

A State-Variable Approach to Submarine Links Capacity Optimization

Alberto Bononi, *Senior Member, IEEE*, Paolo Serena, *Member, IEEE*, and Jean-Christophe Antona

Abstract—We consider the capacity optimization of submarine links when including a realistic model of the gain-flattened constant-pump erbium doped fiber amplifiers (EDFA). While Perin *et al.* [1] numerically attacked this optimization for Constant-Gain (CG) amplified links, we extend the analysis also to more realistic submarine constant power-spectral-density (CPSD) links. As in [1], we concentrate on a single spatial mode of a spatial division multiplexed (SDM) link at low EDFA pump power P_p , and thus consider only the impairments of amplified spontaneous emission noise. Here we adopt a novel semi-analytical approach which consists of fixing the inversion x_1 of the first EDFA (the state-variable of the link) and analytically finding capacity $C(x_1)$ by searching over the x_1 -feasible input wavelength division multiplexed (WDM) PSD distributions. Then the optimum inversion x_1 that maximizes $C(x_1)$ is numerically obtained. This approach enables us to get both approximate (for CG links) and exact (for CPSD links) capacity-maximizing WDM input distributions, which vary inversely with the EDFA gain profile. For CG links the optimal WDM allocation is called the gain-shaped water-filling. Other practical allocations are analyzed, such as the signal to noise ratio equalizing allocation (CSNR), and the constant input power (CIP) allocation which uses a flat WDM distribution. We find that, for typical submarine span attenuations around 10dB and when the link works at the optimal inversion x_1 , CIP and CSNR achieve essentially the same capacity as the optimal allocation. At sufficiently large pump P_p ($\gtrsim 30$ mW) the optimal inversion x_1 is such that the EDFA gain at 1538nm equals the span attenuation, for EDFA emission and absorption as in [1]. When span attenuations increase to 20dB, then we start seeing an advantage of the optimal allocation. Another key finding is that optimized CG and CPSD links behave roughly the same, with a slightly superior capacity for CPSD.

Index Terms—Optical Communications, Optical amplifiers, Submarine transmission, Signal Droop.

I. INTRODUCTION

Submarine systems are fundamentally power constrained. The quest for energy efficiency has become critical, and new industrial cable designs have arisen, leveraging space division multiplexing (SDM) with either fiber bundles combined with pump-farming solutions at the industrial level [2], [3], or investigations of multi-core fiber solutions at the research level [4]. For each spatial-mode, line optimization has also been rethought, with suggestions to optimize amplifiers bandwidth [1], gain shaping filters [5], [6], or

the wavelength division multiplexing (WDM) input power spectral density (PSD) [1], [7], also known as pre-emphasis, most often based on direct numerical or machine-learning optimization.

In particular, Perin *et al.* [1] numerically studied how the capacity of submarine single-mode fiber links with end-span gain-flattened, fixed-pump erbium-doped fiber amplifiers (EDFA) scales with the EDFA pump power. The end-to-end pump requirements of SDM links based on fiber bundles are then inferred from the single-mode fiber results. In [1], the detailed EDFA physics were first introduced into the capacity-maximizing design of submarine pump-power constrained links. The innovative concept was that of bringing the bandwidth-signal to noise ratio (SNR) trade-off of the fixed-pump EDFA into the link design. The results in [1] were obtained by a fully numerical optimization, which limits the understanding of the underlying general concepts.

In this paper, we use a novel semi-analytical approach to both shed light into the results of [1], and extend them to cope with today's submarine links where at all spans the WDM signal PSD, the EDFAs, and their (fixed) gain shaping filters (GSF) are nominally identical. Luckily, these latter links are simpler to analyze, and we present them first. Then we attack the link in [1] and derive results quite close to those in [1], thus both confirming their results and validating our novel method.

The first key enabler of our approach is the use of the classical Saleh EDFA model [8]–[10] extended to include self-saturation by amplified spontaneous emission (ASE) [11]. The Saleh model without ASE self-saturation [8] was also used in [1] to provide the starting conditions for the global optimization search. The reader may review the EDFA model in Appendix 1, whose exposition is tailored to the developments in the main text.

The second and fundamental enabler of our approach is a strategy that allows us to break the complex capacity optimization problem into a set of simpler problems amenable to an analytical solution. Here is the idea. As in [1], our goal is to maximize the single-mode fiber capacity at a given per-EDFA pump power P_p . For each P_p , our original approach is based on assuming knowledge of the inversion x_1 at the first EDFA (the “state-variable” [10] of EDFA 1 and thus of the entire link) and *analytically* finding the capacity-achieving WDM distribution and the capacity value at x_1 . The dependence on all physical link parameters is summarized by the sole state variable x_1 . We finally scan the whole x_1 inversion range and *numerically* obtain the maximum link capacity. Once maximum capacity versus P_p is available, it is possible to carry out optimizations for power-constrained submarine SDM systems [1].

Manuscript received xxxxxxxx xx, xxxx; accepted xxxxxxxx xx, xxxx. Date of publication xxxxxxxx xx, xxxx; date of current version xxxxxxxx xx, xxxx.

A. Bononi and P. Serena are with the Dipartimento di Ingegneria e Architettura, Università di Parma, Parma 43124, Italy (corresponding author e-mail: alberto.bononi@unipr.it). J.-C. Antona is with Alcatel Submarine Networks, Villarsaux, France.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier xx.xxxx/JLT.xxxx.xxxxxx.

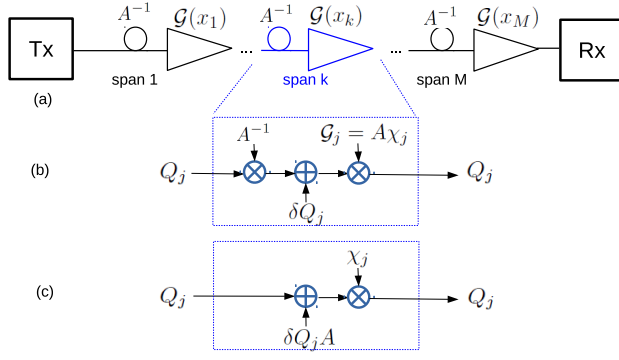


Figure 1. (a) M -span amplified line with span attenuation $A > 1$ and end-span amplification. At each span k , an EDFA with inversion x_k is followed by a GSF, for a total gain at frequency f_j of $\mathcal{G}_j(x_k) = A\chi_j(x_k)$, where $\chi_j \leq 1$ is the span droop. In CPSD mode, the GSF and thus χ_j are tailored such that the span input flux Q_j equals the span output flux at each frequency. (b) block diagram of a span. $\delta Q_j = F_j \Delta f$ is the ASE equivalent EDFA-input flux at frequency f_j over bandwidth Δf . (c) span block-diagram equivalent to (b), obtained by “factoring out” the attenuation, i.e., moving the A multiplier upstream of the addition operator.

At the low per-EDFA pump powers envisaged for submarine SDM, only ASE matters and nonlinear effects can be disregarded [1], [12]. The single-mode fiber is thus a wavelength-parallel additive Gaussian noise channel. If we had a signal power constraint and the received noise power did not depend on the WDM allocation, the capacity-achieving distribution would be the well-known classical water-filling (CW) [13]. Alas, EDFA gain and noise do depend on input power, and the constraint is here on the per-EDFA pump power. For the links considered in [1] we will find an x_1 -dependent quasi-optimal input distribution that we call the *gain-shaped water-filling* (GW) since the sum of signal and noise is not flat as in CW, but shaped as the inverse EDFA gain. For more realistic submarine links with constant PSD at the input/output of each EDFA, which we label constant-PSD (CPSD) links, we manage to get instead a recursive algorithm yielding the optimal WDM PSD and thus capacity. Our treatment of the CPSD links is an extension of the generalized droop model [14] to the wavelength-dependent amplifier gain. This paper is an extension of the conference paper [15], which first introduced the GW allocation for CG links. The rest of the paper is organized as follows: Section II introduces the link model. Section III analyzes the CPSD link. Section IV analyzes the CG link. Finally Section V concludes the paper.

II. SINGLE-MODE FIBER LINK MODEL

As depicted in Fig. 1(a), we consider the transmission of an WDM signal with along a link of M end-amplified single-mode fiber spans. The span attenuation is $A > 1$. Each end-span amplifier consists of an EDFA followed by a Gain-Shaping Filter (GSF) which may either flatten or properly shape the amplifier gain over the amplifier bandwidth.

As described in Appendix 1, the useful bandwidth B_k of the k -th gain-shaped amplifier in the link is defined as the frequency size of the set $\mathcal{B}_k = \{f_j : G_j(x_k) \geq A\}$, i.e., the frequency range over which the EDFA gain exceeds or equals the attenuation. In this study, the amplifier bandwidth is quantized in multiples of the bin size Δf , so that $B_k =$

$N_c(x_k)\Delta f$, where $N_c(x_k)$ is the number of non-overlapping slots covering \mathcal{B}_k . WDM channels may be allocated in all or part of the N_c bins within the bandwidth.

The EDFA length ℓ , the absorption α and emission g^* parameters (see Appendix 1), and the pump power P_p are the same at all EDFAs. As in [14], Fig. 1(b) shows the power-transfer block diagram of the generic span k , where $\delta Q_j(x_k)$ is the ASE equivalent EDFA-input flux (ph/s) and $\mathcal{G}_j(x_k) = A\chi_j(x_k)$ is the shaped-amplifier power gain (EDFA+GSF: $\mathcal{G}_j = G_j G_{SF,j}$) at all bins j centered at frequency f_j and of width Δf , belonging to the set \mathcal{B}_k , with $\chi_j(x_k) \leq 1$ the (span) droop [14]. The droop is thus the net span- k power gain at bin j . The shaped total gain is assumed to be ideally zero out of band. Fig. 1(c) shows a more convenient representation of the span block diagram, where the span loss is “factored out” [14]. $\delta Q_j A$ is the equivalent span-input ASE noise generated at the end-span EDFA.

We will study the following two link types:

1) **Constant Power Spectral Density (CPSD) link:** at all spans k and frequencies f_j in the bandwidth, the shaped-amplifier gain $\mathcal{G}_j(x_k) = A\chi_j$ has a span-independent, frequency-dependent droop $\chi_j \leq 1$ adjusted such that the span-output photon flux Q_j (or equivalently the power P_j) at f_j equals the span-input (i.e., the transmitted (TX)) flux. Hence the whole M -span line attenuates the input WDM channel at f_j by χ_j^M . Such a stabilization of the photon flux profile is achieved by suitable GSFs. Hence the signal+cumulated-ASE PSD at the input/output of every EDFA is preserved. Since all EDFAs have the same pump, input PSD, and physical parameters, they all work at the same inversion x , with identical (frequency-dependent) gain and noise figure. Of course, CPSD implies a constant total output power (COP) at each amplifier, as in [14], but COP does not necessarily imply CPSD. CPSD operation requires an input-PSD-dependent GSF, identical at all amplifiers. We will show numerically that for typical submarine links the combined EDFA+GSF amplifier for CPSD operation, over the range where WDM channels are allocated, is very close to a perfectly flat filter with gain equal to A (to within a tiny fraction of a dB), and has infinite attenuation where no power is allocated.

The CPSD case is studied first since all EDFAs are identical and more analytical insight is possible. CPSD links with a flat input PSD are a typical choice in today’s submarine systems, hence of great practical importance.

2) **Constant Gain (CG) link:**

This is the case considered in ref. [1]. It is less realistic for submarine systems, but better approximates the operation of COP terrestrial links. Here at all spans k and frequency bins in the bandwidth range \mathcal{B}_k the span droop is unity: $\chi_j(x_k) = 1$, hence the shaped-amplifier gain \mathcal{G} is perfectly flat and exactly equals the span attenuation A at all frequencies of the amplifier bandwidth. This implies that amplifier+fiber is a unit-gain block, so that the whole line has unit gain, the transmitted WDM power profile (i.e., the TX signal PSD) flows unchanged down the line, while ASE adds to it and increases down the line, so that the EDFAs inversion decreases down the line for constant-pump EDFAs. Thus the (mostly fixed and thus cheap) GSFs need to be all

different, a serious practical drawback. The next sections detail the analysis in the two cases.

III. CPSD LINK

In this case every EDFA in the line has the *same* input pump flux Q_p (i.e., same power P_p) and the *same* input signal fluxes $\{Q_j/A\}_{j=1}^{N_c}$, thus (per Saleh equation (26), see Appendix 1) the *same* inversion x , hence the same EDFA gain profile $G_j(x)$ at all frequencies $f_j = c/\lambda_j$ (c = speed of light, λ_j the j -th wavelength) and the same frequency-dependent GSF power transfer function:

$$G_{SF,j}(x) = \begin{cases} G_j(x)^{-1} A \chi_j & \forall f_j \in \mathcal{B} \\ 0 & \text{else.} \end{cases} \quad (1)$$

The overall amplifier gain profile is $\mathcal{G}_j \triangleq G_j(x) G_{SF,j}(x)$. With these ideal GSFs there is no need of adjustable gain equalizers regularly spaced along the link, which instead real links have.

As in [14], the droop χ_j at each frequency bin j is found by imposing that the flux Q_j (power P_j) entering the generic span shown in Fig. 1(b), or its equivalent form Fig. 1(c), be equal to the flux/power out of the gain-shaped amplifier:

$$(Q_j + \delta Q_j A) \chi_j = Q_j \quad (2)$$

where (see Appendix 1) the equivalent EDFA-input ASE flux at f_j is $\delta Q_j = F_j \Delta f$, where F_j is the EDFA noise figure at the j -th frequency bin f_j . Thus we find [14]:

$$\chi_j = (1 + SNR_{1j}^{-1})^{-1} \quad (3)$$

where

$$SNR_{1j} \triangleq \frac{Q_j/A}{F_j \Delta f} \quad (4)$$

is the total signal to injected noise ratio at f_j at each EDFA input. Note from (2) that when $SNR_{1j} \gg 1$ then $\chi_j \rightarrow 1$; conversely, when $SNR_{1j} \ll 1$ then $\chi_j \rightarrow 0$.

The received (RX) signal flux after M spans at f_j is $Q_j^{SIG,RX} = Q_j \chi_j^M$, while the RX ASE flux is

$$Q_j^{ASE,RX} = A \delta Q_j [\chi_j^M + \dots + \chi_j^1] = Q_j (1 - \chi_j^M) \quad (5)$$

where we used (3). The SNR at f_j is thus, using (5),(3):

$$SNR_j^{RX} = ((1 + SNR_{1j}^{-1})^M - 1)^{-1} \quad (6)$$

which is the generalized droop (GD) formula [14] at the j -th frequency, and we note that every input flux vector $\underline{Q} = \{Q_j\}_{j=1}^{N_c}$ is associated with a droop vector $\underline{\chi} = \{\chi_j(Q_j)\}_{j=1}^{N_c}$ that ensures the CPSD behavior of the link, as per (2).

The quantity we wish to maximize for this WDM-parallel additive Gaussian noise channel with our signal flux (or power) allocation is the achievable information rate (AIR) per 2-polarization spatial mode at the fixed (Q_p, x) [1]:

$$AIR(Q_p, x, \underline{Q}) = \sum_{j=1}^{N_c} 2\Delta f \log_2(1 + \Gamma SNR_j^{RX}) \quad (7)$$

with an implementation penalty $0 < \Gamma < 1$ known as SNR gap to capacity¹. As in [1], in all calculations we will assume $\Gamma = 0.79$ (-1dB).

¹When $\Gamma = 1$, AIR coincides with capacity.

A. AIR optimization

Our strategy is to fix the value of the (common) inversion x , and among all input flux vectors \underline{Q} compatible with x (or *feasible*) we look for the one that maximizes *AIR*. The set of feasible flux vectors (which are those that solve Saleh equation for the given x and pump Q_p) is a convex simplex, with bounded maximum power. Luckily, in this case the optimum can be found analytically as follows.

We wish to maximize with respect to (w.r.t.) the unknown vector \underline{Q} the function $AIR(Q_p, x, \underline{Q})$ in (7), subject to the positivity constraints for \underline{Q} and Saleh feasibility constraint:

$$\sum_{j=1}^{N_c} \frac{Q_j}{A} (G_j(x) - 1) = K(x, Q_p) \quad (8)$$

where K , given by eq. (27) in Appendix 1, is the useful pump flux.

Before delving into the optimization, let's consider first the AIR of some popular sub-optimal CPSD flux allocations. In the analysis, it is useful to decompose the flux vector as $\underline{Q} = \underline{Q} \underline{s}$, where the scalar total flux is $\underline{Q} \triangleq \sum_{j=1}^{N_c} Q_j$, and the *split vector* is $\underline{s} = \{s_j\}_{j=1}^{N_c}$ with $s_j \triangleq Q_j/\underline{Q} \geq 0$. The split vector is thus a probability mass function (PMF). For any split, the total *feasible* flux \underline{Q} is found from Saleh equation (8),(27) as:

$$\underline{Q} = \frac{AK(Q_p, x)}{\sum_{j=1}^{N_c} s_j (G_j(x) - 1)} \quad (9)$$

and is itself a function of the split vector.

1) *Constant-SNR allocation (CSNR)*: CSNR has been for a long time the standard WDM allocation in submarine links [16]. Let's first search for the TX flux vector that makes the received SNR_j^{RX} in (6) equal at all channels j . It should make the ratio $SNR_{1j} \propto \frac{Q_j}{\delta Q_j}$ equal for all j , i.e., $Q_j \triangleq \underline{Q} s_j \propto F_j$. Since \underline{s} is a PMF, we then must have for all $j = 1, \dots, N_c$

$$s_j = \frac{F_j}{\sum_{k=1}^{N_c} F_k} \quad (10)$$

Hence using (9), for all $j = 1, \dots, N_c$ the CSNR feasible fluxes are explicitly:

$$Q_j = \frac{AK(x, Q_p) F_j(x)}{\sum_{i=1}^{N_c} F_i(x) (G_i(x) - 1)} \quad (11)$$

which means the *WDM signal and the ASE have the same frequency/wavelength power profile*. Since SNR_{1j} is identical for all j , from (3),(4),(11) the droops at all f_j are all equal to:

$$\chi = (1 + \frac{\Delta f \sum_{j=1}^{N_c} F_j (G_j(x) - 1)}{K(x, Q_p)})^{-1} \quad (12)$$

which means that the CSNR allocation produces a *perfectly flat* droop (=net span gain). The equalized RX SNR is then: $SNR_j^{RX} = (\chi^{-M} - 1)^{-1}$, which is the standard wavelength-independent GD formula [14]. The CSNR AIR is finally obtained from (7).

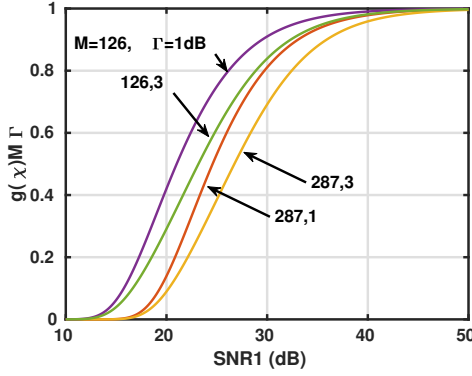


Figure 2. Normalized $g(\chi)$ vs. single span SNR_1 (dB scale) for some values of gap Γ and spans M .

2) *Constant-input-power allocation (CIP)* : CIP is another practical allocation often used in submarine links due to the simplicity by which one can tailor and monitor the needed GSFs that keep a nominally flat WDM spectrum along the line (a special case of CPSD). In the CIP case the input WDM powers (not the fluxes!) are all equal. The power vector is $\underline{P} = P_c \underline{1}$ with P_c the per-channel TX power, and $\underline{1}$ the vector of all ones. So the fluxes are $Q_j = P_c/h/f_j$, the total flux is $Q = \frac{P_c}{h} \sum_{j=1}^{N_c} \frac{1}{f_j}$, and the split vector thus has entries

$$s_j = \frac{Q_j}{Q} \equiv \frac{1/f_j}{\sum_{j'=1}^{N_c} 1/f_{j'}}. \quad (13)$$

Given the above splits, the total flux Q is again given by (9). Hence for all $j = 1, \dots, N_c$ the CIP fluxes are explicitly

$$Q_j = \frac{AK(x, Q_p) \frac{1}{f_j}}{\sum_{i=1}^{N_c} \frac{1}{f_i} (G_i(x) - 1)} \quad (14)$$

which correctly yield a constant per-channel power $P_c = hf_j Q_j$. Since SNR_{1j} is j -dependent, then from (3) the droop is also j -dependent:

$$\chi_j = \frac{1}{1 + SNR_{1j}^{-1}} \equiv \frac{P_c}{P_c + A\Delta f F_j h f_j} \quad (15)$$

The first equality shows that the non-flatness of the span gain χ_j is negligible whenever $SNR_{1j} \gg 1$ for all j . The second equality shows that the χ_j profile is the inverse of that of ASE power. The CIP received SNR and AIR are finally computed as usual per (6), (7).

3) *Optimal allocation (OPT)*: We now tackle the AIR-maximizing signal allocation. We form the Lagrangian: $L(Q) = AIR(Q) - \lambda \sum_{j=1}^{N_c(x)} Q_j (G_j(x) - 1)$, where $\lambda > 0$ is the Lagrange multiplier and AIR is given by (7), and then set the derivative w.r.t. all Q_k , $k = 1, \dots, N_c$, equal to zero. After some algebra we get: $\frac{\partial AIR}{\partial Q_k} = \frac{2\Delta f}{\ln(2)} \frac{\Gamma M}{Q_k} g(\chi_k)$, where we defined the key function

$$g(\chi) \triangleq \frac{\chi^M}{1 - \chi^M} \frac{1 - \chi}{1 - \chi^M(1 - \Gamma)} > 0 \quad (16)$$

which goes to $\frac{1}{M\Gamma}$ as $\chi \rightarrow 1$. Fig. 2 shows the normalized $g(\chi)M\Gamma$ versus the single-span ASE-corrupted $SNR_1 \equiv (\chi^{-1} - 1)^{-1}$ in dB scale for some values of parameters M and Γ .

Now, setting $\frac{\partial L}{\partial Q_k} = 0$ yields:

$$Q_k (G_k - 1) = \theta g(\chi_k) \quad (17)$$

where $\theta \triangleq \frac{2\Delta f \Gamma M}{\lambda \ln(2)}$. The unknown parameter θ is found by summing the solution (17) over all k and plugging into Saleh's constraint: $AK(x, Q_p) \equiv \sum_{k=1}^{N_c} Q_k (G_k - 1) = \theta \sum_{k=1}^{N_c} g(\chi_k)$, yielding $\theta = \frac{AK(x, Q_p)}{\sum_{j=1}^{N_c} g(\chi_j)}$. Plugging back into (17) we finally get the optimal (OPT) fluxes for all $k = 1, \dots, N_c$ as:

$$Q_k = \frac{AK(x, Q_p)}{G_k - 1} \frac{g(\chi_k)}{\sum_{j=1}^{N_c} g(\chi_j)} \quad (18)$$

which is one of the main results of this paper. It is a set of N_c coupled nonlinear equations in the unknown flux vector Q . The positivity constraint on the $\{Q_k\}$ is automatically guaranteed by the form of the solution. To solve (18) for the unknowns Q , the following recursion at epoch $i = 1, 2, \dots$ (based on the assumption that (18) is a contraction map) always worked for us:

$$Q_k^{(i)} = \frac{AK(x, Q_p)}{G_k - 1} \frac{g(\chi_k(Q^{(i-1)}))}{\sum_{j=1}^{N_c} g(\chi_j(Q^{(i-1)}))}$$

where our starting guess $Q^{(0)}$ was the CIP solution (14) at the same (Q_p, x) values. What kind of flux profile vs. frequency do we obtain from (18)? If for simplicity we assume OPT flux is allocated only over the N_b bins where $g(\chi)$ is above 1/10 of its maximum, and the $\{g(\chi_k)\}$ on such bins are all equal, then from (18) the OPT fluxes are $Q_k \cong \frac{AK(x, Q_p)/N_b}{G_k - 1}$ for $k = 1, \dots, N_b$. Hence the OPT allocation has a frequency profile as the inverse EDFA gain, as in the GW allocation that we will tackle later. Although we have no analytical proof, in numerical experiments the above solution is always unique and corresponds to a maximum.

B. CPSD link: numerical results

We analyze a CPSD link, with the same parameters as in [1]: $M = 287$ spans with attenuation $A = 9.5$ dB, in the typical range of submarine links.

For an EDFA length $\ell = 6.27$ m, and pump powers $P_p = [15 \div 180]$ mW, Fig. 3(a) shows AIR vs. x for three different power allocation policies: OPT (solid), CIP (dash-dotted), CSNR (dashed). We observe that AIR versus x grows identically for all three policies up to a maximum at inversion x_o roughly equal to $\cong 0.63$ at all pumps ≥ 30 mW. Beyond the maximum the OPT policy has larger and larger AIR w.r.t. the other policies (which behave almost the same). Reason of the existence of an optimum inversion is the bandwidth-SNR trade-off at play for the AIR (7) when increasing inversion. To highlight this trade-off, Fig. 3(b) shows both bandwidth B and RX SNR (averaged over the WDM channels) versus inversion x . Bandwidth vs. x is seen to increase monotonically independently of the allocation, as discussed in Appendix 1. The optimal x_o turns out to be very close to the “knee” of the B vs. x curve (circle in Fig. 3(b)). As explained in Appendix 1, such a “knee” occurs when the EDFA gain trough at 1538 nm is tangent from above to the A (dB) horizontal line.

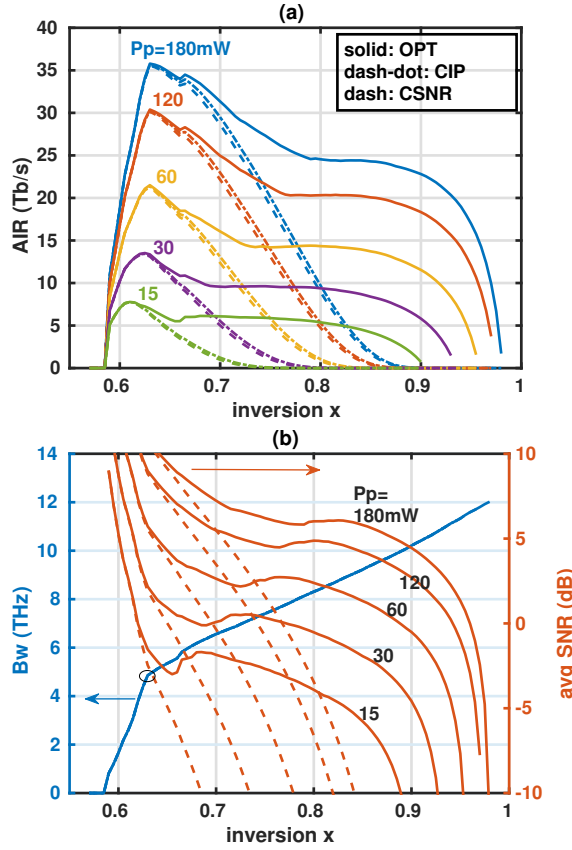


Figure 3. (a) AIR (Tb/s) vs. x . Allocations: solid: OPT; dash-dot: CIP; dash: CSNR. Pump powers $P_p = [15 \div 180]$ mW. Channel bandwidth $\Delta f = 50$ GHz. (b) Corresponding bandwidth and WDM-average SNR versus inversion x . CPSP link, $M = 287$, $A = 9.5$ dB, EDFA length $\ell = 6.27$ m.

SNR instead has normally a decreasing behavior versus x . Reason is that, at fixed pump, increasing x implies more gain and thus smaller *feasible* WDM signals, which implies (normally) a decreasing SNR, even though noise figure (hence ASE) also decreases as x increases. We see that the SNR curves for CIP and CSNR allocations are monotonically decreasing and almost coincident at all x , while the OPT SNR coincides with that of the other allocations up to and a little beyond x_0 (which justifies the common AIR growth for all policies at small x), and then OPT starts to have a much larger SNR than the sub-optimal policies, and thus much larger AIR.

The key observation from 3(a) is that at submarine-typical span loss (e.g. $A = 9.5$ dB) and working at the optimal inversion, where SNR is still “sufficiently” large, the OPT AIR is almost the same as that of the other policies, a fact already known for high-SNR channels where classical waterfilling is optimal [17]. The consequence is that, if pump is large enough and the line inversion is optimally chosen, there is little scope for allocation optimization, and the most practical allocation should be selected [7].

For the CIP allocation, Fig. 4 reports the same AIR values as in Fig. 3(a), but now plotted vs. the total input power $N_c P_c$ for flat CIP allocation, which depends on inversion as per eq. (14). This figure is reminiscent of similar curves with Kerr nonlinearity [19, Fig. 10], although here the maximum is due to the above-described SNR-bandwidth trade-off.

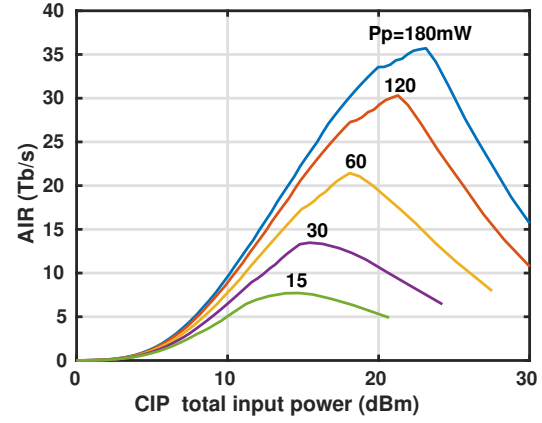


Figure 4. AIR (Tb/s) vs. total input power $N_c P_c$ for CIP allocation. CPSP link, $M = 287$, $A = 9.5$ dB, EDFA length $\ell = 6.27$ m.

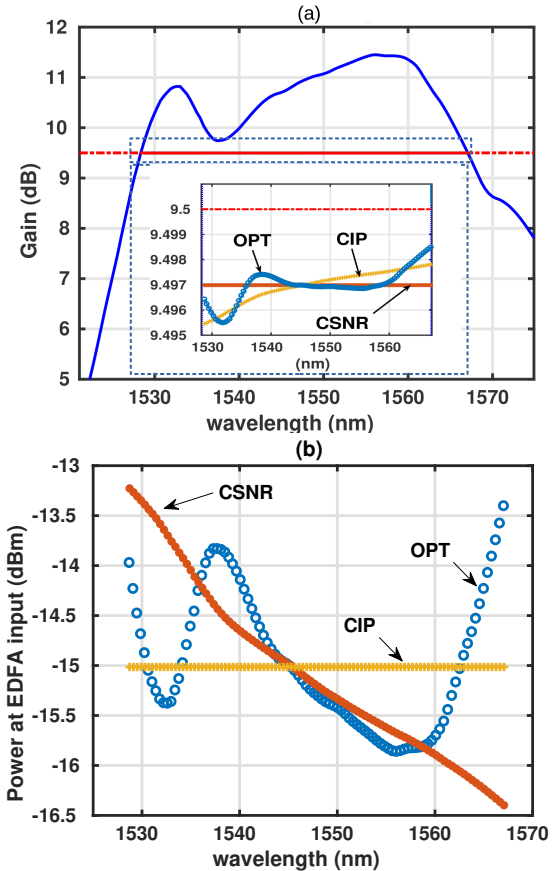


Figure 5. (a) EDFA gain (blue solid) vs. wavelength, and horizontal attenuation line $A = 9.5$ dB (red dash-dot). Inset zooms on the EDFA+GSF equalized gain vs wavelength in a range of 0.006 dB near A (blue 'o': OPT; red '*': CSNR; yellow '+': CIP, all overlapped in main figure). (b) WDM PSD at each EDFA input for OPT, CSNR, and CIP allocations. CPSP link, $M = 287$, $\ell = 6.27$ m, $P_p = 60$ mW, optimal $x_o = 0.63$.

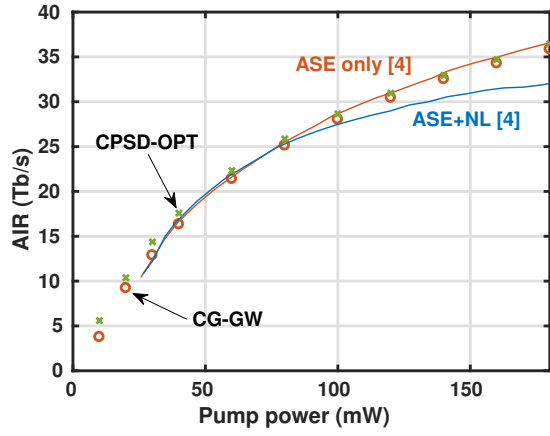


Figure 6. AIR vs. P_p . Solid: CG link AIR [1], Fig.4a, both “ASE only” (red) and “ASE+NL” blue; Circles: CG link with GW allocation; Crosses: CPSD link with OPT allocation. $M = 287$ spans, $A = 9.5$ dB, $\Gamma = 0.79$, EDFA ℓ optimized at every P_p .

At the optimal inversion $x_o = 0.63$ and at pump $P_p = 60$ mW, Fig. 5(a) plots both the EDFA gain (blue solid) and the EDFA+GSF gain (symbols, whose details are given in the inset) versus wavelength for OPT (‘o’), CIP (‘+’), CSNR (‘*’). We see that the equalized gain profiles vs. wavelength (hence the net span gain (i.e., droop) profiles) for the three allocation policies differ by less than 0.006 dB. The only truly flat one is that of CSNR, and the χ_k profile of CIP is the inverse of the ASE power, as already noted in (12),(15). All three droop profiles are thus practically flat and the GSF has essentially the profile of the inverse EDFA gain, as per (1). Given the manufacturing tolerances of the order of a fraction of dB, it is not possible to precisely engineer the three (ideally different) GSFs for the three policies, which is the main reason of the insertion of extra flattening filters placed after a block of spans, whose main task is to restore the TX input PSD.

Fig. 5(b) plots the PSD at each EDFA input versus wavelength for the OPT, CSNR, and CIP allocations. We note that the input PSD for CSNR has the same wavelength profile as the ASE noise, the CIP allocation is flat, while the OPT PSD is shaped as the inverse of the EDFA gain, as noted below (18).

Finally, for $M = 287$ spans and $A = 9.5$ dB, Fig. 6 shows (crosses) the x -maximized AIR of the CPSD link with OPT allocation versus pump power. Here the EDFA length ℓ was optimized at each pump, as explained in the next section. The remaining curves, to be derived later, refer to the CG link in [1]. The figure anticipates the slight AIR superiority (at equal EDFA pump) of the optimized CPSD link over the CG link, mostly visible at the lowest pumps below 50mW, i.e., those of most interest for SDM submarine links [18].

1) *Optimal inversion and EDFA length:* Now the question is: what is the optimal EDFA inversion-length pair (x_o, ℓ_o) ?

We already observed that, at large-enough pump and at typical submarine span attenuations, maximum AIR is normally achieved at an optimal inversion close to the knee of the B vs. x curve, where the gain trough at $\lambda_T = 1538$ nm is tangent from above to the attenuation line and bandwidth is a compact interval. Let α_T, g_T be the absorption and

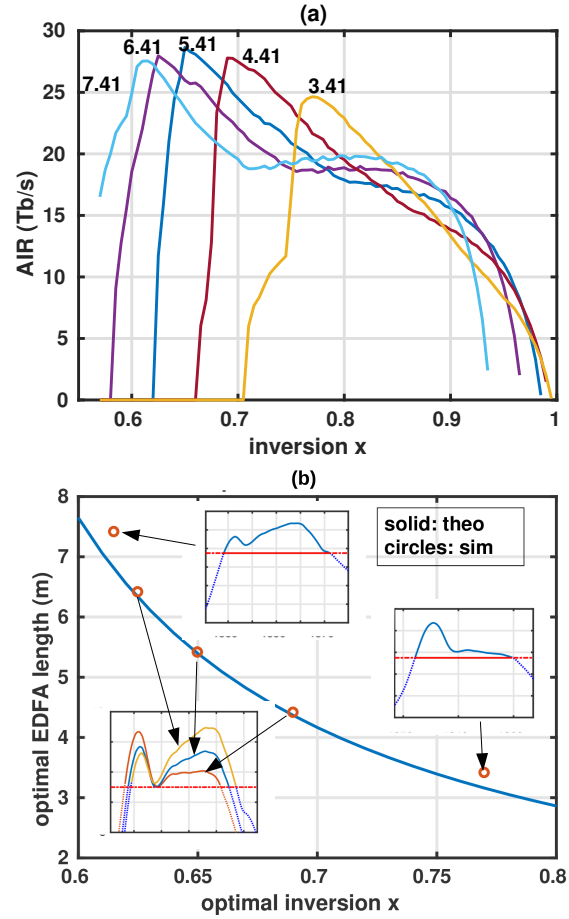


Figure 7. CPSD chain with OPT allocation, $M = 287$, $A = 9.5$ dB, $P_p = 100$ mW (a) AIR vs. inversion x at various EDFA lengths ℓ . (b) (solid) ℓ_o versus x_o for theoretical curve (19). Optimal pairs (ℓ_o, x_o) at top of every $AIR(x)$ curve in (a) are reported as circles in (b). Insets show the corresponding EDFA gain vs. wavelength.

emission coefficients at λ_T . Imposing $G(\lambda_T) = A$ and using (Appendix 1): $G_T = \exp(\ell((\alpha_T + g_T^*)x - \alpha_T))$ yields the following relation between x_o, ℓ_o :

$$\ell_o = \frac{\ln A}{(g_T + \alpha_T)x_o - \alpha_T} \quad (19)$$

Fig. 7(a) shows AIR vs. inversion x at various EDFA lengths ℓ , where we see that $\ell = 5.41$ m yields the largest top AIR (hence the value 6.27m reported in [1] and used above is close to optimum). Fig. 7(b) reports as circles the optimal pairs (x_o, ℓ) from each ℓ in left figure, and shows in the insets the corresponding EDFA gain vs. wavelength. Solid line is the theoretical curve (19). The three circles right on top of theory are for lengths $\ell = [4.41, 5.41, 6.41]$ m which show a cuspid at the top of the AIR vs x curve and have tangency of the gain trough at 1538nm with the attenuation line. The two edge circles refer to lengths $\ell = [3.41, 7.41]$ m at which tangency is not exact, so that circles are slightly off theory and the cuspid is not visible in the AIR curve.

From the bottom inset in Fig. 7(b) we find that the gain averaged over wavelengths is $\bar{G} = [10.3, 10.6, 11.7]$ at lengths $\ell = [4.41, 5.41, 6.41]$ m. Hence among the lengths that cause trough tangency, we see that the AIR-maximizing length $\ell = 5.41$ m (blue) is *not exactly* (but close to) the one with the smallest \bar{G} , i.e., the one

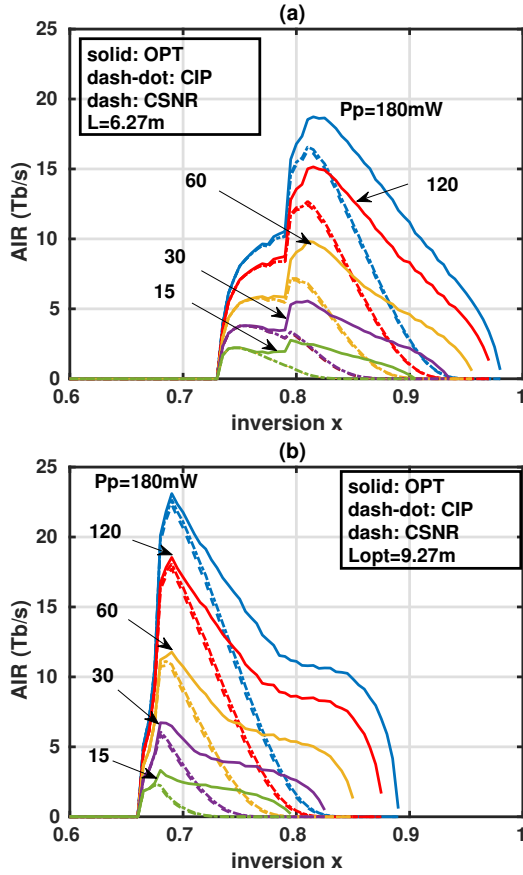


Figure 8. AIR (Tb/s) vs. x in a CPSD chain with $M = 63$, $A = 20$ dB. (a) $\ell = 6.27$ m; (b) $\ell = 9.27$ m (optimal). Lines: solid=OPT; dash-dot=CIP; dash=CSNR.

closest to A , hence the one that minimizes the photon flux wasted by the GSF, i.e., the most *pump-photon efficient*.

We next look at the performance of a $M = 63$ -span link with a typical terrestrial attenuation $A = 20$ dB, which corresponds to roughly a link of half the length of the 287-span link with 9.5 dB attenuation.

As in Fig. 3(a), Fig. 8 shows AIR (Tb/s) vs. x at various pumps, but now for $A = 20$ dB and the sub-figures refer to two different EDFA lengths $\ell = 6.27$ m and $\ell = 9.27$ m. We see that compared to the $A = 9.5$ dB case, at $A = 20$ dB a wider gap between OPT and the other allocations is obtained even at the optimal inversion, especially at the lowest pumps. At EDFA $\ell = 9.27$ m the gap is minimum since that length is optimal. At $\ell = 6.27$ m the inversion needs to be sharply increased right into the range where OPT is much more efficient than the other allocations, and the gap is thus maximum.

2) *Optimal span length*: To straddle a certain total distance, we can either use a few widely spaced high-pump amplifiers, or many narrowly-spaced low-pump amplifiers, and the question is: what is the best number of amplifiers to minimize the required *total* line pump power? For concreteness, we analyze again our $14350 = 7 \cdot 41 \cdot 5 \cdot 5 \cdot 2$ km CPSD link and look for the best among the following 6 combinations: 1) $M=287 \cdot 2=574$ spans of 25km each

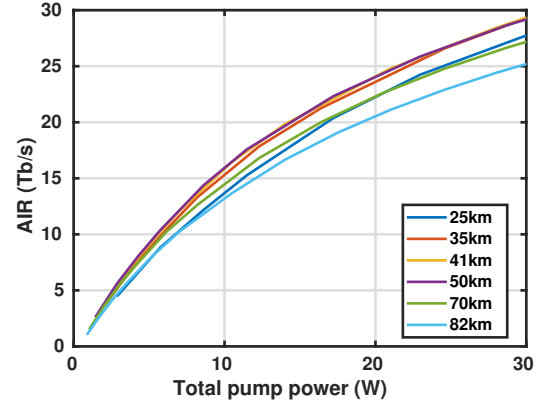


Figure 9. AIR vs. total link pump power MP_p (per single-mode fiber) at optimized EDFA ℓ for CPSD 14350km link with shown span lengths. OPT allocation (but CSNR/CIP yield about same AIR). On the shown power range nonlinearity is negligible [1].

($A = 5.4$ dB); 2) $M=41 \cdot 5 \cdot 2=410$ spans of 35km each ($A = 7.0$ dB); 3) $M=7 \cdot 50=350$ spans of 41km each ($A = 8.0$ dB); 4) $M=287$ spans of 50km each (already shown, $A = 9.5$ dB); 5) $M=41 \cdot 5=205$ spans of length $35 \cdot 2=70$ km each ($A = 12.8$ dB); 6) $M=7 \cdot 25=175$ spans of length $41 \cdot 2=82$ km each ($A = 14.78$ dB).

Using our reduced-complexity extended-Saleh EDFA model, instead of the full EDFA model used in [1], the search is much faster.

Fig. 9 shows the largest AIR (Tb/s) (obtained through EDFA length optimization) versus total link pump power MP_p (W) for the 6 above CPSD systems (as shown in [1], nonlinearities are negligible below $\sim 287 \cdot 100$ mW $\lesssim 30$ W, i.e. over the shown total pump range). AIR is shown for OPT allocation, but is almost identical for both CSNR and CIP allocations. We see that the AIR curve initially improves as the span length decreases (because ASE decreases), but after an optimum span length the AIR curves worsen again because too many amplifiers consume too much of the available total pump power. The 50km/span link ($A = 9.5$ dB) yields the largest AIR per pump watt and is thus the most energy-efficient, with the 41km being just slightly inferior (this is consistent with results in [1]). Clearly, not only power efficiency but also cost must be considered in the final link design [3], [12].

IV. CONSTANT-GAIN (CG) LINK

We next attack the more difficult case considered in [1]. Here the EDFA+GFF amplifier has a flat gain that *exactly equals* the attenuation over all channels in the bandwidth set \mathcal{B} , hence the flattened amplifier (signal) gain is constant at all amplifiers in the link (CG operation). Thus the net span gain (i.e., the droop) is unity at all channels. Hence every span has the equivalent block diagram in Fig. 1(c) with $\chi = 1$. Thus in the CG scenario the RX signal flux after M spans at the j -th frequency bin is $Q_j^{SIG,RX} = Q_j$, the same as the transmitted one, while the RX ASE flux is $Q_j^{ASE,RX} = A \sum_{k=1}^M F_j(x_k) \Delta f$, so the ASE flux increases down the line. Thus EDFAs down the line see increasing input fluxes at the same pump, hence their inversions decrease: $x_1 > x_2 > \dots > x_M$. Since

bandwidth is monotonically increasing with inversion, then the EDFA bandwidth decreases down the line and thus the bandwidth of the whole line is set by that of the last EDFA: $B = N_c(x_M)\Delta f$.

Now, the inversions down the line are a function of the chosen input signal fluxes $\{Q_j\}$ and of the inversion x_1 of the first EDFA (which is taken as a known quantity). TX fluxes $\underline{Q} = \{Q_j\}_{j=1}^{N_c}$ must be *feasible* for the given pump Q_p and inversion x_1 , i.e., they must satisfy Saleh equation (8) at the first EDFA. The inversions x_k of the subsequent EDFAs are recursively found from the known x_1 and the \underline{Q} vector by solving the Saleh equation at the k -th EDFA, for all $k = 2, 3, \dots, M$. Fluxes \underline{Q} depend on the flux allocation strategy, as detailed later.

The RX SNR $Q_j^{SIG,RX}/Q_j^{ASE,RX}$ at frequency j is

$$SNR_j^{RX} = \frac{Q_j}{A \sum_{k=1}^M F_j(x_k) \Delta f} \quad (20)$$

and the AIR per (2-polar.) spatial mode at fixed Q_p and inversion $x \equiv x_1$ of the first EDFA is still given by eq. (7), where $N_c \equiv N_c(x_M)$. The flux vector \underline{Q} to be used in each EDFA Saleh equations is specified depending on the allocation strategy, as detailed next.

A. Gain-shaped Waterfilling allocation (GW)

Define the inversion vector as $\underline{x} = \{x_1, x_2, \dots, x_M\}$, and from (20) define the noise vector as

$$N_j(\underline{x}) \triangleq \frac{A}{\Gamma} \sum_{k=1}^M F_j(x_k) \Delta f \quad (21)$$

for $j = 1, \dots, N_c(x_M)$. Then the optimal \underline{Q} maximizes the AIR

$$AIR(Q_p, x_1, \underline{Q}) = \sum_{j=1}^{N_c(x_M)} 2\Delta f \log_2 \left(1 + \frac{Q_j}{N_j} \right) \quad (22)$$

subject to the Saleh constraint (8) at the first EDFA. The problem differs from that of CPSD links in two respects:

- 1) we have an SNR where the noise flux N_j depends on the inversions only. Alas, the inversions in turn depend on \underline{Q} through Saleh balance equation at each EDFA;
- 2) the total link bandwidth $B(x_M)$ and hence $N_c(x_M)$ are not independent of the actual fluxes \underline{Q} as it was in the CPSD case.

Finding the exact AIR-maximizing \underline{Q} (as we did in the CPSD case) is ruled out, since the feedback dependence of the inversions $\{x_2, \dots, x_M\}$ and $N_c(x_M)$ on the fluxes \underline{Q} is too complicated to be analytically treated. However, if we *ignore* the dependence of the inversions on \underline{Q} , we can find a simple analytical solution as described next, which turns out to be “quasi-optimal”.

When the inversion vector \underline{x} is fixed, and so is $N_c(x_M)$, Appendix 2 derives the optimal fluxes $\underline{Q} \triangleq \underline{Q}_{GW}(\underline{x})$ that maximize AIR (22) subject to constraint (8), whose entries $j = 1, \dots, N_c$ are

$$Q_j = \max(0, \frac{\theta}{G_j - 1} - N_j) \quad (23)$$

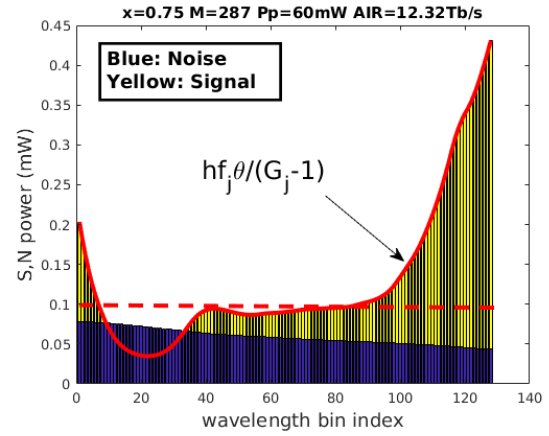


Figure 10. Example of GW signal power allocation for CG link with $M = 287$ spans, $A = 9.5$ dB, EDFA length $\ell = 6.27$ m, pump $P_p = 60$ mW, and inversion $x_1 = 0.75$. Dark shaded: noise power $hf_j N_j$ (mW), eq. (21). Solid: water-level $hf_j \theta / (G_j(x_1) - 1)$. GW signal power $S_j = hf_j Q_j$ (light shaded) is allocated between the noise level and the water-level solid curve. For CW instead the water-level is the flat dashed line.

Algorithm 1 Iterative CG link algorithm

Start from the uniform guess $\underline{x}^{(0)} = \{x_1, x_1, \dots, x_1\}$ where all EDFAs have the same (known) inversion as EDFA-1; Then repeat at all epochs n until convergence:

- 1) update GW as: $\underline{Q}^{(n)} = \underline{Q}_{GW}(\underline{x}^{(n-1)})$, as per eq. (29).
- 2) update the entries $x_k^{(n)}$, $k = 2, \dots, M$ of the new vector $\underline{x}^{(n)}$ by sequentially solving the Saleh equation at each EDFA k :

$$\sum_{j=1}^{N_c(x_M^{(n-1)})} \left(\frac{Q_j^{(n)}}{A} + \sum_{l=1}^{k-1} F_j(x_l^{(n)}) \Delta f \right) (G_j(x_k^{(n)}) - 1) = K(Q_p, x_k^{(n)}). \quad (24)$$

where $\theta / (G_j - 1)$ is the *inverse-gain-shaped water-level*, and theta a parameter to be found from Saleh equation at EDFA 1 (30). We call this flux allocation the (inverse-)Gain-shaped waterfilling (GW), in analogy to the well-known classical water-filling (CW) where the water-level is flat. Fig. 10 gives an example of both GW and classical water-filling (CW) signal power allocations for a CG line, $M = 287$ span, $A = 9.5$ dB, and EDFAs with length $\ell = 6.27$ m and pump $P_p = 60$ mW, working at inversion $x_1 = 0.75$. The solid curve is the non-flat water-level $hf_j \theta / (G_j(x_1) - 1)$, the dark shaded level corresponds to the noise power $hf_j N_j$ (mW), and the GW signal power $S_j = hf_j Q_j$ (light shaded) is allocated between the noise level and the water-level curve, as per (23). For CW instead the water-level is the flat dashed line.

The inversion \underline{x} is next iteratively found by algorithm 1. By exhaustive search over the feasible input simplex, we verified that the sub-optimal GW practically coincides with the optimal allocation at all but the largest inversions. GW will be validated also by comparison with the results in [1]. In Appendix 2 we also consider the classical water-filling allocation (CW) with a flat water level. The inversions vector

\underline{x} in the CW case is found by the same iterative algorithm as above, when updating $\underline{Q}^{(n)} = \underline{Q}_{CW}(\underline{x}^{(n-1)})$ by using CW fluxes (31).

For CG links, we next outline the two sub-optimal flux-allocation strategies CSNR and CIP already considered in the CPSD case.

B. CSNR allocation

To equalize the SNR (20), the ratio $\frac{Q_j}{\sum_{k=1}^M F_j(x_k)}$ must be equal for all j , i.e., $Q_j \equiv \theta s_j \propto \sum_{k=1}^M F_j(x_k)$ using the power split $s_j \triangleq Q_j/\theta$ (θ is here the total TX signal flux²) and since \underline{s} is a PMF we then must have for all $j = 1, \dots, N_c(x_M)$

$$s_j(\underline{x}) = \frac{\sum_{k=1}^M F_j(x_k)}{\sum_{j=1}^{N_c(x_M)} \sum_{k=1}^M F_j(x_k)} \quad (25)$$

which explicitly depends on the inversion vector \underline{x} . These are the splits used in EDFA-1 Saleh equation (8) that allow us to get the total flux θ from (9). So we now have the CSNR flux vector $\underline{Q}_{CSNR}(\underline{x})$ for given \underline{x} . The inversions vector \underline{x} is found by the same iterative algorithm as above, when updating $\underline{Q}^{(n)} = \underline{Q}_{CSNR}(\underline{x}^{(n-1)})$. At convergence, we get the SNR^{RX} from (20) and the AIR from (7).

C. CIP allocation

As in the CPSD case, the CIP split vector entries $\{s_j\}$ are obtained from (13) and they do not depend on inversions \underline{x} . With these, the total flux θ is obtained from (9) and only depends on the known x_1 . So we have directly the CIP flux vector $\underline{Q}_{CIP}(x_1)$. The rest of the inversions vector \underline{x} is found by just one run of step 2 of the standard iterative algorithm. We finally get the SNR^{RX} from (20) and the AIR from (7).

D. CG link: numerical results

Very similar numerical results to those obtained for the CPSD links are obtained for CG links, and we next present them. However note that the shaping filters in the CG link need to be matched to the corresponding inversion and are thus all different. We present it for sanity checks with the seminal work in [1].

Fig. 11(a) shows the achievable information rate AIR (Tb/s) versus inversion x at various pump powers from 15 to 180mW. We see that all allocation policies yield essentially the same AIR (differences below 2%) up to a little beyond the optimal inversion x_1 , which occurs close to the “knee” of the corresponding B vs. x_1 curve, cfr. Fig. 13(b). The global AIR vs x_1 behavior is quite similar to that of the corresponding CPSD link, shown in Fig. 11(b) (same curves as Fig. 3(a) with added GW curve).

We note that the CPSD link has always a slightly larger AIR than the CG link at same pump, and the differences become more noticeable at lower pumps. Also note that for the CPSD link we added the GW allocation (solid) as

² θ plays here the role of a flux vector parameter, exactly as it was for the GW and CW allocations.

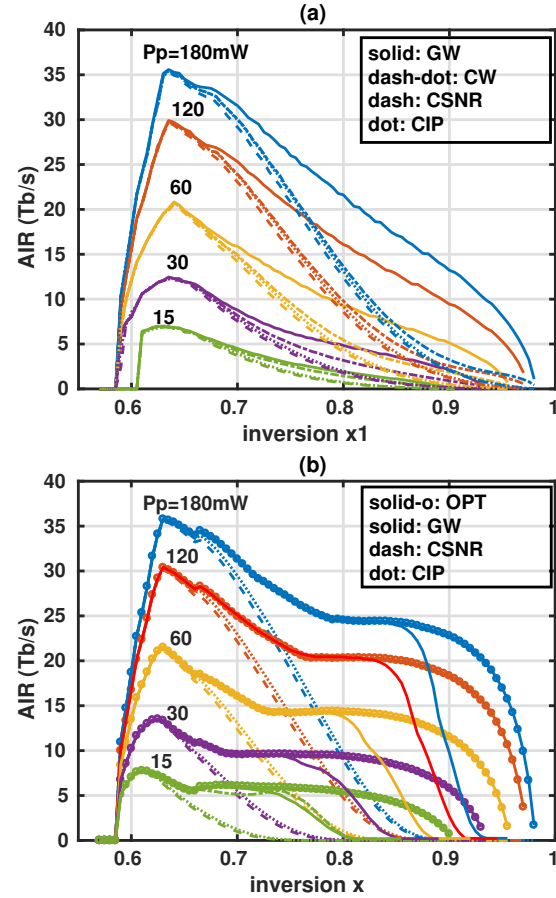


Figure 11. AIR (Tb/s) versus inversion x_1 at various pump powers for (a) CG link and (b) CPSD link, with $M = 287$, $A = 9.5\text{dB}$, $\ell = 6.27\text{m}$. Solid-o: CPSD-OPT; Solid: GW; dash-dot: CW; dashed: CSNR; dot: CIP).

described in Appendix 4, and GW turns out to be coincident with the optimal allocation (solid with circles) up to large inversions.

For a global comparison between optimized CPSD and CG links, please refer to Fig. 6, which shows (circles) the x_1 -maximized AIR of the CG link with GW allocation versus pump power. Here the EDFA length ℓ was optimized at each pump and for each allocation. The solid lines are the AIR results for the same link from [1, Fig.4(a)], both without (red) and with nonlinear effects (blue), which shows that at P_p below 100mW nonlinear effects are negligible [1]. With only ASE, we see that our GW AIR is practically coinciding with the optimal values found in [1].

For the same link, Fig. 12 shows the signal powers at every EDFA input at various pumps and at the optimal inversions, where symbols are our CG GW values and solid curves are the results in [1, Fig. 2a]. We see that we are able to get similar curves. Note that our bandwidth calculation considers all bins with EDFA gain above attenuation, without any smoothing of the bandwidth edge bins, while the work in [1], in order to perform particle swarm maximization, introduces artificial smoothing, which is evident at the bandwidth edges in Fig. 12.

V. SUMMARY AND CONCLUSIONS

We started our exploration of this topic when trying to reproduce the fully-numerical CG link results in the

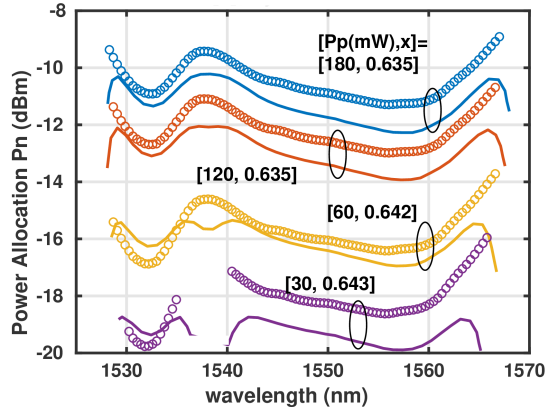


Figure 12. WDM signal power at input of each EDFA versus wavelength at various pump powers and at optimal inversions. Symbols: our CG-GW model. Solid: results from [1, Fig. 2a]. Data: CG link, $M = 287$ $\ell = 6.27$ m, $A = 9.5$ dB.

seminal paper [1]. Key to our novel approach was the use of the Saleh EDFA model enriched with ASE self-saturation (Appendix 1), and the understanding that knowledge of the state variable of the first EDFA [10] and thus of the entire link allows breaking the whole optimization problem into a series of much simpler optimizations, amenable to analytical solution. We thus discovered the GW allocation (Appendix 2) as the quasi-optimal allocation in CG links [15], which is a new twist in optical communications of the classical capacity-achieving water-filling allocation in additive Gaussian noise channels [13]. Today's submarine links, however, are not CG links, but rather CPSD links with CSNR or CIP input WDM PSDs. The PSD is preserved along the line by (nominally) identical GSFs at all EDFAs. Their role is not exactly gain flattening, but rather PSD enforcing. In this paper we explored variations of the CPSD link with non-flat input PSDs. Our analytical approach to the study of CPSD links is an extension of the GD model [14] to the case where the amplifier channel gains are not all equal. Using the GD model we discovered the optimal PSD, eq. (18), that maximizes capacity (more correctly the AIR, since we introduce a “gap to capacity” Γ), whose analytical formulation formally does not resemble the GW allocation. However, we proved that the GW allocation in CPSD links provides essentially the same AIR as the optimal allocation, up to almost the largest inversions (Cfr. Fig. 11(b)). The same thing happens in CG links.

The main findings from this study are:

1) optimized GC and CPSD links behave roughly the same, with a slightly larger optimal AIR for CPSD links (Cfr Fig. 6);

2) with pumps ≥ 30 mW and typical submarine span attenuations in the order of 10dB, sensible allocations such as the flat one (CIP) or the SNR-equalizing allocation (CSNR) achieve AIR values to within a few percent of the optimal AIR (Cfr Fig. 11). There is thus no point in using non-flat allocations [7]. However, when span losses are increased to 20dB (typical of terrestrial systems), the optimal allocation provides sizable AIR improvements, especially at lower pumps or when EDFA length is sub-optimal (Cfr Fig. 8);

3) at submarine span losses, the optimal EDFA inversion

(for absorption and emission coefficients as in [1, Fig. 7]) is very close to the one that makes the EDFA gain at 1538nm equal to the span attenuation (Cfr. Fig. 7);

4) the fixed gain shaping filters to implement CPSD for different PSD allocations differ from the inverse EDFA gain profile by a fraction of a dB (Cfr Fig. 5). Since it is unfeasible to implement them with such tolerances, then periodically spaced adjustable gain/PSD equalization is needed along the link;

5) power-efficiency optimal span attenuations are in the range 8-9.5 dB, corresponding to span lengths 40-50km (Cfr. Fig. 9), in agreement with [1], [3].

An interesting question is how to possibly estimate the inversion x of a black-box commercial amplifier just from input/output measurements, and thus reconnect that estimate to the picture presented in this paper. Some recent black-box EDFA models point in that direction [20], [21], but the problem remains open.

Acknowledgments: A. Bononi and P. Serena acknowledge support from Ministero dell'Istruzione, dell'Università e della Ricerca (FIRST, PRIN 2017).

APPENDIX 1: SALEH EDFA MODEL

While Perin's work [1] solved the complete EDFA rate equations, we will use here the much faster Saleh EDFA model [8]–[10], made more accurate by including self-saturation by ASE as in [11].

Saleh model can be interpreted as an hydraulic system consisting of a single tank (in the *homogeneous broadening* assumption [8], [24]) that gets filled by the pump photon flux Q_p at wavelength λ_p (usually 980 nm or 1480 nm) and whose contents we call the reservoir r , which physically represents the total number of excited Erbium ions in the doped fiber. Ions in the reservoir are ready to be converted to signal photons. The reservoir, normalized to the total number of ions in the doped fiber r_M , is called the average inversion: $x \triangleq r/r_M$. The gain seen by the input WDM photon fluxes Q_j^{in} at the various wavelengths $j = 1, \dots, N_c$ exponentially increases with the doped fiber length ℓ and the inversion x as: $G_j^{dB} = 10 \log_{10}(e) \ell ((\alpha_j + g_j^*)x - \alpha_j)$, and involves the wavelength-dependent emission g_j^* and absorption α_j coefficients (m^{-1}) [22]. In all our calculations we used the values reported in [1, Fig. 7], and a pump at 980 nm.

Saleh Balance equation : At equilibrium, the reservoir (thus the inversion x) depends on the balance of fluxes entering and exiting the tank, as expressed by the Saleh balance equation:

$$\sum_{j=1}^{N_c} Q_j^{in} (G_j(x) - 1) = K(x, Q_p) \quad (26)$$

with

$$K(x, Q_p) \triangleq Q_p - (Q_p G_p(x) + \frac{r_M}{\tau} x + Q_{ASE}^{F+B}(x)) \quad (27)$$

where the left-hand side of (26) represents the fluxes drawn by the input WDM signals out of the tank, and the term K represents the useful pump flux, i.e., the input pump flux Q_p minus the “leakage from the tank”, namely: i) the unused pump flux $Q_p G_p$ (the pump gain is $G_p < 1$); ii)

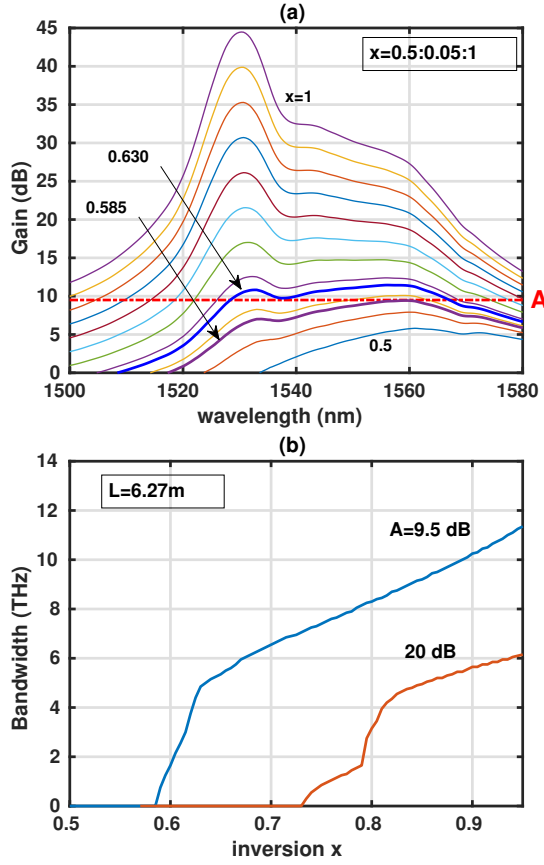


Figure 13. (a) EDFA gain versus wavelength at inversion $x = [0.5 : 0.05 : 1]$; (b) bandwidth B (THz) at span attenuations $A = [9.5, 20]$ dB versus inversion x . “Cutoff” and “knee” for $A = 9.5$ dB are at $x = [0.585, 0.63]$, while for $A = 20$ dB they are at $x = [0.73, 0.81]$. EDFA length $\ell = 6.27$ m.

the fluorescence $\frac{\tau}{\tau}$, with τ the fluorescence time, iii) the forward and backward ASE flux Q_{ASE}^{F+B} , which causes ASE self-saturation, and is analytically included as [11]:

$$Q_{ASE}^{F+B} = 2 \sum_{j=1}^L 2n_{sp,j}(G_j - 1)\Delta f \quad (28)$$

with $n_{sp,j} = \frac{g_j^* x}{(g_j^* + \alpha_j)x - \alpha_j}$ the spontaneous emission factor [22], and Δf the ASE resolution bandwidth over bins $j = 1, \dots, L$ tiling the significant wavelength range. The first multiplier 2 gives equal strength to forward and backward ASE, since (28) is obtained in the assumption of a spatially uniform inversion at value x [11].

EDFA Noise Figure : The EDFA noise figure is defined for all frequency bins j as [23] $F_j = 2n_{sp,j} \frac{G_j - 1}{G_j}$. The equivalent-input (forward) ASE flux to the EDFA at λ_j is $F_j \Delta \nu$.

Amplifier Bandwidth: In Fig. 13(a) we report the Gain versus wavelength at various inversions, where we also show a horizontal line at level $A = 9.5$ dB, i.e., the span attenuation considered in [1]. The gain-flattened amplifiers in the main text consist of an EDFA followed by an ideal gain flattening filter that slices off all gain in excess of the span attenuation A and blocks all wavelengths whose EDFA gain is below A . N_c WDM signals of bandwidth Δf and spacing Δf completely occupy the available amplifier bandwidth B . Such a bandwidth (THz) is plotted in Fig. 13(b) versus

inversion x for attenuation A both 9.5 and 20 dB. For $A = 9.5$ dB, we note that B is zero at all inversions below the cutoff value $x = 0.585$ which is the largest x for which the gain curve is fully below A (see the thick purple curve in Fig. 13(a)). Above cutoff, B increases vs. x initially fast, and then more slowly after a knee (at $x = 0.63$ for $A = 9.5$ dB), which from Fig. 13(a) (thick blue curve) corresponds the smallest x for which the EDFA gain trough at 1538 nm equals A , so that the bandwidth comes from a compact wavelength set, instead of disjoint segments as at lower inversions. Here is why the slope of B vs. x is large at the knee and becomes abruptly small after the knee. The bandwidth change dB due to a change dx can be expressed as $dB = \sum_{i=1}^I df_i$, where df_i is the change in bandwidth at the i -th gain intersection point f_i with the attenuation line, $i = 1, \dots, I$. By multiplying and dividing by the induced gain change $dG_i^{dB} = 10 \log_{10}(e) \ell (\alpha_j + g_j^*) dx$, we write $dB = \sum_{i=1}^I 10 \log_{10}(e) \ell (\alpha_j + g_j^*) \frac{dG_i^{dB}}{df_i}$. Whenever the derivative dG_i^{dB}/df_i is small (as at 1538 nm and $x = 0.63$), then dB is large. Conversely when the G^{dB} vs. f curve is steep at the intersection points, then dB is small, as for all x above the knee.

APPENDIX 2: GW AND CW ALLOCATIONS

Starting from section IV-A on CG links, we now derive the optimal fluxes $\{Q_j\}$ that maximize AIR (22) subject to constraint (8). Assume we know the inversions vector $\underline{x} = \{x_1, \dots, x_M\}$. If we neglect the positivity constraints on the fluxes, we form the Lagrangian $L = \sum_{j=1}^{N_c(x_M)} \frac{2\Delta f}{\ln(2)} \ln(1 + \frac{Q_j}{N_j(x)}) - \lambda \sum_{j=1}^{N_c} Q_j (G_j(x_1) - 1)$, where $N_c(x_M)$ and the noise flux $N_j(x)$ in (21) are fixed, and $\lambda > 0$ is the Lagrange multiplier. We now set the derivative of the Lagrangian w.r.t. all Q_k , $k = 1, \dots, N_c$, equal to zero:

$$0 \equiv \frac{\partial L}{\partial Q_k} = \sum_{j=1}^{N_c} \left[\frac{2\Delta f}{\ln(2)} \frac{1}{N_j + Q_j} \frac{\partial Q_j}{\partial Q_k} - \lambda \frac{\partial Q_j}{\partial Q_k} (G_j - 1) \right].$$

Now, $\frac{\partial Q_j}{\partial Q_k} = \delta_{jk}$ is a Kronecker delta, i.e., 1 if $j = k$ and zero else. Thus we get: $0 = \frac{2\Delta f}{\ln(2)} \frac{1}{N_k + Q_k} - \lambda (G_k - 1)$, hence $(N_k + Q_k)(G_k - 1) = \theta$, with constant $\theta \triangleq \frac{2\Delta f}{\ln(2)\lambda}$, and thus get the signal fluxes for all $k = 1, \dots, N_c$ as: $Q_k = \frac{\theta}{G_k - 1} - N_k$, with a k -dependent *inverse-gain-shaped* water-level $\theta_k \triangleq \frac{\theta}{G_k - 1}$ equal to the sum of signal and noise fluxes on frequency bin k . The Lagrange optimization fails when some of the signal fluxes are negative. In this case we must include the flux non-negativity in the constraints and solve the resulting Karush-Kuhn-Tucker (KKT) problem [25]. As in classical water-filling, we find that the correct KKT solution for all $k = 1, \dots, N_c$ is [13, Sec. 9.4]:

$$Q_k = \left(\frac{\theta}{G_k - 1} - N_k \right)^+ \quad (29)$$

where $x^+ \triangleq \max(x, 0)$. The expression for the water-level parameter θ is found by forcing compliance with the Saleh constraint (8) at EDFA 1:

$$\sum_{k=1}^{N_c(x_M)} \left(\frac{\theta}{G_k - 1} - N_k \right)^+ (G_k - 1) = AK(Q_p, x_1). \quad (30)$$

Contrary to the classical water-filling for the parallel additive Gaussian channel, the water-level $\theta_k \triangleq \frac{\theta}{G_k - 1}$ is now k -dependent, and water-filling is not in the powers but in the photon fluxes. In solving the above nonlinear equation in θ , as a starting guess we may choose the lower bound $\theta_{LB} = \frac{AK(Q_p, x_1) + \sum_{k=1}^{N_c} N_k(G_k - 1)}{N_{ch}} \leq \theta$ that we get in absence of $(\cdot)^+$.

Classical Water-filling (CW)

If instead of condition (29) we impose

$$Q_k = (\theta - N_k)^+ \quad (31)$$

then we obtain the classical water-filling allocation (although using fluxes instead of power) with flat water level θ , obtained by solving Saleh equation at EDFA 1:

$$\sum_{k=1}^{N_c(x_M)} (\theta - N_k)^+ (G_k - 1) = AK(Q_p, x_1). \quad (32)$$

As a starting guess we may use the lower-bound value $\theta_{LB} = \frac{AK(Q_p, x_1) + \sum_{k=1}^{N_c} N_k(G_k - 1)}{\sum_{k=1}^{N_c} (G_k - 1)}$.

APPENDIX 3: GW ALLOCATION IN CPSD LINK

We can also impose allocations such as GW and CW (See Appendix 2) in the CPSD scenario. Here we explain GW.

The RX SNR at every WDM frequency f_j is $\Gamma SNR_j \triangleq Q_j/N_j$, where from (7) the TX-equivalent ASE noise is

$$N_j(x, Q_j) \equiv \frac{Q_j}{\Gamma} (\chi_j^{-M} - 1) \quad (33)$$

and $\lim_{Q_j \rightarrow 0} N_j(x, Q_j) = \infty$. Using this expression in the calculations in Appendix 2 we get that the GW allocated fluxes correspond to equation (29), where the “water-level” parameter θ satisfies eq. (30). Such last equation is nonlinear in the unknown flux vector \underline{Q} . To solve for \underline{Q} , we can use the following recursion. If we have guesses \underline{Q}_{i-1} and θ_{i-1} at epoch $i = 1, 2, \dots$, then from (29) we update all entries $k = 1, \dots, N_{ch}$ of \underline{Q}_i as: $Q_{k,i} = (\frac{\theta_{i-1}}{G_k(x) - 1} - N_k(x, Q_{k,i-1}))^+$, then update $N_k(x, Q_{k,i})$ per (33), and finally update θ_i as the solution of: $\sum_{k=1}^{N_c(x)} (\theta_i - N_k(x, Q_{k,i})(G_k(x) - 1))^+ = AK(Q_p, x_1)$, as per (30). In our Matlab software we use as starting guess \underline{Q}_0 the CIP solution (11) at the same (Q_p, x) values, and use as a starting guess: $\theta_0 = \frac{AK(Q_p, x) + \sum_{k=1}^{N_p} N_k(x, Q_{k,0})(G_k(x) - 1)}{N_p}$, where $N_p \leq N_{ch}(x)$ is the number of channels where N_k is finite.

REFERENCES

- [1] J. K. Perin, J. M. Kahn, J. D. Downie, J. Hurley, and K. Bennett, "Importance of Amplifier Physics in Maximizing the Capacity of Submarine Links," *J. Lightw. Technol.* vol. 37, pp. 2076-2085, May 2019.
- [2] P. Pecci *et al.*, "Pump farming as enabling factor to increase subsea cable capacity," in Proc. *Suboptic* 2019, paper OP 14-4, New Orleans, Apr. 2019.
- [3] M. A. Bolshtyansky *et al.*, "Single-mode fiber SDM submarine systems," *J. Lightw. Technol.* vol. 38, pp. 1296-1304, Mar. 2020.
- [4] K. Saitho, "Few-mode Multi-core Fibres: Weakly-coupling and Randomly-coupling," in Proc. *ECOC'20*, tutorial Tu4A-1, Brussels, Dec. 2020.

- [5] J. Cho *et al.*, "Optimizing Gain Shaping Filters with Neural Networks for Maximum Cable Capacity under Electrical Power Constraints" in Proc. *ECOC'20*, paper Th1H-1.
- [6] M. Ionescu, A. Ghazisaeidi, J. Renaudier, P. Pecci and O. Courtois, "Design Optimisation of Power-Efficient Submarine Line through Machine Learning," in Proc *CLEO* 2020.
- [7] A. Kam *et al.*, "Pre-emphasis-based equalization strategy to maximize cable performance," in Proc. *SubOptic*, New Orleans, LA, USA, Apr. 2019, Paper OP14-1.
- [8] A. A. M. Saleh *et al.*, "Modeling of gain in erbium-doped fiber amplifiers," *IEEE Photon. Technol. Lett.* vol. 2, pp. 714-717, Oct. 1990.
- [9] Y. Sun, J. L. Zyskind and A. K. Srivastava, "Average inversion level, modeling, and physics of erbium-doped fiber amplifiers", *J. Sel. Topics Quantum Electron.*, vol. 3, pp. 991-1007, Aug. 1997.
- [10] A. Bononi and L. A. Rusch, "Doped-fiber amplifier dynamics: a system perspective," *J. Lightw. Technol.*, vol. 16, pp. 945-956, May 1998.
- [11] T. Georges, and E. Deleuaque, "Analytic modeling of high-gain erbium-doped fiber amplifiers," *Opt. Lett.* vol. 17, pp. 1113-111, Aug. 1992.
- [12] R. Dar *et al.*, "Cost-Optimized Submarine Cables Using Massive Spatial Parallelism," *J. Lightw. Technol.*, vol. 36, no. 18, pp. 3855-3865, Sep. 2018.
- [13] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. 2nd ed., Hoboken, NJ, USA: Wiley, 2006.
- [14] A. Bononi, J.-C. Antona, M. Lonardi, A. Carbo-Meseguer, and P. Serena, "The Generalized Droop Formula for Low Signal to Noise Ratio Optical Links," *J. Lightw. Technol.*, vol. 38, no. 8, pp. 2201-2213, Apr. 2020.
- [15] A. Bononi, P. Serena and J.-C. Antona, "Gain-shaped waterfilling is Quasi-optimal for Constant-pump Flattened-EDFA Submarine Links" in Proc. *ECOC* 2020, Paper Tu1F-5, Brussels, Belgium, Sep. 2020.
- [16] F. Forghieri, R. W. Tkach, and D. L. Favin, "Simple Model of Optical Amplifier Chains to Evaluate Penalties in WDM Systems," *J. Lightw. Technol.* vol. 16, n. 9, pp. 1570-1576, Sep. 1998.
- [17] Wei Yu and J. M. Cioffi, "On constant power water-filling," in Proc. *ICC* 2001, vol.6, pp. 1665-1669, Helsinki, Finland, 2001.
- [18] J. D. Downie *et al.*, "Experimental Characterization of Power Efficiency for Power-Limited SDM Submarine Transmission Systems" in Proc. *ECOC* 2020, paper Tu2E-5.
- [19] P. Poggiolini *et al.*, "The GN-Model of Fiber Non-Linear Propagation and its Applications," *J. Lightwave Technol.* vol. 32, pp. 694-721, Feb. 2014.
- [20] J. Cho *et al.*, "Data-Driven Characterization of EDFA In Constant Current Operation," in Proc. 2020 Opto-Electronics and Communications Conference (OECC), Taipei, Taiwan, Oct. 2020.
- [21] A. Carbo-Meseguer *et al.*, "Highly Accurate Measurement-Based Gain Model for Constant-Pump EDFA for non-Flat WDM Inputs," submitted to OFC'21.
- [22] C. Giles and E. Desurvire, "Modeling erbium-doped fiber amplifiers," *J. Lightw. Technol.*, vol. 9, no. 2, pp. 271-283, Feb. 1991.
- [23] H. A. Haus, "The Noise Figure of Optical Amplifiers," *IEEE Photon. Technol. Lett.*, vol. 10, no. 11, pp. 1602-1604, Nov. 1998.
- [24] E. Desurvire, *Erbium-doped fiber amplifiers: principles and applications*. Wiley, New-York, 1994.
- [25] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.