Nonlinear fiber capacity

(article starts on next page)

Nonlinear Fiber Capacity

Erik Agrell

Department of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden. ✉ agrell@chalmers.se

Abstract *In this semi-tutorial presentation, we review fundamental information theory for links with and without memory, in the linear and nonlinear regimes. A comparison between channel models with long (but finite) memory and infinite memory yields an unexpected result.*

Introduction

Shannon, the father of information theory, proved that for a given channel, it is possible to achieve an arbitrarily small error probability, if the transmission rate in bits per symbol is small enough. A rate for which virtually error-free transmission is possible is called an *achievable rate* and the supremum over all achievable rates for a given channel, represented as a statistical relation between its input X and output Y, is defined as the *channel capacity* [1], [2, p. 195]. A capacity-approaching transmission scheme operates in general by grouping the data to be transmitted into blocks, encoding each block into a sequence of information symbols, transmitting this sequence over the channel, and decoding it in the receiver. A long block length is essential to obtain an arbitrarily small error probability, even for memoryless channels, and for channels with memory, the block length should typically be much larger than the channel memory.

The GN model and its capacity

For coherent long-haul fiber-optical links without dispersion compensation, the most popular discrete-time channel models represents the nonlinear interference (NLI) as Gaussian noise (GN), whose statistics depend on the transmitted signal power [3–6]. Thus, the complex single-channel output Y_n is modeled as

$$Y_n = X_n + Z_n, \tag{1}$$

where X_n is the complex channel input and Z_n is a complex, white, Gaussian random sequence, independent of X_n. Although Z_n is independent of the actual transmitted process X_n, the *distribution* of Z_n depends on the distribution of X_n. Splett *et al.* [3], Poggiolini *et al.* [5], and Beygi *et al.* [6] have all derived models where Z_n is a zero-mean, circularly symmetric sequence, whose power $P_Z = \mathbb{E}[|Z_n|^2]$ depends on the transmit power $P_X = \mathbb{E}[|X_n|^2]$ as

$$P_Z = c_0 + c_3 P_X^3, \tag{2}$$

where c_0 and c_3 quantify the amplified spontaneous emission noise of the optical amplifiers and the NLI, resp. The cubic relation (2) holds for both lumped and distributed amplification schemes, and it extends to multiple-wavelength [3, 5] and dual-polarization [5, 6] systems. For uncoded transmission with traditional modulation formats, the GN model has been shown to be very accurate in experiments and simulations.

Since Z_n in (1) is additive and statistically independent of X_n, the channel capacity of the GN model (1)–(2) can be calculated exactly as [3, 7]

$$C = \log_2\left(1 + \frac{P_X}{P_Z}\right) \tag{3}$$

using Shannon's well-known capacity expression [1, Sec. 24], [2, Ch. 9]. Considered as a function of the transmitted signal power P_X, the capacity has the peculiar behavior of reaching a peak and eventually decreasing to zero at high enough power, since the denominator of (3) increases faster than the numerator. The phenomenon, sometimes called the "nonlinear Shannon limit" in the optical communications community, seems to convey the message that reliable communication over nonlinear optical channels becomes impossible at high powers. In the last part of this paper, we question this pessimistic conclusion.

Capacity and mutual information

By Shannon's *channel coding theorem,* the channel capacity of a discrete-time memoryless channel, in bit/symbol, can be calculated as [1], [2, Ch. 7]

$$C = \sup_{f_X} I(X;Y),$$

where $I(X;Y)$ is the *mutual information*

$$I(X;Y) = \iint f_{X,Y}(x,y)\log_2\frac{f_{X,Y}(x,y)}{f_X(x)f_Y(y)}dxdy, \tag{4}$$

$f_{X,Y}$ is the joint distribution of X and Y, and f_X and f_Y are its respective marginal distributions. The capacity is often studied as a function of the transmitted signal power, $P_X = \mathbb{E}[|X|^2]$. The corresponding expression for channels with memory is, under certain assumptions on information stability [8],

$$C = \lim_{n\to\infty} \sup_{f_{X_1^n}} \frac{1}{n} I(\boldsymbol{X}_1^n; \boldsymbol{Y}_1^n), \tag{5}$$

where $\boldsymbol{X}_i^j = (X_i, X_{i+1}, \dots, X_j)$ and $I(\boldsymbol{X}_i^j; \boldsymbol{Y}_i^j)$ is defined as a multidimensional integral analogous with (4).

In this work, we are interested in *channels with finite memory.* Such channels have the property that the channel output Y_n depends on a finite number N of past input symbols, but not on the entire history, i.e.,

$$f_{Y_n|\boldsymbol{X}_1^n}(y_n \mid \boldsymbol{x}_1^n) = f_{Y_n|\boldsymbol{X}_{n-N}^n}(y_n \mid \boldsymbol{x}_{n-N}^n), \tag{6}$$

where N is the *channel memory.* By expanding the right-hand side of (5) using the chain rule for mutual information [2, p. 24], it can be shown that for channels that satisfy (6) and any stochastic process X_n,

$$C \geq I(X_n; Y_n \mid \boldsymbol{X}_{n-N}^{n-1}), \tag{7}$$

where $I(X_n; Y_n \mid \boldsymbol{X}_{n-N}^{n-1})$ is the *conditional mutual information* [2, p. 23].

A finite-memory GN model

Even highly dispersive optical fibers have a finite memory. The output does not depend what was transmitted into the fiber yesterday, or even a second ago. To account for this important feature, we propose a *finite-memory GN model.* The input–output relation is still given by (1), but the *statistical* transmit power P_X in (2) is replaced with an *empirical* transmit power, which is a function of the actual transmitted sequence X_n during a finite window of N past input samples. Assuming a constant weight for all samples in this window, (2) is replaced by

$$P_Z = c_0 + c_3 P_n^3, \tag{8}$$

$$P_n = \frac{1}{N} \sum_{i=n-N}^{n-1} |X_i|^2. \tag{9}$$

In the limit $N \to \infty$, the empirical power in (9) converges to the statistical power $P_X = \mathbb{E}[X_n]$, for any stationary, ergodic input sequence X_n. Hence, for large enough N, assuming uncoded transmission or coded schemes with short or moderate block lengths, the finite-memory GN model is equivalent to the regular GN model in the previous section, which can be regarded as an infinite-memory model.

For a suitable (large but finite) choice of N, the finite-memory model is more physically relevant than the traditional GN model, but much harder to analyze, since every output depends on a vector of inputs. A reasonable choice of N is in the order of $N_{\mathrm{sys}} = 2\pi|\beta_2|LR_{\mathrm{s}}^2$, where β_2 is the group velocity dispersion parameter, L is the fiber length, and R_{s} is the symbol rate [9]. The finite-memory model is not accurate for small values of N, since the GN assumption relies on the central limit theorem [5, 6].

Channel capacity results

An exact expression for the channel capacity of the finite-memory GN model (1), (8)–(9) is unfortunately not available. Shannon's formula (3) does not apply, because the sequences X_n and Z_n are dependent. Furthermore, capacity estimation via (5) is numerically infeasible, since it would involve maximization over a high-dimensional space. We therefore resort to the lower bound (7), and furthermore constrain the input distribution to a ring constellation [4], with two discrete amplitudes r_1, r_2 and uniform phase. In this case, the right-hand side of (7) is

$$I(X_n; Y_n \mid \boldsymbol{X}_{n-N}^{n-1}) = \sum_{s\in\mathcal{S}} \Pr\{P_Z = s\} \frac{2}{s} \int_0^\infty \rho b(g(\rho, s))d\rho - \log_2 e, \tag{10}$$

where $b(u) = -u \log_2 u$, $\mathcal{S}$ is the set of $N+1$ possible values that P_Z in (8)–(9) can take on when $|X_i| \in \{r_1, r_2\}$,

$$g(\rho, s) = \sum_{r\in\{r_1,r_2\}} \Pr\{|X_n| = r\} \cdot \exp\left(-\frac{\rho^2 + r^2}{s}\right) I_0\left(\frac{2\rho r}{s}\right),$$

and $I_k(u)$ is the modified Bessel function of the first kind. The proof of (10) is omitted due to space constraints.

The radii and probabilities of the two rings are varied to maximize $I(X_n; Y_n \mid \boldsymbol{X}_{n-N}^{n-1})$ in (10), for given channel parameters c_0, c_3, and N and a given transmit power P_X. In our numerical example, we use $c_0 = 3.27\ \mu$W and $c_3 = 1.83$ mW^{-2}. The same values were used by Bosco *et al.* [7, 10], representing a standard single-mode fiber with length $L = 8000$ km, parameters $\alpha = 0.22$ dB/km, $\beta_2 = -21.7$ ps^2/km, and $\gamma = 1.27$ (W km)$^{-1}$, ideal distributed amplification with $K_T = 1$, WDM transmission with center frequency $\nu = 190$ THz and bandwidth $B_{\mathrm{WDM}} = 4$ THz, and symbol rate $R_{\mathrm{s}} = 32$ Gbaud.

The lower bound (10) is shown in Fig. 1 as a function of transmit power and channel memory N. We can see that as N increases, the curves approach

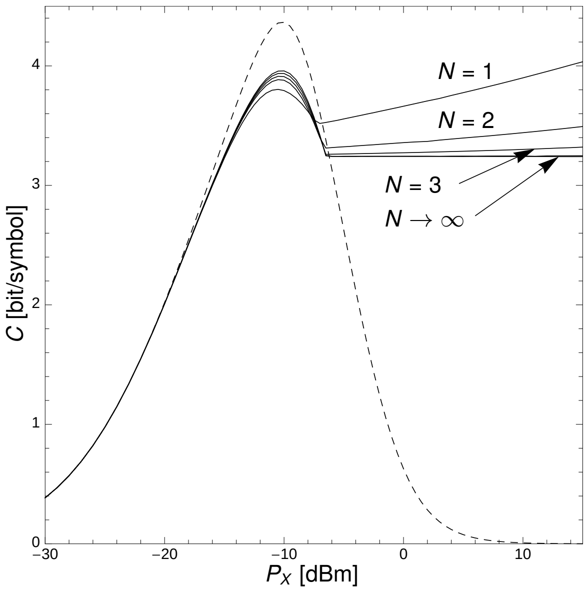

Fig. 1: Lower bounds (10) on the channel capacity C vs. transmit power P_X of the finite-memory GN model for varying channel memory N (solid) and the exact capacity (3) of the regular (infinite-memory) GN model (dashed).

an asymptotic bound, marked $N \to \infty$. This asymptotic bound is plotted for $N = 5$ and 10 (the two curves overlap), and it would look the same for, say, $N = N_{\text{sys}} \approx 1100$ or even higher. It has a peak at 3.96 bit/symbol, after which it decreases to 3.24 bit/symbol.

The rudimentary lower bounds in Fig. 1 represent two-ring constellations only and can be improved by using other input distributions. For example, using so-called satellite constellations, a higher lower bound can be obtained that flattens out without decreasing [11, 12].

The exact channel capacity (3) of the infinite-memory GN model (1)–(2) is included in Fig. 1 for reference. This is the 8000-km curve shown in [10, Fig. 1(a)], rescaled by a factor of two to represent a single polarization. It is striking, and at first glance counterintuitive, that the channel capacity of the infinite-memory model is so different from the asymptotic capacity of the finite-memory model in the nonlinear regime. A mathematical explanation is that in general, $\lim_{N\to\infty} \sup_f C_N(f) \neq \sup_f \lim_{N\to\infty} C_N(f)$. For a more intuitive explanation, we recall that a capacity-approaching transmission scheme should involve coding over a long block length, typically much longer than the channel memory N. Designing such long codes is possible, at least theoretically, for channel models with any finite memory, but not for infinite-memory models. Therefore, the two types of models have different channel capacities, not only for the example studied here but also for other nonlinear fiber channels, including dual polarization, wavelength multiplexing, multimode fibers, etc.

Conclusions

We extended the popular GN model for nonlinear fiber channels with a parameter to account for the channel memory. The new channel model is given by (1) and (8)–(9). For any finite memory, its channel capacity is quite different from that of the regular (infinite-memory) GN model in the nonlinear regime. Hence, infinite-memory channel models, although accurate for uncoded transmission, should not be used in capacity analysis. Their capacities do not bound the achievable rates of the underlying physical links.

We intentionally avoid using the concept "nonlinear Shannon limit." It is an artifact of the use of infinite-memory channel models and has no known analogy for finite-memory models, which are more physically meaningful. The real fiber link does not suffer from the same vanishing capacity at high transmit powers as the regular GN model and other infinite-memory models do. The true capacity for nonlinear, dispersive channels remains an open problem.

Acknowledgments

The author wishes to thank Magnus Karlsson and other members of the FORCE center at Chalmers for many helpful discussions. Funding from SSF and VR is gratefully acknowledged.

References

1 C. E. Shannon, *Bell Syst. Tech. J.*, **27**, 379, 623 (1948).
2 T. M. Cover *et al.*, *Elements of Information Theory*, 2nd ed., Wiley (2006).
3 A. Splett *et al.*, *Proc. ECOC*, 41 (1993).
4 R.-J. Essiambre *et al.*, *J. Lightw. Technol.*, **28**, 662 (2010).
5 P. Poggiolini *et al.*, *IEEE Photon. Technol. Lett.*, **23**, 742 (2011).
6 L. Beygi *et al.*, *IEEE Trans. Commun.*, **60**, 3440 (2012).
7 G. Bosco *et al.*, *Optics Express*, **19**, B440 (2011).
8 S. Verdú *et al.*, *IEEE Trans. Inf. Theory*, **40**, 1147 (1994).
9 E. Ip *et al.*, *J. Lightw. Technol.*, **25**, 2033 (2007).
10 G. Bosco *et al.*, *Optics Express*, **20**, 19 610 (2012).
11 E. Agrell, http://arxiv.org/abs/1209.2820, (2012).
12 E. Agrell *et al.*, *Proc. IPC*, 316 (2012).