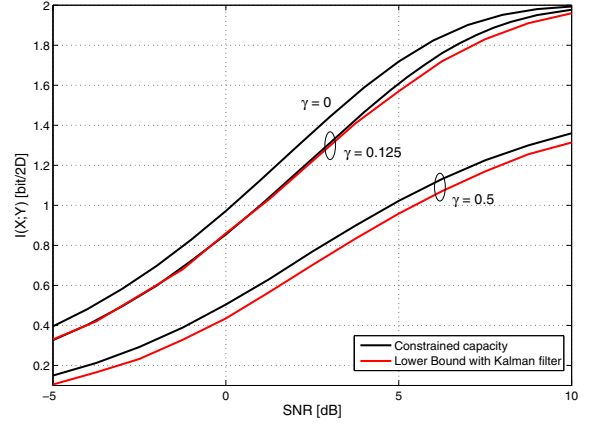


Fig. 1. Actual channel capacity and lower bound for 4-QAM and two values of γ . Capacity of the pure AWGN channel is also reported.

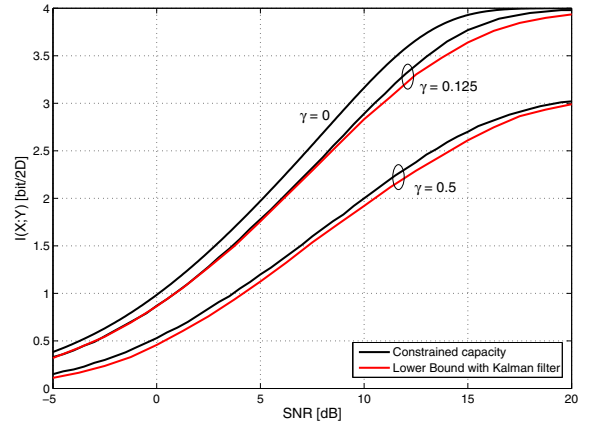


Fig. 2. Actual channel capacity and lower bound for 16-QAM and two values of γ . Capacity of the pure AWGN channel is also reported.

Although being fairly close to the actual capacity, the lower bound is less accurate for $\gamma = 0.5$, because, with a so large value of γ , frequent cycle slips affect the performance of Kalman carrier recovery. A similar analysis holds for the results obtained with 16-QAM and reported in Fig. 2. Also in this case, the lower bound virtually gives the actual channel capacity for $\gamma = 0.125$ and coding rate below 0.75.

VI. CONCLUSION

In the paper, a new lower bound below the information rate transferred through the Wiener phase noise channel has been presented. The results, compared to the actual channel capacity obtained with the computationally demanding method of [14], show that the bound is accurate in many cases of practical interest. Before concluding the paper, a remark is in order about phase noise of order higher than one, which is out of the scope of the present paper and that will be subject of future research. Extension of the Kalman-based approach to

phase noise with memory of order higher than one, as the second-order case studied in [19], is feasible by extending the state space of the Kalman filter. In contrast, quantizing a multidimensional state space according to the approach of [13], [14] would lead to an exponential increase of the number of states of the trellis, making computation unfeasible.

REFERENCES

- [1] A. G. Armada, "Understanding the Effects of Phase Noise in Orthogonal Frequency Division Multiplexing (OFDM)," *IEEE Trans. Broadcast.* vol. 47, no. 2, pp. 153–159, Jun. 2001.
- [2] M. Magarini, L. Barletta, and A. Spalvieri, "Efficient computation of the feedback filter for the hybrid decision feedback equalizer in highly dispersive channels," *IEEE Trans. Wireless Commun.*, vol. 11, no. 6, pp. 2245–2253, June 2012.
- [3] M. Asim, M. Ghogho, and D. McLemon, "Mitigation of phase noise in single carrier frequency domain equalization systems," in *Proc. of Wireless Communications and Networking Conference (WCNC)*, pp. 920–924, April 2012.
- [4] M. Magarini, A. Spalvieri, F. Vacondio, M. Bertolini, M. Pepe, and G. Gavioli, "Empirical modeling and simulation of phase noise in long-haul coherent optical systems," *Optics Express*, vol. 19, issue 23, pp. 22455–22461, Nov. 7, 2011.
- [5] G. J. Foschini and G. Vannucci, "Characterizing filtered light waves corrupted by phase noise," *IEEE Trans. Inform. Theory*, vol. 34, no. 6, pp. 1437–1448, Nov. 1988.
- [6] M. Peleg, S. Shamai (Shitz), and S. Galan, "Iterative decoding for coded noncoherent MPSK communications over phase-noisy AWGN channel," *Proc. IEE Commun.*, vol. 147, pp. 87–95, Apr. 2000.
- [7] G. Colavolpe, A. Barbieri, and G. Caire, "Algorithms for iterative decoding in the presence of strong phase noise," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 9, pp. 1748–1757, Sept. 2005.
- [8] A. Barbieri and G. Colavolpe, "Soft-output decoding of rotationally invariant codes over channels with phase noise," *IEEE Trans. Commun.*, vol. 55, no. 11, pp. 2125–2133, Nov. 2007.
- [9] A. Spalvieri and L. Barletta, "Pilot-aided carrier recovery in the presence of phase noise," *IEEE Trans. Commun.*, vol. 59, no. 7, pp. 1966–1974, July 2011.
- [10] M. Magarini, L. Barletta, A. Spalvieri, F. Vacondio, T. Pfau, M. Pepe, M. Bertolini, and G. Gavioli, "Pilot-symbols-aided carrier-phase recovery for 100-G PM-QPSK digital coherent receivers," *IEEE Photonics Technol. Letters*, vol. 24, issue 9, pp. 739–741, May 01, 2012.
- [11] L. Barletta, M. Magarini, and A. Spalvieri, "Staged demodulation and decoding," *Optics Express*, vol. 20, issue 21, pp. 23728–23734, Oct. 8, 2012.
- [12] M. Magarini, L. Barletta, A. Spalvieri, A. Leven, M. Pepe, and G. Gavioli, "Impact of non-ideal phase reference on soft decoding of differentially encoded modulation," *IEEE Photonics Technol. Letters*, preprint available for download at IEEEExplore website.
- [13] L. Barletta, M. Magarini, and A. Spalvieri, "Estimate of information rates of discrete-time first-order Markov phase noise channel," *IEEE Photonics Technology Letters*, vol. 23, no. 21, pp. 1582–1584, Nov. 1, 2011.
- [14] L. Barletta, M. Magarini, and A. Spalvieri, "The information rate transferred through the discrete-time Wiener's phase noise channel," *IEEE J. Lightw. Technol.*, vol. 30, no. 10, pp. 1480–1486, May 15, 2012.
- [15] A. Barbieri and G. Colavolpe, "On the information rate and repeat-accumulate code design for phase noise channel," *IEEE Trans. on Commun.*, vol. 59, no. 12, pp. 3223–3228, Dec. 2011.
- [16] L. Barletta, M. Magarini, and A. Spalvieri, "A New Lower Bound below the Information Rate of Wiener Phase Noise Channel Based on Kalman Carrier Recovery," *Optics Express*, accepted for publication.
- [17] D. Simon, *Optimal State Estimation*. New York: Wiley, 2006.
- [18] A. Pataputian, "On phase-locked loops and Kalman filters," *IEEE Trans. Commun.*, vol. 47, no. 5, pp. 670–672, May 1999.
- [19] A. Spalvieri and M. Magarini, "Wiener's analysis of the discrete-time phase-locked loop with loop delay," *IEEE Trans. Circuits and Systems II*, vol. 55, pp. 596–600, June 2008.