



Optimization strategy of power control for C+L+S band transmission using a simulated annealing algorithm

HUAIJIAN LUO,¹  JIANING LU,^{1,*}  ZHUILI HUANG,^{1,2} 
CHANGYUAN YU,^{1,3}  AND CHAO LU^{1,3}

¹Photonics Research Centre, Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Hong Kong SAR, China

²Key Laboratory of Optoelectronic Technology & Systems (Ministry of Education), College of Optoelectronic Engineering, Chongqing University, Chongqing 400044, China

³The Hong Kong Polytechnic University Shenzhen Research Institute, Shenzhen 518057, China
*jianing.lu@connect.polyu.hk

Abstract: To increase the transmission capacity, ultra-wideband wavelength-division multiplexing (UWB WDM) has been exploited to enlarge the spectral range. However, inter-channel stimulated Raman scattering (ISRS) results in power transition from high-frequency channels to low-frequency channels in wideband scenarios, which degrades the Q-factor of signals. Hence, we modify the optimization method of power control by applying the simulated annealing (SA) algorithm to search for the optimal power slopes and offsets of three bands to construct an optimum distribution of launch powers over channels. High transmission capacity can be reached by carrying 384 channels (96+96+192) in the C+L+S band with the consideration of dynamic Raman gain and channel-dependent parameters. We show that compared to using brute-force searching (BFS), a comparable and even higher transmission capacity can be achieved by the SA algorithm. Meanwhile, the searching speed of the SA algorithm is much faster. Also, different optimizing strategies can be selected to balance the trade-off between capacity and spectral flatness. This method can be used for designing arbitrary optical fiber UWB WDM systems before practical testing.

© 2021 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

1. Introduction

With the development of 5G communication, many applications have been emerging in recent years, such as the Internet of things (IoT), ultra-high-definition (UHD) video streaming, and cloud computing. These applications require high bandwidth of transmission. This has brought huge challenges and opportunities to the backbone network of modern communication. To further develop the transmission capacity on the digital coherent communication system, efforts can generally be made in two directions: space-division multiplexing (SDM) [1] and ultra-wideband wavelength-division multiplexing (UWB WDM) [2]. For SDM, the multi-core fiber (MCF) [3] and mode-division multiplexing (MDM) [4,5] are two research hot spots. Although SDM is widely used in the field of wireless communication, its development in the field of optical communication still faces many hurdles, such as the high cost of deploying brand-new fibers and difficulties in designing new fibers and amplifiers. Therefore, the SDM still has a long way to go before commercialization [6]. In contrast, expanding the bandwidth of WDM is more practical because it can be adopted in the current optical coherent systems. Many studies based on the UWB WDM have been emerging during recent years. As reported in [7], 94.9 Tbps capacity with average 9.75 b/s/Hz net spectral efficiency (SE) in the C+L band over 1900 km using Erbium-doped fiber amplifiers (EDFAs) has been demonstrated. In [8], 150.3 Tbps transmission was achieved over 40 km with three bands (S, C, and L bands) by using 272 channels with

45-Gbaud PDM-128QAM signals. One Thulium-doped fiber amplifier (TDFA) and two EDFAs were applied for amplifying the S-, and C+L bands, respectively, and a higher SE (11.05 b/s/Hz) was obtained. In 2019, the first-ever 5 bands (E-, O-, S-, C-, and L-band) with 625 channels are transmitted over 60 km [9], and the total capacity of 106.77 Tbps with a 23.5-THz spectral range was demonstrated.

With the exploitation of multiple bands, it is significant to take into account the effect of stimulated Raman scattering (SRS). This is because it can induce the power transfer of WDM signals which cover 100-nm or wider bandwidth [10]. In other words, SRS always affects channels within two or more bands. Influence of SRS was experimentally studied over the S+L band, L+C band, and S+L+C band [9,11,12]. Although these works show that power transition hardly affects the quality of the signal, the power imbalance results from power transfer over channels is still an issue. Since the channel at low frequency gains power from the channel at high frequency, more nonlinear interference (NLI) noise is generated in this channel due to the Kerr effect during transmission. Meanwhile, the needed gain from the optical amplifier (OA) for the low-frequency channel is lower and thus the generated amplified spontaneous emission (ASE) noise power is lower as well. Note that here it is assumed that the noise figure (NF) of OA is flat. Therefore, distinct power thresholds exist for channels at the different center frequencies. When the launch power of one channel exceeds the specific threshold, the degradation of the Q factor of the signal overpasses the improvement degree of the ASE signal-noise-ratio (SNR), which reduces general SNR (GSR) shown in Eq. (1). This is undesirable for increasing the capacity.

$$GSR = \frac{P_{ch}}{P_{ASE} + P_{NLI}} \quad (1)$$

Instinctively, the overall transmission capacity can be improved by optimizing the power distribution between channels. In [13], an iteration optimization method is reported to control the power of the S+C+L band. F. Hamaoka et al. optimized the average launch powers of the S-band and C-band separately for single-band transmission by sweeping the power. Then they optimized the average power of the L- and S-band iteratively with the three-band transmission based on the optimum average powers of S- and C-band found until powers of three bands converged to stable values. A high capacity record, shown above [8], was obtained by using this method. In their work, flat power distribution in bands was assumed. Different from this, B. Correia et al. proposed a multiband power control scheme [14,15] based on the hypothesis that the power distribution in the band is adjustable and a strategy called local optimization leads to global optimization (LOGO) [16]. LOGO states that the optimization over one span decided the overall optimization. Therefore, it can be used to simplify the process by optimizing the performance of the link with the transmission of one span, which decreases the computational complexity a lot. By sweeping combinations of six parameters of launch power profile (pre-tilts and offset powers of S, C, and L band) on the LOGO strategy, they optimized the GSR value to increase the capacity. Although these two works made significant improvements in capacity, iterative searching and brute-force searching (BFS) are time-consuming, especially in practical network planning. Meanwhile, the fixed slope of the Raman gain profile is not applicable when the total bandwidth is over 15 THz, which especially causes deviations of the gain spectrum at the outermost channels.

In this paper, we propose a power control method by leveraging the simulated annealing (SA) algorithm along with the LOGO strategy to quickly search a set of parameters to construct the multi-band launch power profile to increase the transmission capacity. The set of optimized parameters are power slopes and power offsets of three bands. We demonstrate that the searching period of the SA algorithm is faster about 8 times than that of the conventional BFS under the same optimizing strategy, and the average capacity ripple is much lower than that of BFS when achieving similar total capacities. We also find that a trade-off exists between the total capacity and its flatness over channels by comparing three distinct optimizing strategies. The proposed

technique can be applied before the experiments since each parameter in the optical configuration can be adjusted according to the real model, which is also promising for the optimization of optical fiber links.

2. Principle

2.1. SRS in UWB WDM system

As mentioned before, UWB WDM is applied to greatly enlarge the transmission capacity, which makes the total bandwidth of the signals extremely wide, reaching about 10 to 20 THz. And the overall launch power reaches a high degree when most channels are loaded over the wideband. Therefore, in addition to the self-phase modulation (SPM) and cross-phase modulation (XPM), another nonlinear effect, SRS, will accumulate greatly through signal propagation. In a multi-band scenario, SRS will cause the power transition from high frequency to low frequency, thereby influencing the distribution of NLI power. For example, if the launch power of each channel keeps the same for the S+C+L band at the transmitter side, the average power of the channel in L-band will be higher than these of channels in C- and S-band after propagation. Then the power distribution is approximately linear in frequency, i.e. the triangular approximation as shown in Fig. 1, and the power difference between the two outermost channels can be expressed as [17]

$$\Delta\rho [dB] = 4.343 \cdot P_{tot} C_r L_{eff} B \quad (2)$$

where P_{tot} is the total launch power, C_r is the linear slope of the normalized Raman gain spectrum, L_{eff} is the effective length of the fiber, and B is the total bandwidth. The $\Delta\rho$ increases with the rise of total launch power. However, this linear triangular assumption of the Raman profile is only valid when the total bandwidth is less than 15 THz and the launch power distribution is uniform [18]. Eq. 2 should be corrected to fit for the proposed UWB transmission. A more accurate Raman profile can be obtained by solving the Raman equations [19] as below

$$r_n^{(k)} = \exp\left(\sum_{m=1}^N \frac{g_R(w_m - w_n)}{A_{eff}} P_m[k-1] L_{eff,m}(L_{step})\right) \quad (3)$$

$$P_m[k] = \exp(-\alpha L_{step}) r_m^{(k)} P_m[k-1] \quad (4)$$

$$r_n [dB] = 10 \cdot \log_{10}\left(\prod_{k=1}^{\frac{L_{span}}{L_{step}}} r_n^{(k)}\right) \quad (5)$$

where N represents the number of total channels, $P_m[k-1]$ is the power of the m -th channel at the start of k steps, and L_{step} denotes the length of step. In the process of solving the Raman equation, the iterative number depends on the length of the step. It should be noted that the attenuation factor is a vector that varies with frequency. Although the Raman profile is unable to obey the linearity strictly over such a wide bandwidth (about 20 THz) after calibration, the trend of power transfer is still valid. More details about the Raman gain profile will be illustrated in the Section 3.2. Note that the closed-form GN model can still work when parts of channels are under utilization, because the accuracy of closed-form GN model on the mesh network analysis is high enough [18].

2.2. Power control over UWB WDM system

To compensate for this gain imbalance, we intuitively think it may work by adjusting the signal launch power spectrum to make the channel power increase as the frequency increases, thereby counteracting this power transfer. The schematic diagram of the power profile before launching is shown in Fig. 2(a). For each band, there are two parameters to denote the linear power distribution.

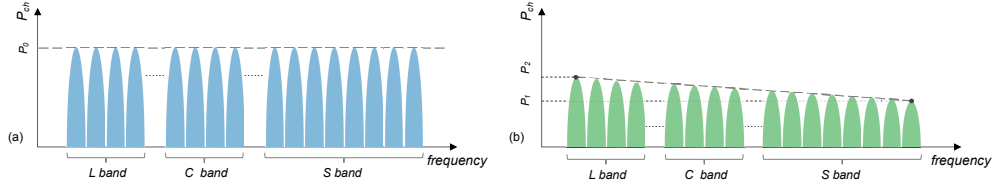


Fig. 1. Schematic diagram of power profile over L+C+S band when the signal (a) is launched at fiber and (b) is transmitted after one span

The first one is the slope of power spectrum, and the second is the offset power of the center channel. Hence there are 6 parameters to be controlled over three bands for maximizing the capacity, and the power distribution can be expressed as

$$P_i(f) = Slope \cdot (f - f_{center_i}) + P_{offset_i} \quad (i = L, C, S) \quad (6)$$

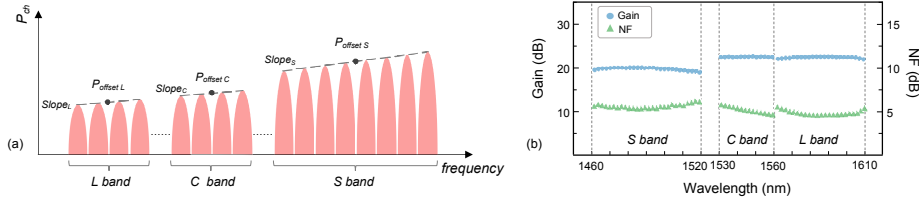


Fig. 2. (a) Schematic diagram of power profile before launching and (b) typical gain and noise figure of commercial OAs over S+C+L band

We use OAs to shape the power profile of UWB WDM channels. Two EDFAs are used for the amplification of the C+L band, respectively, and the TDFA is more suitable for the wavelength range of the S-band. The typical characteristics of gain and NF of commercial OAs over three bands are illustrated in Fig. 2(b). All the gains here work for launch power as high as ~ 0 dBm, and the gain will increase over 30 dB as input power drops. And it should be noted that the covering wavelength range of the current commercial TDFA cannot fully fill the S-band from 1460 nm to 1520 nm. Therefore, we consider two amplification schemes in Section 3.1. One is that the channels in S-band are amplified by only one TDFA, and the other is that two TDFA are applied. Detailed comparison of two schemes on the performance of optimization on power control will be discussed in Section 3.2. It is worthwhile to be noted that the Raman amplifier (RA) is another kind of potential amplifier to construct the hybrid amplifying scheme with EDFA to further decrease the ASE noise. In that case, the additional Raman gain profile and corresponding NF should be taken into consideration.

2.3. Analysis of linear and nonlinear noise

For an UWB WDM system, the theory capacity limit for single polarization can be decided by total bandwidth and GSNR, as shown in Eq. (7).

$$C = B \cdot \log_2(1 + GSNR) \quad (7)$$

To improve the capacity, the essential goal is improving the GSNR of every channel. We exploit the Gaussian noise (GN) model [20] which takes SRS into account and ASE equation to

comprehensively analyze the effects of nonlinear and ASE noise on the GSNR. The ASE noise power is mainly generated from the process of amplification, which can be expressed as

$$P_{ASE} = hf \cdot NF(f)(G(f) - 1)B_{ch} \quad (8)$$

where the h denotes the Planck's constant, B_{ch} represents the bandwidth of the channel, and $NF(f)$ and $G(f)$ are noise figure and gain from OAs which are frequency-dependent. Because of the Raman gain, the required gain of OA can be expressed by

$$G(f) = \alpha(f) \cdot L - G_{Raman}(f) \quad (9)$$

Hence, the SRS effect not only transfers power from high frequency to low frequency, but also causes the distribution of ASE noise unequal. Not to mention the influence of the frequency-dependent attenuation factor. The NLI noise accumulates from propagation in fiber link, which is related to the total power and transmitted distance. It can be calculated with the closed-form of the GN model as below. By leveraging the closed-form GN model, a relatively accurate NLI power distribution can be obtained within a short time period [18], which is much faster than calculating the NLI using the integral form [20].

$$P_{NLI} = \eta_n P_i^3 \quad (10)$$

where P_i is the power of i -th channel in the first span the η_n represents the NLI coefficient after n spans, which consists of self- and cross-channel NLI power terms and can be written as

$$\eta_n(f_i) \approx \sum_{j=1}^n \left[\frac{P_{i,j}}{P_i} \right]^2 \cdot [\eta_{SPM,j}(f_i)n^\epsilon + \eta_{XPM,j}(f_i)] \quad (11)$$

where the $P_{i,j}$ denotes the power of the i -th channel in the j -th span, ϵ is the coherent factor, which is dependent on the frequency and all system parameters [21], and $\eta_{SPM,j}$ and $\eta_{XPM,j}$ are contributions of SPM and XPM effects in the j -th span, which are given by [18].

2.4. Optimization algorithm by using the SA algorithm

According to the LOGO strategy mentioned before, we focus on optimization by considering one-span transmission. After calculating distributions of the ASE noise and the NLI noise, the GSNR and capacity can be obtained successively. In other words, once the distribution of launch power is set given knowing all the parameters of optical configuration, the achievable transmission capacity can be obtained in a fast way. To optimize the set of parameters including power slopes and power offsets of three bands, we define two metrics in the target function below

$$y = w_1 \cdot \frac{N}{\sum_{i=1}^N C_i} + w_2 \cdot [(C_{maxL} - C_{minL}) + (C_{maxC} - C_{minC}) + (C_{maxS} - C_{minS})] \quad (12)$$

where w_1 and w_2 represent the weights of these two terms, respectively, and C_i denotes the capacity of i -th channel among N channels. The first term is set as the reciprocal of the average capacity, and the second term represents the sum of ripples over L-, C-, and S-band. Each ripple denotes the difference between the highest and lowest capacity within one band. By changing values of two weights, w_1 and w_2 , distinct optimization strategies are generated to get the corresponding desired distribution of capacity. For instance, the strategy of maximizing the total capacity is achieved by keeping the first term (i.e. $w_1 = 1, w_2 = 0$). On the contrary, leveraging the second ripple term only (i.e. $w_1 = 0, w_2 = 1$) leads to the flattest capacity distribution. And keeping these two terms with proper weights (i.e. $w_1 = 1, w_2 = 10$) will make the distribution both high and flat. It is important to introduce the weight of the ripple term because controlling

the capacity ripple of a band within a reasonable small range means fairness to different terminal users. This combination of two weights ($w_1 = 1, w_2 = 10$) is selected by pre-simulation to make sure that both terms are comparable to achieve the target of high and flat capacity.

Simulated annealing (SA) algorithm [22] is a method that can assist to approach the global optimal solution of one given function. Compared to the BFS technique, SA is a more effective way, especially when the variable space and ranges of constraints are large. As for other optimization algorithms like genetic algorithm and particle swarm optimization [23,24], they are inferior to the SA in terms of complexity and robustness. One special feature of the SA algorithm is that it will accept the worse solution with a certain probability in the process of the optimization, it can escape from the local optima and approach the global optima. The probability to accept worse solutions temporarily is related to the temperature parameter in the algorithm, and the higher the temperature, the higher the probability. The detailed pseudo-code of the SA algorithm is as follows

Algorithm 1 Process of SA algorithm

Input: $up, low, T_{max}, T_{min}, L, maxstay$

Output: $generation_x_{best}, generation_y_{best}$

```

1: Random initialization:  $x = \{x_i, x_i \in [low_i, up_i], i \in [1, 6]\}$ 
2: Initialization:
    $x_{current}, x_{best} = x$ 
    $y_{current}, y_{best} = Capacity(x)$ 
    $stay = 0, epoch = 0, T = T_{max}$ 
    $generation\_x_{best} = [], generation\_y_{best} = []$ 
3: while ( $T \geq T_{min}$ ) and ( $stay \leq maxstay$ ) do
4:   for  $j = 1 \rightarrow L$  do
5:     Update  $x_{new}$  via Eq. 13
6:      $y_{new} = Capacity(x_{new})$ 
7:      $df = y_{new} - y_{current}$ 
8:     if  $exp(-\frac{df}{T}) > rand(0, 1)$  then
9:        $x_{current} = x_{new}$ 
10:       $y_{current} = y_{new}$ 
11:      if  $df < 0$  then
12:         $x_{best} = x_{new}$ 
13:         $y_{best} = y_{new}$ 
14:      end if
15:    end if
16:  end for
17:   $epoch + = 1$ 
18:  Update  $T$  via Eq. 14
19:   $generation\_x_{best}.append(x_{best})$ 
20:   $generation\_y_{best}.append(y_{best})$ 
21:  if  $|generation\_y_{best}(end) - generation\_y_{best}(end - 1)| < 1e - 3$  then
22:     $stay + = 1$ 
23:  else
24:     $stay = 0$ 
25:  end if
26: end while

```

We develop the SA algorithm in MATLAB by referring to the python package, scikit-opt [25]. It is worthy to note that the x in the initialization process represents a random-generated

six-dimensional vector. The first three elements in x represent slopes of the power profile of three bands and they range from -1.5 dBm/THz to 1.5 dBm/THz. The last three are corresponding average power offsets, which vary from -13 dBm to -1 dBm. From the updating function in Eq. (13) we can see these two boundary terms strictly limit the generating range of parameters within the range mentioned above. The function $Capacity(x)$ represents using the GN model to calculate the NLI noise power and further to obtain the capacity with the input of variable, x , which is the set of parameters for optimization. The history of target optimization is stored in the variable of $generation_y_{best}$. And we use the Eq. (14) below to update the temperature

$$x_{new} = x_{current} + sign(r) \times T \times \left[\left(1 + \frac{1}{T}\right)^{|r|} - 1 \right] \times (up - low) \quad (13)$$

$$T = T_{max} \cdot exp[-exp(-1) \cdot epoch] \quad (14)$$

where the r in Eq. 13 is a randomly generated parameter, ranging from -1 to 1, which controls the sign of the updating increment of the input vector. And the T_{max} is the initial temperature, which should be an appropriate value because it takes a lot of time to converge if the T_{max} is set too high. Conversely, a lower T_{max} prevents accepting bad temporary updates. Thus we set $T_{max} = 300$ for its good performance.

3. Results and discussions

3.1. Setup

In this work, we leverage L (96 channels) + C (96 channels) + S (192 channels) band to carry signals. Based on the ITU-T G.694.1 [26], the frequency ranges of three bands are set as 185.975 THz ~190.775 THz, 191.275 THz ~196.075 THz, and 196.575 THz ~206.175 THz, respectively. Since the Nyquist-WDM system is applied, the symbol rate of signals is set as same as the channel spacing, 50 GHz. The guard bands between L- and C-band, C- and S-band are 500 GHz. For fiber network topology, we consider all the spans are identical with standard single-mode fiber (SSMF), and the length of each span is 80 km. The nonlinear coefficient $\gamma = 1.2$ W/km. Based on the ITU-T G.652D [27], the attenuation factors of L-, C-, and S-band are set as 0.21 dB/km, 0.20 dB/km, and 0.21 dB/km, respectively. The group velocity dispersion (GVD) parameter D is 17 ps/(nm · km) at 1550 nm which varies with the lambda with the dispersion slope S of 0.091 ps/(nm² · km), as shown in Eq. 15.

$$D(\lambda) = D_{1550} + S_{1550}(\lambda - 1550) \quad (15)$$

As mentioned in Section 2.1, two amplification schemes are used respectively. The first one possesses 2 EDFAs for L- and C-band and 1 TDFAs for S-band, separately. The second scheme uses two TDFAs and each amplifies half of the S-band beside two EDFAs. In the optimization process, we assume the NFs of OAs over bands are spectrally flat and their average values are 5 dB, 5 dB, and 6.5 dB for L-band, C-band, and S-band, separately. All parameters are listed in Table 1.

3.2. Results and discussions

As stated in Section 2.1, it is better to apply the dynamic iteration method to calculate the Raman gain. To observe its difference with the triangular approximation, we generate three Raman gain profiles with distinct input power distributions shown in Fig. 3. For the uniform power input with -10 dBm per channel shown in Fig. 3(a), we can see from Fig. 3(b) that the middle parts of Raman gain generated by the triangular assumption and dynamic iteration are quite close and keep good linearity. The error happens at outer band channels if the fixed Cr is used. When we input the tilt power shown in Fig. 3(e) with similar total power with the uniform distribution, the

Table 1. System parameters

Parameters	Values
Loss of L-band (α_l) [dB/km]	0.21
Loss of C-band (α_c) [dB/km]	0.20
Loss of S-band (α_s) [dB/km]	0.25
Dispersion at 1550 nm (D) [ps/nm/km]	17.0
Dispersion slope (S) [ps/nm ² /km]	0.091
Nonlinear coefficient (γ) [1/W/km]	1.2
Length of span (L_{span}) [km]	80
Symbol rate [Gbaud]	50
Channel bandwidth (B_{ch}) [GHz]	50
Channel spacing [GHz]	50
Bandwidth of guard bands (B_{gb}) [GHz]	500
Number of channels	384
Optical bandwidth (B_{tot}) [THz]	19.2
Reference wavelength [nm]	1454.2
Noise figure of L-band (NF_l) [dB]	5
Noise figure of C-band (NF_c) [dB]	5
Noise figure of S-band (NF_s) [dB]	6.5

Raman profile in Fig. 3(f) is basically the same as that shown in Fig. 3(b). We also increase the per-channel power from -10 dBm to -8 dBm as shown in Fig. 3(c), then the total power increased from 15.84 dBm to 17.84 dBm. From Fig. 3(d), it can be observed that the shape of Raman gain maintains the same. Meanwhile, the gain is equivalent to multiplying a rescaling factor. Although the deviations are quite low (about 0.5 dB at the outermost channels), it is better to apply this dynamic iteration method to decrease the error of the GN model.

Then, we test the performance of the SA on searching optimal parameters. The optimization history of five random trials is shown in Fig. 4(a). It can be seen that all the target values converge to almost the same value, although their initial values are different. Even more, we select the best set of parameters by finding the lowest target value and its corresponding capacity distribution. We then can compare this distribution with which generated by the parameters found by BFS, and illustrate the result in Fig. 4(b). For BFS, we set the searching grid of power offset and the power slope with the step of 2 dBm and 0.5 dBm/THz to control the traversing time in a reasonable period. Thus, 117649 sets of parameters are finally generated and tested. Even so, it is clear from Fig. 4(c) that the BFS is inferior to the SA algorithm in terms of time required and achievable average capacity ripple. It is indeed that the overall performance using BFS will be higher if the sweeping granularity becomes finer. However, the number of combinations to be swept will exponentially rise, and the required searching time is so long and impractical for optimization. Therefore, higher capacity and less optimization time validate the excellence of the SA algorithm.

In Section 2.1 and Section 2.3, it is mentioned that we perform the optimization by using two amplification schemes with 3 OAs and 4 OAs, respectively, and apply three kinds of optimization strategies by changing two weights shown in Eq. 12. Therefore, we compare the performance between two amplifying schemes and three strategies over three bands (96+96+192 channels) in Fig. 5(a). In terms of amplification scheme, results of 3 OAs schemes are represented by solid lines, and dash lines denote schemes with 4 OAs. Note that to optimize schemes with 4 OAs, the length of the parameter vector increases from 6 to 8 due to the newly added power slope and power offset. It can be seen that the capacity distributions obtained by these two schemes

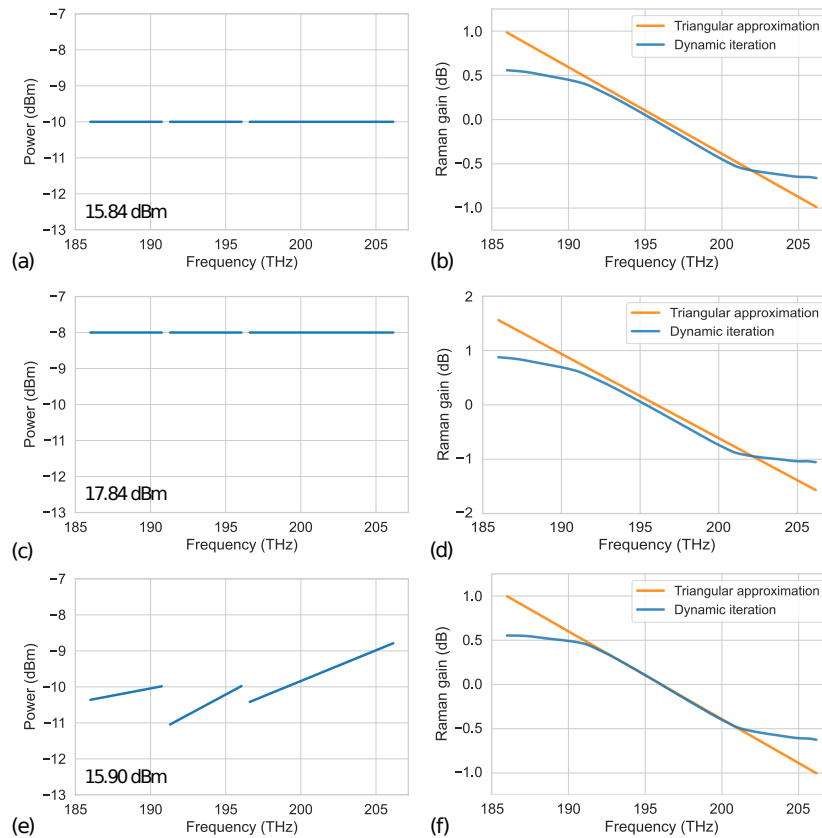


Fig. 3. (a) Uniform launch power with -10 dBm per channel and its corresponding (b) Raman gain profile generated by the triangular approximation (fixed C_r , orange line) and dynamic iteration (blue line); (c) Uniform launch power with -8 dBm per channel and its corresponding (d) Raman gain profile; (e) Tilt launch power and its corresponding (f) Raman gain profile

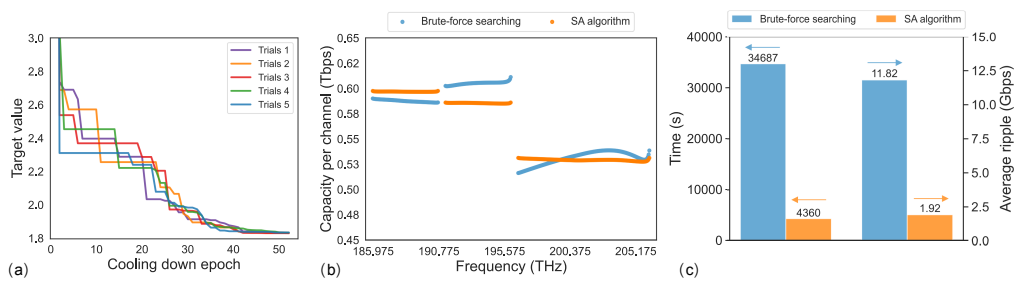


Fig. 4. (a) The optimization history of five trials, (b) the capacity distributions generated from two optimization algorithms, and (c) the comparison of time and average capacity ripple between BFS and SA algorithm

over different strategies are basically the same. We compared the total capacities between two amplifying schemes in Fig. 5(b). And the similar results over three strategies indicate that the number of OA has almost no effect on the overall capacity. Therefore, using one T DFA is

more cost-effective as long as the wavelength range of commercial TDFA can cover the whole S-band. And the power control strategy can be applied with different amplifying schemes based on practical requirements.

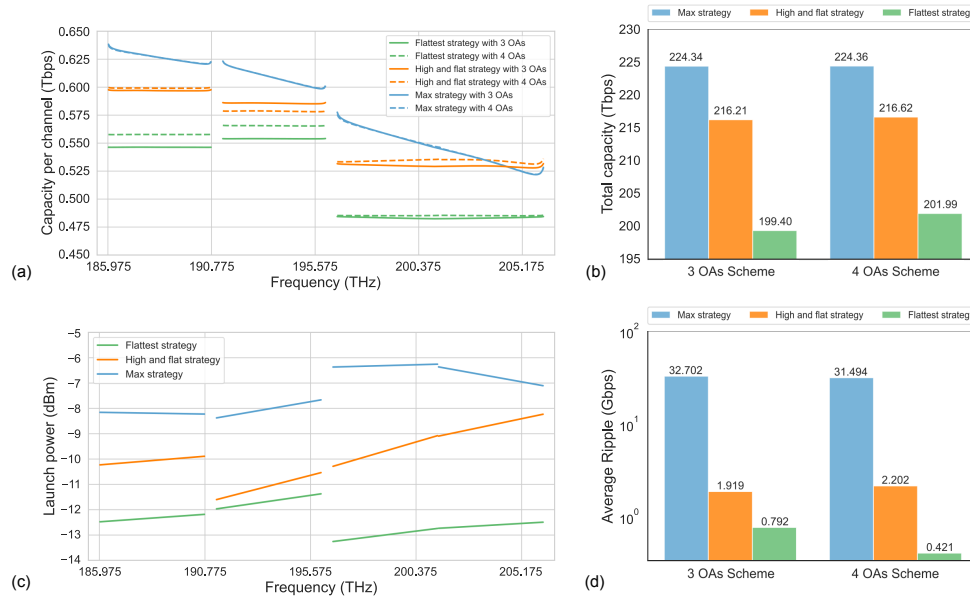


Fig. 5. (a) The capacity distributions and (b) the comparison of total capacity and (d) the average ripples with two kinds of amplification schemes, and (c) the launch power distribution (4 OAs scheme) generated by three strategies

For the SA algorithm, the optimization strategy is more significant due to its huge effect on the distribution of capacity over channels. Take the performance of three strategies on the 3 OAs scheme as an example, it is obvious that although the maximizing strategy gets the highest total capacity (224.34 Tbps in Fig. 5(b)), the capacity distribution over channels is tilted and the average ripple is large (32.7 Gbps in Fig. 5(d)) which is undesired in the WDM system because it is unfair for users at different terminals. On the contrary, applying the flattening strategy satisfies the fairness principle within the band (0.792-Gbps ripple), but it can only get a much lower total capacity (199.4 Tbps). The third strategy combines the advantages of the first two strategies with only 8 Tbps capacity decrease (216.21 Tbps) compared to the maximizing strategy and relatively good flatness (1.92-Gbps ripple), which can be seen in Fig. 5(d).

Finally, we show the distributions of launch power over the whole band generated by the parameters found from three strategies with the 4 OAs scheme in Fig. 5(c). It is observed that for power distributions found by high and flat strategy (orange line) and flattest strategy (green line), all slopes are positive which agrees with the conception in Section 2.1 about amplifying high-frequency channels to higher launch powers. However, the power distribution generated by the max strategy is counter-intuitive with a negative slope at the right half of the S-band. This indicates optimizing power distribution by BFS at a positive range of slope doesn't work for maximizing capacity strategy. Besides this, the irregular mean power of each band also violates our initial idea of resisting Raman linear gain by setting a high-frequency channel with a higher power. This proves the capability of SA based on the target function without relying on unreliable presuppositions.

Due to the limitation of experimental conditions, this work is mainly based on simulation. But the ability to optimize a set of parameters for launch power control with arbitrary optical

system architecture in a few minutes is valuable. Indeed, more works need to be done under scenarios with reconfigurable optical add-drop multiplexer (ROADM) over longer transmission mesh networks, which are left for future research.

4. Conclusions

In this work, we propose a power control technique by leveraging the SA algorithm. A set of optimum parameters, describing the launch power profile, can be searched to make the total capacity high and meanwhile keep the capacity distribution flat over bands. We demonstrate that using the SA algorithm is much faster than using the BFS scheme and higher achievable transmission capacity can be obtained as well. It is observed that a trade-off exists between the total capacity and its spectral flatness over bands by comparing three optimization strategies. Finally, the launch power profile is illustrated whose pattern can guide the design of the optical line system. That is to say, this method can be applied with arbitrary configurations of the UWB WDM systems according to the practical requirements, which is time-saving and promising when planning new point-to-point links and mesh networks.

Funding. National Key Research and Development Program of China (2018YFB1800902); National Natural Science Foundation of China (62005030); Research Grants Council, University Grants Committee (15200718, 15211619).

Acknowledgment. The authors would like to thank Dr. Qirui Fan and Chuang Xu for their kind help.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

1. H. Chen and A. T. Koonen, "Spatial division multiplexing," in *Fibre Optic Communication*, (Springer, 2017), pp. 1–48.
2. M. Nebeling and H. J. Thiele, *Coarse wavelength division multiplexing: technologies and applications* (CRC press, 2018).
3. H. Yuan, M. Furdek, A. Muhammad, A. Saljoghei, L. Wosinska, and G. Zervas, "Space-division multiplexing in data center networks: on multi-core fiber solutions and crosstalk-suppressed resource allocation," *J. Opt. Commun. Netw.* **10**(4), 272–288 (2018).
4. G. Rademacher, B. J. Puttnam, R. S. Luís, T. A. Eriksson, N. K. Fontaine, M. Mazur, H. Chen, R. Ryf, D. T. Neilson, P. Sillard, F. Achten, Y. Awaji, and H. Furukawa, "Peta-bit-per-second optical communications system using a standard cladding diameter 15-mode fiber," *Nat. Commun.* **12**(1), 4238 (2021).
5. G. Rademacher, B. J. Puttnam, R. S. Luís, J. Sakaguchi, W. Klaus, T. A. Eriksson, Y. Awaji, T. Hayashi, T. Nagashima, T. Nakanishi, T. Taru, T. Takahata, T. Kobayashi, H. Furukawa, and N. Wada, "10.66 Peta-Bit/s Transmission over a 38-Core-Three-Mode Fiber," in *Optical Fiber Communication Conference (OFC) 2020*, (Optical Society of America, 2020), p. Th3H.1.
6. D. J. Richardson, J. M. Fini, and L. E. Nelson, "Space-division multiplexing in optical fibres," *Nat. Photonics* **7**(5), 354–362 (2013).
7. J.-X. Cai, H. Batshon, M. V. Mazurczyk, C. R. Davidson, O. V. Sinkin, D. Wang, M. Paskov, W. Patterson, M. A. Bolshtyansky, and D. G. Foursa, "94.9 Tb/s single mode capacity demonstration over 1,900 km with C+L EDFAs and coded modulation," in *2018 European Conference on Optical Communication (ECOC)*, (IEEE, 2018), pp. 1–3.
8. F. Hamaoka, K. Minoguchi, T. Sasai, A. Matsushita, M. Nakamura, S. Okamoto, E. Yamazaki, and Y. Kisaka, "150.3-Tb/s Ultra-Wideband (S, C, and L Bands) Single-Mode Fibre Transmission over 40-km Using > 519 Gb/s/A PDM-128QAM Signals," in *2018 European Conference on Optical Communication (ECOC)*, (IEEE, 2018), pp. 1–3.
9. S. Okamoto, K. Minoguchi, F. Hamaoka, K. Horikoshi, A. Matsushita, M. Nakamura, E. Yamazaki, and Y. Kisaka, "A study on the effect of ultra-wide band WDM on optical transmission systems," *J. Lightwave Technol.* **38**(5), 1061–1070 (2019).
10. J. Renaudier, A. C. Meseguer, A. Ghazisaeidi, P. Tran, R. R. Muller, R. Brenot, A. Verdier, F. Blache, K. Mekhazni, B. Duval, H. Debregeas, M. Achouche, A. Boutin, F. Morin, L. Letteron, N. Fontaine, Y. Frignac, and G. Charlet, "First 100-nm Continuous-Band WDM Transmission System with 115Tb/s Transport over 100km Using Novel Ultra-Wideband Semiconductor Optical Amplifiers," in *2017 European Conference on Optical Communication (ECOC)*, (2017), pp. 1–3.
11. K. Minoguchi, S. Okamoto, F. Hamaoka, A. Matsushita, M. Nakamura, E. Yamazaki, and Y. Kisaka, "Experiments on stimulated Raman scattering in S-and L-bands 16-QAM signals for ultra-wideband coherent WDM systems," in *Optical Fiber Communication Conference*, (Optical Society of America, 2018), pp. Th1C-4.

12. G. Saavedra, D. Semrau, M. Tan, A. Iqbal, D. J. Elson, L. Galdino, P. Harper, R. I. Killey, and P. Bayvel, "Inter-channel stimulated Raman scattering and its impact in wideband transmission systems," in *2018 Optical Fiber Communications Conference and Exposition (OFC)*, (IEEE, 2018), pp. 1–3.
13. F. Hamaoka, M. Nakamura, S. Okamoto, K. Minoguchi, T. Sasai, A. Matsushita, E. Yamazaki, and Y. Kisaka, "Ultra-wideband WDM transmission in S-, C-, and L-bands using signal power optimization scheme," *J. Lightwave Technol.* **37**(8), 1764–1771 (2019).
14. B. Correia, R. Sadeghi, E. Virgillito, A. Napoli, N. Costa, J. Pedro, and V. Curri, "Power control strategies and network performance assessment for C+ L+ S multiband optical transport," *J. Opt. Commun. Netw.* **13**(7), 147–157 (2021).
15. B. Correia, R. Sadeghi, E. Virgillito, A. Napoli, N. Costa, J. Pedro, and V. Curri, "Optical Power Control Strategies for Optimized C+ L+ S-bands Network Performance," in *2021 Optical Fiber Communications Conference and Exhibition (OFC)*, (IEEE, 2021), pp. 1–3.
16. P. Poggiolini, G. Bosco, A. Carena, R. Cigliutti, V. Curri, F. Forghieri, R. Pastorelli, and S. Piciaccia, "The LOGON strategy for low-complexity control plane implementation in new-generation flexible networks," in *2013 Optical Fiber Communication Conference and Exposition and the National Fiber Optic Engineers Conference (OFC/NFOEC)*, (IEEE, 2013), pp. 1–3.
17. M. Zirngibl, "Analytical model of Raman gain effects in massive wavelength division multiplexed transmission systems," *Electron. Lett.* **34**(8), 789–790 (1998).
18. D. Semrau, R. I. Killey, and P. Bayvel, "A closed-form approximation of the Gaussian noise model in the presence of inter-channel stimulated Raman scattering," *J. Lightwave Technol.* **37**(9), 1924–1936 (2019).
19. I. Roberts, J. M. Kahn, J. Harley, and D. W. Boertjes, "Channel power optimization of WDM systems following Gaussian noise nonlinearity model in presence of stimulated Raman scattering," *J. Lightwave Technol.* **35**(23), 5237–5249 (2017).
20. D. Semrau, R. I. Killey, and P. Bayvel, "The Gaussian noise model in the presence of inter-channel stimulated Raman scattering," *J. Lightwave Technol.* **36**(14), 3046–3055 (2018).
21. P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, "The GN-model of fiber non-linear propagation and its applications," *J. Lightwave Technol.* **32**(4), 694–721 (2013).
22. S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science* **220**(4598), 671–680 (1983).
23. S. Mirjalili, "Genetic algorithm," in *Evolutionary algorithms and neural networks*, (Springer, 2019), pp. 43–55.
24. M. Clerc, *Particle swarm optimization*, vol. 93 (John Wiley & Sons, 2010).
25. F. Guo, "Scikit-opt: simulated annealing algorithm," Github, (2021), <https://github.com/guofei9987/scikit-opt/blob/master/sko/SA.py>.
26. International Telecommunication Union, G.694.1 : Spectral grids for WDM applications: DWDM frequency grid (2020). <https://www.itu.int/rec/T-REC-G.694.1-202010-I/en>.
27. International Telecommunication Union, G.652 : Characteristics of a single-mode optical fibre and cable (2016). <https://www.itu.int/rec/T-REC-G.652-201611-I>.