Available online at www.sciencedirect.com
SCIENCE DIRECT®
C. R. Physique 4 (2003) 11–28
COMPTES RENDUS PHYSIQUE

Optical telecommunications/Les télécommunications optiques

Quantum and nonlinearity limitations of the optical communication channel

Limites de bruit quantique et de non-linéarité du canal de communication optique

Emmanuel Desurvire

Alcatel Submarine Networks, Centre de Villarceaux, 91625 La Ville du Bois, France

Received 2 December 2002

Presented by Guy Laval

Abstract

We analyze the information-capacity limitations of the optical communication channel, as determined by noise accumulation from optical amplification and nonlinear wave-mixing. We review the concepts of signal-to-noise ratio and entropy for binary-coded and continuous communications, leading to a definition of ultimate capacity for the optically-amplified channel. A unified quantum model, describing both amplification and nonlinearity limitations, makes possible to determine the power transmission window within which the channel capacity can be maximized. ***To cite this article: E. Desurvire, C. R. Physique 4 (2003).***

Résumé

Nous analysons les limites en capacité d'information des communications par canal optique, telles que déterminées par l'accumulation du bruit d'amplification optique et de mélange à quatre-ondes non-linéaire. Nous revoyons les notions de rapport signal-à-bruit et d'entropie concernant les communications à codage binaire ou continu, lesquelles conduisent à la capacité ultime d'un canal optiquement amplifié. Un modèle quantique unifié, décrivant les limitations dues à l'amplification et à la non-linéarité, permet de déterminer la fenêtre de puissance signal à l'intérieur de laquelle la capacité du canal peut être maximisée. ***Pour citer cet article : E. Desurvire, C. R. Physique 4 (2003).***

Keywords: Optical communications; Optical fibers; Fiber dispersion; Erbium-doped fiber amplifiers; Distributed amplification; Quantum noise; Information theory; Entropy; Equivocation; Shannon–Hartley theorem; Nonlinear Schrödinger equation; Fiber nonlinearity; Four-wave mixing; Wavelength-division multiplexing; All-optical regeneration

Mots-clés : Communications optiques ; Fibres optiques ; Dispersion ; Amplificateurs à fibre dopée à l'erbium ; Amplification distribuée ; Bruit quantique ; Théorie de l'information ; Entropie ; Équivocation ; Théorème de Shannon–Hartley ; Équation de Schrödinger non-linéaire ; Nonlinéarité des fibres ; Mélange à quatre ondes ; Multiplexage en longueur d'onde ; Régénération tout-optique

E-mail address: Emmanuel.Desurvire@alcatel.fr (E. Desurvire).

1631-0705/03/$ – see front matter
doi:10.1016/S1631-0705(02)00002-6

1. Introduction

At this current and very advanced stage of optical communications, a relevant issue is the identification of ultimate system performance limits, as defined by *fundamental physics principles* and *information theory* considerations. In this paper, we review both aspects and their combined conclusions. The underlying concept is the '*optical communication channel*' (OCC), which must be revisited in order to take into account the combined effects of *quantum noise* (as coming from optical amplification) and *nonlinearity* (as coming from the transmission fiber). As described here, the new OCC concept departs from the original one described in 1948 by C.E. Shannon, the 'noisy communication channel', which consisted in a linear channel perturbed by additive noise [1]. In fact, both quantum and nonlinearity noise are *non-additive*, since their magnitudes are functions of transmitted signal power. The power being too low, quantum noise sets the OCC limit. The power being too high, nonlinearity noise sets the OCC limit. Given quantum and nonlinearity noise constraints, the two limits thus define a power window for the OCC [2–4]. The analysis must take into account the interplay between system nonlinearity and dispersion. The highest transmission bandwidth (in terms of capacity $\times$ distance) is achieved through a careful trade-off between the nonlinearity and the dispersion, within other limits caused by cumulated amplifier-noise [5,6], and data-encoding performance associated with error-correction coding (ECC) capabilities [7]. It is noteworthy that in this trade-off regime OCC signals are neither purely-linear pulses nor exact nonlinear 'solitons', representing as many possible solutions for the so-called *nonlinear Schrödinger equation* (NLSE) [8,9]. To complete this inventory of new OCC features, one must finally consider *all-optical in-line regeneration* [4,10, 11]. As the name indicates, signals can be periodically 'regenerated' as they propagate through the line. Interestingly from the physics viewpoint, optical regeneration makes possible to *remove* quantum and nonlinearity noises from the channel, allowing 'infinite' OCC transmission distances without any degradation of signal quality, and in some conditions, even with signal quality improvement. This approach should not be confused with periodic *opto-electronic* regeneration for which OCC noise is cumulative.

The above inventory illustrates how different is the current OCC from its original and classical formulation of a purely linear channel with additive noise. In this paper, we present the state of the art in the analysis of OCC limitations due to quantum (amplification) noise and nonlinearity noise. The analysis is based upon a unified quantum model [3] which encapsulates the two concepts through the same noise-source formalism. The paper is divided into seven sections (including this introduction).

In Section 2, we first recall the definitions of *signal-to-noise ratio* (SNR), *bit-error-rate* (BER), as applying to the linear and nonlinear systems. We discuss the different strategies which have been used so far to minimize nonlinearity. We then introduce (in a tutorial-like way) the concept of *entropy* for discrete channels. This background leads to the definition of *information capacity* (*or information spectral density*, ISD) for noisy OCCs. As an illustration, we present an original result which links channel capacity and BER in the case of ON–OFF binary OCCs. Considering continuous channels, we then introduce the *Shannon–Hartley theorem* (SHT), relating the ISD to SNR.

In Section 3, we briefly recall the quantum description for amplifier noise, showing that it originates from two independent vacuum-field couplings, consistent with the model of a imperfectly-inverted laser system. We show that amplifiers can also be modeled by a Langevin-like thermal noise source, as a conceptual derivation of the first model. This description allows one to obtain exact definitions for photon-number mean and variance, leading to a formulation of SNR and BER in periodically-amplified OCC. The effect of amplifier spacing is taken into account, leading to an original definition of the ultimate channel capacity for ideal distributed-amplification systems.

In Section 4, we consider limitations introduced by nonlinearity, in the case of multi-channel (WDM) transmission. We first recall the origin of nonlinear wave-mixing in a tutorial-like way, leading to the definition of the NLSE. We then recall the results obtained in [2] from the NLSE leading to an analytical definition of WDM nonlinearity noise linking all relevant system parameters (nonlinear coefficient, dispersion, loss, signal poser, channel spacing, bit rate, ...). We then describe a quantum model [3] where nonlinearity noise is generated by a Langevin-like chaotic source. This allows one to model simultaneously amplifier and nonlinearity noises. The approach leads to a re-formulation of the SHT, talking into account both noises types.

In Section 5, we use the previous results to assess the ultimate OCC capacity, as limited by amplification and nonlinearity. Finally, we discuss in Section 6 the implications of all-optical regeneration to overcome the OCC limits in the case of binary systems. We show that optical regeneration is equivalent to introduce *negative entropy* in the OCC. The conclusion is provided in Section 7.

2. The noisy optical-communication channel

The field of information theory [1] views a 'communication channel' as comprising the following elements (Fig. 1):

- an information source (messages to be transmitted);
- a transmitter (means to encode messages into *symbols* by modulating a carrier source;

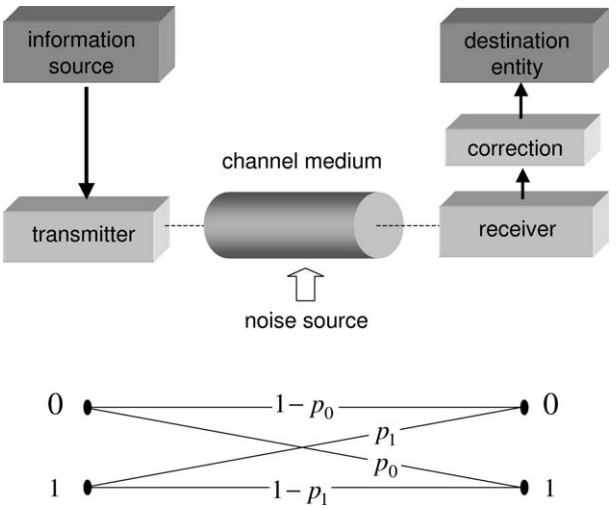

Fig. 1. Elements involved in a communication channel (top), and diagram showing different possible events associated with the transmission of two symbols '0' and '1' (successful or failed reception).

- a channel medium (spanning some finite transmission distance);
- a receiver (decoding and reconstructing the original message);
- a corrective signal-processing device (improving signal quality);
- a destination entity (receiving the message).

The channel may be polluted by an internal noise source (noisy channel), which introduces *uncertainty* in the reception of symbols. This noise is generally assumed stationary, meaning that the uncertainty does not depend upon the system's history and transmitted symbol sequences, as referred to as a *memory-less* channel. In the classic view, the noise is assumed *additive*, meaning that the channel output is the linear superposition of the signal with a noise background. Here, we focus on an optical communication channel (OCC), where the transmission medium is an optical fiber, the symbols are light pulses, and the first noise source is *quantum noise* generated by line optical amplifiers. In the following subsections, we consider the *linear* and the *nonlinear* (dispersive) cases.

2.1. Linear case

Considering binary, intensity-modulated light signals (referred to as ON–OFF keying), the two possible symbols are '0' (no pulse) and '1' (pulse). Assuming that the 0/1 symbols are not perfectly transmitted, there is a finite probability that the receiver (including posterior correction) outputs a '1' when a '0' was emitted/transmitted and the reverse. The associated error probabilities are $p_0 = p(1|0)$ and $p_1 = p(0|1)$ respectively. Correct 0/1 symbol receptions have the probabilities $p(0|0) = 1 - p_0$ and $p(1|1) = 1 - p_1$, respectively, as illustrated by the diagram in Fig. 1. If the 0/1 symbols have the probability of emission, $p(0)$ and $p(1)$, the BER can be expressed as:

$$BER = p(0)\,p(1|0) + p(1)\,p(0|1) \equiv \frac{1}{2}\big[p(1|0) + p(0|1)\big]. \tag{1}$$

The approximation in Eq. (1) assumes equal symbol probabilities, $p(0) = p(1) = 1/2$, as is the case with actual long coding sequences. It can be shown that, assuming Gaussian noise, the minimal BER is obtained by an optimal 'symbol decision' for which the conditional probabilities, are equal, i.e., $p(1|0) \approx p(0|1)$. A low BER (e.g., 10^{-9}, or one error out of a billion received bits) corresponds to high transmission quality. By use of powerful error-correcting codes (ECC), as based upon redundancy-bit algorithms, it is actually possible to restore BERs as high as 10^{-4}–10^{-3} to low values of 10^{-12}–10^{-9} [2,7].

In the above picture, the channel is assumed *linear*. This implies that sending multiple channels at different carrier wavelengths (and power levels) in the OCC leaves symbol uncertainty unchanged, as modeled by a single common/external noise source. Considering separate carrier-wavelength sub-channels or a single global OCC lead to the same analysis. However, two OCC impairments are the medium dispersion (group-velocity dependence in carrier wavelength) and the medium transmission loss. Dispersion causes symbol pulses to spread in the time domain. In order to recover the original 0/1 pulse shapes, dispersion can be compensated by use of dispersive elements (such as *dispersion-compensating fibers* or DCF) with

opposite effect periodically inserted in the signal path [5,6,12]. The medium transmission loss causes the pulse energy to vanish, bringing the '1' signals closer to the receiver noise background an increasing the probability $p_1 = p(0|1)$. Loss can be exactly compensated by optical in-line amplifiers (such as *erbium-doped fiber amplifiers*, or EDFA [4,13,14]).

If one then overlooks receiver noise, the only limitation of the OCC is that imposed by the common noise source, namely the *quantum noise* generated by optical amplification. Note that the term 'quantum noise' used here points to its vacuum-field fluctuations origin, as described in Section 3, which is unexplained through classical theory (in spite of semi-classical assumptions). The word 'quantum' is also consistent with microwave engineering practice, i.e., at optical carriers for which $h\nu/k_B T \gg 1$, 'quantum noise' dominates over 'thermal noise' in wave amplifiers [4,14]. On the other hand, this quantum feature applies to signals having relatively large photon numbers, which conceptually distinguish this field from true 'quantum communications' [15].

Since amplified systems contain several amplifiers in cascade so as to form a transparent line, the amplifier noise accumulates along the signal path, causing the SNR to decrease with distance. In turn, the SNR decay causes the BER to *exponentially increase*, according to the following definition [4,14]:

$$BER = \frac{1}{Q\sqrt{2\pi}} \exp\left(-\frac{Q^2}{2}\right) \tag{2}$$

with the Q-factor being defined by

$$Q = 2\sqrt{2}\frac{SNR}{\sqrt{1+\sqrt{1+4SNR}}} \approx \sqrt{2SNR}. \tag{3}$$

The above SNR definition corresponds to the ratio of mean/time-averaged signal power to un-polarized amplifier noise power, as is customary in optical fiber communications. The high-SNR limit in Eq. (3) together with Eq. (2) shows that the BER exponentially decays as

$$BER \approx \frac{\exp(-SNR)}{SNR\sqrt{4\pi}}. \tag{4}$$

For a given per-wavelength signal power launched into the OCC, the maximum transmission distance (at required BER) and minimum BER (at required distance) are thus intrinsically limited by SNR decay due to quantum noise (see analytical definition in Section 3). Note that in this case, all wavelength carriers experience the same limitations and have therefore identical BER performance.

2.2. Nonlinear case

Assume next that the channel medium is *nonlinear*. This means that the symbol probabilities must also be functions of the signal power in each wavelength carrier. Fig. 2 shows that nonlinearity causes the carriers to act both as self-imposed and mutually-imposed noises sources within the OCC. Such a diversification and cross-coupling of noise sources comes from several factors:

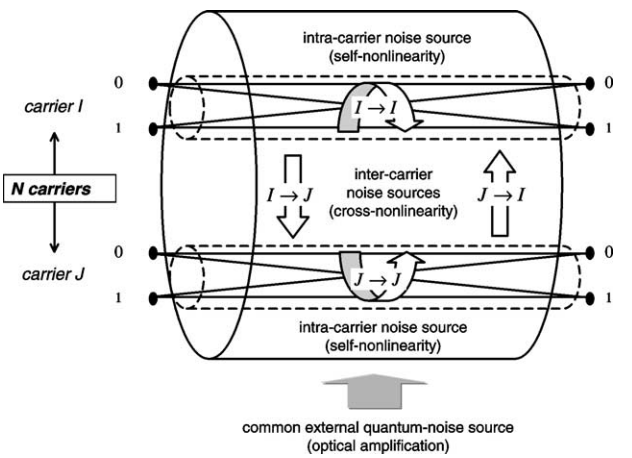

Fig. 2. Representation of the nonlinear optical communication channel with multiple input wavelength carriers, showing a common external quantum-noise source (optical amplification), and the variety of nonlinearity-noise sources, as being self-imposed (intra-carrier) or mutually-imposed (inter-carrier).

– symbol values 0/1 are random, causing random nonlinear interference events between pulses X of a carrier and pulses Y of another carrier (namely 0–0, 0–1, 1–0 and 1–1);
– wavelength carriers propagate at different group velocities (dispersion), causing the strength of the nonlinear interaction to randomly depend upon relative velocity difference (interaction being the strongest for small velocity differences);
– nonlinear interactions between carriers depend upon their mutual polarization states, but initial carrier polarizations are arbitrary and are then randomly scrambled during fiber propagation, which randomizes the interaction strength;
– tens to hundreds of wavelength carriers can be mutually interacting at any time and medium location, according to the above random conditions.

Unlike the linear OCC case, it no longer possible to exactly compensate the combined effect of nonlinearity and dispersion, since its history is the sum of random or unpredictable intra- and inter-wavelength-carrier occurrences. Reducing signal power under the nonlinearity threshold can suppress the associated noise. But this reduces the SNR, and hence increases the BER. The only way out from this dilemma is to make the OCC operate as closely as possible to linear, keep the local dispersion high (see below) while maximizing the carrier power and minimizing the resulting BER. From such a picture, it is clear that *the nonlinear-dispersive/noisy* OCC obeys to far more complex laws and optimization principles than its *linear-dispersive/noisy* counterpart previously described. The BER is still defined as in Eqs. (2) and (3), except that the evaluation and optimization of the Q-factor must rely upon sophisticated and intensive numerical computations. As mentioned in the introduction, such computations are modeled by the NLSE, in fact as many NLSEs as there are wavelength-carriers (with two polarization states), forming a highly complex nonlinear coupled-equation set (see Section 4).

The above limitations of the nonlinear-dispersive OCC led investigators to develop a variety of performance-optimization strategies. These can be listed as follows:

– reduce the signal *intensity* (power/surface) by increasing the guided-mode effective area [12];
– locally increase dispersion with periodically-opposite signs (dispersion management) [5,6,12];
– improve BER by error-correction-coding (ECC) algorithms [7];
– constructively exploit dispersion and nonlinearity through soliton-like effects [8,9].
– use in-line optical regeneration to restore signal integrity inside channel (Section 6, [11]).

The above strategies, which are all mutually compatible, have been investigated since only a decade. Only the first three are being exploited in terabit/s systems currently deployed. Pure soliton transmission (referred to as *Schrödinger solitons*) has not been up to expectations, because of additional (soliton/soliton and soliton/amplifier-noise) impairments and relatively lower performance in multi-wavelength implementation. However, dispersion management introduced new soliton-like effects, which in some cases come up very close to pure linear transmission [9]. When all the above resources have been exhausted, only the last solution of periodic in-line optical regeneration, is capable to lock the transmission quality (BER) near some asymptotic level, regardless of channel distance, as further discussed in Section 6.

2.3. Channel information capacity

In order to model the channel *information capacity*, as expressed in terms of 'number of successfully transmitted symbols per channel use', the concept of *entropy* was developed. It is beyond the scope of this paper to recall the theoretical grounds of [1], even as a brief introduction [4]. Here, we shall only use the entropy concepts which are necessary for this paper's purposes. In the process, we present an original expression linking channel capacity and BER in binary ON–OFF communications.

Entropy is a measure of *uncertainty* contained in the signal information. The higher this uncertainty, the higher the information contained in the communication. If one knows for sure the result of a communication, one does not 'learn' anything from it, no information is conveyed. It would be like a channel where all symbols received are identical (called the useless channel). The uncertainty in the choice of symbols made by the source is therefore a true characterization of the information to be transmitted (source entropy). If the channel was noiseless, the same uncertainty would characterize the signal received (receiver entropy). In case of a noisy channel, more uncertainty is introduced by the noise, but it carries no information. Here, we shall briefly introduce the definition of entropy, which leads to (maximum) channel capacity.

Assume that the source uses a discrete K-symbol alphabet X $(x_1, \ldots, x_K)$. There is no reason for the receiver to use the same alphabet with a one-to-one correspondence or even not to use a continuous-symbol alphabet. For simplicity, assume that the receiver alphabet Y is also discrete with N symbols $(y_1, \ldots, y_K)$. The simple case $K = 2$ thus corresponds to the usual binary-symbol transmission. Each symbol $z_i = x_i$ (coded by the source) or y_i (obtained at the receiver) is characterized by a finite probability $p(z_i)$. By definition, the associated source/receiver *entropy* is

$$H(Z) = -\sum_i p(z_i) \log\big[p(z_i)\big], \tag{5}$$

where the logarithm is in base two. The entropy is thus defined as the statistical mean of the function $-\log[p(z_i)]$, which is the concept introduced by L. Boltzmann in statistical mechanics. For the source, the entropy is the measure of the information conveyed. If only one symbol is used (i.e., all probabilities being zero, but one), it is easy to see from Eq. (5) that the entropy is identical to zero. It can also be easily established that if all source symbols have identical probabilities (i.e., $p(x_i) = 1/K$), the entropy is maximum, i.e., $H = \log K$. The unit of entropy is 'symbol per channel use'. For binary signals ($K = 2$) there is only one bit per (1/0) symbol, so $H = \log_2 2 = 1$ bit-per-channel-use. If the channel is used at a rate B (bits/s or s^{-1} or Hz) the channel capacity is $C = HB = B$ bit/s and per unit bandwidth $C' = C/B = 1$ bit/s/Hz. Thus, noiseless channels with ON–OFF binary sources can theoretically convey no more than 1 bit of information per cycle of carrier frequency. The bit/s/Hz figure should be referred to as the *information spectral density* (ISD), sometimes wrongly called spectral efficiency (since not bounded to unity in the general case).

Consider next the case of noisy channels. The received entropy is defined according to Eq. (5) using $z_i \equiv y_i$. As previously mentioned, this entropy should reflect extra uncertainty introduced by the channel noise, which is not information-related. Since there is no way to establish a one-to-one correspondence between transmitted/received symbols and the original source ones, we can only resort to a probabilistic approach and separate in the received entropy what is real information from what is introduced by the channel noise. To do this, one can conceive of an experiment in which the same symbol x_i is transmitted a sufficient number of times, with corresponding receiver measurements $y_1, \ldots, y_K$. Repeated for all source symbols, all this information provides us with the conditional-probability matrix $p_{ij} = p(y_j|x_i)$. Define then the *conditional entropy*:

$$H(Y|X) = -\sum_j \sum_i p(x_i)\, p(y_j|x_i) \log\bigl[p(y_j|x_i)\bigr] \tag{6}$$

which is also referred to as *equivocation*. Although its definition looks complicated, it immediately appears that equivocation is the full measure of the information-less uncertainty introduced by channel noise. This is because it measures all possible received symbol uncertainty for any source symbol knowingly input to the channel. The minimum equivocation is zero, corresponding to $p(y_j|x_i) = \delta_{ij}$ for any source symbol x_i. The above notions having been introduced, it is clear that the noisy-channel *capacity* C (in units of 'number of symbols successfully transmitted by channel use') is given by the difference between received entropy and equivocation, i.e.:

$$C = \bigl\{H(Y) - H(Y|X)\bigr\}_{\max\{X,Y\}}. \tag{7}$$

In the above definition the subscript $\max\{X, Y\}$ refers to a maximization problem which takes into account the fact that one should find a choice of alphabets X, Y that also maximize the received/equivocation entropies difference, and hence the channel capacity.

It is interesting to apply this result to a binary communication channel. As stated in Section 2.1, the minimal BER is given by setting the receiver decision threshold such that $p(1|0) \approx p(0|1)$, or after Eq. (1), $p(1|0) = p(0|1) = BER$. Using the property $p(0) = p(1) = 1/2$, it is a simple exercise to obtain the equivocation from Eq. (6):

$$H(Y|X) = -(1 - BER)\log(1 - BER) - BER\log(BER) \approx BER(1 - BER) \approx BER\,\mathrm{e}^{-BER}, \tag{8}$$

where the last two approximations assume $BER \ll 1$. From Eq. (5), using $p(y = 0) = p(y = 1) = 1/2$, the received entropy is $H(Y) = 1$. Thus, according to these results and Eq. (7), also using the fact that no further optimization is possible, the noisy/binary channel capacity at given BER is:

$$C = H(Y) - H(Y|X) = 1 + (1 - BER)\log(1 - BER) + BER\log(BER) \approx 1 - BER \approx \exp(-BER) \tag{9}$$

thus providing a direct relation between binary-channel capacity and BER (to the best of our knowledge, this simple result was never reported in the literature, nor is found in textbooks). The result shows that at low BERs, the maximum noisy-channel capacity is very close to the noiseless case, i.e., $C = 1$ bit per channel use, or $ISD = 1$ bit/s/Hz. Even for BERs as high as $BER = 10^{-3}$ (the limit of ECC possibilities to restore signal integrity), the loss of channel capacity is only 1/1000 or 0.1%. What happens in the limiting case where the BER is maximum, i.e., $BER = 5 \times 10^{-1} = 1/2$? This corresponds to the 'useless' channel situation where the uncertainty in received 1/0 bits is maximum and equal to 1/2, meaning that the same result can be obtained by turning off the receiver and flipping a coin! Replacing $BER = 1/2$ in Eq. (9), without approximation, quite nicely yields $C = 0$, meaning that the channel has indeed zero capacity, consistently with the assumptions.

The above description was meant to introduce one of the most important results of information theory, which leads to expressing the maximum channel capacity through the well-known *Shannon–Hartley theorem* (SHT). Consider now that symbol alphabets can be made continuous, using an infinity of signal levels for each symbol. The input/output symbol probabilities, as well as that associated with channel noise are also continuous. In [1], it is shown through a Lagrange-optimization method (see also [4]) that if the channel additive noise is Gaussian with variance $N = \sigma_{\mathrm{ch}}^2$, and under an input signal power constraint

$S = \sigma_{\text{in}}^2$, the optimum symbol-probability distribution is also Gaussian. The total output noise power containing both symbol information and additive channel noise is then $\sigma_{\text{out}}^2 = \sigma_{\text{in}}^2 + \sigma_{\text{ch}}^2$. The received entropy and equivocation are given by the integrals

$$H(Y) = -\int p(y)\log\big[p(y)\big]\,\mathrm{d}y \equiv \frac{1}{2}\log\big(2\pi\,\mathrm{e}\sigma_{\text{out}}^2\big) \tag{10}$$

and

$$H(Y|X) = -\iint p(x)p(y|x)\log\big[p(y|x)\big]\,\mathrm{d}x\,\mathrm{d}y \equiv \frac{1}{2}\log\big(2\pi\,\mathrm{e}\sigma_{\text{ch}}^2\big) \tag{11}$$

respectively. The maximum channel-capacity-use per second is given by subtracting the above results and multiplying by $2B$ (*Nyquist sampling rate* [4]), which yields:

$$C_{\text{bit(/channel-use)/s}} = B\big[H(Y) - H(Y|X)\big] \equiv B\log\left(1 + \frac{S}{N}\right) \equiv B\log(1 + SNR) \tag{12}$$

or

$$C_{\text{bit/s/Hz}} = \frac{C_{\text{bit/s}}}{B_{\text{Hz}}} = \log(1 + SNR). \tag{13}$$

The result in Eq. (13), known as SHT, elegantly relates the maximum achievable channel capacity (ISD) to the SNR (ratio of source signal power S over channel noise power N). This fundamental result will be used later, with refinements introduced by a proper definition of 'signal' and 'noise' powers in the case of quantum noise for periodically-amplified OCCs (Section 3), and in the case of OCCs having both quantum and nonlinearity noises (Section 4).

The only conclusion one can draw at this point is that, according the SHT, the OCC capacity could apparently be increased to any arbitrary level by two methods: (i) reduce channel noise N, and (ii) increase signal power S. The first method is intrinsically limited by the fact that after all reduction possibilities (system design, ideal source and receiver, coding, coherent detection, error correction, ...) the noise has a lower bound that is ultimately defined by quantum laws. If optical amplifiers are used to increase the OCC distance, the minimum noise also increases (Section 3). The other approach is to increase the signal power. However, nonlinearity introduces noise above some threshold (Section 4). Thus the SNR is bounded between minimum channel noise and maximum allowed channel power, which represents a fundamental design concept. This issue is further analyzed in Section 5.

3. Quantum noise in optically-amplified OCC

Optical amplifier noise can be accurately described only through quantum-physics principles, i.e., assuming a *quantized* electromagnetic field represented by photon creation/annihilation operators and interacting atoms with *quantized* energy levels. Even in the 'classical limit' of large photon numbers, where the electromagnetic field is truly a classical wave, the noise associated with electric-field amplification has a quantum-origin signature, as described in this section. The term 'quantum', however, should not be interpreted as describing effects specific to low photon-number physics [15], even if the theory developed here also applies to this regime. We consider *laser* amplifiers (such as *erbium-doped fiber amplifiers* or EDFAs), but the quantum theory also applies to *Raman fiber amplifiers* (RFA) [5,13], *semiconductor optical amplifiers* (SOA) and *non-degenerate parametric amplifiers*, also used in optical telecommunications.

3.1. Quantum beam-splitter model for linear amplifiers

From the quantum perspective, it is a well-known feature that passive attenuators (field transmission $\sqrt{T} < 1$) and ideal amplifiers (field gain $\sqrt{G} > 1$, full medium inversion) are four-port beam-splitter devices, which couple the vacuum-field into a second input port [4,15]. The vacuum-field interference is responsible for uncertainty in the output field and associated photon number. In the case of passive attenuators, this uncertainty reflects the fact that photons cannot be split between the two paths corresponding to absorption or survival. With ideal amplifiers, the uncertainty comes from the fact that photons cannot be 'exactly' multiplied by a number G, even if G is an integer, because it would violate Heisenberg's uncertainty principle [14]. Photon multiplication by stimulated emission is thus a random process (of which G is the mean result), whose details are analyzed in [16]. Another contribution to uncertainty in the amplifier's output port is *spontaneous emission*, i.e., the fact that photons can be spontaneously 'created' inside the beamsplitter. As shown in [16], *both* effects of spontaneous and stimulated emission are responsible for amplifier noise, when noise is considered not only as mean power but actual photon-number uncertainty (variance), see further below.

For the attenuator, the output photon-number expectation value (or mean) is $\langle n\rangle = T\langle n_0\rangle$, and for the amplifier it is $\langle n\rangle = G\langle n_0\rangle + N$. In these expressions, $\langle n_0\rangle$ is the mean input signal photon number and $N = G - 1$ the mean number of

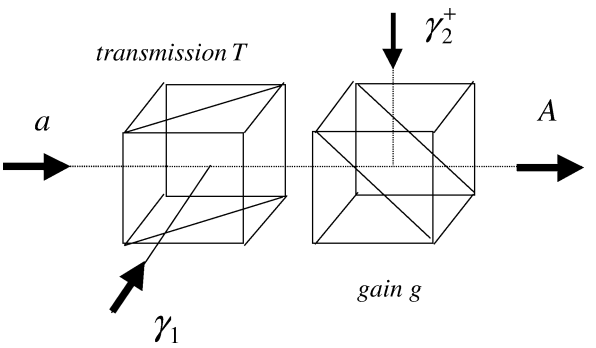

Fig. 3. Concatenation of two quantum-field beamsplitters, modelling non-ideal amplifiers with power gain $G = gT$: a passive attenuator (power transmission $T < 1$) vacuum-field coupling γ_1, is followed by an ideal amplifier (power gain $g > 1$) with vacuum-field coupling γ_2^+.

photons spontaneously produced by the amplifier in a single polarization mode, also called *amplified spontaneous emission* or ASE.

We consider the general case of *non-ideal* amplifiers, in which medium inversion is incomplete. In [17,18], we showed that any non-ideal amplifier (with net gain $G = gT$) can be described by the concatenation of a passive attenuator (transmission $T < 1$) followed by an ideal amplifier (of gain $g > 1$), as shown in Fig. 3. The system is equivalent to a *three-dimensional quantum beamsplitter* (3D-QBS), as shown in Fig. 4. In the following, we derive the *mean* and *variance* of the output photon number through this 3D-QBS model.

Let a and A be the input and output photon-annihilation operators, respectively, and γ_1, γ_2^+ the vacuum-field operators associated with the passive attenuator and the ideal amplifier and, respectively. Their input/output relation can be put under the form [4,17]:

$$A = \sqrt{G}a + \sqrt{N}\gamma_2^+ + \sqrt{P}\gamma_1, \tag{14}$$

where the relation $P = N - (G-1)$ makes possible for A to satisfy to the boson commutation rule, i.e., $[A, A^+] = [a, a^+] = 1$. Using Eq. (14), it is then possible to calculate the expectation values $\langle n\rangle = \langle A^+A\rangle$, $\langle n^2\rangle = \langle A^+AA^+A\rangle$ and the corresponding photon noise variance $\sigma^2 = \langle n^2\rangle - \langle n\rangle$. The details of the calculation are found in [4,17]. We find:

$$\langle n\rangle = G\langle n_0\rangle + N \tag{15}$$

and

$$\sigma^2 = G\big[\sigma_0^2 - \langle n_0\rangle\big] + G\langle n_0\rangle + N + 2G\langle n_0\rangle N + N^2, \tag{16}$$

where σ_0^2 is the input signal variance.

The results contained in Eqs. (15) and (16) are of great significance. Indeed, consider first the case of an ideal amplifier, which corresponds to $P = 0$ in Eq. (14). The boson commutation rule then imposes that $N = G - 1$. Looking at Eq. (15), we observe that N is the mean number of photons generated by the amplifier, also referred to as ASE. Thus, ideal amplifiers produce a minimum output ASE of $G - 1$ photons in average.

In the case of non-ideal amplifiers, we have $P = N - (G-1) > 0$ or $N > (G-1)$. One can write $N = n_{\rm sp}(G-1)$ with $n_{\rm sp} > 1$ being a 'spontaneous emission factor'. Thus, non-ideal amplifiers produce more ASE than in the ideal case, which is measured by the parameter $n_{\rm sp}$. It can be easily shown [14] that for amplifiers with uniform (coordinate-independent) population inversion $N_2 - N_1$, we have $n_{\rm sp} = N_2/(N_2 - N_1)$, which reduces to $n_{\rm sp} = 1$ in the case of the ideal, fully-inverted amplifier ($N_1 = 0$). In the general case (such as with EDFAs), there is no analytical definition for $n_{\rm sp}$, and the ASE is defined by an integral over the amplifier length. Most generally, one can write the ASE as $N = G - 1 + n_{\rm ex}G$, where $n_{\rm ex} = \int[\alpha(z)/G(z)]\,{\rm d}z$ is an 'excess noise factor' reflecting the effect of internal absorption $\alpha(z)$ in the amplifier [2,14].

Consider next the noise variance definition in Eq. (16), which shows several noise contributions. In the case of coherent input signals (Poisson statistics), we have $\sigma_0^2 = \langle n_0\rangle$ and the first *excess noise* term, $G[\sigma_0^2 - \langle n_0\rangle]$, vanishes. The second group of terms, $G\langle n_0\rangle + N$, correspond to the shot noise of the amplified signal and of the ASE, respectively. The rest of the terms are referred to as 'signal–ASE' and 'ASE–ASE beat noises', respectively [4,14]. The fact that these terms exist on top of the shot noise is an indication that amplified signal has lost some of its initial coherence, as is a well-known feature. It can be shown that contrary to common belief (from classical analysis), the signal–ASE beat noise is not an interference effect with ASE (unlike the ASE–ASE term), but is rather attributable to thermal fluctuations in stimulated emission [4,16].

The above decomposition of non-ideal amplifiers into two quantum-beamsplitter stages (loss followed by ideal gain) with independent vacuum-field couplings makes physical sense as it models the simultaneous effect of ground-level absorption

and stimulated emission by the atomic system. The order of these two operations (loss followed by ideal gain) is however fundamental [4,17]. The reverse order (ideal gain followed by loss) cannot explain non-ideal amplifiers. This is intuitive, since it is not possible to reduce an ideal amplifier (with minimal ASE) to a non-ideal device (with excess ASE) by placing a passive attenuator (which does not generate ASE) at the output. One should note that the loss/ideal-gain decomposition is not unique. Non-ideal amplifiers can also be decomposed into a discrete chain of 3D-QBS with an arbitrary number of elements, except infinity [4,18].

3.2. Langevin-operator model for linear amplifiers

The above showed that two vacuum-field couplings are required to describe non-ideal amplifiers. However, it is possible to express the amplifier input/output field relation under the alternate form [3,4]:

$$A = \sqrt{G}a + F, \tag{17}$$

where F is a *Langevin*-like quantum operator commuting with a. In this model, that F is not a Boson operator ($[A, A^+] = 1$ imposes that $[F, F^+] = 1 - G \neq 1$) and must obey the following properties: (i) $\langle F^+F\rangle = N$ and (ii) $\langle (F^+)^2(F)^2\rangle = 2!N$. The first condition sets the amplifier ASE to the correct value N, which is not predictable by the model. The second corresponds to a property of thermal/chaotic light sources. As shown in [3,4], the output mean and variance resulting from the computation of $\langle n\rangle = \langle A^+A\rangle$ and $\langle n^2\rangle = \langle A^+AA^+A\rangle$ are strictly identical to that obtained in Eqs. (15) and (16). In fact, the correspondence between the previous 3D-QBS model is easily established by setting $F = \mu\gamma_1 + \nu\gamma_2^+$, which yields the same coefficients as in the quantum-field decomposition of Eq. (14), i.e., $\mu = \sqrt{P}$ and $\nu = \sqrt{N}$. Since the 3D-QBS model only involves Boson operators, makes minimal assumptions, and self-predicts the ASE, it can be considered more fundamental than the Langevin-operator model, even if the results for noise are strictly equivalent. In Section 4, the Langevin-operator model is applied to describe WDM nonlinearity noise, this choice being justified by the need of a formalism more concise than in the 3D-QBS model.

3.3. The linear amplified optical-communications channel

Long-haul optical communications systems are based upon the principle of fiber loss compensation by in-line amplification. For a fiber span of length $L = L_{\text{tot}}/k$ (with L_{tot} = total system length, k = number of amplified spans), the transmission is $t = \exp(-\alpha L)$ where α is the fiber attenuation coefficient (typically, $\alpha = 0.20$ dB/km or $\alpha = 0.046$ km^{-1}). Each amplifier is set to a gain $g = 1/t$ so that each span (hence the whole system) is made 'transparent', i.e., $gt = (gt)^k = 1$. Overlooking nonlinearity, such a system configuration corresponds to a linear, noisy/amplified OCC picture. It is easily shown [14] that a k-span amplifier chain is physically equivalent to an single *non-ideal* amplifier having a net gain $G = 1$ and an output ASE $N' = ktN$ (if amplification precedes loss, configuration A) or $N' = kN$ (if amplification follows loss, configuration B), where $N \equiv n_{\text{sp}}(G - 1)$ is the output ASE of a single amplifier. After Eqs. (15) and (16), the mean output signal and variance of the OCC (e.g., configuration A) are then:

$$\langle n\rangle = \langle n_0\rangle + ktN \tag{18}$$

and

$$\sigma^2 = \left[\sigma_0^2 - \langle n_0\rangle\right] + \langle n_0\rangle + ktN + 2\langle n_0\rangle ktN + (ktN)^2. \tag{19}$$

The above results show that the cumulated ASE grows linearly with the number of amplified spans k, Eq. (18), and that the noise variance includes both linear and quadratic contributions. The quadratic contribution, $(ktN)^2$, is usually small since in most telecom applications the signal power dominates the cumulated ASE, such that $\langle n_0\rangle \gg kN$, making the signal/ASE beat noise, $2\langle n_0\rangle ktN$, the dominant contribution.

3.4. Signal-to-noise-ratio and channel capacity of linear amplified OCC

From the results in previous subsection, we can express the OCC output SNR as follows:

$$\begin{aligned} SNR_{\text{out}}^{(k)} &= \frac{\langle n_0\rangle^2}{\sigma^2} = \frac{\langle n_0\rangle^2}{\langle n_0\rangle + ktN + 2\langle n_0\rangle ktN + (ktN)^2} \\ &= \frac{\langle n_0\rangle}{1 + 2ktN + ktN(1 + ktN)/\langle n_0\rangle} \approx \frac{\langle n_0\rangle}{1 + 2ktN} = \frac{\langle n_0\rangle}{1 + 2kn_{\text{sp}}(1 - \mathrm{e}^{-\alpha L})}. \end{aligned} \tag{20}$$

In Eq. (20), it is assumed that the OCC input signal is coherent ($\sigma_0^2 = \langle n_0 \rangle$) and that any thermal noise from the receiver ($\sigma_{\text{th}}^2 = 4k_B T/R$) is negligible compared to the amplifier beat noise. The last two terms correspond to the high-signal approximation $\langle n_0 \rangle \gg k$. Consistently, the SNR at input is $SNR_{\text{in}} = SNR_{\text{out}}^{(0)} = \langle n_0 \rangle$, which is the minimal (shot-noise-limited) value for coherent (classical-light) signals.

It is seen from Eq. (20) that the output SNR asymptotically decays with the number of spans k, according to a factor $1/(1+\beta k)$ where $\beta = 2n_{\text{sp}}(1-\exp^{-\alpha L})$. For long amplification spans ($L \gg 1/\alpha$), and long system lengths ($k = L_{\text{tot}}/L > 10$), the decay factor becomes $1/(2kn_{\text{sp}}) = L/(2n_{\text{sp}}L_{\text{tot}})$. This result suggests that for a given OCC distance (L_{tot}) and amplifier inversion (n_{sp}) it is possible to compensate the SNR decay by boosting the input signal power by the inverse factor (i.e., $2n_{\text{sp}}L_{\text{tot}}/L$). However, such power compensation could be impractical and prohibitive, for both considerations of amplifier saturation and system nonlinearities. A correct design approach consists in keeping the average signal power between two amplifiers under some threshold value. The path-average power is given by (configuration A):

$$\langle n \rangle_{\text{path}} = \frac{1}{L}\int_0^L G\langle n_0 \rangle \, \mathrm{e}^{-\alpha z}\, \mathrm{d}z \equiv G\langle n_0 \rangle \frac{1-\mathrm{e}^{-\alpha L}}{\alpha L} \equiv G\langle n_0 \rangle \frac{1-1/G}{\log G} = \langle n_0 \rangle \frac{G-1}{\log G}. \tag{21}$$

With this new definition, we can rewrite the definition of output SNR in Eq. (20) under the form:

$$SNR_{\text{out}}^{(k)} \approx \frac{\langle n \rangle_{\text{path}} \log G}{(G-1)2kn_{\text{sp}}(1-1/G)} \equiv \frac{\langle n \rangle_{\text{path}} G \log G}{2kn_{\text{sp}}(G-1)^2} \equiv \frac{\langle n \rangle_{\text{path}} G}{2\alpha L_{\text{tot}} n_{\text{sp}}} \left(\frac{\log G}{G-1}\right)^2. \tag{22}$$

The generic result obtained in Eq. (22), shows that for amplified-OCCs with given length (L_{tot}), amplifier/loss parameters (n_{sp}, α) and power constraint ($\langle n_{\text{path}} \rangle$), the SNR is bound to the gain-dependent figure of merit $f(G) = G[\log G/(G-1)]^2$ [19]. This figure is minimal ($f = 1$) for $G \to 1$, corresponding to *distributed amplification*. In practice, optical amplifiers and related components introduce excess loss, and the ideal case of distributed amplification can only be approached.

In the following, we shall derive an original expression for channel capacity in ideal systems.

According to the above, the minimal SNR is given by $f = n_{\text{sp}} = 1$ (ideal distributed amplification), i.e., $SNR_{\text{min}} = \langle n \rangle_{\text{path}}/(2\alpha L_{\text{tot}}) \equiv SNR_{\text{in}}/(2\alpha L_{\text{tot}})$, where we use the fact that, in distributed amplification, $\langle n \rangle_{\text{path}} = \langle n_0 \rangle$. Thus, the minimum SNR decay (or penalty) experienced by signals is $\eta = 1/(2\alpha L_{\text{tot}})$. For an OCC with $L_{\text{tot}} = 1000$ to 10 000 km (or $\alpha = 0.046\ \text{km}^{-1}$) we have $\eta = 1.0 \times 10^{-2}$ to 1.0×10^{-3} (or -20 dB to -30 dB).

Let briefly analyze the consequence of the above SNR limit on the OCC capacity. Consider first binary ON–OFF signals. In Section 2.1, the SNR definition corresponds to the ratio of mean 1/0 bit power to un-polarized ASE, which is 1/4 the value defined by SNR_{min}. However, with this alternate SNR definition, the decay factor is the same. We can thus express from Eq. (3):

$$Q_{\text{min}} = \sqrt{\frac{2SNR}{2\alpha n_{\text{sp}} L_{\text{tot}}}} \equiv \frac{Q_{\text{in}}}{\sqrt{2\alpha n_{\text{sp}} L_{\text{tot}}}}, \tag{23}$$

where Q_{in} is the Q-factor of the OCC source signal. If Q_{in} is chosen high enough so that Q_{min} falls in the range $Q_{\text{min}} = 3$–6 ($BER = 10^{-3}$–10^{-9}, after Eq. (2)), then the impact on OCC capacity ($C \approx 1 - BER$, Eq. (9)) is fully negligible.

The case of continuous-channel for linear OCC capacity is given by the SHT in Eq. (13), together with the general result in Eq. (22), i.e.,

$$C_{\text{bit/s/Hz}} = \log_2(1 + SNR_{\text{out}}) \approx \log_2\left[1 + \frac{\langle n \rangle_{\text{path}} f(G)}{2\alpha L_{\text{tot}} n_{\text{sp}}}\right]. \tag{24}$$

In the ideal case ($f(G) = n_{\text{sp}} = 1$), we have

$$C_{\text{bit/s/Hz}}(\text{ideal}) \approx \log_2\left(1 + \frac{\langle n_0 \rangle}{2\alpha L_{\text{tot}}}\right). \tag{25}$$

Taking OCC lengths of $L_{\text{tot}} = 1000$–10 000 km as representative examples for terrestrial and undersea systems, we have $\eta = 1/(2\alpha L_{\text{tot}}) = 10^{-2}$–$10^{-3}$, corresponding to minimum capacities of $C_1 = \log_2(\langle n_0 \rangle/100)$ to $C_2 = \log_2(\langle n_0 \rangle/1000)$. To provide a practical example, assume that OCC nonlinearity limits the path-average power to $\langle n_0 \rangle = 10^6$ photons (corresponding to 1.3 mW in 10 GHz bandwidth at 1.55 μm wavelength). The corresponding capacities are $C_1 = 13.2$ bit/s/Hz and $C_2 = 9.9$ bit/s/Hz, respectively. Taking the EDFA bandwidth as being approximately $B = 32$ nm, or 4 THz (1 nm = 125 GHz), the number of wavelength channels would be 4 THz/10 GHz = 400, and the corresponding system bit-rates would be $BC_1 = 53$ Tbit/s ($L_{\text{tot}} = 1000$ km) and $BC_2 = 40$ Tbit/s ($L_{\text{tot}} = 10\,000$ km). For both capacity and bit-rates, these figures are between one or two orders of magnitude above those already achieved experimentally with ON–OFF binary keying [5,6]. Channel nonlinearity has, however, been overlooked in the analysis, and moreover the amplified-OCC characteristics have been assumed ideal. Finally, we note that one would be allowed a ten-fold signal power increase ($\langle n_0 \rangle = 10^7$ or 13 mW) this would only result in a 'marginal' capacity improvement of $\log_2(10) = 3.3$ additional bit/s/Hz, i.e., $C_1 = 16.5$ bit/s/Hz and $C_2 = 13.2$ bit/s/Hz.

4. The nonlinear (amplified) OCC

The previous section concerned the analysis of a linear OCC where channel noise is exclusively due to in-line optical amplification. In this section, we consider the case of a *nonlinear* OCC, where nonlinearity is a cause of channel noise. Since nonlinear transmission systems generally include in-line optical amplification, it is sensible to directly analyze the combination of both noise types, i.e., of amplifier (quantum) origin and of nonlinearity origin, rather than first considering a nonlinear OCC without loss, or with loss but without amplification. These last two academic cases are however analyzed in detail in [3].

The most general OCC picture is that shown in Fig. 2: we assume a common 'external' noise source, due to in-line amplification, and a multiplicity of 'internal' noise sources, due to intra- and inter-channel WDM nonlinearity. In the following subsection, we shall first review the origin of such a nonlinearity, then proceed with the model, first as a classical description, then as a unified quantum description.

4.1. Origin of OCC nonlinearity

It is worth recalling the physical origin of nonlinearity, in particular in WDM optical transmission systems. As a well-known feature, the electric field E associated with a light wave induces a vector polarization P in the dielectric medium, which is proportional to the field intensity, i.e., $P = \chi^{(1)}E$. In turn, this linear polarization acts as a field source which slows down the wave propagation (the refractive index being defined as $n = \sqrt{1+4\pi\chi^{(1)}} \geqslant 1$, unity standing for vacuum). Under high electric-field magnitudes, as is the case in optical fibers, the medium polarization becomes nonlinear, following a complex vector expansion of the form $P = \chi^{(1)}E + \chi^{(2)}E:E + \chi^{(3)}E:E:E + \cdots$, where colons (:) indicates multiple-order tensorial products. Unless specially prepared, the usual glass fibers do not have preferential symmetry, which causes second-order effects to vanish (i.e., $\chi^{(2)}E:E \approx 0$). Thus, nonlinear polarization in fibers are of third-order, i.e., of the form $P^{\mathrm{NL}} = \chi^{(3)}E:E:E$. Assume next that the electric field E is a superposition of different fields E_k ($k = 1, \ldots, N$), each having optical carrier frequencies, i.e., $E = \sum_k E_k(\omega_k)$, as in a WDM system. Each carrier field can be expressed under the real form $E_k = u_k\,\mathrm{e}^{\mathrm{i}\Phi_k} + u_k^*\,\mathrm{e}^{-\mathrm{i}\Phi_k}$, where u_k is the spatial (longitudinal and transverse) amplitude and $\Phi_k = \omega_k t + \varphi_k$ the phase. If one reduces the analysis to a polarization mode, the nonlinear polarization takes the general scalar form:

$$\begin{aligned} P^{\mathrm{NL}} &= \sum_{j,k,l} \chi^{(3)}_{(\omega_j,\omega_k,\omega_l)} E_j E_k E_l \equiv \chi^{(3)} \sum_{j,k,l} \chi^{(3)}_{(\omega_j,\omega_k,\omega_l)} \big(u_j\,\mathrm{e}^{\mathrm{i}\Phi_j} + u_j^*\,\mathrm{e}^{-\mathrm{i}\Phi_j}\big)\big(u_k\,\mathrm{e}^{\mathrm{i}\Phi_k} + u_k^*\,\mathrm{e}^{-\mathrm{i}\Phi_k}\big)\big(u_l\,\mathrm{e}^{\mathrm{i}\Phi_l} + u_l^*\,\mathrm{e}^{-\mathrm{i}\Phi_l}\big) \\ &\equiv \chi^{(3)} \sum_{j,k,l} \big(u_j\,\mathrm{e}^{\mathrm{i}\Phi_j} + u_j^*\,\mathrm{e}^{-\mathrm{i}\Phi_j}\big)\big(u_k\,\mathrm{e}^{\mathrm{i}\Phi_k} + u_k^*\,\mathrm{e}^{-\mathrm{i}\Phi_k}\big)\big(u_l\,\mathrm{e}^{\mathrm{i}\Phi_l} + u_l^*\,\mathrm{e}^{-\mathrm{i}\Phi_l}\big) \\ &\equiv \chi^{(3)} \sum_{j,k,l} \big(u_j u_k u_l\,\mathrm{e}^{\mathrm{i}\Phi_{jkl}} + u_j u_k u_l^*\,\mathrm{e}^{\mathrm{i}\Phi_{jk-l}} + u_j u_k^* u_l\,\mathrm{e}^{\mathrm{i}\Phi_{j-kl}} + u_j u_k^* u_l^*\,\mathrm{e}^{\mathrm{i}\Phi_{j-k-l}} + \mathrm{cc}\big), \end{aligned} \tag{26}$$

where $\Phi_{k\pm l\pm m} = \Phi_k \pm \Phi_l \pm \Phi_m$, and 'cc' means the complex conjugate of all previous terms. For simplicity, it is assumed that the third-order susceptibility, $\chi^{(3)}$, is independent of the frequency combinations and exhibits no associated resonances.

The above equation illustrates how complex third-order nonlinearity is, even in this simplified scalar case. Each of the terms $u_j u_k^{(*)} u_l^{(*)}$ represent *field-mixing products* with associate phase mismatchs $\Phi_{k\pm l\pm m}$ and corresponding oscillating frequencies $\omega_{k\pm l\pm m} = \omega_k \pm \omega_l \pm \omega_m$. Considering only two interacting fields ($m = 1, 2$ with $j, k, l = m$) here exists as many as 32 mixing products in the above development. With N fields, the number of mixing products is $4N^3$, representing for instance about 10^6 terms for 64 WDM channels. Fortunately, most of the terms are either oscillating outside the frequency bandwidth of interest (e.g., third harmonics ω_{3j}, ω_{3k}, ω_{3l} and frequency sums ω_{2j+k}, ω_{2k+l}, ω_{2l+j}) and/or *phase-mismatched*, meaning that the nonlinear polarization is oscillating too rapidly for any substantial effect to build up and coherently interfere with any signal.

Three categories of mixing products of interest remain. The first two are of the type $u_j|u_j|^2$ and $u_j|u_k|^2$ ($k \neq j$), which oscillate at frequency ω_j. They correspond to the effects of *self-phase modulation* (SPM) and *cross-phase modulation* (XPM), respectively. For the carrier at ω_j, SPM and XPM correspond to nonlinear refractive-index changes of the form $n = n_0 + n_2|u_{j,k}|^2$. The third category of mixing products is of the type $u_j u_j u_k^*$ (or cc), which oscillates at frequency $2\omega_j - \omega_k$ (or $2\omega_k - \omega_j$). If the carriers are equally spaced, i.e., $\omega_k = \omega_j + \Delta\omega$ (and so on), it is simple to establish that the resulting frequencies match those of the interacting carriers. If the channel spacing is narrow, the dispersion does not significantly vary over the frequency interval, and *phase-matching* occurs, corresponding to constructive/destructive interference and nonlinear energy transfer between channels. Thus, two interacting channels (at ω_1, ω_2) generate mixing products in the two adjacent ones (at $\omega_1 - \Delta\omega$, $\omega_2 + \Delta\omega$), which is a source of nonlinear noise for all four channel involved (nonlinear loss for ω_1, ω_2, nonlinear gain for $\omega_1 - \Delta\omega$, $\omega_2 + \Delta\omega$, or the reverse, depending upon phase-matching conditions). This nonlinear effect is known as *four-wave mixing* (FWM).

The starting point for analyzing FWM in WDM transmission systems is the *nonlinear Schrödinger equation* (NLSE) [8,9, 14]. Considering lossless, dispersive media, the most famous analytical solution of the NLSE is the *Schrödinger soliton*, a light pulse for which nonlinearity (pulse chrip due to SPM) and positive dispersion (pulse broadening) exactly balance each other, hence generating a stationary, particle-like pulse, with all types of periodic behaviors [9]. In general, the NLSE is the most comprehensive propagation equation for any type of nonlinear/dispersive medium.

For a given carrier at ω_j, the field-amplitude evolution over time and longitudinal coordinate (z) is a solution of the general NLSE:

$$\frac{\partial u_j}{\partial z} - \beta_j \frac{\partial u_j}{\partial t} - \frac{\mathrm{i}}{2}\frac{\partial^2 u_j}{\partial t^2} = \mathrm{i}\left\{|u_j|^2 + 2\sum_{k\neq j}|u_k|^2\right\}u_j + \mathrm{i}\left\{u_j^* \sum_{k\neq j} u_k^2\, \mathrm{e}^{\mathrm{i}\phi_{2k-j}} + u_j^2 \sum_{k\neq j} u_k^*\, \mathrm{e}^{\mathrm{i}\phi_{2j-k}}\right\} + f(u_j). \tag{27}$$

In Eq. (27), the left-hand side models pulse propagation in a linear/dispersive medium (first-order dispersion β_j), under a slowly-varying envelope approximation. The RHS include different perturbative terms. These correspond to the effects of SPM and XPM (first braces), then of FWM (second braces), then of periodic amplification/attenuation (function $f(u_j)$, e.g., $f(u_j) = [g(z)/2 - \alpha/2]u_j$, $g =$ gain coefficient), respectively. The above forms a generic background from which a classical analysis of nonlinear noise in WDM systems was developed, as described next.

4.2. Classical analysis of nonlinear OCC

The previous subsection was sufficient to illustrate that, under conditions of high channel power, low dispersion and long interaction lengths, FWM acts as a source of nonlinearity noise in the OCC (SPM, XPM, dispersion, periodic amplification and channel coding contributing to further randomization, as discussed in Section 2.2). To model FWM noise, the analysis in [2] assumes that the output carrier field is of the form:

$$u_j^{\mathrm{out}} = u_j + \beta \sum_k h_{jk} u_k + \varepsilon_j, \tag{28}$$

where h_{jk} is a scattering matrix, β is a nonlinear strength coefficient and ε is a random-noise field associated with amplification. Other assumptions are $\langle h_{jk}\rangle = \langle \varepsilon_j\rangle = 0$, $\langle h_{ij}h_{kl}\rangle = \delta_{ij}\delta_{kl}/q$, and $\langle \varepsilon_i \varepsilon_j\rangle = \delta_{ij}N/q$ where q is the number of WDM carriers and N is the total ASE power. Using these assumptions, it is straightforward to obtain the output power variance $\sigma^2_{j,\mathrm{out}} = \langle u_j^{\mathrm{out}} - \langle u_j^{\mathrm{out}}\rangle^2\rangle$, which can be written:

$$\sigma^2_{j,\mathrm{out}} = |u_j|^2 + \frac{\beta^2 S + N}{q}, \tag{29}$$

with $S = \sum_k |u_k|^2$ being the total WDM power. Summing Eq. (29) over all the channels 'j' yields the total OCC output noise $\sigma^2_{\mathrm{out}} = S + \beta^2 S + N$. In this result, the extra contribution $\beta^2 S$ and N correspond to FWM and ASE noises, respectively. Since $\beta^2 S$ corresponds to the power scattered by nonlinearity, power conservation requires that we reformulate the output OCC noise as $\sigma^2_{\mathrm{out}} = S(1-\beta^2) + \beta^2 S + N$, where the actual information-carrying signal power is $\sigma^2_{\mathrm{out,signal}} = S(1-\beta^2)$.

The nonlinear OCC capacity can now be defined according to the SHT, i.e., following Eq. (13):

$$C_{\mathrm{bit/s/Hz}} = \log\left(1 + \frac{S(1-\beta^2)}{N + \beta^2 S}\right). \tag{30}$$

The second step of the analysis consists in relating the nonlinear strength β^2 to actual WDM system parameters. The reader can refer to [2–4] for the essential aspects of this complex derivation, of which we shall directly provide the result.

Because increased signal power means increased scattering and signal loss, it is sensible that β^2 be also power-dependent, and rapidly converge to unity. The information power $S(1-\beta^2)$ would thus vanish for $S \to \infty$, and so would the channel capacity in Eq. (30). Therefore, we define:

$$\beta^2 = \left(\frac{P_S}{P_{\mathrm{th}}}\right)^2 \approx 1 - \exp\left[-\left(\frac{P_S}{P_{\mathrm{th}}}\right)^2\right], \tag{31}$$

where P_S is the OCC signal power and P_{th} its nonlinearity threshold. In customary telecom units, the threshold is defined by [2–4]:

$$P_{\mathrm{th}}(W) = 10^{-1.5}\sqrt{\frac{B_{\mathrm{GHz}}|D|_{\mathrm{ps/nm\cdot km}}\Delta\lambda_{\mathrm{nm}}}{2\gamma^2_{\mathrm{W^{-1}\cdot km^{-1}}} k L^{\mathrm{eff}}_{\mathrm{km^{-1}}} \log(q/2)}}. \tag{32}$$

In this definition, B is the OCC signal bandwidth (in GHz), $|D|$ is the local absolute dispersion (in ps/nm/km), $\Delta\lambda$ is the WDM carrier-wavelength spacing (in nm), γ the nonlinear index coefficent (in W^{-1}/km), k is the number of amplification spans, $L^{\text{eff}} = 1/\alpha$ the effective nonlinear length and $q > 2$ is the number of WDM channels. Replacing definition in Eq. (31) into that in Eq. (30) yields:

$$C_{\text{bit/s/Hz}} = \log\left(1 + \frac{P_S\, e^{-(P_S/P_{\text{th}})^2}}{P_N(1 + (P_S/P_N)[1 - e^{-(P_S/P_{\text{th}})^2}])}\right), \tag{33}$$

where P_N is the cumulated ASE noise power in OCC bandwidth B. Since the capacity is finite at low powers ($P_S \ll P_{\text{th}}$) but vanishes at high powers ($P_S \to \infty$), it must pass through a maximum at some optimum signal power P_S^{opt}. Such an optimum is easily found by approximating the exponentials in Eq. (33), which yields (as closely approximate values):

$$P_S^{\text{opt}} = \left(\frac{P_N P_{\text{th}}^2}{2}\right)^{1/3} \tag{34}$$

corresponding to the maximum capacity:

$$C_{\text{bit/s/Hz}}^{\max} \approx \frac{2}{3}\log_2\left(\frac{2}{3\sqrt{3}}\frac{P_{\text{th}}}{P_N}\right). \tag{35}$$

Note that since all powers involved in the definitions of Eqs. (33)–(35) are expressed as dimensionless ratios P_S/P_N or P_S/P_{th}, one can also use *per-carrier* powers in bandwidth $B_c = B/q$, including this change also in definition of Eq. (32).

We illustrate the above results through a practical system example. Consider for instance a an OCC with $L_{\text{tot}} = 5 \times 100$ km ($k = 5$) with $q = 100$ channels at $\lambda = 1.55$ µm, $B_c = 10$ GHz, $|D| = 1$ ps/nm/km, $\Delta\lambda = 0.2$ nm, $\gamma = 1$ W^1/km, and $L^{\text{eff}} = 20$ km, which from Eq. (32) yields $P_{\text{th}} = 1.6$ mW. The cumulated ASE noise is assumed to be $P_N = 2kn_{\text{sp}}(G-1)h\nu B' = 5$ µW ($B' = 0.2$ nm·125 GHz/nm = 25 GHz, $G = 100$, $n_{\text{sp}} = 1.6$). The optimum signal power is then $P_S^{\text{opt}} = 0.185$ mW (−7 dBm), corresponding to a peak capacity of $C_{\max} = 4.6$ bit/s/Hz. Fig. 5 shows plots of the amplified/nonlinear OCC as function of the signal power (in dBm or decibel-mW, as defined by the decimal log scale $10\log_{10}[P_S^{\text{mW}}/1\text{ mW}]$). The other curves are obtained in the same conditions, except that the local dispersion is increased by powers of two ($|D| = 2$–4–8 ps/nm/km), corresponding to power thresholds of $P_{\text{th}} = 3.2$–4.8–6.4 mW, respectively. The straight lines correspond to the purely linear case ($P_{\text{th}} \to \infty$) and the purely nonlinear case ($P_N \to 0$), respectively. It is seen that a capacity maximum (as predicted) is defined near the cross-point of the purely linear or nonlinear regimes. It increases as the power threshold is increased, but only logarithmically, according to Eq. (35). A doubling of the threshold only provides a 1 bit/s/Hz channel capacity improvement.

The classical theory of amplified/nonlinear OCC thus determines a *power window* through which maximum OCC capacity can be achieved. The peak channel capacity performance is determined by a variety of key WDM system parameters (local dispersion, bandwidth, carrier-wavelength spacing, nonlinear coefficient, amplifier spacing, number of spans, and number of channels), all of which being nicely contained in a closed-form definition of nonlinear threshold power (Eq. (32)).

4.3. Unified quantum model for amplified/nonlinear OCC

As we have seen the noise characteristics, and hence the capacity of the nonlinear/amplified OCC can be fully described by the combination of two models: a quantum-field operator model for in-line amplification (Section 3) and a classical-field model for nonlinear wave-mixing (previous subsection). The second model also introduces linear amplifier noise in a semi-classical way, i.e., by assuming ASE as an ad-hoc parameter. A main assumption consists in introducing ASE noise in the total OCC noise variance only under the form of mean power N, as seen in Eq. (29), which is equivalent to a shot noise effect. As described in Section 3, the noise in amplified light signals is dominated by the signal–ASE beat component, which is not additive and is proportional to the signal power. Another issue is the effect of signal–ASE and ASE–ASE wave-mixing, overlooked in the previous analysis. It is sensible therefore to refine the analysis in order to obtain a noise expression where these different effects appear. It also makes sense to unify the two analysis of amplification and nonlinearity into a single quantum model. This does not imply that the effect of nonlinear FWM must be of quantum origin or could be only satisfactorily explained by quantum principles. Rather, it is the matter of modeling FWM noise through quantum formalism so that it is fully compatible with the quantum model for amplifiers. In the following, we briefly describe the background assumptions and results the unified quantum model for amplification and nonlinearity, originally developed in [3].

Consider nonlinear transmission without amplification in a short, lossless OCC, meaning that the channel transparent in the linear/low-power regime. An assumption central to the model, is that the OCC can be characterized by the same evolution equation as in Eq. (17), i.e.,

$$A = \eta a + H, \tag{36}$$

where $\eta \leqslant 1$ is the nonlinear transmission and H a Langevin-like field operator associated with nonlinearity, with properties being identical to that of F (Eq. (17)), and with mean noise power $\langle H^+ H\rangle = P$ (this term P not to be confused with that used in Section 3.1). The same formal developments as in Section 3.2 lead to the definition of mean signal output power:

$$\langle n\rangle = \eta^2\langle n_0\rangle + P. \tag{37}$$

Since there is no external power source, energy conservation requires from Eq. (37) that

$$P = \big(1-\eta^2\big)\langle n_0\rangle \equiv \chi\langle n_0\rangle, \tag{38}$$

where χ is defined as a 'nonlinear scattering' or NLS coefficient. Thus, the action of the nonlinear segment is convert a fraction of input signal energy into output noise energy. Following the same approach as in Section 3.1, we can decompose the noise operator H into:

$$H = \mu\gamma_1 + \nu\gamma_2^+, \tag{39}$$

where γ_1, γ_2 are boson fields from two independent or uncorrelated noise sources, with zero power $\langle \gamma_i^+\gamma_i\rangle = 0$. According to the same commutation and bracketing conditions as used for F and A in the amplifier model, one obtains:

$$\mu = \sqrt{1-\eta^2+P} = \sqrt{\chi\big(1+\langle n_0\rangle\big)}, \qquad \nu = \sqrt{P} = \sqrt{\chi\langle n_0\rangle}. \tag{40}$$

Substituting the results of Eqs. (39) and (40) into Eq. (36) we get:

$$A = \sqrt{1-\chi}\,a + \sqrt{\chi\big(1+\langle n_0\rangle\big)}\gamma_1 + \sqrt{\chi\langle n_0\rangle}\gamma_2^+ \tag{41}$$

which fully defines the quantum-operator properties of the nonlinear lossless OCC. Thus, the nonlinear/lossless OCC can be modeled by a 3D-QBS (Fig. 4), which couples two vacuum-fields to the transmitted signal, similarly to the amplification case. Because two vacuum fields are necessary to model the nonlinear transmission, the effect is not equivalent to passive attenuation, which requires only one vacuum field (Section 3.1).

One then notices that the input/output relation in Eq. (41) has the same functional form as that of the amplifier in Eq. (14), according to the substitution:

$$G \to 1-\chi, \qquad N \to \chi\langle n_0\rangle. \tag{42}$$

Therefore, there is no need to carry out again the tedious derivation leading to the output moment $\langle (b^+b)^2\rangle$ and associated variance σ^2. Indeed, suffices it to make the substitution defined in Eq. (42) into the result of Eq. (16), which yields:

$$\sigma^2 = \big\{(1-\chi)^2\big(\sigma_0^2 - \langle n_0\rangle\big)\big\} + \langle n_0\rangle + \chi(2-\chi)\langle n_0\rangle^2 \tag{43}$$

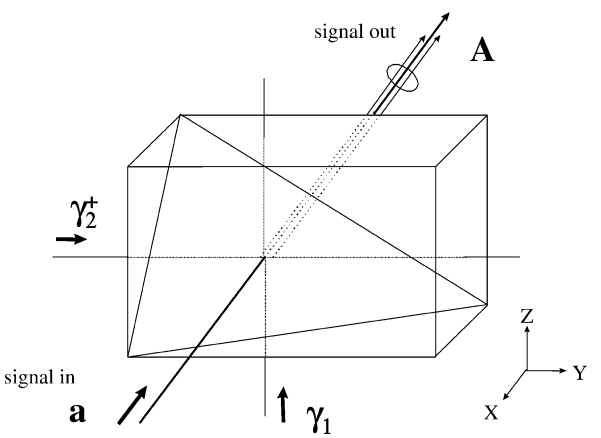

Fig. 4. Three-dimensional quantum beam-splitter model for non-ideal amplifiers, equivalent to that of Fig. 3 (see text for coupling-coefficient definitions).

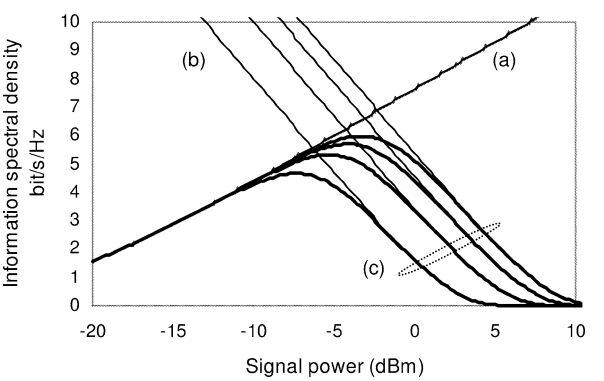

Fig. 5. Amplified/nonlinear optical communication channel capacity (information spectral density) as function of total signal power, corresponding to the example of a 5×100 km WDM transmission system: (a) linear amplified case, (b) nonlinear case with no amplifier, (c) general case, in full line. The system dispersion is increased from left to right by powers of two from an initial value of $D = 1$ ps/nm·km. See text for other parameters. After [4], © Wiley, 2002.

or for coherent input signals ($\sigma_0^2 = \langle n_0 \rangle$):

$$\sigma^2 = \langle n_0 \rangle + \chi(2-\chi)\langle n_0 \rangle^2. \tag{44}$$

In the result of Eq. (44), the first term in the RHS is identified as a combined shot noise contribution from both signal ($(1-\chi)\langle n_0 \rangle$) and NLS ($\chi\langle n_0 \rangle$) sources. The second term corresponds to the sum of two beat noises, namely signal–NLS ($2\chi(1-\chi)\langle n_0 \rangle^2$) and NLS–NLS ($\chi^2\langle n_0 \rangle^2$) beat noises, respectively.

Since nonlinearity can be modeled with this new Langevin source H, we can now proceed to combine it with the effects of loss and amplification. This can be done for instance by considering the concatenation of an amplifier with a lossy nonlinear segment, with the gain compensating the segment loss, as described in [3]. While the resulting derivation is tedious, it leads to a result that can be directly obtained by assuming at once the following input/output field relation:

$$A = \sqrt{\lambda}a + F + H. \tag{45}$$

In this relation, the two Langevin-like operators, F for the amplifier and H for the nonlinearity are characterized as follows:

- $\langle F^{(+)} \rangle$, $[a^{(+)}, F^{(+)}]$ and $\langle a^{(+)} F^{(+)} \rangle$ are all zero (parenthesis in exponent indicating optional Hermitian conjugation), and $\langle F^+ F \rangle = N$, $\langle (F^+)^2 (F)^2 \rangle = 2N^2$ (property of thermal/chaotic sources) and $\langle (FFF^+)^{(+)} \rangle = 0$, where N is the ASE;
- H is assumed to have the same properties as F, except for a mean power given by $P = \langle H^+ H \rangle = \chi\langle n_0 \rangle$, where $\langle n_0 \rangle = \langle a^+ a \rangle$ is the mean throughput signal power and χ a nonlinear scattering parameter which will be specified later. Thus we have $\langle (F^+)^2 (F)^2 \rangle = 2\chi^2 \langle n_0 \rangle^2$, meaning that nonlinearity noise is thermal/chaotic, corresponding to maximum randomness. Since the ASE and NLS processes are independent, the quantities $[F^{(+)}, H^{(+)}]$, $\langle H^{(+)} F^{(+)} \rangle$, $\langle H^{(+)} F^{(+)} F^{(+)} \rangle$ and $\langle F^{(+)} H^{(+)} H^{(+)} \rangle$ are all identical to zero.

The mean output signal power, as calculated from Eq. (45) and applying the above properties, takes the form:

$$\begin{aligned} \langle n \rangle = \langle A^+ A \rangle &= \left\langle (\sqrt{\lambda}a^+ + F^+ + H^+)(\sqrt{\lambda}a + F + H) \right\rangle \equiv \lambda\langle a^+ a \rangle + \langle F^+ F \rangle + \langle H^+ H \rangle \\ &= \lambda\langle n_0 \rangle + N + P \end{aligned} \tag{46}$$

which is the expected result. An important condition is that A also describes a boson field, i.e., $[A, A^+] = [a, a^+] = 1$. Calculation from Eq. (45) yields $[F, F^+] + [H, H^+] = 1 - \lambda$, which is useful in the computation of the second moment $\langle (A^+ A)^2 \rangle$. Using Eq. (45) we obtain first:

$$\left\langle (A^+ A)^2 \right\rangle = \left\langle (A^+ A)(A^+ A) \right\rangle = \left\langle (\sqrt{\lambda}a^+ + F^+ + H^+)(\sqrt{\lambda}a + F + H)(\sqrt{\lambda}a^+ + F^+ + H^+)(\sqrt{\lambda}a + F + H) \right\rangle \tag{47}$$

which decomposes into no less than 81 bracket terms! It is then a patient exercise to identify the non-vanishing brackets and regrouping them, then to apply the above commuting rules and properties [3,14]. However, such a task is rewarded by a nicely tractable result:

$$\sigma^2 = \lambda^2\left(\sigma_0^2 - \langle n_0 \rangle\right) + \lambda\langle n_0 \rangle\left[1 + 2(N+P)\right] + (N+P)\left[1 + (N+P)\right]. \tag{48}$$

The first term in the RHS of Eq. (48) is the usual excess noise, which vanishes for coherent input signals. The other terms in the RHS exactly correspond to those of Eq. (16) with the substitution $N \to N + P$, which serves as a calculation proof. One can also group the noise terms from Eq. (48) as follows:

$$\sigma^2 = \sigma_{\text{amp}}^2 + \sigma_{\text{NL}}^2, \tag{49}$$

$$\sigma_{\text{amp}}^2 = \lambda\langle n_0 \rangle(1 + 2N) + N(N+1), \tag{50}$$

$$\sigma_{\text{NL}}^2 = \lambda\langle n_0 \rangle 2P + P(1 + 2N + P). \tag{51}$$

One recognizes in Eq. (50) the noise from optical amplification (shot and beat noises) in single ASE polarization (generalization to two polarizations can be done by doubling the last term in $N(N+1)$, but this is not necessary as this contribution remains negligible compared to the other beat-noises). The second noise source, Eq. (51) decomposes into four terms. The first, $(2\lambda\langle n_0 \rangle P)$, is related to an effect of *signal–NLS beat noise*. The remaining ones (P, $2PN$ and P^2) correspond to *NLS shot noise*, *NLS–ASE* and *NLS–NLS beat noise*, respectively. Classically, these four noise terms make physical sense if one conceives NLS as an internal OCC source generating an electric field incoherent with both ASE and amplified signal.

5. Nonlinear/amplified OCC capacity according to unified quantum model

Having obtained the expression for the nonlinear/amplified OCC mean and variance, we can define the SNR:

$$SNR = \frac{\lambda^2 \langle n_0 \rangle^2}{\sigma_{\text{amp}}^2 + \sigma_{\text{NL}}^2}. \tag{52}$$

Replacing into this SNR the definitions in Eqs. (50) and (51) yields:

$$SNR = \frac{\lambda \langle n_0 \rangle}{1 + 2N + N(N+1)/(\lambda \langle n_0 \rangle) + 2P + P((1+2N)/(\lambda \langle n_0 \rangle) + P/(\lambda \langle n_0 \rangle))}. \tag{53}$$

Using the high-signal approximation ($\langle n_0 \rangle \gg 1,\ N,\ N^2$) and substituting $P = \chi \langle n_0 \rangle$:

$$SNR \approx \frac{(1-\chi)\langle n_0 \rangle}{N' + \chi(1 + 1/(1-\chi))\langle n_0 \rangle}. \tag{54}$$

Finally, we can express the OCC capacity, $C = \log(1 + SNR)$, introducing the notations $\langle n_0 \rangle \equiv S$, $\chi \equiv \beta^2$, which gives:

$$C_{\text{bit/s/Hz}} = \log\left[1 + \frac{S(1-\beta^2)}{1 + 2N + \beta^2 S(1 + 1/(1-\beta^2))}\right]. \tag{55}$$

Comparison with result from the classical theory (Section 4.2, Eq. (30)) shows that the new capacity definition in Eq. (55) is very similar. There are only three differences in the SNR denominator:

- the term '1', which traces back to signal shot noise as a non-vanishing contribution when amplification and nonlinearity are turned off ($N = 0$, $\beta^2 = 0$);
- the term $2N$, which traces back to signal–ASE beat noise, thus introducing a factor of '2' penalty with respect to the classical model (where only N appears);
- the extra term $1/(1-\beta^2)$, which traces back to the effects of signal–NLS and NLS–NLS beating effects, not taken into count in the classical model.

The impact of these three corrective terms turns out to be small. Indeed, the maximum SNR 'penalty' introduced by the definition involved in Eq. (55), with respect to that involved in Eq. (30) corresponds to only 20% or 1 dB [20]. The maximum SNR/capacity is also independent of nonlinearity threshold and noise powers. This can be showed by considering the coordinates of the new optimum point (easily obtained from Eq. (55), after substituting the definition of β^2 in Eq. (31)):

$$P_S^{\text{opt}} = \left(\frac{P_{N'} P_{\text{th}}^2}{4}\right)^{1/3}, \tag{56}$$

$$C_{\text{bit/s/Hz}}^{\text{max}} \approx \log_2\left[\frac{2}{3}\left(\frac{P_{\text{th}}}{2P_{N'}}\right)^{2/3}\right] \tag{57}$$

with $P_{N'} = (1 + 2N)h\nu B$ (or $P_{N'} = (1 + 2ktN)h\nu B$ for cumulated ASE in configuration A). Note that for consistency, the comparison between the two SNR models [20] must assume the same *linear* noise $P_{N'}$, i.e., making the substitution $P_N \to P_{N'}$ in Eq. (33). It is easily checked that the OCC capacity predicted by the unified quantum model (Eq. (57)) corresponds to 0.33 bit/s/Hz (or a 'maximum SNR' penalty of 20%) with respect to the classical model (Eq. (35)). The unified quantum model for the amplified/nonlinear OCC thus introduces a relatively small correction in both SNR and capacity.

6. All-optical regeneration

Several strategies are possible to optimize the performance of amplified/nonlinear OCCs, as discussed in previous Section 2.6. These include system-design optimization (use of special transmission fibers, dispersion management, reduction of amplifier span, power pre-emphasis, ...), and the use of optimized/robust modulation formats with error correction coding/decoding (see [5–7]). When all these approaches have exhausted their potential to either increase the transmission distance (at fixed bit-rate), or increase the bit-rate (at fixed transmission distance), a final recourse is to regenerate the signals by use of opto-electronic (OE) transceivers. Such transceivers perform the three functions of signal *re-amplification*, *re-shaping* and *re-timing*, hence the name '3R'. The transmission distance can then be expanded by concatenating several OCCs to form a chain, each stage requiring OE regeneration followed by electro-optic (EO) re-conversion (referred to as *OEO regeneration*).

For WDM systems, such OEO regeneration requires a complex and expensive per-carrier-wavelength implementation. An alternative approach, is *all-optical in-line regeneration*, which consists in by-passing the electronic-processing stage and performing 3R signal processing through all-optical or hybrid EO modulation. As shown by recent reviews [4,10,11], there exists a wealth of different techniques for implementing all-optical 3R regeneration, including the *simultaneous processing* of WDM channels.

From the academic standpoint, one of the most interesting features of all-optical regeneration is that the OCC transmission distance can be made 'infinite', without any SNR/BER degradation. The transmission distance is indefinitely increased because the SNR (hence the BER) asymptotically stabilizes itself to some level, as regeneration periodically removes, in a converging mode, the noise cumulated in the line. Here, 'infinite transmission' means in practice that optical signals could be made to re-circulate in a fiber loop with a regenerator, while keeping their information integrity *ad-infinitum*. This could not be the case with OEO regeneration, where the BER is cumulative. This is because each OEO stage involves a finite probability of symbol decision error. In contrast, all-optical processing does not 'decide' on the symbol information contents ('0' or '1' bit), but only removes the associated noise. Furthermore, we have also shown that, under certain conditions, the SNR can be *improved* by in-line optical regeneration (not just asymptotically stabilized), corresponding to an effect of *eye-re-opening* without symbol error [21]. The effect of eye re-opening can be compared to *amplitude squeezing*, but only by analogy since the signal electric field is classical. However, squeezing-like effects in both amplitude and phase are produced by all-optical regeneration [22], which remains to be further studied for ultimate limits and potential coherent-transmission/regeneration applications.

The above observations point to new directions in the physics of noise and associated information theory for noisy channels, as restricted so far to binary, ON–OFF signaling. The stabilization of SNR or its improvement through all-optical regeneration means that the intrinsic channel *equivocation entropy* can be kept under control, regardless of noise origin (amplification, nonlinearity) and their cumulative distance, or even removed in certain cases. In this paper, we have shown (Eq. (8)) that the channel equivocation takes the form $H(Y|X) \approx BER \exp(-BER)$, which asymptotically becomes linear with BER. We have also shown that the BER is asymptotically of the form $BER \approx \exp(-SNR)/\sqrt{4\pi\, SNR}$, Eq. (4). Assume that in the regenerated OCC the SNR is finally stabilized at a value SNR^*, corresponding to a constant equivocation $H^*(Y|X) \approx BER^* \exp(-BER^*)$, where $BER^* \approx \exp(-SNR^*)/\sqrt{4\pi\, SNR^*}$. It is then possible to define the difference between the regenerated and non-regenerated OCC entropies according to:

$$\Delta H = H^*(X|Y) - H(X|Y) = BER^* \,\mathrm{e}^{-BER^*} - BER\,\mathrm{e}^{-BER} \tag{58}$$

which (by convention) corresponds to negative entropy. To provide an example, assume typical values of $BER = 10^{-3}$ (regeneration OFF) and $BER = 10^{-6}$ (regeneration ON). From Eq. (58), we get $\Delta H = -0.9980 \times 10^{-3}$ bit/channel use (or -0.9980 bit/s at 1 KHz channel-use rate). One can interpret the result in terms of an increase of channel capacity of 1 bit/s at 1 KHz user rate. This means that the loss of 1 bit per thousand is removed. Since BERs are usually much smaller than unity, the 'negative entropy' introduced by optical regeneration is only of academic interest in this case. However, the realistic implementation of optically-regenerated systems concerns 'useless' channels, i.e., where $BER \leqslant 0.5$. As discussed in Section 2.3, the (binary) channel capacity vanishes in this case. Assuming that the regenerated BER is much smaller than unity, $BER^* \ll 1$, and using the exact definition in Eq. (8), we obtain for the negative entropy:

$$\Delta H \approx -H(X|Y) = (1 - BER)\log_2(1 - BER) + BER\log_2(BER) = \log_2\left[BER^{BER}(1 - BER)^{1-BER}\right] \tag{59}$$

which reaches a maximum ($\Delta H = -1$) when $BER = 0.5$ (full channel capacity restoration). The interest of the 'negative entropy' concept is therefore to qualify the improvement introduced by optical regeneration in 'useless' binary channels, on a $[-1; 0]$ measurement scale.

7. Conclusion

In this paper, we have reviewed the key concepts associated with information capacity in the optical communications channel. First, we have recalled the definitions of signal-to-noise ratio, entropy and channel capacity (Shannon–Harley theorem) concerning both binary and continuously-coded channels. An original expression linking equivocation entropy and bit-error-rate in binary channels was derived. Next, we have considered the limitations caused by optical amplification noise and its accumulation with distance, as based upon a quantum-field operator model. The model shows that two independent vacuum-field couplings are necessary to describe non-ideal amplification. However, its also shown that an equivalent and more concise description of non-ideal amplifiers can be made through a Langevin-operator formalism. Based upon this analysis, the case of *linear* channels with discrete and distributed in-line amplification was then analyzed, which lead to an original definition of ultimate channel capacity for ideal systems. Then we have considered the nonlinear optical communication channel. First we recalled the origin of nonlinearity due to wave-mixing between wavelength carriers and how it is modeled through the

nonlinear Schrödinger equation. A recent classical model for noise accumulation in nonlinear communication channels was then described, leading to a reformulation of the SHT which takes into account both amplifier and nonlinearity limitations. In particular, this classical model shows that it is possible to define a power transmission window where the amplified/nonlinear channel capacity is maximized. We presented then a unified quantum model for both amplification and nonlinearity, as based upon Langevin-operator formalism. The model is shown to predict small corrections from the classical approach, which leads to a mode detailed and accurate definition of the SHT. Finally, we have considered the case of all-optically regenerated channels, which makes possible to realize infinite-distance transmission and even eye re-opening by stabilization or improvement of the signal-to-noise ratio. It was shown that all-optical regeneration is in fact equivalent to introduce negative entropy in the communication channel, which corresponds to the channel equivocation entropy. For binary channels, this negative entropy is measured on a $[-1; 0]$ scale, which determines the transition from useless- to error-free channel operation.

In these times where technologies for broadband and global networking have reached a very high degree of maturity, analyzing ultimate capacity limits in the optical-communication channel has become a most relevant task. Such limits are not only determined by innovation engineering and technology evolutions, but also by fundamental principles from quantum physics and information theory. With continued improvements in our understanding of channel noise and coding limitations, much progress can be anticipated in this field.

References

[1] C.E. Shannon, Bell Systems Tech. J. 27 (1948) 379–423 and 623-656 and references [1,2] therein.
[2] J.B. Stark, P. Mitra, A. Sengupta, Optical Fiber Technology 7 (4) (2001) 289.
[3] E. Desurvire, Optical Fiber Technology 8 (4) (2002) 210.
[4] E. Desurvire, B. Desthieux, D. Bayart, S. Bigo, Erbium-Doped Fiber Amplifiers, Device and System Developments, Wiley, New York, 2002.
[5] S. Bigo, Y. Frignac, J.-C. Antona, G. Charlet, S. Lanne, C. R. Physique 4 (2002).
[6] O. Gautheron, M. Suyama, C. R. Physique 4 (2003).
[7] O. Ait-Sab, H. Bissessur, C. R. Physique 4 (2003).
[8] A. Hasegawa, Optics & Photonics News 13 (2) (2002) 33.
[9] S. Turitsyn, E.G. Shapiro, S.B. Medvedev, M.P. Fedoruk, V.K. Mezentsev, C. R. Physique 4 (2003).
[10] O. Leclerc, B. Lavigne, D. Chiaroni, E. Desurvire, in: I.P. Kaminow, T. Koch (Eds.), Optical Telecommunications IV, Academic Press, 2002, Chapter 15.
[11] O. Leclerc, B. Lavigne, E. Balmefrezol, P. Brindel, L. Pierre, D. Rouvillain, F. Seguineau, C. R. Physique 4 (2003).
[12] P. Nouchi, L.-A. de Montmorillon, P. Sillard, A. Bertaina, P. Guenot, C. R. Physique 4 (2002).
[13] D. Bayart, C. R. Physique 4 (2003).
[14] E. Desurvire, Erbium-Doped Fiber Amplifiers, Principles and Applications, Wiley, New York, 1994.
[15] I. Abram, P. Grangier, C. R. Physique 4 (2003).
[16] E. Desurvire, Optical Fiber Technology 5 (1) (1999) 40.
[17] E. Desurvire, Optical Fiber Technology 5 (1) (1999) 82.
[18] E. Desurvire, Optical Fiber Technology 6 (2) (2000) 199.
[19] J.P. Gordon, L.F. Mollenauser, IEEE J. Lightwave Technology 9 (2) (1991) 170.
[20] E. Desurvire, Electron. Lett. 38 (17) (2002) 983.
[21] E. Desurvire, O. Leclerc, Optical Fiber Technology 6 (3) (2000) 230.
[22] O. Leclerc, E. Desurvire, Opt. Lett. 23 (18) (1998) 1453.