

# Optimization of 125- $\mu\text{m}$ Heterogeneous Multi-Core Fibres Design using AI

Xun Mu, Alessandro Ottino, Filipe M. Ferreira, Georgios Zervas

**Abstract**—We propose an automated heterogeneous trench-assisted multi-core fibre (MCF) design method. This method uses neural networks to speed up coating loss estimation by  $\sim 10^6$  times and using particle swarm optimization (PSO) algorithm to explore the optimal MCF design under various objectives and properties constraints. The latter reduces the permutation evaluations by ten orders of magnitude compared with the brute force method. The artificial intelligence (AI)-based method is used to design MCFs on two objectives: minimizing crosstalk (XT) and maximizing effective mode area ( $A_{\text{eff}}$ ). By optimizing XT with different  $A_{\text{eff}}$  and cutoff wavelength constraints combinations for 6-core fibres, we achieved -92.1 dB/km ultra-low XT for C+L band fibre and -64 dB/km for E+S+C+L-band fibre. Meanwhile, we explored the upper limit of  $A_{\text{eff}}$  given different bandwidth constraints resulting in a 6.82 relative core multiplicity factor. We performed capacity analysis of fibres for two transmission lengths. It is shown that bandwidth is the dominant factor while the increase brought by  $A_{\text{eff}}$  and the penalty caused by XT are relevantly small. Our fibres exceed the cutoff-limited capacity of the 7-core fibre in literature by 35.1% and 84.8% for 1200 km and 6000 km transmission respectively.

**Index Terms**—Space division multiplexing (SDM), Multi-core fibre (MCF), Particle swarm optimization (PSO), crosstalk, wide-band transmission, long-haul transmission

## I. INTRODUCTION

SPACE division multiplexing (SDM), as one of the methods to improve the transmission capacity of optical networks, has attracted intensive research interest in the past decade [1]. The use of SDM with multi-core fibres (MCFs) has revealed its ability to improve the network performance while delivering increased energy and cost efficiency [2], [3]. MCF interconnect systems have shown the potential to increase the front panel density [4] with silicon photonic on-board transceivers [5] without fan-in/out or core pitch conversion devices. Some uncoupled MCFs

X.Mu (e-mail: xun.mu.17@ucl.ac.uk), A.Ottino, F.M.Ferreira and G.Zervas are with the Optical Network Group, University College London, London, WC1E 7JE,UK

Manuscript received...

in standard cladding diameter are reported to have higher capacity than conventional single-mode fibre (SMF). For example, the use of a 4-core fibre in 125  $\mu\text{m}$  cladding diameter has led to 118.5 Tbit/s over 316 km in [6] which is around 4 times the capacity of a SMF calculated with the same transmission parameters (26.22 Tbit/s).

While there has been significant research on MCFs design, design strategies are often presented in an ad-hoc fashion by omitting critical reasoning and underlying strategies such as in [7]. Also, it is often the case that optimisation only considers a subset of the parameters that analytically define the refractive-index profile. For example, in [8], for a W-type pure silica core, layer widths are pre-selected and set as constants. Only the relative refractive index difference between core and the depressed layer ( $\Delta n$ ), and core pitch (CP), which is the distance between the center of two cores, are independent variables to calculate the inter-core crosstalk (IC-XT) and the excess loss. Then, according to the requirement of IC-XT and the excess loss threshold, a 2-dimensional search space ( $\Delta n$ , CP) is defined, over which optimal fibre design search takes place. A *core selection method* is proposed in [9] where the refractive-index profile of the core at the outer and inner layers is decided one by one.

However, to approach an optimal MCF structure all the fibre parameters need to be considered together. Taking all the parameters into consideration while controlling the optical properties makes the MCF design multi-dimensional and multi-constraint problem. Targeting such complex problem with a brute force method would be time-consuming and impractical especially for heterogeneous MCFs (Hete-MCFs). Instead, we propose using artificial intelligence to optimize the design of Hete-MCFs over all refractive-index profile parameters. Particle swarm optimization (PSO) algorithm [10], [11], as one of the most important swarm intelligence paradigms, is chosen to realize this work. PSO has been recently proven to optimize the on-off switch operation of semiconductor optical amplifiers, also a

multi-dimensional problem [12] and to have better performance and generalisability compared to Genetic Algorithm (GA) and Ant Colony Optimization (ACO).

In our previous work [13], we designed several 6-core fibres for minimizing IC-XT using AI-based technique. Herein we further developed our AI-based technique and make thorough comparison and analysis of the results on both XT and transmission:

- The design process is explained in details, including PSO algorithm, coating loss estimation, the objective function design and the SNR and capacity calculation.
- The AI-based technique is tested on MCF design with two different objectives: *XT-optimization* and *A<sub>eff</sub>-optimization*.
- SNR results are shown for different cutoff wavelength and two objectives.
- The capacity of the optimized MCFs and the capacity per core are compared with the fabricated MCF and the SMF.
- The capacity of the MCFs in *A<sub>eff</sub>-optimization* are calculated for 1200 km and 6000 km. MCFs performance with different transmission lengths are explored.
- The influence of *A<sub>eff</sub>* and XT on the capacity are analyzed quantitatively.

In this paper, we explore the design of heterogeneous trench assisted 6-core MCFs on a standard 125  $\mu\text{m}$  cladding diameter. A single-ring layout is chosen to prevent the cutoff wavelength elongation and IC-XT degradation associated with having a central core. To reduce the design and fabrication complexity we consider a symmetric fibre design with  $N/2$  same pairs of dissimilar cores - with even  $N$ . Each core (in the pair) can be described by five design parameters defining the refractive-index profile while fixing the cladding as pure silica. The parameters are optimised within the domain compatible with current fabrication constraints and limited resolution is assumed for easy fabrication as shown in Table. I. Despite of this simplification, brute force would be unpractical, for instance, considering discrete values per parameter leads to more than  $3.7 \times 10^{15}$  permutations. By adjusting the objective function and corresponding fitness value, PSO can be utilized to search for near-optimal Hete-MCF structures with various conditions and constraints of mechanical and optical properties while significantly reduce the number of permutations evaluated compared to the brute force method (from  $\sim 10^{15}$  to  $\sim 10^5$ ). Moreover, the computation of the attenuation degradation caused by

coating [7] which we noted as coating loss (CL), is speed up (from  $\sim 100$  s to  $\sim 50$   $\mu\text{s}$ ) using a statistical classifier and a regressor based on neural networks.

Up to now, the capacity of the MCFs has not been taken into account quantitatively during the design process but rather measured experimentally upon fabrication. Capacity of a single-mode MCF link is directly related to the bandwidth of each fibre core, the number of cores supported and the signal-to-noise ratio (SNR) limited by linear and nonlinear effects. The achievable SNR is limited by the nonlinear interference (NLI) noise and IC-XT introduced by the link fibre and by the accumulative spontaneous emission (ASE) noise introduced by the link optical amplifier. To optimize the capacity of MCF is a multi-dimensional problem. It involves complex relationships among the fibre parameters that either directly or indirectly influence the three main factors: core number, cutoff wavelength (i.e. bandwidth), and SNR, as described in Section III-D. Specifically, it is the refractive-index profile of fibre core and the core distribution which directly determine the cutoff wavelength, the modal effective area ( $A_{\text{eff}}$ ) ( $\propto 1/\text{NLI}$ ) and IC-XT. Furthermore, we explore how these influence the theoretical maximum capacity. In this work, using fabrication viable constraints we designed MCFs using PSO to optimize IC-XT or  $A_{\text{eff}}$  with different cutoff wavelengths while satisfying key constraints (non-overlap and CL) such that how the capacity is influenced by the parameters is explored.

In the following, we will be focusing on heterogeneous trench-assisted MCF (Hete-TA-MCF). The trench-assisted refractive-index profile has stronger light confinement ability than the conventional step-index while it is easier to fabricate compared with the hole-assisted or rod-assisted structures. For the heterogeneous structure, the effective refractive index ( $n_{\text{eff}}$ ) difference between the cores leads to propagation constant difference. The influence of this propagation constant difference on the power coupling between a core pair depends on the bending radius ( $R_b$ ). When  $R_b$  exceeds the critical bending radius, it can benefit from the non-phase matching and get a low XT.

The automated method we propose here uses the PSO algorithm to explore the optimal MCF structures for two objectives, a) minimal overall XT under certain constraints on  $A_{\text{eff}}$  and cutoff wavelength, and b) maximal  $A_{\text{eff}}$  with certain cutoff wavelength constraint for MCFs in standard cladding diameter with silica cladding. The output of the PSO feeds the inter-channel stimulated Raman scattering (ISRS) Gaussian noise (GN) model [14] to calculate the SNR

and in turn the capacity of the fibre. The obtained results are used to showcase the influence of individual parameters and are also compared to the existing MCFs to validate the effectiveness of the method.

The rest of this paper is structured as follows: in Section II, MCF's layout and the corresponding refractive-index profile utilized in this work are described in detail. Then, Section III introduces the whole automatic MCF design process for two objectives. The PSO performance and optimization results' optical properties for 6-core fibres are presented in Section IV. Their transmission performance are illustrated and discussed in term of SNR and capacity in Section V. Finally, the work conclusions are included in Section VI.

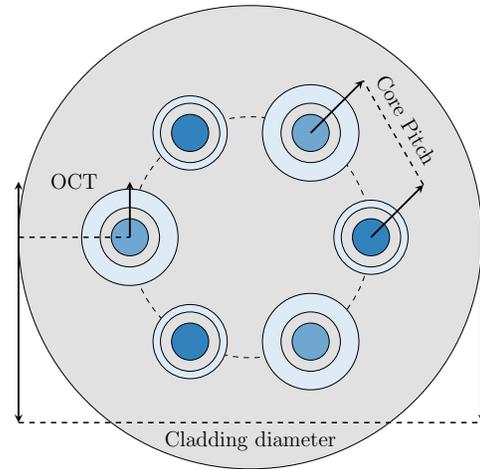
## II. MCF PROFILE AND DESIGN STRATEGIES

The schematic diagrams of Hete-TA-MCF and its refractive-index profile are shown in Fig. 1, for the 6-core fibre. Making use of the symmetry to reduce the search space and the complexity of fibre design, we assumed two types of cores. They all have neighbouring cores non-identical to themselves so that it can benefit from the  $n_{\text{eff}}$  difference. The cores have fixed position on the ring uniformly. The core pitch is set at  $32.5 \mu\text{m}$  for 6-core fibres since the outer cladding diameter is set to be  $30 \mu\text{m}$  [15], larger values would lead to non negligible coating loss. With various layer widths, the fibre should meet the mechanical feasibility conditions:

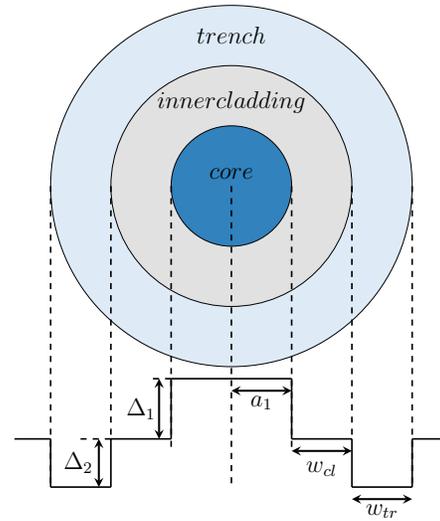
- 1) The distance between the adjacent trenches' edges is larger than or equal to  $2 \mu\text{m}$  to prevent the cores overlapping with each other [16].
- 2) The coating loss needs to be less than  $0.001 \text{ dB/km}$  [17] which will be estimated by classifier and regressor in Section III (B).

In Fig. 1(b),  $a_1$ ,  $w_{\text{cl}}$  and  $w_{\text{tr}}$  denote core radius, width of inner cladding and trench width, respectively.  $a_2$  is the distance from the core center to the inner cladding edge and  $a_3$  is the distance from the core center to the trench edge.  $n_{\text{core}}$ ,  $n_{\text{cl}}$  and  $n_{\text{tr}}$  stand for the refractive index of the core, cladding and trench, respectively.  $\Delta_1$  and  $\Delta_2$  are the relative refractive index difference between core and cladding, trench and cladding, respectively.

We assume that the core is doped with germanium, the trench is doped with fluorine [18], and both the inner and outer cladding index are fixed as the index of pure silica. There are the five refractive-index profile parameters which become variables for each core, including the layer widths ( $a_1$ ,  $w_{\text{cl}}$ ,  $w_{\text{tr}}$ ) and the relative refractive index differences ( $\Delta_1$  and  $\Delta_2$ ). Since



(a)



(b)

**Figure 1:** (a) 6-core heterogeneous trench-assisted fibre layout with two types of core; (b) Refractive-index profile diagram in heterogeneous trench-assisted multi-core fibre (OCT: outer cladding diameter)

it is assumed that two neighbouring cores have non-identical refractive-index profiles, there are 10 variables in total to be optimized for each MCF structure. According to the fabricated MCF core index profiles [8], [19]–[21], the parameters are limited in the certain ranges as shown in Table I.

IC-XT in MCFs is regarded as the power leakage from the adjacent cores to the target core, which can be affected by the coupling coefficient [7]. As for the Hete-TA-MCF, the mode coupling coefficient between two TA non-identical cores can be analytically

**Table I:** Index Profile Parameters Range

Parameters	Range	Step	Choices
$a_1$ [ $\mu\text{m}$ ]	4 ~ 6	0.1	21
$w_{cl}$ [ $\mu\text{m}$ ]	2.5 ~ 7.5	0.1	51
$w_{tr}$ [ $\mu\text{m}$ ]	2.5 ~ 7.5	0.1	51
$\Delta_1$	0.3% ~ 0.6%	0.01%	31
$\Delta_2$	-0.7% ~ -0.35%	0.01%	36

expressed as Eq. (33) in [22]. Considering a MCF which is bent at a constant radius and is twisted continuously at a constant rate, its average PCC ( $h_{pq}$ ) between a core pair (e.g. core  $p$  and core  $q$ ) over a twist pitch is expressed as Eq. (12) in [23]. When the MCF is homogeneous,  $h_{pq}$  are the same to all the neighbouring core pairs.

Taking the coupled-power theory into consideration, the IC-XT of core  $p$  for Hete-MCFs with all the neighbouring cores excited is given by [1]:

$$IC-XT_p = \frac{n - \sum_{i=1}^n e^{-(n+1)h_{pq}L}}{1 + \sum_{i=1}^n e^{-(n+1)h_{pq}L}}, \quad (1)$$

where  $p$  is the target core,  $q$  is any of its  $n$  neighbors, and  $L$  is the fibre length. This XT model has been verified against experimental values for fabricated MCFs [19]. The mean of the IC-XT in Eq. 1 over all the cores is considered as the optimization indicator denoted as overall XT.

### III. AUTOMATED FIBRE DESIGN PROCESS

One of fibre design strategies is the constraint to single-mode operation for preventing the inter-modal crosstalk — defined by the cutoff condition. The cutoff wavelength can be assessed by checking the cutoff condition at different wavelengths, for instance, 1530 nm, 1460 nm and 1360 nm. Thus, bending loss of LP01 should be smaller than 0.5 dB/100turn when  $R_b=30$  mm at 1625 nm in ITU-T recommendations G.655 and G.656 while bending loss of LP11 need to be larger than 1 dB/m when  $R_b=140$  mm [24], [25]. The bending loss is calculated with the Maxwell equations' solutions described in [26]. The  $n_{\text{eff}}$  and  $A_{\text{eff}}$  are calculated by solving the Maxwell equations numerically using the method described in [27].

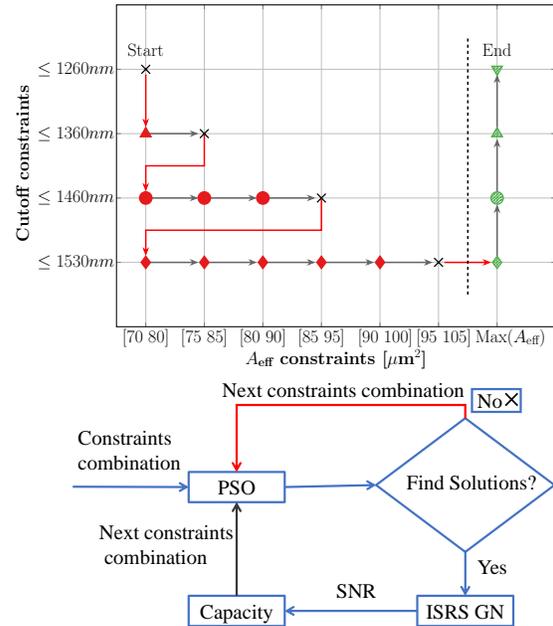
We set four cutoff wavelength targets:

1) : cutoff  $\leq 1530$  nm, MCF can cover C+L band;

- 2) : cutoff  $\leq 1460$  nm, MCF can cover S+C+L band;  
 3) : cutoff  $\leq 1360$  nm, MCF can cover E+S+C+L band;  
 4) : cutoff  $\leq 1260$  nm, MCF can cover O+E+S+C+L band.

We conduct *XT-optimization* with PSO using cutoff wavelength constraints and  $A_{\text{eff}}$  constraints. In the XT-optimization, we use five  $A_{\text{eff}}$  target constraints in the range of 70-100  $\mu\text{m}^2$ . We initialised the process by running the PSO under the shortest cutoff wavelength constraint and the smallest  $A_{\text{eff}}$  constraint, this is: cutoff  $\leq 1260$  nm and  $A_{\text{eff}} \in [70 \ 80]$   $\mu\text{m}^2$ . Once the *XT-optimization* is complete, the *A<sub>eff</sub>-optimization* follows with PSO using the associated objective function. The maximum  $A_{\text{eff}}$  is searched with PSO one by one for four cutoff wavelength constraints. The design process traverses the constraint combinations to optimize the MCF with PSO.

As illustrated in Fig. 2, the black cross markers indicate those cutoff wavelength and  $A_{\text{eff}}$  constraints combinations at which PSO has not found solutions after ten runs (e.g. cutoff  $\leq 1360$ nm and  $A_{\text{eff}} \in [75 \ 85]$   $\mu\text{m}^2$ ) satisfying all the constraints. The PSO then



**Figure 2:** Automated fibre design process through different cutoff and  $A_{\text{eff}}$  constraint combinations. The flowchart illustrates the actions taken for each combination. Red and black arrows (both in flowchart and plot) indicate the transition following a unsuccessful and successful fibre design respectively.

considers the next constraint combination without SNR and capacity calculation. The green markers represent the successful runs using *XT-optimization* while the red markers are representing the successful runs using the *A<sub>eff</sub>-optimization*. Following this, the optimal fibre structure is then passed to ISRS GN model for SNR and capacity calculation.

#### A. Particle Swarm Optimization Algorithm

Each particle  $\mathbf{x}_p = (x_{p1}, x_{p2}, \dots, x_{pm})$  in the swarm can be regarded as one potential solution (an MCF structure) with includes  $m = 10$  dimensions (variables) as we mentioned in Section. II. With each solution, there is a corresponding fitness value calculated with the objective function. It is utilized to indicate how this solution performs. By comparing the fitness values, firstly the particle needs to think as an independent individual about the best solution that the particle itself finds so far which is denoted as  $\mathbf{p}_{best}$ . Simultaneously, as a social group, particles communicate with each other. The best solution obtained with the global group's discussion is called  $\mathbf{g}_{best}$ . Both self-thinking and social communication influence the final decision on how to optimize. The optimization direction is described by the velocity  $\mathbf{v}_p = (v_{p1}, v_{p2}, \dots, v_{pm})$ . During the optimization process, each particle adjusts appropriately its optimization direction at each iteration. In [28], Y. Shi et al. introduced an inertia factor,  $\omega$ , which changes with the iteration dynamically to improve the local search precision. In this work, it decreases from 1 to 0.1 linearly from the beginning to the maximal iterations.

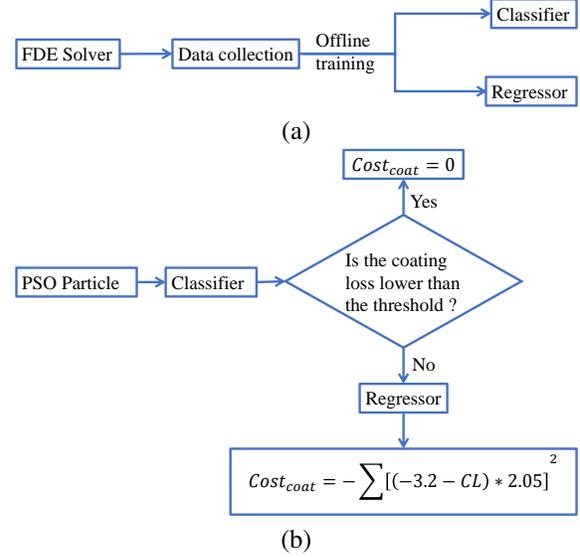
$$\omega_i = \omega_{max} - \frac{i(\omega_{max} - \omega_{min})}{i_{max}}, \quad (2)$$

$$\mathbf{v}_p^i = \omega \mathbf{v}_p^{i-1} + c_1 \cdot \mathbf{s} \cdot (\mathbf{p}_{best}^{i-1} - \mathbf{x}_p^{i-1}) + c_2 \cdot \mathbf{s} \cdot (\mathbf{g}_{best}^{i-1} - \mathbf{x}_p^{i-1}), \quad (3)$$

$$\mathbf{x}_p^i = \mathbf{x}_p^{i-1} + \mathbf{v}_p^i. \quad (4)$$

$(\mathbf{p}_{best}^{i-1} - \mathbf{x}_p^{i-1})$  is cognition component and  $(\mathbf{g}_{best}^{i-1} - \mathbf{x}_p^{i-1})$  is social component. Let  $\mathbf{s}$  be a  $1 \times m$ -dimension matrix of realisations of the random variable  $S \sim \mathcal{N}(0, 1)$ . The constants  $c_1$  and  $c_2$  determine the weight and in turn the influence of the two components on the velocity. They are termed as learning rate. Generally  $c_1 = c_2 = 2$ .  $\mathbf{v}_{max}$  is used to constrain the velocity [29].  $\mathbf{v}_{min} = -\mathbf{v}_{max}$  and  $\mathbf{v}_{max} = 0.2 * (ub - lb)$  for providing 20% dynamic range for particles' activity.

In *XT-optimization*, particle number is set as 100 while the maximal iteration number is 500. In *A<sub>eff</sub>-optimization*, particle number is set as 200 while



**Figure 3:** (a) FDE solver is used to collect coating loss data. The data are used to offline train the classifier and regressor; (b) Coating loss estimation process: the  $Cost_{coat}$  is used in Eq. (5) in section. III(C).

the maximal iteration number keeps the same. We adopted a convergence criterion to judge the particles' status: if more than 99% particles hold the same value as the  $\mathbf{g}_{best}$ , we consider the swarm loses its exploration ability and converges then the optimization process is stopped.

#### B. Coating Loss Classifier and Regressor

The coating loss is defined as the outermost core's bending loss of LP01 at 1625 nm with  $R_b=140$  mm with coating index as 1.475 [17]. Using the finite difference eignemode (FDE) solver in Lumerical [30] to calculate coating loss, each calculation would take around 100 s. FDE solver can solve the mode and calculate the loss by comparing  $n_{eff}$  with or without the bending. To speed up the coating loss computation, we trained a statistical classifier and a regressor to replace FDE solver in the design process.

7000 coating loss data are collected with FDE solver for training the neural networks (both classifier and regressor) as illustrated in Fig. 3(a). The data are split into three parts: 50% for training, 25% for validation and 25% for test. Both the classifier and the regressor were implemented using a 4-layer fully-connected neural network with the relu activation function [31] and the Adam optimizer [32].

The four layers of the classifier have 20, 10, 5 and 1 neurons respectively. The green part illustrates the

region where the coating loss is lower than the acceptable threshold and the core structure should be classified as 1. The orange represents where the coating loss is above the threshold and the core structure should be classified as 0 in Fig. 4(a). The classifier demonstrated a 99.92% accuracy for the whole dataset. All 0.08% errors merely appear in the green region as illustrated in Fig. 4 (b) translating to few false negative results. Two red dashed line indicates the error interval  $[-3.08 -3.03]$ . And there is no error in orange region meaning 0% false positives. Thus, there are no output from the classifier that appear to be valid (below the acceptable coating loss) but are actually above it. The four layers of the regressor have 64, 32, 16 and 1 neurons respectively. The regressor resulted in a mean square error of  $2.66e-04$ .

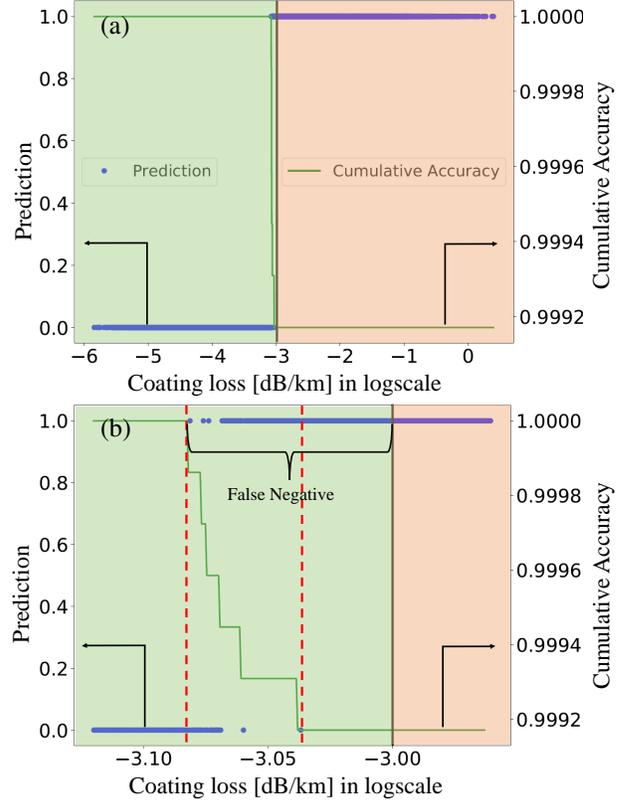
The coating loss of the PSO designed fibres is also cross-verified with Lumerical. All the cores hold the coating loss smaller than the threshold. Classifier inference time is  $14 \mu s$  and regressor inference time is  $32 \mu s$  on Nvidia V100 GPU for single calculation. Compared with estimation time using FDE solver, the classifier and regressor are  $10^6$  times faster.

In the coating loss estimation in Fig. 3(b), the classifier checks whether or not the coating loss of the core is higher than 0.001 dB/km. If it is higher than the threshold (being classified as 1 (negative) by the classifier), the regressor is utilized to calculate the magnitude of the coating loss and the output of the regressor is used in the fitness value calculation to make the coating loss penalty  $Cost_{coat}$  in the objective function continuous. The  $Cost_{coat}$  is only used in the fitness value calculation for the invalid particles in Eq. 5 that appears in the following section. For the valid particles, since the  $Cost_{coat} = 0$ , the fitness value is not influenced by the coating loss.

### C. Objective Function

Objective function contains all the design requirements. It returns the fitness value to PSO for every particle so that PSO can adjust the optimization direction properly according to the fitness value. We use two objective functions, 1) focusing on optimizing XT with strict constraints on  $A_{eff}$  and cutoff wavelength that we call it *XT-optimization* and 2) purely aiming at maximizing  $A_{eff}$  for certain cutoff wavelength and we call it  *$A_{eff}$ -optimization*.

The objective function checks the following constraints in order. The first two are the essential conditions: non-overlap and coating loss. Hence, if they



**Figure 4:** (a) The prediction performance and the cumulative accuracy of the coating loss classifier; (b) Zoom-in of (a) around the threshold.

are not satisfied, the particle will be treated as invalid and the fitness value will be returned as

$$Fitness\ value = Cost_{overlap} + Cost_{coat} \quad (5)$$

When they are both satisfied, the valid particles have different objective functions in *XT-optimization* and  *$A_{eff}$ -optimization*. The fitness value of PSO in the *XT-optimization* case is

$$Fitness\ value = XT + Cost_{A_{eff}} + Cost_{cutoff} \quad (6)$$

The  *$A_{eff}$ -optimization* Fitness value is

$$Fitness\ value = - \sum A_{eff} + Cost_{cutoff} \quad (7)$$

The detail of the cost terms are described below.

- 1) If the distance between the adjacent trench is smaller than  $2 \mu m$  [22],  $overlap_{pq} = CP_{pq} - (a_{3,p} + a_{3,q} + 2)$  will be smaller than zero.  $p$  and  $q$  are the core ID of the neighbouring core pair. If  $overlap_{pq}$  is below zero, the fitness value suffers a penalty:  $Cost_{overlap} = (overlap_{pq} * 1e5)^2$ . Otherwise,

$$Cost_{overlap} = 0.$$

- 2) We firstly use the classifier to check if the outer core has a coating loss higher than the threshold. If the output of classifier is higher than 0.5, we use the regressor to estimate the magnitude of coating loss and the fitness value suffers a penalty:  $Cost_{coat} = (\sum (-3.2 - CL) * 2e5)^2$  in which  $CL$  is the regressor's output for each core.
- 3) If the cutoff wavelength of the core is higher than the design constraints, the fitness value suffers a penalty: if  $cutoff > 1530$  nm,  $Cost_{cutoff} = (\sum cutoff - 1530 * 2.05)^2$  taking 1530 nm as the example. Otherwise,  $Cost_{cutoff} = 0$ .
- 4) The  $A_{eff}$  at 1550 nm is constrained in a range, for instance, between 75 and 85  $\mu m^2$ , to offer choices for heterogeneous structure while keep  $A_{eff}$  close to each other. If the core holds a  $A_{eff}$  higher than 85  $\mu m^2$ ,  $Cost_{A_{eff}} = ((A_{eff} - 85) * 50)^2$ . If the core holds a  $A_{eff}$  smaller than 75  $\mu m^2$ ,  $Cost_{A_{eff}} = ((75 - A_{eff}) * 50)^2$ .

#### D. SNR and Capacity

The SNR and capacity performance of the designed MCFs are calculated using ISRS GN model [14] while taking XT [33] into consideration. Since some of our MCFs have bandwidths wider than 15 THz, the Raman gain coefficient ( $C_r$ ) cannot be approximated as a constant. Therefore, we solve the Raman equations to obtain the actual power profile and then match it to these power profile for getting  $C_r$  [33]. Under the assumption that the channels have uniform launch power, SNR in the presence of XT is approximated by [33]:

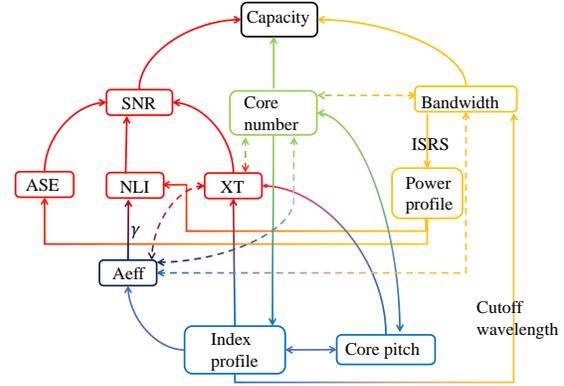
$$SNR = (SNR_0^{-1} + XT)^{-1}, \quad (8)$$

in which  $SNR_0$  is the SNR calculated without XT, where XT refers to the crosstalk with neighboring cores. The capacity of a MCF is calculated as

$$Capacity = \sum_j^{N_c} \sum_i^N B_{ch} \cdot \log_2(1 + SNR_i), \quad (9)$$

in which  $N_c$  is the number of cores in the MCF,  $B_{ch}$  is the bandwidth of each channel and  $N$  is the total number of wavelength channels in each core.

There are numerous complex and non-linear relationships between the fibre design parameters and the theoretical maximum capacity as shown in Fig. 5. The solid line denotes direct dependency and impact of a parameter while the dashed line denotes indirect influence. These indirect relationships bring limitations



**Figure 5:** The main factors which influence the capacity of a single-mode MCF and their relationships.

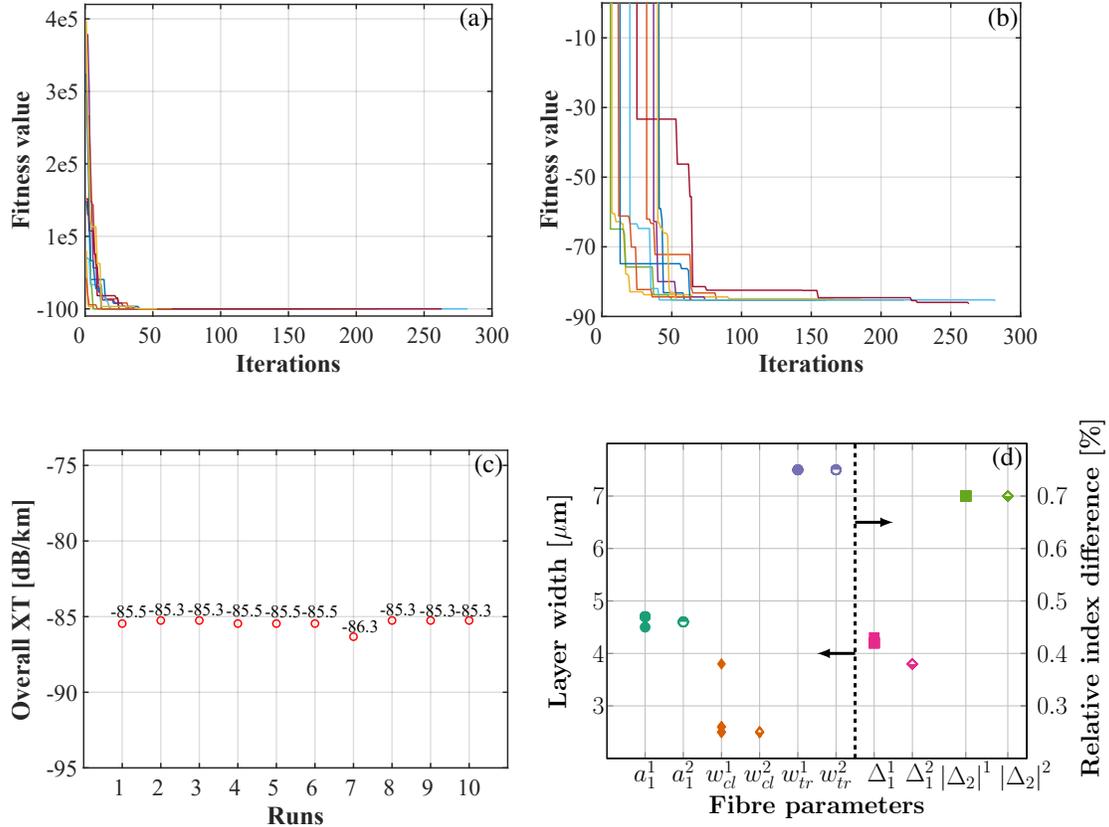
to the elements. Trade-off is needed when optimization is conducted on the parameters influenced by the indirect relationships.

The parameters that can be optimized are the index profile of the each core and the core pitch between adjacent cores in the MCF design illustrated in the light-blue box at the bottom layer. They directly determine the fibre cores' optical properties: cutoff wavelength which indicates the bandwidth,  $A_{eff}$  and XT.

Apropos of the direct dependencies, wider bandwidth leads to more wavelength division multiplexing (WDM) channels. However, wider bandwidth impacts the contribution of NLI and ASE to SNR via the power profile which is highly influenced by the energy transfer of ISRS over the whole bandwidth. Meanwhile, the nonlinear coefficient ( $\gamma$ ) which is inversely proportional to the  $A_{eff}$  which influences NLI. XT brings a penalty on the SNR. As for the number of cores in the standard cladding diameter, higher core density means more spatial channels but will limit the choices of the index profiles and reduce the CP between cores.

The indirect relationships are nonlinear. The wider the bandwidth requirement is, the more limited is the range of the parameters that define the index profile. Then it becomes difficult to reach large  $A_{eff}$  and low XT, simultaneously. If large  $A_{eff}$  is required, XT will stay high because of the weaker light confinement. The more cores one tries to fit in the MCF, the higher will be the IC-XT, in this way forcing/favoring solutions with narrower cores and closer CP.

Taking all these factors into consideration, the capacity of MCFs is influenced by several parameters which interact with each other directly or non-linearly rendering the MCF capacity-optimization design a



**Figure 6:** (a) Fitness value decreases with iterations in the  $[70\ 80]\ \mu\text{m}^2$   $A_{\text{eff}}$  constraint and  $\leq 1460$  nm cutoff wavelength constraint; (b) Zoom in of (a); (c) The final overall XT which PSO obtained after all the iterations; (d) Parameters in  $\mathbf{g}_{\text{best}}$  of ten runs obtained in cutoff  $\leq 1360$  nm case (the superscript indicates which core the parameter belongs to);

multi-dimensional complex problem.

#### IV. OPTIMIZED MCF DESIGN

We collected MCF designs with PSO under various fibre optical properties constraints' combinations for 6-core fibres. In this section, we evaluate PSO's performance and the designed MCFs' optical properties. PSO showed its reliability in different constraints combinations. MCFs have lower XT and good  $A_{\text{eff}}$  compared with the state-of-the-art.

##### A. PSO Performance

Compared to  $\sim 10^{15}$  permutations needed in the brute force, PSO only needs  $\sim 10^5$  permutation calculations to converge to a good fitness value. This ten orders of magnitude reduction means significant computation improvement. The PSO convergence process and finale design result are illustrated for 1460 nm cutoff constraint and  $[70\ 80]\ \mu\text{m}^2$   $A_{\text{eff}}$  constraint as an example in Fig. 6.

We analyzed the PSO performance from three aspects: whether it obtains valid solutions, standard deviation among the valid solutions and how close the fibre structure given by the valid solutions. In the *XT-optimization*, except the case —  $[90\ 100]\ \mu\text{m}^2$   $A_{\text{eff}}$  with 1530 nm cutoff constrains, ten runs of PSO all found valid solutions which satisfy the constrains. In Fig. 6(a) and (b) which is the zoom-in of (a) in small fitness value region, it can be seen that though the maximal iteration number is set as 500, the fitness value reaches the minimum with less than 300 iterations because of the convergence criterion mentioned in Section. III-A. The ten runs converged to similar fitness value and overall XT of the ten runs is shown in Fig. 6(c). They are the same as the final fitness value PSO reached, which means that the MCF designs satisfy all the constrains. The exceptional case holds 6 runs which found the valid solutions. In the  *$A_{\text{eff}}$ -optimization*, all the cases have ten to ten runs obtaining valid solutions.

The standard deviation of the ten PSO results in *XT-optimization* cases is between 0 to 1.39 for overall XT. In case of *A<sub>eff</sub>-optimization*, it is between 0 to 4.02 for *A<sub>eff</sub>*. Meanwhile, PSO results obtained in different runs are similar for each constraints combinations. Each result contains ten fibre parameters, five for core 1 and five for core 2, shown in Fig. 6(d). For each parameter, the values obtained in different runs converge together with very small variance.

For the fibre structure illustrated in Fig. 6(d), an example of their PCC curves to the bending radius is illustrated in Fig. 7. For the non-identical core pair in the 6-core fibre, the PCC linearly increases with bending radius when the bending radius is smaller than the critical bending radius. When it exceeds the threshold, the PCC decreases sharply and the MCF enters the non-phase matching area.

In total, PSO proves its ability to find out MCF designs satisfying the constraints we set and to optimize two different fitness value focusing on XT and *A<sub>eff</sub>* respectively. Meanwhile, in the ten runs in each case, the similar fitness values and the similar final fibre designs which PSO converges to in ten runs demonstrate PSO's reliability in MCF design.

### B. XT and RCMF Performance Comparison

In Fig. 8, *XT-optimization* and *A<sub>eff</sub>-optimization* results are illustrated in green and red markers respectively same as in Fig. 2. The blue markers are the fabricated fibres with silica cladding and standard cladding diameter in the previous literature. Relative core multiplicity factor (RCMF) illustrates the performance of MCF compared with standard single mode fibre (SSMF) in terms of core density and *A<sub>eff</sub>*. For our MCFs with standard cladding diameter, the RCMF can be calculated as

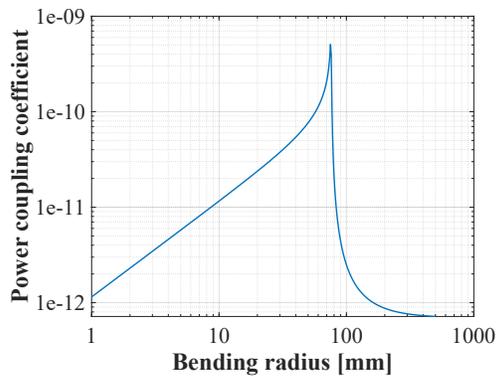


Figure 7: The PCC corresponding to the result in Fig. 6(d).

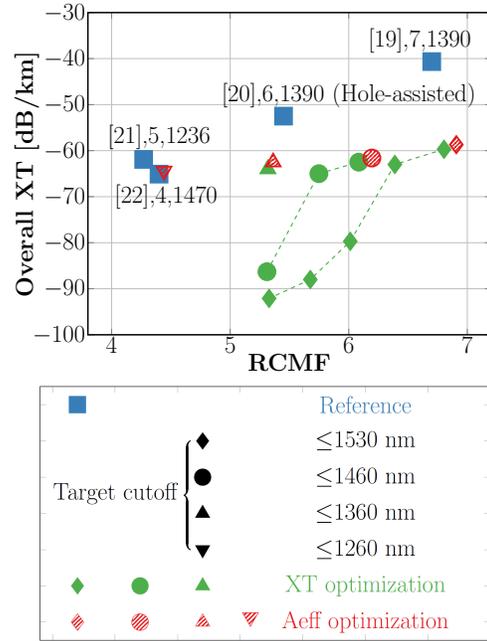


Figure 8: Minimal overall crosstalk obtained with various constraints comparing with references.

$$RCMF = \frac{\sum_i^{N_c} A_{\text{eff}}^i}{A_{\text{eff-SSMF}}} \quad (10)$$

The aim of MCF design is to achieve higher RCMF and lower XT. Our MCF designs perform significantly better in terms of overall XT even for similar RCMF. Meanwhile, the RCMF in our result has exceeded that of state-of-the-art.

*XT-optimization* objective function aims to to minimize the overall XT of MCF while constraining the optical properties as is explained in Section III(C). Considering the four cutoff wavelength targets starting from the lowest, there are 0, 1, 3 and 5 cases that satisfy the *A<sub>eff</sub>* constraint respectively. In each cutoff case, with the increasing RCMF the PSO-optimized overall XT increases. It can be concluded that *A<sub>eff</sub>* and XT has an inverse-proportional relationship when the cladding diameter is fixed.

The lowest overall XT reached with cutoff  $\leq 1530$  nm is  $-92.1$  dB/km. The XT is 25 dB lower than the lowest XT of the reference points,  $-67$  dB/km. In [34], a universal specification of IC-XT per unit length, around  $-60$  dB/km is proposed which is suitable for different systems from metropolitan ( $\sim 100$  km) to trans-pacific ( $\sim 10\,000$  km). Our 6-core fibres all satisfy this universal specification including the results of *A<sub>eff</sub>-optimization* of which the highest overall XT is

**Table II:** Simulation Parameters

Parameters	
Channel Bandwidth [GHz]	49.5
Channel Spacing [GHz]	50
Channel Launch Power [dBm]	-2
Noise Figure [dB]	5
Transceiver penalty [dB]	23
Span Length [km]	60
Span Number	20
Total Length [km]	1200

-58.7 dB/km.

In  $A_{\text{eff}}$ -optimization, the upper limits of  $A_{\text{eff}}$  were explored without XT constraint. The highest sum of  $A_{\text{eff}}$  is searched with PSO while constraining cutoff wavelength. Compared to the  $XT$ -optimization, the PSO results sacrificed a small overall XT performance for a higher  $A_{\text{eff}}$ . Since different  $A_{\text{eff}}$  constraints have been tested in the previous section, the  $A_{\text{eff}}$ -optimization just shows a small improvement on  $A_{\text{eff}}$  compared to the  $XT$ -optimization results.

By optimizing  $A_{\text{eff}}$ , RCMF of our 6-core fibres is up to 6.82 which is higher than the highest RCMF of the reference points, 6.7, in [19]. Comparing the fibres with similar RCMF, the wider bandwidth the worse XT. That is because higher bandwidth requirement leads to limitation of the core choices. This also limits the  $A_{\text{eff}}$  choice in each cutoff case.

## V. TRANSMISSION PERFORMANCE OF OPTIMIZED MCFS

PSO designed MCFs which perform well in both RCMF and XT. In this section, we further explore their SNR and capacity performance taking both XT and ISRS into consideration. As we will see in the following, XT is not a limiting factor for our work, its influence on SNR performance is insignificant even for ultra-long transmission (for XT smaller than or around -60 dB/km). Differently, ISRS shows strong impact on SNR in MCFs with various bandwidths. As for the corresponding capacity, bandwidth plays the determinant role while the increase of  $A_{\text{eff}}$  has small benefit to the capacity.

### A. SNR Performance

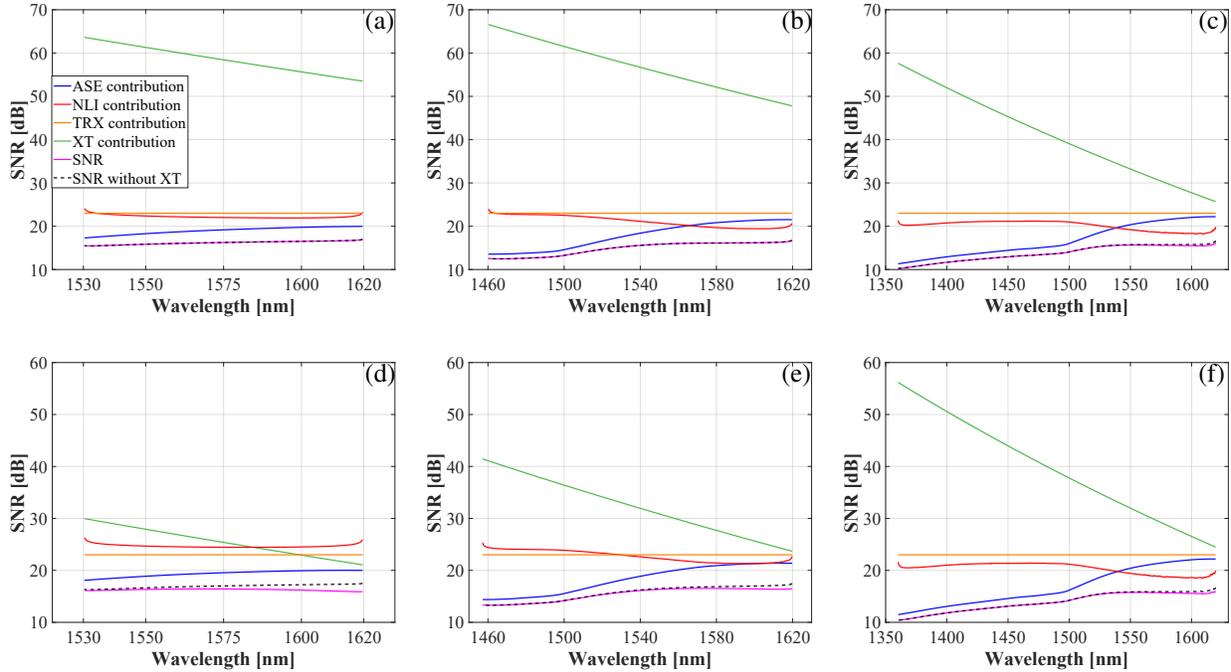
The simulation parameters for SNR calculation of single core are listed in Table II. First, we compare the SNR and its various contributions for both  $XT$ -optimization and  $A_{\text{eff}}$ -optimization illustrated in Fig. 9. Because the cutoff wavelength constraints are set as smaller than or equal to a certain value, in the low  $A_{\text{eff}}$  constraint cases, some of the optimized MCF structures hold two largely different cutoff wavelengths. For example, in the [70 80]  $\mu\text{m}^2$   $A_{\text{eff}}$  constraint and  $\leq 1530$  nm cutoff wavelength constraint, the two types of core in the optimized fibre have 1356 nm and 1530 nm as cutoff wavelength, respectively. When calculating the XT penalty on the SNR, only the wavelengths which both cores guide are considered. But the cores shown in Fig. 9 are with the cutoff close to the constraint for comparing the bandwidth influence.

Fig. 9(a)-(c) showcases the cores obtained with  $XT$ -optimization objective function. They cover different optical bands with low overall XT. Fig. 9(d)-(f) shows the cores optimized  $A_{\text{eff}}$  but with slightly higher overall XT. Comparing the cores designed with  $XT$ -optimization and  $A_{\text{eff}}$ -optimization objective functions – that cover the same bandwidth – we see the following behaviour. The  $XT$ -optimization method leads to lower XT which proved to negligibly or marginally influence the SNR. Specifically Fig. 9(a) shows a XT induced penalty smaller than 0.001 dB while in Fig. 9(d) the penalty is up to 1.6 dB for 1200 km transmission. On the other hand, the  $A_{\text{eff}}$ -optimization method offers higher  $A_{\text{eff}}$  that leads to 2.71 dB improvement in SNR from NLI contribution compared to the former. In Section V.C, the impacts of XT and  $A_{\text{eff}}$  on capacity will be further discussed. The O+E+S+C+L-band fibre, has a 12.3  $\mu\text{m}^2$  smaller  $A_{\text{eff}}$  and 107.6 THz wider bandwidth compared to E+S+C+L-band fibre. This leads to a reduced SNR by 3.18 dB (average) due to NLI contribution as shown in Fig. 10.

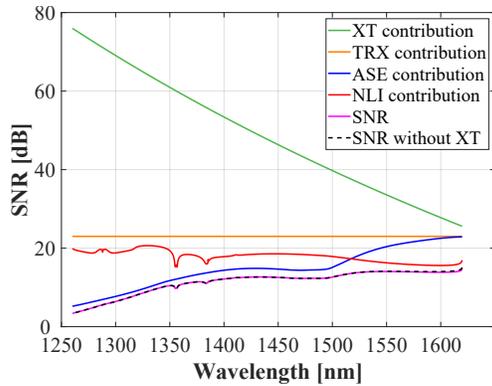
The cores with wider bandwidth have worse SNR performance in the short wavelength range. The ISRS transfers energy from short wavelengths to long wavelengths during the transmission. Thus, at the end of each span the power at the short wavelengths is lower than that of the long wavelengths. The unbalance in the actual power profile highly influences the ASE contribution and NLI contribution in SNR.

### B. Capacity performance

The capacity performance of each MCF is calculated and compared against various parameters for better



**Figure 9:** (a),(b),(c) are the cores picked from *XT-optimization*: (a) Fibre core with  $A_{\text{eff}} \in [70 \ 80] \mu\text{m}^2$  and cutoff  $\leq 1530$  nm; (b) Fibre core with  $A_{\text{eff}} \in [70 \ 80] \mu\text{m}^2$  and cutoff  $\leq 1460$  nm; (c) Fibre core with  $A_{\text{eff}} \in [70 \ 80] \mu\text{m}^2$  and cutoff  $\leq 1360$  nm; (d),(e),(f) are cores picked from *A<sub>eff</sub>-optimization*: (d) Fibre core with cutoff  $\leq 1530$  nm; (e) Fibre core with cutoff  $\leq 1460$  nm; (f) Fibre core with cutoff  $\leq 1360$  nm.



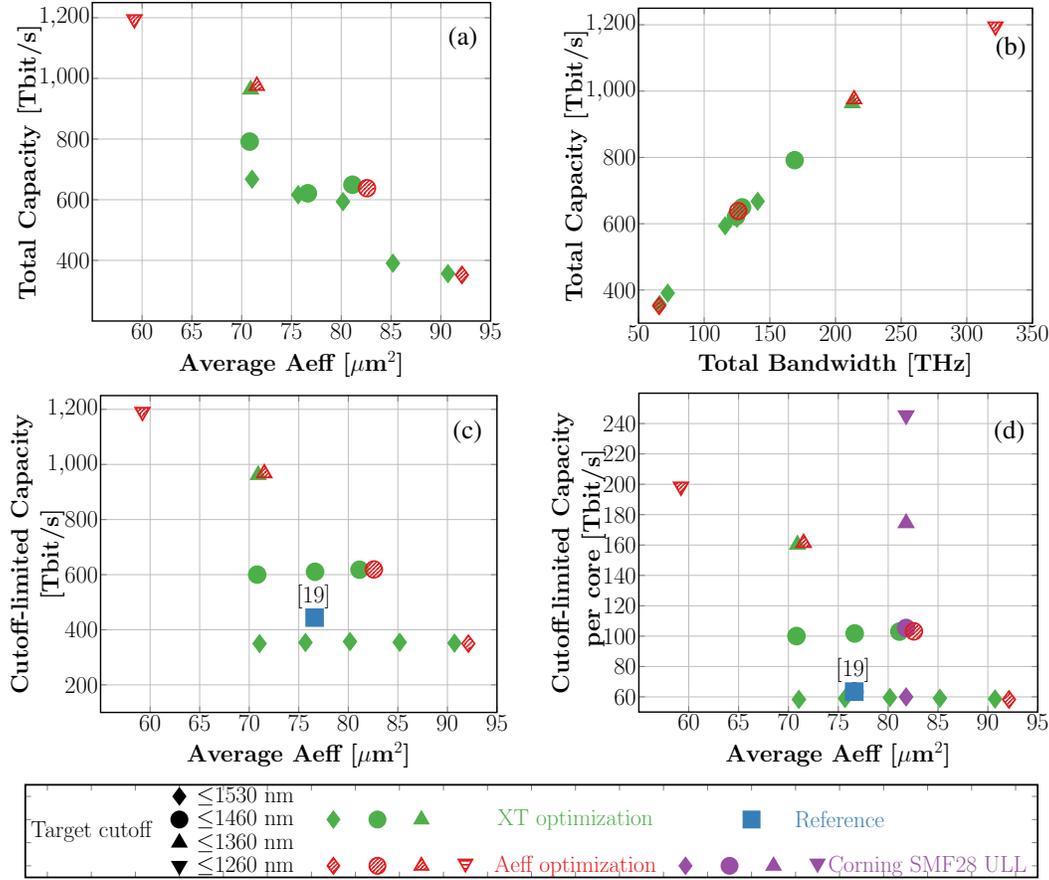
**Figure 10:** The SNR and its contributions of O+E+S+C+L-band fibre in the *A<sub>eff</sub>-optimization*.

understand their influence as illustrated in Fig. 11. When considering the total capacity across the *entirety of the supported spectrum of each core* then the performance degraded with increasing  $A_{\text{eff}}$  in Fig. 11(a), which is opposite to what we expect. Larger  $A_{\text{eff}}$  (smaller  $\gamma$ ) should result in smaller NLI, thus better SNR and larger capacity. When the relationship between total capacity and the total bandwidth of the whole MCF,  $\sum_i^N B_{ch}$ , is illustrated in Fig. 11(d), it

becomes clear that the total bandwidth dominates the capacity performance. The capacity increases almost linearly to the total bandwidth with a slope as  $\sim 4.08$  Tbit/s/THz.

If we limit the bandwidth to the optical bands which we designed to cover, the influence of the total bandwidth difference can be eliminated. For example, cores with  $\leq 1530$  nm cut-off wavelength constraint support the C+L bands and their associated capacity noted as cutoff-limited capacity. The impacts of XT and  $A_{\text{eff}}$  on cutoff-limited capacities can be investigated among the fibres having the same cutoff-limited. Our fibres which cover S+C+L band have about 35.1% more cutoff-limited capacity for 1200 km transmission than the 7-core fibre in [19] which also covers S+C+L band. The fibres supporting O+E+S+C+L-band fibres have 1.68-times higher cutoff-limited capacity than [19]. As for capacity per core, that of [19] is only comparable to that of our 6-core C+L-band fibres.

The average of the cutoff-limited capacity per core in the fibres designed in this paper can be compared with that of the Corning Ultra-low loss (ULL) SMF28 as illustrated in Fig. 11(d). 6-core fibres have similar capacity per core to the SMF28 with very small decrease due to the  $A_{\text{eff}}$  and XT. The 7-core fibre in [19] has much



**Figure 11:** (a) Total capacity of all the optimized MCFs calculated over the whole bandwidth of each fibre versus the  $A_{\text{eff}}$ ; (b) Total capacity versus the total bandwidth of the MCF; (c) Cutoff-limited Capacity calculated when bandwidths are limited in the optical bands for 1200km transmission; (d) Cutoff-limited capacity in (c) averaged by core number for 1200km transmission.

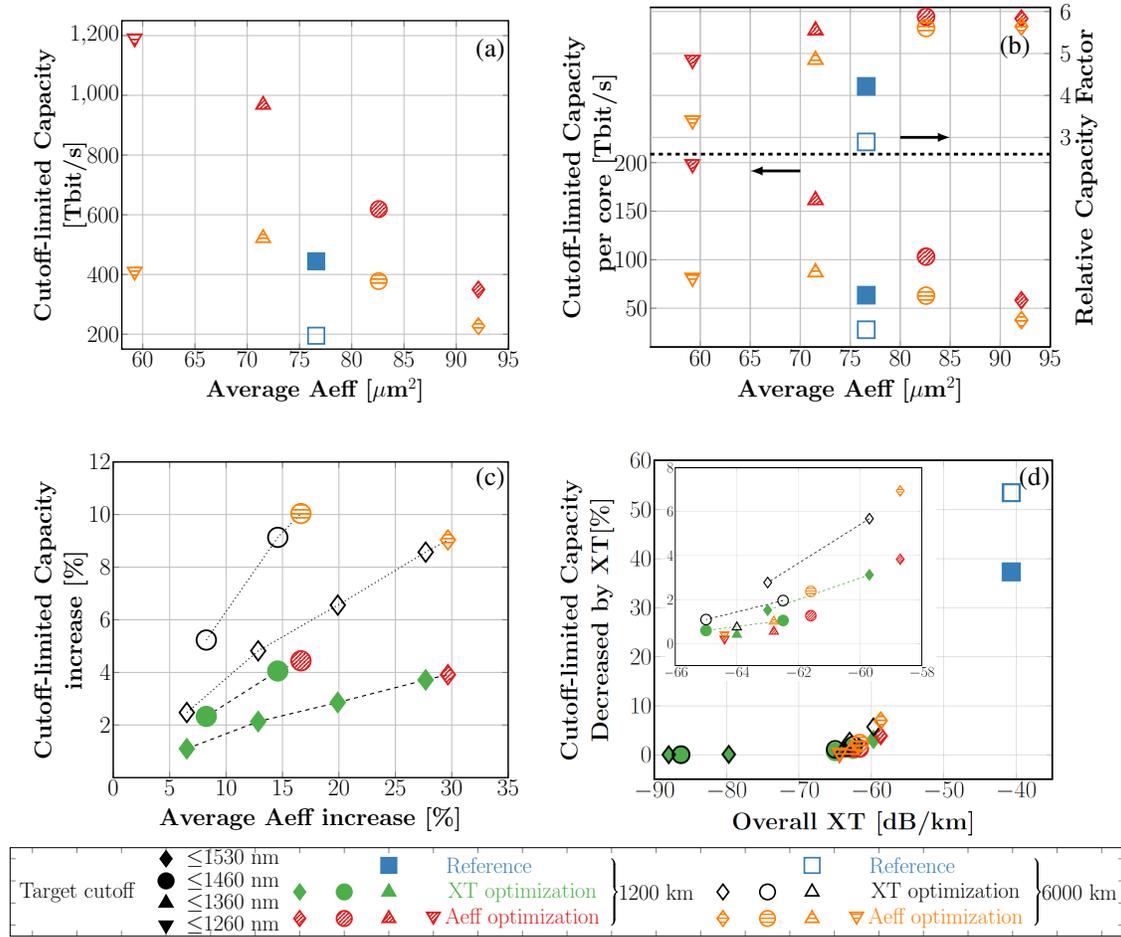
smaller cutoff-limited capacity per core compare to the SMF28 when it is limited to S+C+L band.

According to the analysis and discussion above about various capacity metrics, the O+E+S+C+L-band 6-core fibre has the highest total capacity. But in the practical applications, it might face the lack of amplifiers in O-Band and E-band as well as low transceiver efficiency. If we want to apply the fibre without any concern about the amplifier, the 1460 nm fibres can be the one which cover S+C+L bands and have good capacity simultaneously. If the application highly depends on the  $A_{\text{eff}}$ , then the  $A_{\text{eff}}$ -optimized 6-core fibre with 1530 nm cutoff wavelength will be the best choice.

### C. Capacity Analysis and Comparison

The transmission length is elongated to 6000 km with 60 km/span for exploring 6-core fibres' capacity in transoceanic transmission. We pick the three

$A_{\text{eff}}$ -optimization results with different optical bands to compare them for two transmission lengths. For 1200 km transmission, O+E+S+C+L-band fibre holds a cutoff-limited capacity up to  $\sim 1.2$  Pbit/s while for 6000 km, it is 409.7 Tbit/s with a 65.6% drop. Fibres with narrower bandwidth have smaller drop, 46.1%, 39.0% and 35.2% for E+S+C+L-, S+C+L- and C+L-band fibres, respectively, as illustrated in Fig. 12(a). This is because the narrower the fibre bandwidth the larger  $A_{\text{eff}}$  and XT. To some extent, the benefit of  $A_{\text{eff}}$  compensates the XT penalty increase due to the transmission length elongation in the fibre with narrower bandwidth. For 6000 km transmission, due to the large capacity drop of O+E+S+C+L-band MCF, E+S+C+L-band fibre has the highest capacity, 521.8 Tbit/s. The cutoff-limited capacity of [19] is further degraded while the transmission length increases to 6000 km because of poor SNR. Our



**Figure 12:** (a) Cutoff-limited capacity of  $A_{eff}$ -optimization results for two transmission lengths; (b) Cutoff-limited capacity per core and relative capacity factor comparison between 1200 km and 6000 km transmission ; (c) The benefit of  $A_{eff}$  increase to cutoff-limited capacity; (d) The penalty that cutoff-limited capacity suffers from XT.

S+C+L-band fibre has around 84.8% more cutoff-limited capacity.

By normalizing their cutoff-limited capacity per core with the SMF28 capacity in Fig. 11 which is calculated with the same bandwidth and transmission length to the MCFs, the ratio is noted as relative capacity factor (RCF) shown in Fig. 12(b). O+E+S+C+L-band fibre has around  $20 \mu\text{m}^2$  smaller  $A_{eff}$  than the ULL SMF. Thus, it has smaller RCF, 4.8 and 3.4 for 1200 km and 6000 km transmission, respectively. The others are slightly smaller than 6 due to XT penalty. In other words, our 6-core fibres have comparable capacity per core to the SMF28. Moreover [19] is 7-core fibre, its cutoff-limited capacity per core is lower than that of our 6-core fibres, and therefore, its RCF is 4.2 and 2.9 for 1200 km and 6000 km transmission, respectively.

Under these two transmission lengths, we investigate

the impacts of  $A_{eff}$  and XT on the capacity performance separately. Firstly, the cutoff-limited capacity without considering XT increases slowly with the average  $A_{eff}$ . To analyse further their relationship, we calculated the increment in percentage by comparing the difference between the data of each point to the one which has the lowest average  $A_{eff}$  and smallest capacity. A 5-30% increase of  $A_{eff}$  leads to just 1-4.5% increase in total capacity for 1200 km transmission. When the transmission is up to 6000 km, the same increment of  $A_{eff}$  results in around two times increase in cutoff-limited capacity as illustrated in Fig. 12(c).

When taking XT into the SNR calculation, the cutoff-limited capacity is reduced. The XT penalty on cutoff-limited capacity increases with XT. When XT increases from  $\sim 65$  dB/km to  $\sim 59$  dB/km, the penalty increases

from around 1% up to 7% as illustrated in the zoom-in Fig. 12(d). The longer the transmission length, the higher accumulated XT and therefore the more reduction of the cutoff-limited capacity. Since our works all have ultra-low XT, the benefit of  $A_{\text{eff}}$  is offset by the penalty due to XT, and therefore the cutoff-limited capacity in Fig. 11(d) is similar for the same optical bands. But in the higher XT region, XT will have non-negligible penalty in capacity as the reference point shows.

## VI. CONCLUSIONS

In this paper, we have proposed an automated MCF design method based on the PSO algorithm to optimize MCF optical properties and capacity. The MCFs design is simplified to optimize the neighbouring core pair taking advantages of the symmetry and heterogeneous MCF structure. In this work, the MCF design took all the refractive-index profile parameters as variables for the first time to search for the optimized MCF structure. The AI-based method can design MCFs with multi-objective functions and different constraints autonomously and efficiently compared to the brute force methods with ten orders of magnitude reduction of permutation evaluations. Meanwhile, by using the statistical classifier and regressor to estimate the CL instead of a FDE solver, the computation time is reduced by six orders of magnitude.

Taking the minimization of XT as the main purpose, we have achieved MCFs with ultra-low XT from -92.1 dB/km to -58.7 dB/km, covering up to E+S+C+L band, which is comparable or much lower than the lowest XT in literature, -65.1 dB/km. When taking the maximization of  $A_{\text{eff}}$  as the main purpose, we achieved a RCMF of 6.82 for a 6-core fibre which is better than the 6.7 achieved in the literature. The successful designs with two objective functions demonstrate the reliability of the automated AI-method to realize the multi-dimensional MCF design problem and its ability to optimize different MCF optical properties.

With ISRS GN model, the SNR and capacity performance can be checked upon MCF designing. By investigating the relationships among the optical properties of PSO designed MCFs and the SNR and capacity, it can be concluded that the bandwidth dominants the capacity performance.  $A_{\text{eff}}$  increase brings almost linear cutoff-limited capacity increase without XT influence while XT results in exponential-increasing penalty to cutoff-limited capacity.

PSO-designed MCFs have similar capacity per core for the S+C+L-band and C+L-band fibres compared with the Corning ULL SMF28 (RCF approaching 6) in both

1200 km and 6000 km transmission. The E+S+C+L-band fibre results in total capacity of ~1 Pbit/s for 1200 km transmission and ~500 Tbit/s for 6000 km transmission. The O+E+S+C+L-band fibre delivers up to ~1.2 Pbit/s for 1200 km transmission and RCF approaching 5.

Compared with the 7-core fibre in [19], our fibres have 35.1% and 84.8% more capacity considering the S+C+L optical bands, respectively, for 1200 km and 6000 km transmission. From the above, it can be concluded that the proposed AI-based MCF design method, with capacity-check post-process, is a promising tool-set for the design of future MCFs with single-ring or double-ring structure in the same or wider cladding diameter for the purpose of controlling or optimizing XT,  $A_{\text{eff}}$ , cutoff wavelength or other optical property requirements.

## APPENDIX

Considering the coupled-power theory, the mean IC-XT of one homogeneous trench-assisted MCF (Homo-TA-MCF) core can be expressed as Eq. 11 [1].

$$IC-XT = \frac{n - ne^{-(n+1)hL}}{1 + ne^{-(n+1)hL}}. \quad (11)$$

The numerator is the sum of the power in a certain core originating from the other cores while the denominator is that from the certain core, where  $n$  denotes the number of adjacent cores of the target core and  $L$  stands for the fiber length.

In a heterogeneous trench-assisted MCF (Hete-TA-MCF), assuming that every core has an unique index profile, the power coupling coefficient (PCC,  $h$ ) between each core pair is different from each other. The PCC between the interested core  $p$  and one of its possible neighbouring cores  $q$  is denoted as  $h_{pq}$ . When core  $p$  is excited with power  $P_p(0) = 1$ , while all neighbouring cores are not excited, the power in core  $p$  is expressed as [35]:

$$P_{p,p}(z) = \frac{1 + \sum_{i=1}^n e^{-(n+1)h_{pq}z}}{n + 1}. \quad (12)$$

When all the neighbouring cores are all excited while core  $p$  is not excited, the power in core  $p$  from core  $q$  is expressed as:

$$P_{p,q}(z) = \frac{1 - e^{-(n+1)h_{pq}z}}{n + 1}. \quad (13)$$

Then, the mean crosstalk of core  $p$  from all  $n$  possible neighbouring cores is the ratio between the sum of the

power coupled from all  $n$  neighbouring cores and the power in core  $p$  itself:

$$\begin{aligned}
 IC\text{-}XT_p &= \frac{\sum_{i=1}^n P_{p,q}(z)}{P_{p,p}(z)} = \frac{\sum_{i=1}^n \frac{1-e^{-(n+1)h_{pq}z}}{n+1}}{1+\frac{\sum_{i=1}^n e^{-(n+1)h_{pq}z}}{n+1}} \\
 &= \frac{n - \sum_{i=1}^n e^{-(n+1)h_{pq}L}}{1 + \sum_{i=1}^n e^{-(n+1)h_{pq}L}} \quad (14)
 \end{aligned}$$

#### ACKNOWLEDGEMENT

This work is supported by EPSRC, UK EP/L015455/1, OptoCloud EP/T026081/1, TRANSNET EP/R035342/1 & UKRI FLF MR/T041218/1. The author would like to thank Dr. Daniel Semrau for useful discussion.

#### REFERENCES

- [1] G. M. Saridis, D. Alexandropoulos, G. Zervas, and D. Simeonidou, "Survey and evaluation of space division multiplexing: From technologies to optical networks," *IEEE Commun. Survey and Tutorials*, vol. 17, no. 4, pp. 2136–2156, Aug. 2015.
- [2] Y. Liu, H. Yuan, A. Peters, and G. Zervas, "Comparison of SDM and WDM on direct and indirect optical data center networks," in *Proc. of Eur. Conf. Opt. Commun. (ECOC)*, Dusseldorf, Germany, M.1.F.2, Sep. 2016.
- [3] A. Saljoghei, H. Yuan, V. Mishra, M. Enrico, N. Parsons, C. Kochis, P. De Dobbelaere, D. Theodoropoulos, D. Pnevmatikatos, D. Syrivelis, A. Reale, T. Hayashi, T. Nakanishi, and G. Zervas, "MCF-SMF hybrid low-latency circuit-switched optical network for disaggregated data centers," *J. Lightw. Technol.*, vol. 37, no. 16, pp. 4017–4029, Aug. 2019.
- [4] 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.*, vol. 10, no. 4, pp. 272–288, Apr. 2018.
- [5] T. Hayashi, A. Mekis, T. Nakanishi, M. Peterson, S. Sahni, P. Sun, S. Freyling, G. Armijo, C. Sohn, D. Foltz *et al.*, "End-to-end multi-core fibre transmission link enabled by silicon photonics transceiver with grating coupler array," in *Proc. of Eur. Conf. Opt. Commun. (ECOC)*, Gothenburg, Sweden, Th.2.A, Sep. 2017.
- [6] T. Matsui, T. Kobayashi, H. Kawahara, E. L. T. de Gabory, T. Nagashima, T. Nakanishi, S. Saitoh, Y. Amma, K. Maeda, S. Arai, R. Nagase, Y. Albe, S. Aozasa, Y. Wakayama, H. Takeshita, T. Tsuritani, H. Ono, T. Sakamoto, I. Morita, Y. Miyamoto, and K. Nakajima, "118.5 tbit/s transmission over 316 km-long multi-core fiber with standard cladding diameter," in *Proc. of Opto. Commun. Conf. (OECC)*, Singapore, Aug. 2017.
- [7] T. Hayashi, T. Taru, O. Shimakawa, T. Sasaki, and E. Sasaoka, "Design and fabrication of ultra-low crosstalk and low-loss multi-core fiber," *Opt. Express*, vol. 19, no. 17, pp. 16 576–16 592, Aug. 2011.
- [8] Y. Sagae, T. Matsui, and K. Nakajima, "Ultra-low-XT multi-core fiber with standard 125- $\mu$ m cladding for 10000km-class long-haul transmission," in *Proc. of Opto. Commun. Conf. (OECC)*, Fukuoka, Japan, TuC3-4, Jul. 2019.
- [9] J. Tu, K. Long, and K. Saitoh, "An efficient core selection method for heterogeneous trench-assisted multi-core fiber," *IEEE Photon. Technol. Lett.*, vol. 28, no. 7, pp. 810–813, 2016.
- [10] J. Kennedy, "Particle swarm optimization," *Encyclopedia of machine learning*, pp. 760–766, 2010.
- [11] J. Mata, I. de Miguel, R. J. Durán, N. Merayo, S. K. Singh, A. Jukan, and M. Chamanía, "Artificial intelligence (ai) methods in optical networks: A comprehensive survey," *Optical Switching and Networking*, vol. 28, pp. 43 – 57, 2018.
- [12] C. W. F. Parsonson, Z. Shabka, W. K. Chlupka, B. Goh, and G. Zervas, "Optimal control of soas with artificial intelligence for sub-nanosecond optical switching," *J. Lightw. Technol.*, vol. 38, no. 20, pp. 5563–5573, 2020.
- [13] X. Mu, F. M. Ferreira, A. Ottino, and G. Zervas, "Design optimization of uncoupled six-core fibers in standard cladding diameter using artificial intelligence," in *Proc. of Opt. Fiber Commun. Conf. and Exhib. (OFC)*, Th1A.34, Jun. 2021.
- [14] D. Semrau, L. Galdino, E. Sillekens, D. Lavery, R. I. Killey, and P. Bayvel, "Modulation format dependent, closed-form formula for estimating nonlinear interference in s+c+l band systems," in *Proc. of Eur. Conf. Opt. Commun. (ECOC)*, Dublin, Ireland, W.1.D, Sep. 2019.
- [15] K. Takenaga, Y. Arakawa, Y. Sasaki, S. Tanigawa, S. Matsuo, K. Saitoh, and M. Koshihba, "A large effective area multi-core fiber with an optimized cladding thickness," *Opt. Express*, vol. 19, no. 26, pp. B543–B550, Dec. 2011.
- [16] S. Jiang, L. Ma, M. N. Velazquez, Z. He, and J. K. Sahu, "Design of 125- $\mu$ m cladding diameter multicore fibers with high core multiplexing factor for wideband optical transmission," *Opt. Fiber Technol.*, vol. 50, pp. 55–61, Jul. 2019.
- [17] T. Hayashi, T. Taru, O. Shimakawa, T. Sasaki, and E. Sasaoka, "Characterization of crosstalk in ultra-low-crosstalk multi-core fiber," *J. Lightw. Technol.*, vol. 30, no. 4, pp. 583–589, Feb. 2011.
- [18] W. Hermann and D. Wiechert, "Refractive index of doped and undoped pcvd bulk silica," *Materials Research Bulletin*, vol. 24, no. 9, pp. 1083 – 1097, 1989.
- [19] K. Takenaga, Y. Arakawa, S. Tanigawa, N. Guan, S. Matsuo, K. Saitoh, and M. Koshihba, "Reduction of crosstalk by trench-assisted multi-core fiber," in *Proc. of Opt. Fiber Commun. Conf. and Exhib. (OFC)*, Los Angeles, CA, USA, OWJ4, Mar. 2011.
- [20] T. Sakamoto, K. Saitoh, N. Hanzawa, K. Tsujikawa, L. Ma, M. Koshihba, and F. Yamamoto, "Crosstalk suppressed hole-assisted 6-core fiber with cladding diameter of 125  $\mu$ m," in *Proc. of Eur. Conf. Opt. Commun. (ECOC)*, London, UK, Mo.3.A.3, Sep. 2013.
- [21] T. Gonda, K. Imamura, R. Sugizaki, Y. Kawaguchi, and T. Tsuritani, "125  $\mu$ m 5-core fibre with heterogeneous design suitable for migration from single-core system to multi-core system," in *Proc. of Eur. Conf. Opt. Commun. (ECOC)*, Dusseldorf, Germany, W.2.B.1, Sep. 2016.
- [22] J. Tu, K. Saitoh, M. Koshihba, K. Takenaga, and S. Matsuo, "Design and analysis of large-effective-area heterogeneous trench-assisted multi-core fiber," *Opt. Express*, vol. 20, no. 14, pp. 15 157–15 170, Jun. 2012.
- [23] M. Koshihba, K. Saitoh, K. Takenaga, and S. Matsuo, "Analytical expression of average power-coupling coefficients for estimating intercore crosstalk in multicore fibers," *IEEE Photon. J.*, vol. 4, no. 5, pp. 1987–1995, Oct. 2012.
- [24] Y. Sasaki, Y. Amma, K. Takenaga, S. Matsuo, K. Saitoh, and M. Koshihba, "Trench-assisted low-crosstalk few-mode multicore fiber," in *Proc. of Eur. Conf. Opt. Commun. (ECOC)*, London, UK, Tu.1.F.3, Sep. 2013, pp. 1–3.
- [25] X. Xie, J. Tu, X. Zhou, K. Long, and K. Saitoh, "Design and optimization of 32-core rod/trench assisted square-lattice structured single-mode multi-core fiber," *Opt. Express*, vol. 25, no. 5, pp. 5119–5132, Mar. 2017.
- [26] J. ichi Sakai and T. Kimura, "Bending loss of propagation modes in arbitrary-index profile optical fibers," *Appl. Opt.*, vol. 17, no. 10, pp. 1499–1506, May 1978.

- [27] C. Yeh and G. Lindgren, "Computing the propagation characteristics of radially stratified fibers: an efficient method," *Appl. Opt.*, vol. 16, no. 2, pp. 483–493, Feb. 1977.
- [28] Y. Shi and R. C. Eberhart, "Parameter selection in particle swarm optimization," in *Proc. of Int. Conf. on Evolutionary Programm (ICET)*, San Diego, CA, USA, Mar. 1998, pp. 591–600.
- [29] R. Eberhart and J. Kennedy, "Particle swarm optimization," in *Proc. of IEEE Int. Conf. on Neural Networks*, vol. 4, Perth, WA, Australia, Nov. 1995, pp. 1942–1948.
- [30] Lumerical, <https://www.lumerical.com/products/>.
- [31] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *Proc. of International Conf. on Machine Learning (ICML)*, Madison, WI, USA, Jun. 2010, p. 807–814.
- [32] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2017.
- [33] R. S. Luís, G. Rademacher, B. J. Puttnam, D. Semrau, R. I. Killey, P. Bayvel, Y. Awaji, and H. Furukawa, "Crosstalk impact on the performance of wideband multicore-fiber transmission systems," *IEEE J. Select. Topics Quantum Electron.*, vol. 26, no. 4, pp. 1–9, 2020.
- [34] J. M. Gené and P. J. Winzer, "A universal specification for multicore fiber crosstalk," *IEEE Photon. Technol. Lett.*, vol. 31, no. 9, pp. 673–676, 2019.
- [35] K. TAKENAGA, Y. ARAKAWA, S. TANIGAWA, N. GUAN, S. MATSUO, K. SAITOH, and M. KOSHIBA, "An investigation on crosstalk in multi-core fibers by introducing random fluctuation along longitudinal direction," *IEICE Transactions on Communications*, vol. E94.B, no. 2, pp. 409–416, 2011.

**Xun Mu** is a PhD student under the supervision of Dr. Georgios Zervas in Optical Network Group in University College London. She received the B.S. in Optoelectronic Information Engineering from the Huazhong University of Science and Technology in 2015 and the MSc degree in Nanophotonics from University of Strasbourg in France in 2017. Then, she joined UCL and received the MRes degree in Integrated Photonic and Electronic System. From October 2018, Xun has been a part of the Optical Networks Group at UCL as a PhD candidate. Her research interests includes SDM in fibre optical communication and artificial intelligence.

**Alessandro Ottino** is a PhD student under the supervision of Dr. Georgios Zervas in in the Optical Networks Group at University College London. He receive a B.Eng in Electrical and Electronic Engineering with outstanding academic achievement at university College London in 2019. From October 2019, Alessandro has been part of the Optical Networks Group at UCL as a PhD candidate. His research interest include Distributed Deep Learning Systems, HPC, mathematical modelling and artificial intelligence aided optical device design and optimisation.

**Dr. Filipe Ferreira** is currently a UKRI Future Leaders Fellow in the Department of Electronic and Electrical Engineering at University College London – exploring massive spatial parallelism in optical fibres by merging spatial light modulation with machine learning. He received his PhD in Electrical and Computer Engineering from University of Coimbra, in 2014, for his work in high-capacity optical transmission in multipath fibres with the aim of exceeding the limits of conventional single-mode fibres. During that time, he worked as a research engineer at Nokia Siemens Networks leading several research advances on mode-division multiplexing in the EU-FP7 project MODEGAP. He was then awarded a Marie Skłodowska-Curie Fellowship to continue his research at Aston University. In 2019, Filipe joined University College London, and during the 2019-2020 period he worked on the EPSRC TRANSNET programme in the development of intelligent transceivers capable for better approaching the Shannon limit by adapting to the characteristics of the transmission link. He is the author and co-author of over 96 international peer-reviewed journal and conference papers with numerous prestigious invited papers.

**Dr. Georgios Zervas** is an Associate Professor and EPSRC Fellow of optical networked systems at UCL. Georgios is leading a group of researchers on optical networked systems for data centers and high-performance computing. His research contributions and interests span across optical interconnects, switching, SDM/WDM/TDM multiplexing, networking, intelligent and fast network control, device and system design and optimisation, system architectures and their dependencies for building cloud and heterogeneous systems. His work has been published in over 250 peer-reviewed papers with over 4000 citations and has an h-index of 34. Dr Zervas has been the PI and Co-I in numerous national (EPSRC) and European (H2020) projects worth more than £2.75M and £8.5M to his institution as well as many industrial-funded projects. He is featured on the world Top 2% scientists for 2019 published on Mendeley Data by Stanford University researchers.