

Nonlinear communication channels with capacity above the linear Shannon limit

Konstantin S. Turitsyn^{1,*} and Sergei K. Turitsyn²

¹*Department of Mechanical Engineering, MIT, Cambridge, Massachusetts 02139, USA*

²*Aston Institute of Photonic Technologies, Aston University, Aston Triangle, Birmingham B4 7ET, UK*

*Corresponding author: turitsyn@mit.edu

Received May 28, 2012; accepted June 29, 2012;

posted July 9, 2012 (Doc. ID 169459); published August 24, 2012

We prove that, under certain conditions, the capacity of an optical communication channel with in-line, nonlinear filtering (regeneration) elements can be higher than the Shannon capacity for the corresponding linear Gaussian white noise channel. © 2012 Optical Society of America

OCIS codes: 060.2330, 060.4370, 190.4370, 060.1660, 000.6800.

The modern information theory was developed mostly for channels in which conditional probability of signal input-to-output transformation can be described in a simple treatable manner, such as, e.g., in the classical linear additive noise channel [1]. The information capacity of a linear channel with additive white Gaussian noise (AWGN) is often called the Shannon limit, stressing the fact that it is the maximum error-free data rate achievable in such channel [1] and in any channels sub optimal to AWGN. The optical fiber channel characterized by additive amplifier noise demonstrates nonlinear properties for realistic signal powers as a result of the intensity-dependent refractive index (Kerr effect). Recent studies have shown that spectral efficiency of optical systems is limited by fiber nonlinearity [2–8]. As a result, the capacity of fiber channels as a function of signal-to-noise-ratio (SNR) is commonly considered as upper-bounded by the Shannon capacity for linear AWGN and limited by nonlinear effects at high SNR values.

There is a widely spread misconception that nonlinearity can only limit the capacity of communication channels. As a matter of fact, nonlinearity in transmission may be both destructive, leading to nonlinear impairments, and constructive, e.g., assisting nonlinear control of signal transmission and providing an effective suppression of noise. It is important to stress that there is no single universal “nonlinear fiber channel” because any new arrangement of a power distribution along the communication link or introduction of control elements would correspond to a different interplay between nonlinear effects and noise and, thus, to different nonlinear fiber channels.

The destructive role of nonlinearity is much more broadly known among engineers and it contributes to a common perception in the optical communications community that the effect of nonlinearity is “always to limit the capacity and quality of a communications channel.” Although it may sound counterintuitive, the nonlinearity, however, might improve system performance beyond the limitations of linear channels. Mastering the nonlinear effects can translate into a significant increase in capacity of communications systems.

In this work, we aim to change the common perception and show the optical community the potential of

specially designed nonlinear channels to unlock the capacity of future communications systems. We will demonstrate that channels using nonlinear filters can have a capacity exceeding the seminal Shannon limit for an AWGN channel. Before we present strict mathematical results, we explain in simple terms the key physical idea. The logarithmic dependence of capacity C on SNR S/N in the limit of large SNR in the Shannon formula can be understood as an estimate of the entropy of the maximal space-filling packing of the area characterized by the scale S (signal power) with objects of the size N (constellation points diffused by the noise power). The space filling corresponds to a number of nonoverlapping areas of size N qualitatively having meaning of a maximum number of constellation points that can be transmitted error-free. The number of bits corresponding to S/N of constellations levels is $\log_2(S/N)$.

Although this relation is strict only for uniformly distributed noise, it also is correct for AWGN channels through a more accurate analysis using soft (probabilistic) sphere packing [1]. The idea of exceeding the Shannon limit via nonlinear filtering can be explained using the same intuition. Whenever the nonlinear filter transformation has multiple fixed points, consequent interleaving the noise with nonlinear filter will produce effective suppression of the noise N improving performance compared to linear channel.

The Shannon capacity C (per unit of bandwidth B) of arbitrary communication channel $P(Y|X)$ (including nonlinear ones) is defined as a maximum of the functional (called mutual information) with respect to the statistics $P(X)$ of an input signal X [1]:

$$C = \max_{P(X)} \iint DXDY P(X, Y) \log_2 \frac{P(X, Y)}{P(X)P(Y)}. \quad (1)$$

Here, $P(X, Y) = P(Y|X)P(X)$ is the joint distribution function of input X and output Y ; the specific properties of a communication channel are given by the conditional probability $P(Y|X)$. We would like to stress once more that the capacity C is defined as a maximum over all possible signal distributions $P(X)$ (e.g., constellations). The optimal distribution $P(X)$ determines the signal modulation and coding that lead to the highest achievable transmission rate over given channel $P(Y|X)$. Therefore, results based on analysis of particular classes of signal

modulation should be understood only as an estimate of capacity from below and can even be misleading in some situations.

Recent studies [2–8] have shown that, in numerous practical situations, nonlinearity degrades the spectral efficiency of transmission. The current common perception is that, at a certain SNR level, nonlinear effects inevitably degrade fiber channel capacity. However, in terms of the Shannon capacity (as defined above), sub-optimal considerations can give only a lower bound of capacity. The capacity is defined as a maximum over $P(X)$, therefore, even when all tested signal constellations show a decline in transmission rate with growing SNR, it does not prove that *capacity* decreases. Moreover, information theory analysis [9] (see also [10]) proves that the capacity cannot decrease with SNR.

To simply illustrate for optical community the very possibility of exceeding the Shannon capacity limit (capacity of AWGN channel) in nonlinear channels, we consider as a proof of principle the chain of nonlinear filters that transform the signal X to the closest element of some finite alphabet: X_k . Each of the elements of the alphabet is associated with “region of attraction” S_k such that whenever $X \in S_k$ the result of transformation is simply X_k . Formally, this nonlinear transformation can be expressed as $Y = F(X) = \sum X_k I_{S_k}(X)$, where $I_k(X)$ is an indicator function for the set S_k : $I_k(X) = 1$ if $X \in S_k$ and $I_k(X) = 0$ otherwise.

In the following, we assume that the union of the sets S_k completely covers the space of possible signals X . Whenever the set S_k is chosen to be the set of points that are closer in some norm to X_k in comparison to any other point in alphabet, the corresponding decomposition is a so-called Voronoi diagram. For simplicity, in this Letter we focus on the analysis of the alphabets in a complex plane that are direct products of equally spaced, one-dimensional alphabets. In this case, the problem of finding the capacity and spectral efficiency of the channel can be reduced to a one-dimensional one. However, the general procedure can be straightforwardly applied to any alphabet and two-dimensional (2D) transformation $F(X)$. The example of such decomposition (corresponding to quadrature amplitude modulation) is shown in Fig. 1. In this case the alphabet consists of $K^2 = 25$ points in the complex plane.

Here we consider a model transmission system that consists of M spans with linear additive noise interleaved with nonlinear regenerator. For simplicity, we assume that all the regenerators are the same, and all the transmission line segments are the same. Given the conditional probability of the individual channel segment $P_C(Y|X)$ one can straightforwardly find the probability distribution corresponding to a serial composition of the channel C and regenerator R . The probability of having Y at the output of the generator R given the signal X at the input of channel C is

$$P(Y|X) = \int dX' \delta(Y - F(X')) P_C(X'|X). \quad (2)$$

For the ideal regenerator introduced above, the distribution of Y becomes discrete, with Y taking values only in X_k and the expression (2) simplifies to

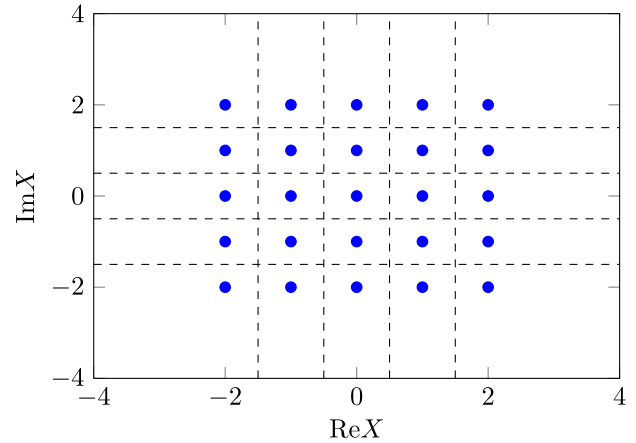


Fig. 1. (Color online) Schematic depiction of the modulation format and nonlinear transformation for $K = 5$. Dashed lines represent the Voronoi diagram of the system, and also are the boundaries of the regions of attraction of ideal nonlinear filters.

$$P(Y = X_k|X) = \int_{S_k} dX' P_C(X'|X). \quad (3)$$

Note that the nonlinear filter does not suppress the errors completely because there is generally a finite probability that the initial signal X_k will result in the output Y lying outside of the domain S_k . However, the suppression of all small errors that kept the signal X' within the domain S_k can lead to the increased capacity of the channel. To illustrate this point, we consider a simple system where we represent each segment of the original system as an AWGN channel $Y = X + \xi + i\eta$ with ξ, η being independently and identically distributed Gaussian variables each with zero mean and the variance given by N/M . In the absence of the regenerators, this channel corresponds to the classical one introduced by Shannon in [1]. In what follows, we present calculations only for one of two independent channels denoting $x = \text{Re} X$ and $y = \text{Re} Y$. The output y_n of the n th regenerator can be found using the following recurrence relation: $y_{n+1} = f(y_n + \xi_n)$, $y_0 = x$. Here $f(x)$ is the one-dimensional projection of the transformation F , and ξ_n is the random noise added in the original line segment n . In our system, both the initial signal x and all intermediate values y_n are parts of the alphabet, so the conditional probability is fully characterized by the matrix $W_{kl} \equiv P(y_n = x_k | y_{n-1} = x_l)$ that according to (3) is found as

$$W_{kl} = \frac{1}{2} [\text{erf}(x_{kl}^+) - \text{erf}(x_{kl}^-)], \quad (4)$$

where $x_{kl}^\pm = (x_k + x_{k\pm 1} - 2x_l) / \sqrt{8N/M}$; here we assume that the values of x_k are sorted in increasing fashion: $x_1 < x_2 < \dots < x_K$ and we define $x_0 = -\infty$, $x_{K+1} = +\infty$. The expression for the conditional probability for M regenerators can be found by taking the M th power of the matrix W : $P(y_M = x_k | x = x_l) = (W^M)_{kl}$. One can see that for the single segment the probability of error $1 - W_{kk}$ can be estimated as $\sim \exp[-M(\delta x)^2/N]$, where δx is the characteristic spacing between the alphabet elements. Assuming the single segment error probability

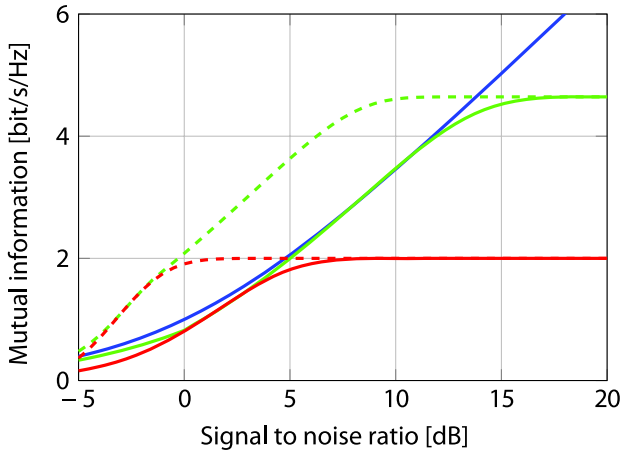


Fig. 2. Mutual information is shown as a function of SNR: the solid red ($K = 2$) and green ($K = 5$) lines are capacities of the linear discrete memoryless channels with K^2 elements. The corresponding dashed lines depict the mutual information in the system with 10 regenerative nonlinear filters with the same alphabets of 4 and 25 elements. The solid blue line shows for comparison the classical Shannon capacity of the linear 2D AWGN channel.

to be small, the aggregated error rate for the M segment line can be estimated as $1 - (W^M)_{kk} \sim M \exp[-M(\delta x)^2/N]$. This function decays very rapidly with the number of segments and quantifies the effect of the regenerators on the error rate improvement.

To demonstrate the beyond-Shannon performance of the system with nonlinear filters, we have calculated the mutual information of the several sample schemes ($K = 2$ and $K = 5$) for $P(x)$ satisfying the power constraint $\int dx P(x) x^2 = S$. The resulting value of mutual information has dimension of spectral efficiency and provides the lower bound on the actual capacity of the channel. We use as a trial function the maximum entropy distribution that is given by $P(x) = \sum_k w_k \delta(x - x_k)$ with $w_k = \alpha \exp(-\beta x_k^2)$ where the constants α and β were chosen to satisfy the normalization and power constraints and optimization was carried out over the spacing between the points of the alphabet. The results of these simulations for alphabets with $K = 2$ and $K = 5$ are presented in Fig. 2.

Figure 2 shows that the capacity of a system with regenerative nonlinear filters can exceed the Shannon limit for AWGN (shown by the solid blue line) without any optimization of $P(X)$. Exceeding the Shannon capacity for AWGN channel is possible for a broad class of nonlinear

filters that satisfy the following properties: (i) it should have a large number of stable fixed points (alphabet) that guarantees there is enough entropy to transmit information using only these points as the alphabet and (ii) the region of attraction of each of the fixed point should be larger than the characteristic noise amplitude on the segment ($\sqrt{N/M}$) to guarantee that the probability of moving Y to a different fixed point is very low.

Optical regeneration [11,12] and active nonlinear control providing for separation of the signal from noise can dramatically change the way optical noise is accumulated along the fiber link leading to capacity above the linear Shannon limit for AWGN channel. To simplify, capacity depends on the SNR and continuous regeneration of signal means an effective removal of optical noise from the system—an effective nonlinear filtering. As a result, the overall noise is smaller compared to corresponding linear system with the same number of amplifiers.

In conclusion, we have demonstrated that the capacity of an optical communication channel with regenerating in-line elements under certain conditions can be higher than the Shannon capacity for linear AWGN channel. We hope that our result will stimulate further research on the practical implementation of optical channels with capacity above the linear Shannon limit.

We would like to acknowledge the financial support of the European Research Council (ERC) and Engineering and Physical Sciences Research Council (EPSRC) project UNLOC.

References

1. C. E. Shannon, *Bell Syst. Tech. J.* **27**, 379 (1948).
2. P. P. Mitra and J. B. Stark, *Nature* **411**, 1027 (2001).
3. E. Desurvire, *J. Lightwave Technol.* **24**, 4697 (2006).
4. R.-J. Essiambre, G. Foschini, G. Kramer, and P. Winzer, *Phys. Rev. Lett.* **101** (2008).
5. A. D. Ellis, J. Zhao, and D. Cotter, *J. Lightwave Technol.* **28**, 423 (2010).
6. D. J. Richardson, *Science* **330**, 327 (2010).
7. R. Killey and C. Behrens, *J. Mod. Opt.* **58**, 1 (2011).
8. A. Mecozzi and R.-J. Essiambre, *J. Lightwave Technol.* **30**, 2011–2024 (2012).
9. E. Agrell, “The channel capacity increases with power,” <http://arxiv.org/abs/1108.0391> (2011).
10. K. Turitsyn, S. Derevyanko, I. Yurkevich, and S. Turitsyn, *Phys. Rev. Lett.* **91**, 1 (2003).
11. M. Nakazawa, E. Yamada, H. Kubota, and K. Suzuki, *Electron. Lett.* **27**, 1270 (1991).
12. T. Lakoba and M. Vasilyev, *Opt. Express* **15**, 10061 (2007).