The nonlinear Shannon limit and the need for new fibres

A.D. Ellis

Tyndall National Institute & Dept. of Physics, University College Cork, Dyke Parade, Cork, Ireland

ABSTRACT

This paper will review the current understanding of the so called nonlinear Shannon limit, and will speculate on methods to approach the limit through new system configurations, and increase the limit using new optical fibres.

Keywords: Capacity Limit. Nonlinear Optics. Optical Communications

1. INTRODUCTION

Since the first commercial wire-line communication system (telegraph) was deployed by Wheatstone and Cooke in 1839 [1], it may be argued that communications capacity offered to residential customers has increased exponentially at a constant compound annual growth rate of nearly 40% per annum. Today, growth is quantified in terms of the overall internet capacity, where similar overall growth rates are found [2] and is fuelled by a succession of communications applications, each of which evolves through stages of early adoption, increasing penetration and establishment. This was readily seen to be the case of direct dialled telephony, fax and dial up modems, despite transport over a common platform. Similarly, text messaging, image sharing, file sharing and video services represent different applications, sharing a common platform widely referred to as the internet. Recent growth, since the introduction of one of the 1st internet like service, the bulletin board service [3], accessed by dial up modems, through to passive optical networks, still the subject of current research [4] is shown in Figure 1. This figure also illustrates the typical capacity switched within the core, for example today’s WDM systems have a total capacity of up to 8 Tbit/s, but typically a single wavelength may be routed using a reconfigurable optical add drop multiplexer and would have a capacity of between 10 and 100 Gbit/s.

![Figure 1: Evolution of access capacity and core network switching granularity since the introduction of BBS](image-url)
It is interesting to note that the ratio between the switchable capacity in the core of the network to the access rate has remained close to 1,000 over this entire period, despite numerous generations of routing platforms including schemes as diverse as the Plesiochronous Digital Hierarchy (SDH) and Internet Protocol transported by Wavelength Division Multiplexing (IP over WDM). This consistent ratio represents an engineering trade-off between the impact of failures, which favours many separate links, and cost, which favours fewer higher capacity links. Both straw polls of consumer demand and government strategy suggest that the demand for increased consumer bandwidth will be maintained for the foreseeable future as testified by the increasing deployment of optical fibre in the local loop and the design bandwidths of successive generations of wireless technologies. This suggests that we will require the ability to switch capacities of several Tbit/s in the core network, with maximum capacities a few orders of magnitude higher than this.

Figure 2 illustrates a common figure of merit for optical communication system experiments, the “bit rate distance product” which combines the two important features, the total bit rate and the maximum transmission distance in a single figure of merit, along with the actual total fibre capacity. Rapid progress was observed throughout the 1980’s and 1990’s where the decline in growth was observed. This decline was readily attributed to the end of the so called “dotcom boom”, which saw a considerable over deployment of network capacity, in particular of 10 Gbit/s WDM systems. However, once the capacity demand, which continued uninterrupted throughout the crisis, had exhausted this overbuild, whilst total fibre capacities resumed their former growth rate, maximum transmission distances declined, where there is only a marginal increase in the maximum bit rate distance product which may be attributed to the adoption of super channels [5] and coherent detection [6] (green symbols).

2. THE NONLINEAR SHANNON LIMIT

The origins of this “capacity crunch” where user demand (Figure 1) appears to be approaching the limits of an optical fibre (Figure 2) is a trade of between the well known Shannon limit and nonlinearity. Shannon predicted that the information capacity of a fixed bandwidth communication channel without memory is proportional to the bandwidth of the channel, and logarithmically proportional to the signal to noise ratio. On the other hand other effects such as...
chromatic dispersion and nonlinearity induce memory into the channel and induce severe signal distortion. This additional nonlinear distortion limits the maximum power which may be launched into an optical fibre. In the majority of circumstances, the limiting nonlinearity is the optical Kerr effect [7], and accurate predictions may be made using the nonlinear Schrödinger equation as proposed by Zakharov [8]. Such readily predicted degradations may of course be compensated by sufficiently sophisticated signal processing [9, 10,11,12], and indeed, it has been shown that if the ability to compensate for all nonlinearity in the transmission line is assumed, then the capacity of the communications channel should increase monotonically with signal power [13,14]. However, if the entire output of the communication channel is not known, either because there is insufficient processing power available [12], or because individual wavelengths propagate along different paths [15], then the unknown signals must be considered as sources of noise to the known signals. This additional noise increases with signal power and prevents the arbitrary increase of the signal to noise ratio, and thus imposes a maximum capacity on the communication channel. The existence of this maxima was most famously articulated at the end of the dotcom boom [16], and many of the consequences recently summarized in extensive review articles [17,18]. The original formulation assumes large chromatic dispersion, nonlinearity accumulating over many spans, no nonlinear interaction with noise, and intra channel effects that are either compensated or negligible. The expression derived using this approach is reproduced in equation (1)

\[ C \approx \log_2 \left( \frac{1}{B} \text{snr} \right) = \log_2 \left( 1 + \frac{P_s}{P_N + \left( 1 - e^{-\left( \frac{P_s}{P_{NL}} \right)^2} \right)} \right) \]

(1)

where, \( C \) is the total capacity of the system occupying a bandwidth \( B \), \( \text{snr} \) is the signal to noise ratio, \( P_s \) and \( P_N \) the signal and noise power spectral densities respectively and \( P_{NL} \) a “nonlinear power spectral density” scaling factor. For a system with lumped amplifiers the noise and nonlinearity power spectral densities are given by

\[ P_N \approx N_a (G - 1) n_{sp} h \nu \]

(2)

\[ P_{NL} \approx B \frac{1}{\lambda^2} \frac{B.D.}{\Delta f} \frac{\lambda c}{\gamma^2 L_{eff}} \]

(3)

where \( N_a \) represents the number of amplified spans, \( G \) the amplifier gain, \( n_{sp} \) the spontaneous emission noise factor, \( c \) the speed of light \( h \) Planck’s constant, \( \nu \) the carrier frequency (\( \lambda \) the corresponding wavelength), \( D \) the chromatic dispersion of the fibre, \( \gamma \) its nonlinear coefficient and \( L_{eff} \) the nonlinear effective length of the overall system (sum of the conventional effective lengths of each span). \( \Delta f \) represents both the channel spacing and the bandwidth of each of the \( Nch \) WDM channels. Subsequently, many authors have investigated the problem with a variety of different assumptions, including arbitrary dispersion [19], dominance of four wave mixing [20] and WDM channel bandwidths less than the channel spacing [21]. The majority of these results bear close resemblance to equations (1) to (3), especially if the Taylor series expansion is considered [22] and the predicted capacities agree closely. Detailed numerical simulations [23] reveal that each model performs best for systems with designs matching the assumptions of the model (eg OFDM system designs match [20] and Nyquist designs [21]), but even so, except in extreme cases, the trends predicted by the various models are highly correlated. In terms of system configuration, the results are weakly (if at all) dependent on the channel spacing for a large number of channels, and for the fibre parameters, a figure of merit may be derived for a lumped amplifier system;

\[ P_{NL}^{2} \left( \text{Lumped} \right) \propto \left( 1 - e^{-\lambda L_{span}} \right) \left( e^{\alpha L_{span}} - 1 \right)^2 \frac{\gamma^2}{D} \]

(4)

where \( L_{span} \) represents the span length between amplifiers. This illustrates a weak, but monotonic dependence of absolute value of the dispersion and the nonlinear coefficient. Dependence on fibre loss is more complex due to the trade off.
between noise (low loss) and nonlinear effective length (high loss), and care should be taken in generalizing this parameter and the system configuration (repeater spacing) may have a significant impact on the conclusions [17, 18]. For an ideal Raman amplified system, the situation is clearer, with lower loss always favourable.

![Graph showing typical nonlinear Shannon limit curves](image-url)

Figure 3 Typical nonlinear Shannon limit curves showing (top) information spectral density per polarization as a function of the launched power spectral density from each amplifier for noise figures of 4.5 (blue), 3 (green), and 0 (red) dB and (bottom) the total information spectral density (both polarization) as a function of the total launched power for a fixed information bandwidth for doped fibre amplifiers (blue) with 4.5 (solid) and 3 (dashed) dB noise figures, Raman amplifiers (red) and phase sensitive amplifier with simultaneous idler transmission (green) with 0 (solid) and -3 (dashed) dB noise figures.

Typical plots showing the nonlinear Shannon limit are shown in Figure 3. The left hand chart illustrates the impact of varying the amplifier noise figure alone. This has no effect on the nonlinearity, and so at high powers all three curves
converge, however, in the noise limited region, the same signal to noise ratio is obtained at signal launch powers, alternatively, if the noise figure is dropped by 3dB, the signal to noise ratio is increased by a factor of two and hence the maximum achievable information spectral density is increased by 1 b/s/Hz/pol. Note that in order to achieve a noise figure below 3dB, it is necessary to use a phase sensitive amplifier where the idlers have also been transmitted along the fibre [24]. In this case, the total signal power (right hand chart in figure 3) is increased and there is no change in the maximum information spectral density at a fixed total amplifier output power in the signal to noise ratio limited regime. On the other hand, in the nonlinear regime, the output power is spread over more channels, reducing the power spectral density, and hence the nonlinear crosstalk between closely spaced channels. However, with both interpretations (signal power spectral density and total amplifier output power) the increase in maximum ISD due to improvements in amplifier noise figure is small. Similarly, whilst super-channels, either based on OFDM [5, 25] or Nyquist WDM [26] only offer a marginal increase in the ISD due to the elimination of guard bands.

3. OPTICAL REGENERATION

Given that modifications to the transponders and optical amplifiers will have a modest impact on the overall ISD of an optical link, more radical solutions are required. A long standing proposal has been to employ optical regenerators in place of optical amplifiers [27,28,29], however in order to be cost effective, such regenerators must be WDM compatible and should allow net ISD exceeding that of conventional coherent detection. Whilst WDM compatibility has been demonstrated for a range of modulation formats [29, 30], it is unlikely that the simple binary regenerators demonstrated to date will meet this later criterion. The required regenerator performance may be understood by plotting the nonlinear Shannon limit in terms of the maximum reach between forward error correction modules for a given QAM constellation, as shown in Figure 4.

Given that a PM QPSK system is adequate for the majority of applications, in order to allow the reach to be at least doubled (and at most octupled), the optical regeneration system should be compatible with PM-16QAM data and for formats beyond PM-256QAM, more than 20 regenerators will be required to match the reach of a PM-QPSK digital coherent system. This is a highly challenging target, however recent results are promising, showing black box WDM [30] and QPSK [31, 32] regeneration and m-PSK processing [33]. Optical regeneration may therefore offer an attractive means to increase fibre capacity, but calculations such as Figure 4 may also be used to estimate an upper bound on the regenerator cost and/or power consumption. For example, comparing \( M \) parallel systems employing \( N_a \) amplifiers with power consumption \( P_a \) to a regenerated system offering \( M \) times the maximum capacity per wavelength using \( N_r \) regenerators in place of amplifiers, then the power consumption of the regenerator, \( P_r \), should be

\[
P_r \leq P_a \left( M - 1 \right) \frac{N_a}{N_r} + 1
\]

which allows for a regenerator power consumption 19 times greater than that of an amplifier is 20 regenerators are required in place of amplifiers in a 75 span link to enable the link capacity to be increased by factor of 6 (e.g. from PM-QPSK to PM-256QAM). In the limit of replacing each amplifier by a regenerator, the regenerator power consumption should have less than \( M \) times the power consumption of the amplifier to be a power efficient solution.
Figure 4 Relative performance limit of digital coherent transmission using optical amplifiers (red) and a system employing 20 optical regenerators (blue) with 80km fibre spans, normalized to the reach of a PM-QPSK system.

4. NOVEL OPTICAL FIBRES

The nonlinear Shannon limit described above gives a fundamental limit on the information spectral density of the fibre, logarithmically directly dependent on the fibre parameters (equations 1 and 4). Modifications to the amplifier noise figure and modulation format may, in certain circumstances, allow the ISD to be increased by 25-100%, and optical regeneration may enable ISD increases of up to \( \log_2(\frac{L_{\text{sys}}}{L_{\text{span}}}) \) per polarization, where \( L_{\text{sys}} \) is the overall link length, if regenerators ubiquitously replace optical amplifiers (around 8 bit/s/Hz for an ultra long haul system). The remaining alternative is to deploy new fibre cables and a number of options are possible, such as

1. Multiple standard single mode fibres
2. Multi-core fibres
3. Few-mode fibres with multiple-input multiple-output signal processing
4. Novel optical fibres, in this paper we will consider the extreme case of hollow core photonic band gap fibre.

The baseline solution is to employ many parallel systems, also known as spatial multiplexing. When employed over conventional fibres, this offers linear power consumption scaling with network capacity whilst demand continues to grow exponentially. For this solution the resultant exponential growth in energy consumption will inevitably result in the imposition of limitations on network capacity through pricing or regulation. But do proposed new fibre types [e.g. 34,35,36,37,38] offer any substantial benefit in terms of ISD or energy consumption? Figure 5 illustrates the minimum pump power summed over all optical amplifiers in the chain required to achieve a given ISD assuming polarization multiplexing with coherent detection. A challenging 2,000km system with 100km spaced fibre amplifiers (no Raman amplification) is chosen as a reference for comparison, and the fibre parameters which have been assumed are summarized in Table 1 and include two types of multi-mode fibre (step index few-mode fibre with high differential mode delay, low mode coupling, large effective area but slightly higher loss and graded index few mode fibre with reduced differential
mode delay, and loss and effective areas similar to that of conventional fibre). In Figure 5, each curve is initially signal to noise ratio limited and so, as the ISD increases the required power also increases at a fixed rate (see equation 1). However, once the system is impacted by nonlinearity, the optical signal to noise ratio penalty increases the required launch power up to a point where the ISD may no longer be achieved. All four new fibres considered: multi-core fibre (MCF), step index few-mode fibre (SI-FMF), graded index few-mode fibre (GI-FMF) and hollow core photonic band gap fibre (PBGF), offer substantial increases in the achievable ISD, but since the loss of all of the solid core fibres is similar, there is no significant improvement in the signal to noise ratio limited performance of each channel, and so no improvement in the required pump power. On the other hand, the theoretically predicted improvements in loss of hollow core PBGF allow for substantial reductions in the required signal launch power, even when the fibre operates with a single mode. Similar conclusions may be drawn for a wide range of system lengths.

Table 1 Fibre parameters used for calculations used unless otherwise specified

<table>
<thead>
<tr>
<th>Item</th>
<th>Standard fibre</th>
<th>Multi-core fibre</th>
<th>Step Index few-mode fibre</th>
<th>Graded index few-mode fibre</th>
<th>Hollow core fibre</th>
</tr>
</thead>
<tbody>
<tr>
<td>System length</td>
<td>2,000 km</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Repeater spacing</td>
<td>80 km</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Channel spacing</td>
<td>50 GHz</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dispersion</td>
<td>20 ps/nm/km</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loss</td>
<td>.2 dB/km</td>
<td>.2 dB/km</td>
<td>.22 dB/km</td>
<td>.2 dB/km</td>
<td>0.05 dB/km</td>
</tr>
<tr>
<td>Nonlinearity</td>
<td>1.4 /W/km</td>
<td>1.4 /W/km</td>
<td>0.5 /W/km</td>
<td>1.4 /W/km</td>
<td>0.001 /W/km</td>
</tr>
<tr>
<td>Modes</td>
<td>2</td>
<td>2 x 6</td>
<td>6</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Wavelength</td>
<td>1550 nm</td>
<td>1550 nm</td>
<td>1550 nm</td>
<td>1550 nm</td>
<td>2000 nm</td>
</tr>
<tr>
<td>Information loss per span</td>
<td>0%</td>
<td>0%</td>
<td>1%</td>
<td>.3%</td>
<td>0%</td>
</tr>
</tbody>
</table>

Figure 5 minimum total pump power (summed over all amplifiers) to achieve a given ISD over a 2,000km link with 100km amplifier spacing for conventional fibre (black), graded index few mode fire (blue, dotted), step index few-mode fibre (blue, solid), multi-core fibre (green) and hollow core PCF (red) with (dotted) and without (solid) mode multiplexing. Subscript denotes the total number of modes (including polarizations). The purple line represents a constant pump power per bit.
However, as indicated above, predictions for lumped amplifier systems are critically dependent on the amplifier spacing (equation 4), and we would anticipate that the relative merits that spatial multiplexing offers and low loss fibre are determined by this parameter. This is illustrated in Figure 6 which shows the required total pump power as a function of the amplifier spacing for a 2,000 km system with a target ISD of 10 b/s/Hz. At low repeater spacing, information loss from mode coupling, assumed to be proportional to the number of amplifiers and higher for low differential group delay fibres, imposes a large OSNR penalty and dominates the performance of the system. However, for inland networks, where amplifiers are typically spaced 80km apart, the hollow core PBGF offers the lowest required pump power, even when operated as a single mode fibre (two polarizations). If the repeater spacing, may be freely selected, for example in a submarine network, spatial multiplexing offers improved performance below 80km, especially using multi-core fibres (or multiple fibres).

![Figure 6](image_url)

Figure 6 Variation of required total pump power with amplifier spacing for a 2,000km system with an ISD of 10 b/s/Hz. Curves are as specified in Figure 5.

Note that for this system configuration, spatial multiplexing over solid core few-mode fibre offers no significant advantage over multi-core fibres. Again, the conclusion that the energy consumption associated with amplification is minimized by minimizing the fibre loss is true for a wide range of systems lengths and target ISD’s.

Note that in terms of total energy consumption; commercially available fibre amplifiers with output powers in the region of 100mW typically have electrical power consumptions in the region of 10’s of Watts [39]. Consequently we may anticipate that the total power consumption of the transmission line would be between 100W and 1kW for a bi-directional system. This should be compared to the power consumption of digital coherent transponders, typically in the region of 1-3 W/Gb [40]. For a 10 Tbit/s system (eg 100 channels each at 100 Gbit/s) the transponder power consumption would exceed 10kW, over 10 times the power consumption of the transmission line. Furthermore, one may anticipate that the power consumption associated with the MIMO processing for a few-mode fibre system would be significantly higher than a conventional digital coherent receiver.

5. CONCLUSIONS

In this paper we have considered the implications of the nonlinear Shannon limit, and in particular the consequences for transmission link designs which would meet the growing demand for bandwidth without significant increases in energy consumption. Whilst many techniques are found to offer the potential for capacity increases, the logarithmic nature of the limit implies that even optical regeneration only offers an ISD increase of 2 to 3 times. The techniques will delay a capacity crunch associated with energy consumption. In terms of energy consumption, loss is clearly seen to be the
dominant parameter with low loss hollow core PBGF offering the lowest link power consumption for a wide range of system configurations.

6. ACKNOWLEDGEMENTS

This work has been supported by Science Foundation Ireland under Grant 06/IN/I969 and by the EU FP7-ICT MODE-GAP project under grant agreement 228033

7. REFERENCES


