

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



# Global Optimization of Conventional and Holey Double-Clad Fibres by Stochastic Search

Ioannis Dritsas, Tong Sun and Ken Grattan  
City University London  
UK

## 1. Introduction

High power fibre lasers (HPFLs) find applications in the material processing, automotive, medical, telecoms and defence industries. Over 1kW of output power [1] has been demonstrated as the race to scale up the power while maintaining excellent beam quality and achieving impressive power conversion efficiency is ongoing. During the mature stages of the HPFLs technology, the automated simulation-based optimization of HPFLs is expected to contribute significantly to the formulation of optimal designs and to improve intuition for the conception of new fibre lasers. This chapter researches the common ground between computational photonics and direct search optimization methods with the prospect to propose optimized fibre designs for HPFLs.

Published work on the subject of pump light enhancement in the active core of cladding pumped fibres could be categorized as follows:

- a. Pump absorption ion system optimization [2-5]
- b. Pumping techniques focusing on how to couple more power into the inner cladding [6-10]
- c. Fibre designs that focus on maximizing the overlap between the coupled pump light and absorbent core volume [11-15]
- d. Holistic solutions that attempt to address (b) and (c) simultaneously [16-18]

Schemes in (c) are usually compatible with categories (b) and (d) meaning that the special fibres proposed by (c) can be pumped by schemes in (b) or they can be modified for use in the schemes of category (d) to further increase the pump absorption.

In category (b), Koplow *et al* [9] list a set of pumping schemes evaluation criteria and propose an embedded mirror side pumping scheme after discussing the contemporary pumping methods. Their technique initially appeared attractive and for that it was tested numerically within the computation environment of the simulation method proposed in [19]. It was found, however, that it does not benefit from the fibre cross section optimization because it reduces the percentage of higher order modes resulting in absorption degradation. Another side pumping technique which, in contrast with the previous, did not require machining of the pumped fibre was proposed by Polynkin *et al* [8]. A DCF was pumped via evanescent field coupling. This scheme appears to be fully compatible with the incorporation of optimized fibre topologies in place of the conventional circular inner cladding with centred core. Lassila [6] proposed a scalable side pumping scheme that could benefit from tailored axially symmetrical (presumably as far as the inner cladding is concerned) cross sections.

A pump absorption enhancing scheme that could fit in category (c) was proposed by Baek *et al* [14]. The authors incorporated a long period fibre grating (LPFG) in a cladding pumped configuration and measured a 35% increase in pump absorption as a direct result of the LPFG. A similar approach based on the reflection of the residual pump light was reported by Jeong *et al* [15]. The free end of the single-end pumped DCF was shaped into a right-angled cone that reflected more than 55% of the unabsorbed pump light that offered an 18% increase in absorbed pump power. This is one more scheme which could benefit from optimised fibre topologies. Recently, the use of a large area helical core was proposed [11] for the enhancement of pump absorption and simultaneous rejection of high order lasing modes naturally suggesting that optimized helical solid-state holes (that may be fabricated by rotation just like the helical core) could exhibit a similar tapering effect [19] as that reported here. This could avoid the need to increase the core area when increasing the inner clad area [12] to accept more pump power.

In the category of holistic solutions, Kouznetsov and Moloney proposed [16] and modelled analytically [17] the tapered slab delivery of multimode pump light to a small diameter inner cladding. This scheme combines the specially designed pump waveguide and corresponding inner cladding that could also fit in the shallow-angle single pumping category listed in [9]. It benefits highly from the coupling of multimode light into a narrow inner cladding while potential drawbacks are the leakage of high order pump modes and the fabrication difficulties. An alternative approach is demonstrated experimentally by Peterka *et al* [18]. The proposed DCF is single-end pumped and has a double-D cross section with the core at the centre of its half section. The input side is processed so that signal and pump delivery fibres can be spliced on the two specially fabricated facets. Overall, a promising way forward appears to be the development of generic and modular solutions within categories (a), (b) and (c) and then the synergistic combination of the three. This would act as a practical two stage approach that could amplify the pump absorption enhancement and consecutively the laser output power.

The results reported in this chapter fit in the aforementioned second category of pump absorption enhancing schemes. The interpretation of the original NM algorithm [20] as well as the deterministic cross section shape perturbation technique [21] are presented in this chapter in the form of structured pseudocode-functions. The proposed notation serves as the background for the development and validation of improved methods. Furthermore, additional fibre topology encoding schemes at higher dimensions are introduced and a modern interpretation of NM is given prior to the proposal of stochastically enhanced NM forms described in pseudocode syntax. The proposed algorithms are compared with commercial implementations of the genetic algorithm (GA) [22], generalised pattern search (GPS) [23-27] and mesh adaptive direct search (MADS) [28-29] methods that are also tested here for their performance and suitability. All the aforementioned algorithms share a set of common characteristics: they can operate exclusively on the function values (zeroth-order or derivative free or direct search methods) and they were tuned to their most parsimonious instances to the extent that their global convergence properties were not compromised. Here, the term global convergence is used to mean first order convergence to a point far enough from an arbitrary starting point. It does not mean convergence to a point  $x_*: f(x_*) \leq f(x), \forall x \in \mathcal{R}^n$  adhering to the terminology in the extensive review for direct search methods by Kolda *et al*. A third common characteristic is that they all call a 3-dimensional (3-D) fibre simulation method, described and validated in [19], in order to evaluate the objective function.

The contributions made in this chapter are summarized below:

- First reported stochastic simulation-based optimization of DCF topologies (to the best of the authors knowledge)
- Pseudocode descriptions of proposed algorithms for ease of verification and/or use by other researchers
- Benchmarking of several optimization algorithms with an emphasis on their statistical nature
- An optimization problem description scheme that allows the incorporation of inhomogeneous independent variables
- The proposal of perturbed stochastic search patterns as generalizations of the simplex formation pattern with possible applicability in pattern search algorithms
- The concept of implicitly constrained optimization via perturbed pattern search
- The proposal of the enhanced stochastically perturbed Nelder-Mead (ESPNM) method for implicitly constrained global optimization with simple bounds
- The unified description of NM, NM's stochastic forms, GPS and MADS methods based on the pattern search concept
- Mostly globally (as opposed to mostly locally in [21]) optimized DCF designs with an emphasis on manufacturability and modular design

The next section describes a set of optimization schemes on relatively low dimensions and compares NM, NM's stochastic variants (simple sampling Monte Carlo techniques), GA, GPS and MADS methods. Section 3 focuses on optimization schemes and algorithms in higher dimensions, introduces the perturbed patterns for simple and importance sampling Monte Carlo optimization and compares the locally introduced algorithms. The simulation parameters as well as the settings of each algorithm are given in section 4 where the optimization results for DCFs with polymer as well as air outer cladding are also discussed. Finally, section 5 concludes this chapter.

## 2. Bound-constrained zeroth-order optimization algorithms

The optimization problem considered in this chapter is

$$\min_{\mathbf{P} \in \Omega} f(\mathbf{P}), \quad f: \mathcal{R}^n \rightarrow \mathcal{R} \cup \{\infty\} \quad (1)$$

$$\text{where, } f(\mathbf{P}) = -P_{abs, tot}(\mathbf{P}), \quad (2)$$

$\mathbf{P}$  is a point in  $\mathcal{R}^n$ ,  $n$  is the number of variables and  $\Omega$  is the bounded function domain. Equation (2) gives the objective function which maps a DCF topology to the corresponding negative total absorbed pump power value [19]. The current notation partly adheres to that of [23] by assuming that

$$\Omega = \{\mathbf{P} \in \mathcal{R}^n : l \leq \mathbf{P} \leq u\} \quad \text{where } l, u \in \{\mathcal{R} \cup \{\pm\infty\}\}^n. \quad (3)$$

The optimization domain  $\Omega$  constitutes a declaration of the computational bounds and physically meaningful function domain. It acts as a barrier when applying the optimization algorithm not to  $f$  but to  $f_\Omega$  where

$$f_{\Omega} = \begin{cases} f(\mathbf{P}) & \text{if } \mathbf{P} \in \Omega \\ \infty & \text{otherwise} \end{cases} \quad (4)$$

The current work attempts to solve a simulation-based optimization problem where the objective function can be evaluated to only a few significant figures. This observation along with the noise that may be present in the computed function values or the expense of these computations render the calculation of derivatives impossible or impractical. Hence there is a need to treat the optimization problem with direct search methods.

The GA, GPS and MADS methods are implemented here via the commercially available 'genetic algorithm and direct search toolbox' within the MATLAB technical computing environment. The amount of subjective evaluation of the aforementioned algorithms was kept to a minimum by carefully tuning their parameters so that both global convergence and low computational cost are served in a well balanced way. Moreover, directly comparable sets of optimizations were performed by each method in order to gain statistical insight into the benefits of each algorithm and build intuition into their performance for a more objective judgment.

The NM simplicial search method has been comprehensively studied theoretically [30-32], extensively applied mostly in chemistry but also in optics [33-34], criticized for its inadequacies [35], remedied [36], enhanced [37,38] and even stochastically incorporated [39]. However, all theoretical improvements have led to a reduction in its computational efficiency. The main strength of the original algorithm is that when it succeeds it offers the best efficiency indicating that the most successful modifications of the simplex descent concept, with applications in computationally intensive problems, are expected to be those that maintain the number of function evaluations required to a minimum. Due to NM's susceptibility to different interpretations and the need to clearly and concisely describe the NM-based methods proposed here, its current interpretation is crystallized in algorithm NM and associated subalgorithms NM\_SimplexGener, FuncEval, SmxAssessm and NM\_Step. The later follows the modern practice examined by Lagarias [30] and is described in section 3 as a subset of subalgorithm ESPNM\_Step introduced there. Algorithm NM shows distinctively its two main operations being the generation of the initial simplex ( $\mathbf{S}_0$ ) along orthogonal directions around the start point (at line 3) and the line search procedure recursively executed (at line 10) by calling NM\_Step during an iteration (while loop: lines 8-12). The descent path is governed by the descent coefficient set {reflection, expansion, contraction, shrinkage} assigned in line 6. Line 2 of algorithm NM implies that the generation of the initial simplex (a polytope in  $\mathcal{R}^n$  with  $n+1$  vertices - the minimum statistical information required to capture first order information) is essentially a pattern search operation along all  $n$ -directions denoted by the column vectors of the  $n \times n$  pattern matrix ( $\mathbf{\Xi}_{\text{NM}}$ ) which in this case is practically the identity matrix ( $\mathbf{\Xi}_{\text{NM}} \equiv \mathbf{I}_{n \times n}$ ). The initial simplex is generated by the subalgorithm NM\_SimplexGener with respect to the start point and vector  $\mathbf{M}$  where the mesh sizes of the all independent variables are stored. In this way, the simultaneous optimization of inhomogeneous variables (of different physical meaning, units, mesh size) is practically implemented. An example is the case where the diameter and refractive index of an embedded hole are simultaneously optimized. Essentially, this is the integration of a parametric optimization procedure into a more robust non-parametric optimization scheme.



An important implication of subalgorithm NM\_SimplexGener is that it should form a nondegenerate initial ( $j = 0$ ) simplex. That is,

$$\text{vol}(\mathbf{S}_j) = \frac{\left| \det(\mathbf{P}_1^{(j)} - \mathbf{P}_{n+1}^{(j)}, \mathbf{P}_2^{(j)} - \mathbf{P}_{n+1}^{(j)}, \dots, \mathbf{P}_n^{(j)} - \mathbf{P}_{n+1}^{(j)}) \right|}{n!} > 0 \quad (5)$$

The satisfaction of inequality (5) is important in order to conserve the numerical integrity of the 'flexible polytope' when descending in  $\mathfrak{R}^n$  and avoid convergence to a non-minimizer after collapsing one or more of its vertices on the hyperplane of others [35].

A simple sampling Monte Carlo approach is exercised here by means of the stochastic Nelder-Mead method (SNM) with the prospect to increase NM's efficiency and probability to find a global minimizer in low dimensions. The SNM method is partly implemented by substituting line 2 in algorithm NM with

**Algorithm NM.** Interpretation of the modern Nelder-Mead (NM) method:

$$[\mathbf{P}_l, \mathbf{f}_l, \sigma_j] = \text{NM}(\mathbf{P}_1, \mathbf{M}, \sigma_{\text{halt}}, \mathbf{\Omega})$$

*Input:* (start point  $\mathbf{P}_1 = [p_{1,1} \ p_{2,1} \ \dots \ p_{n,1}]^T$  in  $\mathfrak{R}^n$ , mesh size vector  $\mathbf{M} = [m_1 \ m_2 \ \dots \ m_n]^T$ , stopping value for the halting criterion and optimization domain  $\mathbf{\Omega} = [\mathbf{B}_1 \ \mathbf{B}_2 \ \dots \ \mathbf{B}_n]^T$  where  $\mathbf{B}_i = [p_{i,\min} \ p_{i,\max}]^T \mid_{i=1,2,\dots,n}$  (bounds)). *Output:* [optimal point, corresponding function value, standard deviation of  $\{\mathbf{f}_i \mid_{i=1,2,\dots,n+1; i \neq h}\}$  after the last iteration].

```

1   $j := 0$  // iteration index
2   $\Xi_{\text{NM}} := [\xi_1 \ \xi_2 \ \dots \ \xi_n] \equiv \mathbf{I}_{n \times n}$  // initial simplex formation pattern
3  call  $[\mathbf{S}_0] = \text{NM\_SimplexGener}(\mathbf{P}_1, \Xi_{\text{NM}}, \mathbf{M}, n)$  // nondegenerate initial simplex
4  for each  $\{\mathbf{P}_i \mid_{i=1,2,\dots,n+1}\}$  call  $s[\mathbf{f}_i] = \text{FuncEval}(\mathbf{P}_i, \mathbf{\Omega})$  endfor // objective function evaluations
5   $\mathbf{F}_j := [f_1 \ f_2 \ \dots \ f_{n+1}]_{1 \times (n+1)}$  // initial objective matrix
6   $\{r, e, c, s\} := \{1, 2, 1/2, 1/2\}$  // descent coefficients standard values
7  call  $[\mathbf{f}_h, \mathbf{f}_l, \mathbf{P}_h, \mathbf{P}_l, \bar{\mathbf{f}}, \bar{\mathbf{P}}] = \text{SmxAssessm}(\mathbf{S}_j, \mathbf{F}_j)$  // current simplex ( $\mathbf{S}_j$ ) assessment
8  while  $(\sigma_j \geq \sigma_{\text{halt}})$  // where,  $\sigma_j = \left( \left\langle (\mathbf{f}_i - \bar{\mathbf{f}})^2 \mid_{i=1,2,\dots,n+1; i \neq h} \right\rangle \right)^{1/2}$  (descent halting criterion)
9       $j := j + 1$  // increment
10     call  $[\mathbf{S}_j, \mathbf{F}_j, \text{step}] = \text{NM\_Step}(\mathbf{P}_h, \mathbf{P}_l, \bar{\mathbf{P}}, \mathbf{\Omega}, \mathbf{f}_h, \mathbf{f}_l, r, e, c, s, \mathbf{S}_j, \mathbf{F}_j)$  // NM step
11     call  $[\mathbf{f}_h, \mathbf{f}_l, \mathbf{P}_h, \mathbf{P}_l, \bar{\mathbf{f}}, \bar{\mathbf{P}}] = \text{SmxAssessm}(\mathbf{S}, \mathbf{F})$  // simplex assessment
12 endwhile // end of iteration loop
13 return  $\mathbf{P}_l, \mathbf{f}_l, \sigma_j$  // output.
```

**Subalgorithm NM\_SimplexGener.** NM initial simplex generation:

$$[S_0] = \text{NM\_SimplexGener}(\mathbf{P}_1, \Xi_{\text{NM}}, \mathbf{M}, n)$$

*Input:* (start point  $\mathbf{P}_1$ , NM pattern, mesh size vector and length of  $\mathbf{P}_1$ ). *Output:* [initial simplex matrix].

```

1  for each simplex vertex in the set  $\{\mathbf{P}_i \mid i=2,3,\dots,n+1\}$ 
2       $\mathbf{P}_i := \mathbf{P}_1 + (\mathbf{M} \circ \xi_{i-1})$ 
        // where,  $\circ : [a_{ij}]_{m \times n} \circ [b_{ij}]_{m \times n} = [a_{ij}b_{ij}]_{m \times n}$ 
3  endfor
4       $\mathbf{S}_0 := [\mathbf{P}_1 \mathbf{P}_2 \cdots \mathbf{P}_{n+1}]_{n \times (n+1)} \mid \text{vol}(\mathbf{S}_0) > 0$ 
5  return  $\mathbf{S}_0$                                      // output.
```

**Subalgorithm FuncEval.** Objective function evaluation:

$$[f_i] = \text{FuncEval}(\mathbf{P}_i, \Omega)$$

*Input:* (point in  $\mathcal{R}^n$ , bounds). *Output:* [function value].

```

1  if  $\mathbf{P}_i \in \Omega$  then
2       $f_i := f(\mathbf{P}_i)$                                      // compute (simulate)
3  else
4       $f_i := +\infty$                                      // assign a large positive value
5  endif
6  return  $f_i$                                            // output.
```

**Subalgorithm SmxAssessm.** Simplex assessment:

$$[f_h, f_l, \mathbf{P}_h, \mathbf{P}_l, \bar{f}, \bar{\mathbf{P}}] = \text{SmxAssessm}(\mathbf{S}, \mathbf{F})$$

*Input:* (simplex, objective matrices). *Output:* [worse, best function values, corresponding points, mean function value and centroid for  $i \neq h$ ].

```

1   $f_h := \max\{f_i \mid i=1,2,\dots,n+1\}$ ; assign  $\mathbf{P}_h \mid f_h = f(\mathbf{P}_h)$ 
2   $f_l := \min\{f_i \mid i=1,2,\dots,n+1\}$ ; assign  $\mathbf{P}_l \mid f_l = f(\mathbf{P}_l)$ 
3   $\bar{f} := \langle f_i \mid i=1,2,\dots,n+1; i \neq h \rangle$ ;  $\bar{\mathbf{P}} := \langle \mathbf{P}_i \mid i=1,2,\dots,n+1; i \neq h \rangle$ 
4  return  $f_h, \mathbf{P}_h, f_l, \mathbf{P}_l, \bar{f}, \bar{\mathbf{P}}$                  // output.
```

$$\Xi_{\text{SNM}} = (\xi_1, \xi_2, \dots, \xi_n)_{n \times n} = \begin{bmatrix} m_1 & 0 & \cdots & 0 \\ 0 & m_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & m_n \end{bmatrix} \quad (6)$$

where  $\{m_i | i=1,2,\dots,n\} \in [-1,1]$  is a set of uniformly distributed pseudorandom numbers. A short description of the pseudorandom number generator used is given in section 4. SNM can use subalgorithm NM\_SimplexGener to generate the initial simplex after substituting  $\Xi_{\text{NM}}$  with  $\Xi_{\text{SNM}}$  in the set of input arguments. During the generation of  $\mathbf{S}_0$  around  $\mathbf{P}_1$ , on the basis of  $\Xi_{\text{SNM}}$ , a set of randomly signed orthogonal directions are searched while the initial mesh sizes fluctuate randomly as well (between zero and their nominal values stored in  $\mathbf{M}$ ). The second and last part of the SNM implementation is to discard line 6 of algorithm NM and to add the following line

$$\text{assign } \{r, e, c, s\}_j \in \{[0.5, 1], [2, 4], [0.25, 0.5], [0.3, 0.7]\} \quad (7)$$

just before line 10. The later means that the descent coefficients are recursively set to uniformly distributed random values within the designated ranges. With regard to the modern understanding of the Nelder-Mead algorithm, the descent coefficients must satisfy the conditions  $r \in (0, +\infty)$ ,  $e \in \{(1, +\infty) \cap (r, +\infty)\}$ ,  $c \in [0, 1]$  and  $s \in [0, 1]$ . According to Lagarias [30], the condition  $r \in (0, +\infty)$  is not stated explicitly in the original paper by Nelder and Mead but is implicit in the presentation of the original algorithm [20].

A significant role during an optimization is played by the corresponding optimization problem encoding key which orderly stores the independent variables of a fibre topology in a column vector (point  $\mathbf{P}$  in  $\mathfrak{R}^n$ ) that is read by the fibre simulator. The construction of the computation grid and/or the setting of the simulation parameters involved in the evaluation of the objective function are then based on the information encoded in the coordinates of  $\mathbf{P}$ . For fixed perimetric lines of laminae participating in a DCF cross section, the following encoding keys are used in this chapter:

$$\mathbf{P} = (y, z, y_{h,1}, z_{h,1}, y_{h,2}, z_{h,2}, \dots, y_{h,N}, z_{h,N})^T \quad (8)$$

in  $\mathfrak{R}^{2(N+1)}$  for an inner cladding topology embedding  $N$ -holes and a single active core. The first pair  $(y, z)$  of elements represents the coordinates of the core centroid on the cross section plane while each pair in the set  $\{(y_{h,i}, z_{h,i}) | i=1,2,\dots,N\}$ , appearing in  $\mathbf{P}$ , represents the centroid coordinates of the  $i$ -th hole. Equation (8) encodes a fibre topology according to the 'Offset' optimization scheme under which the centroid coordinates of each involved lamina is optimized independently. Following the same notation, the point

$$\mathbf{P} = (y, z, y_{h,1}, z_{h,1}, \dots, y_{h,N}, z_{h,N}, d_1, \dots, d_N)^T \quad (9)$$

in  $\mathfrak{R}^{3N+2}$  encodes the same topology under the 'Offset-Diameter' scheme that, in addition to (8), allows the simultaneous optimization of circular hole diameters or square hole side lengths. Furthermore, the point



$$\mathbf{P} = (y, z, y_{h,1}, z_{h,1}, \dots, y_{h,N}, z_{h,N}, A_1, \dots, A_N, B_1, \dots, B_N)^T \quad (10)$$

in  $\mathfrak{R}^{2(2N+1)}$  allows, in addition to (8), the independent optimization of the horizontal and vertical characteristic dimension of each hole (major-minor axis of an ellipse for initially circular holes and length-height of a parallelepiped for initially square holes). Encoding keys (9) and (10) demonstrate the need for individually defined mesh sizes tailored to the domains within which the search for optimal values is desired. This view is reinforced by

$$\mathbf{P} = (y, z, y_{h,1}, z_{h,1}, \dots, y_{h,N}, z_{h,N}, A_1, \dots, A_N, B_1, \dots, B_N, R_1, \dots, R_N)^T \quad (11)$$

that includes variables representing refractive index values. The independent variables in (11) are inhomogeneous not only in terms of corresponding mesh size and domain but also in physical meaning and units. In this case, point  $\mathbf{P}$  in  $\mathfrak{R}^{5N+2}$  represents the 'Offset-Major-Minor-Index' optimization scheme that expands (9) by including the refractive indices  $\{R_i |_{i=1,2,\dots,N}\}$  of dielectric holes embedded in the inner cladding.

Four groups of thirty optimizations are executed next in each of the  $\mathfrak{R}^{10}$  and  $\mathfrak{R}^{18}$  spaces in order to compare the performance of SNM variants with algorithm NM and in relation to the dimensionality of the optimization space. To avoid fragmentation, it is thought adequate for the current discussion to report that all optimizations were initiated from the same start point. The later represents a double-clad design with a polymer outer cladding and four rods embedded in the inner cladding (solid-state circular dielectric holes) assumed to be made of CBYA alloy glass with a refractive index of 1.430 [40,19].

Figure 1 demonstrates the  $\mathfrak{R}^{10}$  sets of optimizations executed following different strategies. The type of search strategy is denoted by the {Initial simplex, Descent coefficients} pair where the letter D denotes deterministic as opposed to S denoting stochastic implementation. The initial circular inner cladding topology with centred core included four symmetrically embedded circular holes at the corners of a centred square and absorbed  $P_{abs,tot} = 8.60W$ . Due to the high number of optimizations required, lower resolution than in

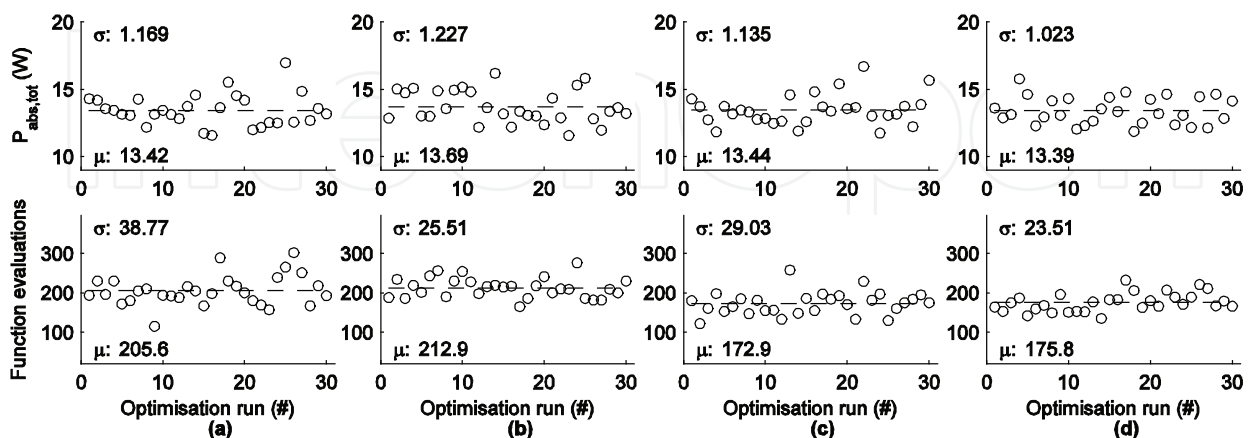


Fig. 1. Four groups of 30 optimizations in  $\mathfrak{R}^{10}$  from the same starting point and under different optimization strategies: (a) SNM{D,D}  $\equiv$  NM (here for variable mesh size). (b) SNM{S,D}. (c) SNM{D,S}. (d) SNM{S,S}.

section 4 was used here after verifying that approximately the same trends were followed. The fibre was 1cm long and 126 rays carried the pump energy while the rest of the parameters were kept constant. The graphs along the first row of figure 1 plot the values of the total absorbed pump power (optimized as a function of the core and hole offsets) while those along the second row present the corresponding number of objective function evaluations recorded prior to convergence. The mean value ( $\mu$  - dashed line) and standard deviation ( $\sigma$ ) of the plotted values is also reported in each graph. Figure 1(a) (1st column) reveals the influence of the mesh size random variance on the NM results. The SNM{S, D} strategy results are shown in figure 1(b) where the initial simplex vertices are formed stochastically while the simplex descent is based on deterministic coefficient values. Figure 1(c) corresponds to the case of constant initial simplex (that of the first optimization in figure 1(a)) but this time the value of each optimization coefficient is recursively and randomly determined prior to each iteration during the simplex descent (SNM{D, S} strategy). Finally, figure 1(d) presents the results for the case where both the initial simplex and descent coefficients are randomly determined (SNM{S, S}). All optimizations in figure 1 were initiated from the same start point. The results variations observed in figures 1(b)-(d) are attributed solely to the stochastic nature of SNM while those in figure 1(a) originate from the mesh size variations. The best performing optimization strategy in  $\mathfrak{R}^{10}$  can be chosen on different criteria serving different applications. The strategy that delivers acceptably optimized objective function values with minimum uncertainty is preferred here. It offers the smallest spread of objective function values for the second lowest mean number of function evaluations.

Figure 2 presents the corresponding study in  $\mathfrak{R}^{18}$  where the area and the ellipticity of the four holes are optimized in addition to the core and hole offsets previously optimized in  $\mathfrak{R}^{10}$ . The four examined strategies are presented here in the same order as in figure 1. Strategy (b) is preferred in this case because it offered the highest mean absorption at the highest certainty. This comes at the cost of the maximum mean number of function evaluations exhibiting this time the strongest spread around their mean value. In both  $\mathfrak{R}^{10}$  and  $\mathfrak{R}^{18}$  spaces it appears that SNM{D, S} offers the lowest number of function evaluations and, more importantly, a slower growth in function evaluations with increasing dimensions [29]. This is a highly desired feature for the optimization of expensive objective functions.

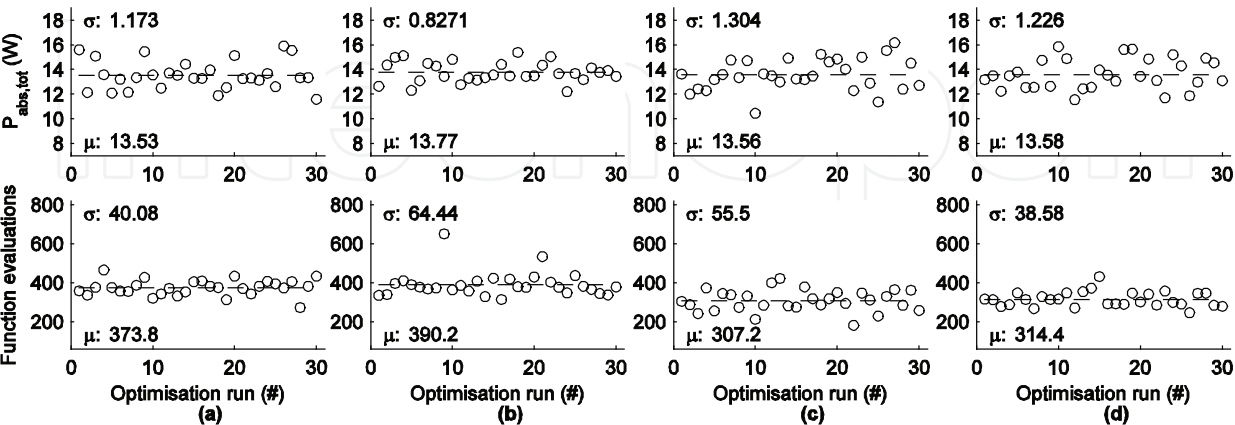


Fig. 2. Four groups of 30 optimizations in  $\mathfrak{R}^{18}$  from the same starting point and under different optimization strategies: (a) SNM{D,D}  $\equiv$  NM. (b) SNM{S,D}. (c) SNM{D,S}. (d) SNM{S,S}.

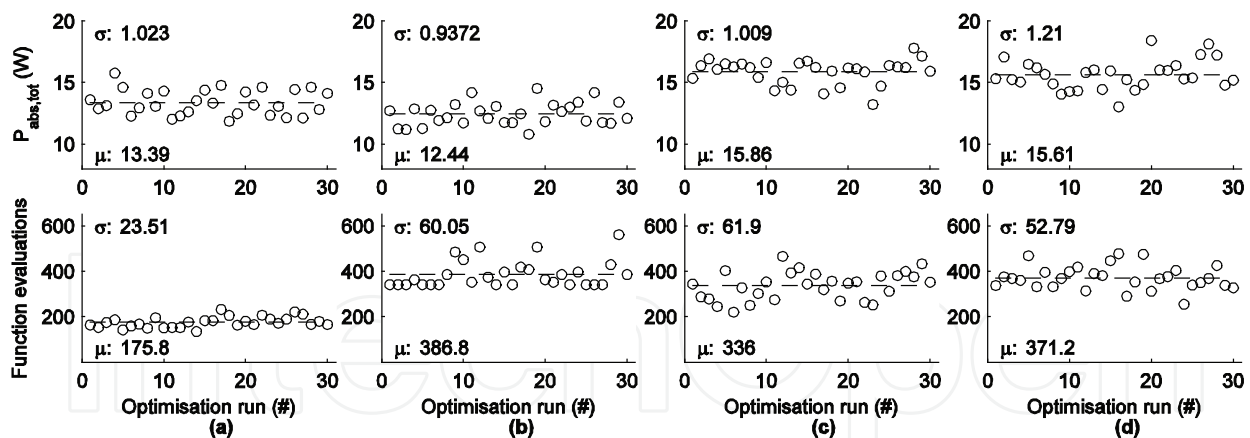


Fig. 3. Four groups of 30 optimizations in  $\mathcal{R}^{10}$  from the same starting point and driven by different algorithms: SNM{S,S}, GA{Np1}, GPS{Np1,2N}, MADS{Np1,2N}.

The fittest SNM strategies are compared next with three global optimization methods operating in  $\mathcal{R}^{10}$  and  $\mathcal{R}^{18}$  in figures 3 and 4 correspondingly. Figure 3(a) plots again the results for algorithm SNM{S, S} while figures 3(b)-(d) report the corresponding results from GA, GPS and MADS methods. The detailed set-up of each method is reported in section 4. The expression GA{Np1} denotes that each GA optimization started with  $(n+1)$  initial population members generated by random sampling of  $\Omega$ . By GPS{Np1, 2N} it is meant that the search pattern includes  $n+1$  directions and that the poll pattern matrix stores  $2n$  directions. GPS and MADS algorithms implement two distinct steps namely the search and poll. The search step can be absent or be a pattern search or any other heuristic or Monte Carlo method [41] or preferably a method that uses inexpensive surrogates to approximate the objective function [42]. The search step adopted in this work implements a pattern search along the directions denoted by the column vectors in the pattern matrix

$$\Xi_{\text{GPS,Search}} \equiv \Xi_{\text{GPS,Np1}} = \begin{bmatrix} 1 & 0 & \cdots & 0 & -1 \\ 0 & 1 & \cdots & 0 & -1 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -1 \end{bmatrix}_{n \times (n+1)}. \quad (12)$$

The poll step is a compulsory pattern search that is closely linked to the convergence theory of pattern search algorithms [29]. The adopted poll patterns are represented by the column vectors in

$$\Xi_{\text{GPS,Poll}} \equiv \Xi_{\text{GPS,2N}} = \begin{bmatrix} 1 & 0 & \cdots & 0 & -1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & -1 \end{bmatrix}_{n \times 2n}. \quad (13)$$

The GPS algorithm invokes the poll step only when the search step fails to produce a point in  $\mathcal{R}^n$  that improves the optimal function value recorded so far. After a poll step, the mesh size is adapted (contracts after an unsuccessful poll and expands after a successful poll) and

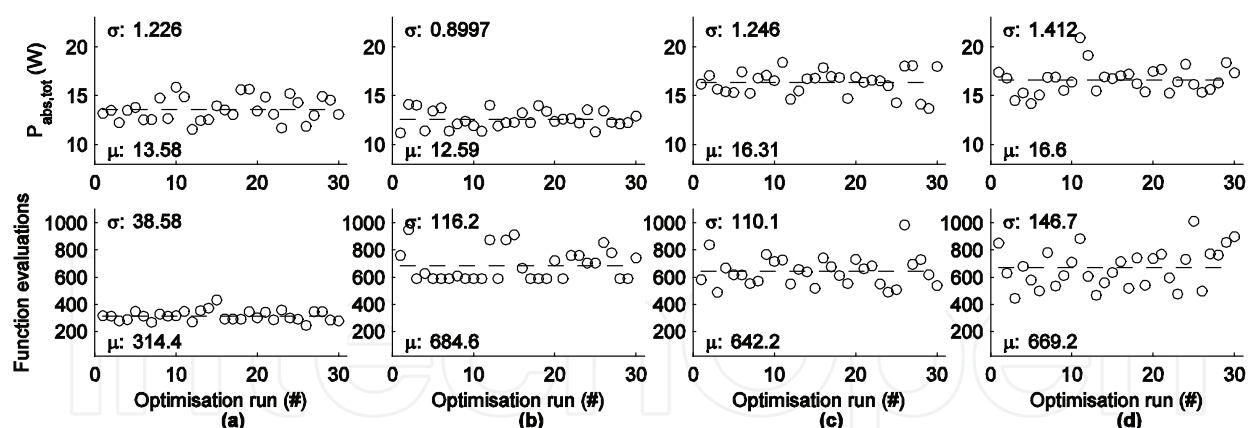


Fig. 4. Four groups of 30 optimizations in  $\mathcal{R}^{18}$  from the same starting point and driven by different algorithms: SNM{S,S}, GA{Np1}, GPS{Np1,2N}, MADS{Np1,2N}.

a new iteration begins. MADS is a stochastic form of GPS. The  $\Xi_{\text{MADS,Search}} \equiv \Xi_{\text{MADS,Np1}}$  pattern matrix stores  $n + 1$  randomly generated column vectors while  $\Xi_{\text{MADS,Poll}} \equiv \Xi_{\text{MADS,2N}}$  is generated using a random permutation of an  $(n \times n)$  linearly independent lower triangular matrix. Both of the above patterns are regenerated prior to each iteration according to MATLAB's documentation.

Before discussing the results in figure 3, it is informative to note that the variations in the GPS optimization results are due to the use of a different mesh size for each optimization whilst all other results exhibit variations originating from the stochastic nature of the corresponding algorithm. SNM, GA, GPS and MADS achieved an average objective improvement of 56%, 45%, 84% and 81% correspondingly. In  $\mathcal{R}^{18}$  (figure 4) the corresponding percentages are 58%, 46%, 90% and 93%. It is obvious at this stage that GPS and MADS managed to find optimizers located in deeper valleys indicating global convergence with higher probability than GA and SNM. On the computational expense front in  $\mathcal{R}^{10}$  the GA, GPS and MADS were correspondingly 121%, 91% and 111% more expensive than SNM while in  $\mathcal{R}^{18}$  they were 118%, 104% and 113% more expensive than SNM. The GA is consistently the most expensive method. The reported results agree with other benchmark results [43,44] and although GA promises global convergence when evolving a large initial population [45], it is not preferred here due to it being unsuitable for the optimization of expensive functions. The above analysis indicates that in the examined dimensions the most efficient strategy would be to use SNM as a first stage optimization tool, a numerical telescope that can relatively inexpensively designate the vicinity that offers the highest probability to contain a global optimizer. A second stage search with the significantly more expensive GPS or MADS methods is then justified in the SNM designated subdomains. Nevertheless, and in agreement with section 4 results, the SNM method offers the best case efficiency when it succeeds in finding a global optimizer.

### 3. Implicitly constrained zeroth-order optimization algorithms with simple bounds

The stochastic forms of NM proposed in this section solve optimization problems in higher dimensions that are difficult to treat or incompatible with GA, GPs and MADS. In addition they achieve global convergence at low computational cost. The 'Offset-Perimeter' encoding

key ([21] gives a schematic representation) is used to map a variable perimetric line shape for each lamina comprising a fibre cross section. Under this scheme, the shape of a given cross section can be fully optimized but at a considerably higher computational cost. The dimensionality of the objective function domain increases by at least an order of magnitude depending on the sampling density of each lamina perimeter included in a cross section. A fibre topology that includes  $N$  -holes in the inner cladding is represented in  $\mathbb{R}^{2(n_c+n_{h1}+\dots+n_{hN}+1)}$  by a single point of the form

$$\mathbf{P} = \left( y, z, y_{c,1}, \dots, y_{c,n_c}, z_{c,1}, \dots, z_{c,n_c}, y_{h1,1}, \dots, y_{h1,n_{h1}}, z_{h1,1}, \dots, z_{h1,n_{h1}}, \dots, y_{hN,1}, \dots, y_{hN,n_{hN}}, z_{hN,1}, \dots, z_{hN,n_{hN}} \right)^T \quad (14)$$

where  $n_c$  is the number of points that sample the inner cladding perimeter and  $n_{hi} \mid_{i=1,2,\dots,N}$  is the  $i$ -th hole perimetric point set population. The aforementioned encoding key includes the core centre coordinates but does not optimize the hole offsets. However, this is a feature that could be included into the coordinates set of  $\mathbf{P}$ .

Even for a low resolution polygonic approximation of a smooth perimeter, all the previously compared algorithms generate trial points that abruptly perturb a smooth start point and lack physical integrity and/or manufacturability. Examples of such perturbations are given in figures 5(a)-(c) showing typical trial points that the corresponding algorithms NM, GPS, MADS may generate during an optimization. Most representative trial points are those of the GA algorithm shown in figures 5(d), 5(e) for two different bounding configurations. It becomes obvious that GA scrambles randomly the start point coordinates failing to produce children or members of the initial population with physical integrity. Figure 5(e) suggests that a scheme capable of generating smooth perturbations is needed. An effort was

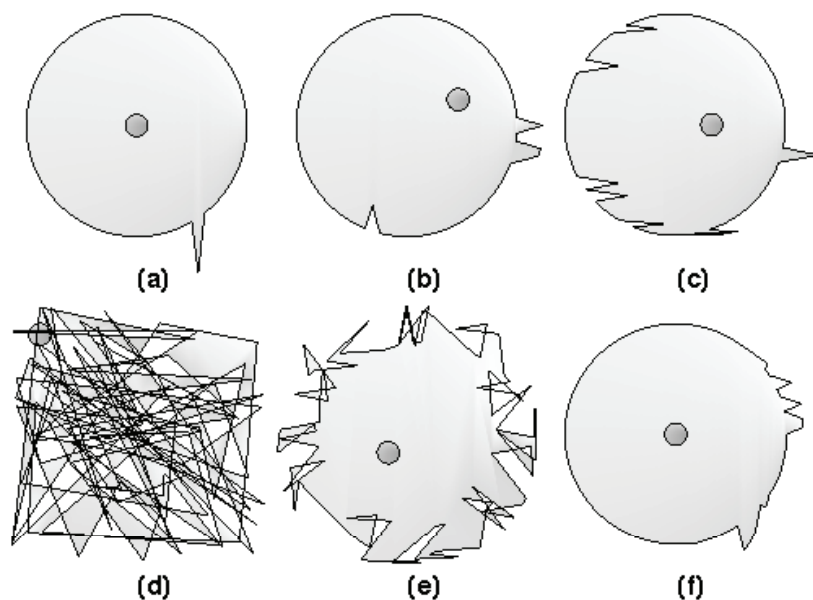


Fig. 5. Trial points (or initial population members). (a) NM. (b) GPS. (c) MADS. (d) GA with own population, bounded within  $[-4,4]$ mm (e) GA with own population, bounded within the  $\pm 50\mu\text{m}$  zone from start point. (f) GA with PNM initial population.



made to construct suitable constraints that would force the mapped coordinates to change in groups forming smooth, local and able to propagate perturbations along the perimeter of a lamina but it appears that this is a non-functioning approach.

Although algorithm NM is not meant for constrained optimization it was found that it can be modified to perform implicitly constrained optimization. The outline of the related process is that after generating a suitable pattern, the vertices of the initial simplex could obey pattern imprinted constraints which propagate all the way to the convergence point at the end of a descent. The simplest implementation of the above concept is implemented via the perturbed Nelder-Mead (PNM) algorithm and by virtue of subalgorithms PNM\_PattGener and PNM\_SmxGener. The former of the subalgorithms generates a pattern of the form

$$\Xi_{\text{PNM}} = \begin{bmatrix} \nu_2 & \nu_1 & 0 & \dots & 0 \\ \nu_3 & \nu_2 & \nu_1 & \dots & 0 \\ 0 & \nu_3 & \nu_2 & \dots & 0 \\ 0 & 0 & \nu_3 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & \nu_1 \\ 0 & 0 & 0 & \dots & \nu_2 \end{bmatrix}_{n \times n} \quad (15)$$

when the perturbed element group population is  $k=3$ . Equation (15) demonstrates essentially the propagation of a constant disturbance involving  $k$ -elements along the length of the additive identity ( $n \times 1$  zero vector). In line 1 of PNM\_PattGener, the set  $\{\nu_q\} \in (0, 1/\sigma_N \sqrt{2\pi})|_{q=1,2,\dots,k}$  with statistical median  $\nu_\mu = \nu_{(k+1)/2}$ , follows the normal distribution  $N(\mu, \sigma_N^2)$  where  $\sigma_N$  is the predefined standard deviation of the distribution with a probability density function shown in the line 1 comment. It is notable that  $\Xi_{\text{PNM}} \equiv \Xi_{\text{NM}} = \mathbf{I}$  when  $k=1$  and  $\sigma_N = 1/\sqrt{2\pi}$ , indicating that  $\Xi_{\text{PNM}}$  is a generalization of

**Algorithm PNM.** The Perturbed Nelder-Mead (PNM) method:

$$[\mathbf{P}_l, f_l, \sigma_j] = \text{PNM}(\mathbf{P}_1, M, \sigma_{\text{halt}}, \mathbf{\Omega}, \sigma_N, k)$$

*Input: (as in algorithm NM but with scalar mesh size and in addition, standard deviation of the normal distribution and perturbed element set population (odd positive integer:  $k = 2\tau + 1; \tau \in \mathbb{Z}^*$ )). Output: [as in algorithm NM].*

```

1                                     // same as in algorithm
    NM
2  call  $[\Xi_{\text{PNM}}] = \text{PNM\_PattGener}(n, \sigma_N, k)$            // PNM pattern
    generation
3  call  $[\mathbf{S}_0] = \text{PNM\_SmxGener}(\mathbf{P}_1, \Xi_{\text{PNM}}, M, n)$          // initial simplex generation
1-13
NM                                     // same as in algorithm
```



**Subalgorithm PNM\_PattGener.** Perturbed Nelder-Mead (PNM) pattern generation:

$$[\Xi_{\text{PNM}}] = \text{PNM\_PattGener}(n, \sigma_N, k)$$

*Input:* (number of variables, standard deviation of the normal distribution, perturbed element set population (odd positive integer:  $k = 2\tau + 1; \tau \in \mathbb{Z}^*$ )). *Output:* [PNM pattern matrix].

```

1   $\mathbf{N} := [\nu_1 \nu_2 \cdots \nu_k]^T \big|_{k=2\tau+1; \tau \in \mathbb{Z}^*}$  // where  $\{\nu_q \mid q=1,2,\dots,k\} = (1/\sigma_N \sqrt{2\pi}) \exp[-(q-\mu)^2 (2\sigma_N^2)^{-1}]$ 
2   $\varepsilon := (k-1)/2$  // number of variables in either bell shape branch excluding the median ( $\mu$ )
3  for each PNM pattern-matrix column in the set  $\{\xi_i \mid i=1,2,3,\dots,n\}$ 
4       $\xi_i := (\xi_{i1}, \xi_{i2}, \dots, \xi_{in})^T \equiv (0, 0, \dots, 0)^T$  // additive identity ( $n \times 1$  zero vector)
5       $\left[ (\xi_{i-\varepsilon}, \xi_{i-(\varepsilon-1)}, \dots, \xi_{i-1}, \xi_i, \xi_{i+1}, \dots, \xi_{i+(\varepsilon-1)}, \xi_{i+\varepsilon})^T \right]_{k \times 1} := \mathbf{N}$  // bell shaped perturbation
6  endfor
7   $\Xi_{\text{PNM}} := (\xi_1, \xi_2, \dots, \xi_n)_{n \times n}$  // PNM pattern matrix
8  return  $\Xi_{\text{PNM}}$  // output.
```

**Subalgorithm PNM\_SmxGener.** PNM initial simplex generation:

$$[\mathbf{S}_0] = \text{PNM\_SmxGener}(\mathbf{P}_1, \Xi_{\text{PNM}}, M, n)$$

*Input:* (start point  $\mathbf{P}_1$ , PNM pattern, mesh size and length of  $\mathbf{P}_1$ ). *Output:* [initial simplex matrix].

```

1  for each simplex vertex in the set  $\{\mathbf{P}_i \mid i=2,3,\dots,n+1\}$ 
2       $\mathbf{P}_i := \mathbf{P}_1 + M(1/\xi_{\max,i-1})\xi_{i-1}$ 
        // where,  $\xi_{\max,i-1} = \max\{\xi_{w,i-1} \mid w=1,2,\dots,n\}$ 
3  endfor
4   $\mathbf{S}_0 := [\mathbf{P}_1 \mathbf{P}_2 \cdots \mathbf{P}_{n+1}]_{n \times (n+1)} \big|_{\text{vol}(\mathbf{S}_0) > 0}$ 
5  return  $\mathbf{S}_0$  // output.
```

$\Xi_{\text{PNM}}$ . Subalgorithm PNM\_SmxGener returns the initial simplex vertices as a result of the superposition between the start point ( $\mathbf{P}_1$ ) and the search directions stored in  $\Xi_{\text{PNM}}$ . The practical outcome is the propagation of a bell-shaped perturbation along the elements in  $\mathbf{P}_1$  and is illustrated in figure 6 which assumes that  $k = 3$  and shows clearly the  $n$ -steps of the perturbation propagation process which generates the initial simplex vertices  $\mathbf{P}_2$  to  $\mathbf{P}_{n+1}$ . Also clearly demonstrated is that a set of vertices created at the start and the end of the process bare a perturbation that is abrupt at one end. This is a drawback of the described technique that has a small overall effect though due to the comparatively small number of vertices bearing such a non-smooth perturbation. Soon after the start of the simplex decent,

the abruptly perturbed vertices are naturally substituted by newly discovered and better performing smoothly perturbed vertices. The only damage made is the comparatively small reduction in the probability to capture optimal vertices right from the start of the process. The height of the bell shape is controlled, in subalgorithm PNM\_SmxGener, via the mesh size  $M$  while its full-width half-maximum is set via  $\sigma_N$ . The factor  $1/\xi_{\max}$ , used in line 2, normalizes the bell-shaped perturbation, stored in the pattern, to the maximum value of 1 in order to scale the perturbation height to the predefined mesh size  $M$ . A set of decoded initial simplex vertices is given in figure 7(a) where the start point was a cross section with circular inner cladding embedding an offset circular hole and an offset core. For completeness, figure 5(f) shows a child produced by GA after having been initiated with the same initial population that comprised the initial simplex vertices in PNM. Here the child's features have been improved compared to figures 5(d), 5(e) but still the GA algorithm appears unable to generate a smooth optimizer.

Following the proposal of SNM method in section 2, algorithm PNM naturally suggests its stochastic version SPNM which can be implemented by the simultaneous random assignment of  $(M, \sigma_N)$  and/or the simplex descent coefficients. The random assignment of  $(\sigma_N, M)$  is implemented just before the generation of  $\Xi_{\text{SPNM}}$  which now stores, as opposed to  $\Xi_{\text{PNM}}$ , a set of directions that still smoothly but this time randomly perturb the

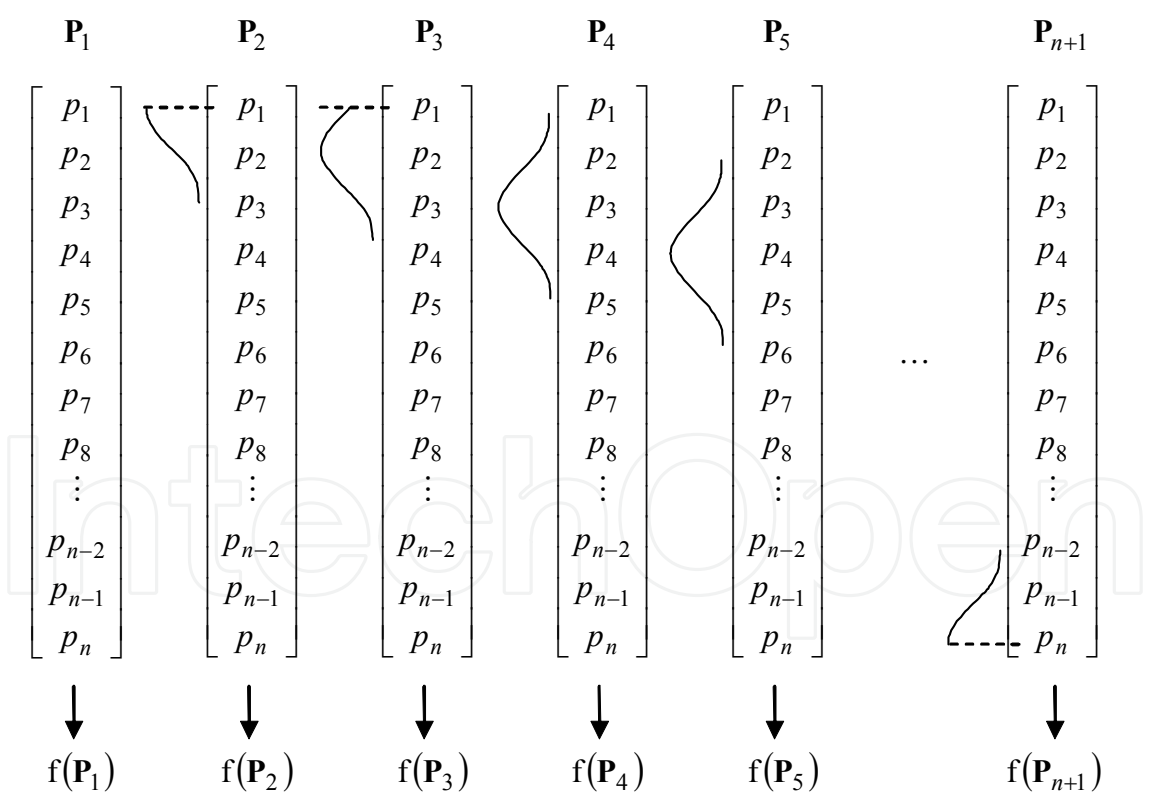


Fig. 6. Illustration of the bell shape propagation in the nonrandomized initial simplex generation scheme for perturbed vertex elements number  $k = 3$ . Under this scheme, the shape of the perturbation propagates along the whole vertex in  $n$ -steps while preserving its shape.

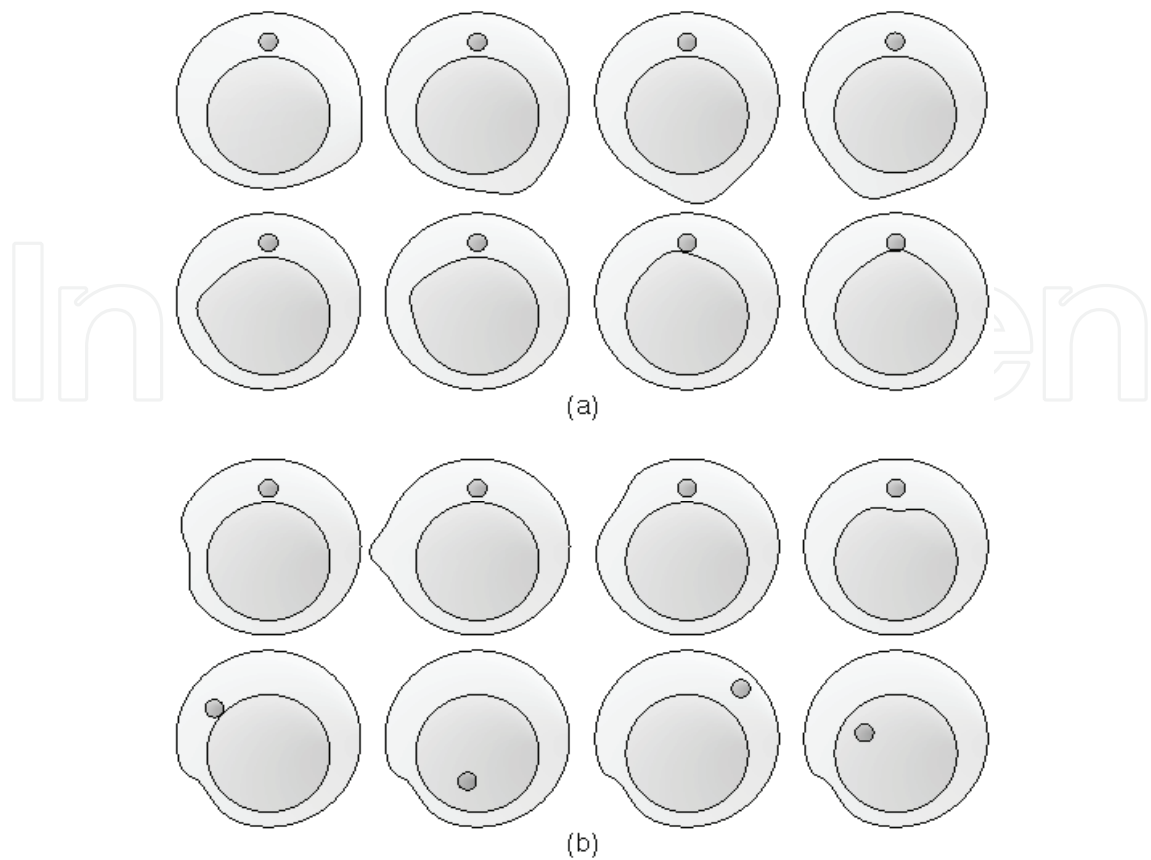


Fig. 7. Propagation instances of a perturbation envelope. (a) PNM method. (b) ESPNM method stochastic envelope (1st row) and random core offsets inside a selected vertex (importance sampling- 2nd row).

additive identity. Both PNM and SPNM algorithms still call the same iteration subalgorithm (NM\_Step) as NM and SNM do. The assignment of the random  $M, \sigma_N$  parameter values is implemented as in algorithm ESPNM (enhanced stochastically permuted NM) proposed next. Algorithm ESPNM enhances SPNM method by dynamically and preferentially forming the initial and also intermediate (during a descent) simplices as well as conditionally and adaptively regenerating the intermediate simplices. The implementation of ESPNM method is described in algorithm ESPNM and associated subalgorithms ESPNM\_PattGener, ESPNM\_SmxObjGener and ESPNM\_Step. Subalgorithm ESPNM\_PattGrner generates a simplex formation pattern of the type

$$\Xi_{\text{ESPNM}} = \begin{bmatrix} v_{1,1} & 0 & \dots & 0 & 1 & \dots & 1 \\ \vdots & v_{1,2} & \dots & 0 & 1 & \dots & 1 \\ v_{k,1} & \vdots & & 0 & 0 & \dots & 0 \\ 0 & v_{k,2} & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & v_{1,n-(k-1)} & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & 0 & \dots & 0 \\ 0 & 0 & \dots & v_{k,n-(k-1)} & 0 & \dots & 0 \end{bmatrix}_{n \times n} \cdot \quad (16)$$

**Algorithm ESPNM.** Enhanced stochastically perturbed Nelder-Mead (ESPNM) method:

$$[\mathbf{P}_l, f_l, \sigma_j] = \text{ESPNM}(\mathbf{P}_1, M, \sigma_{\text{halt}}, \mathbf{\Omega}, \sigma_N, k, v_m)$$

*Input: (as in algorithm PNM, max consecutive shrinkages( $\geq 2$ )). Output: [as in algorithm NM].*

```

1   $j := 0$  // iteration index
2   $[\mathbf{\Xi}_{\text{ESPNM}}] = \text{ESPNM\_PattGener}(n, \sigma_N, k)$  // stochastically permuted pattern
3  call  $[\mathbf{S}_j, \mathbf{F}_j] = \text{ESPNM\_SmxObjGener}(\mathbf{P}_1, \mathbf{\Xi}_{\text{ESPNM}}, \mathbf{\Omega}, M, n, k)$  // stochastic simplex
4   $v_c := 0$  // consecutive shrinkages number initialization
5  call  $[f_h, f_l, \mathbf{P}_h, \mathbf{P}_l, \bar{f}, \bar{\mathbf{P}}] = \text{SmxAssessm}(\mathbf{S}_j, \mathbf{F}_j)$  // current simplex ( $\mathbf{S}_j$ ) assessment
6  while  $(\sigma_j \geq \sigma_{\text{halt}})$  // where,  $\sigma_j = \left( \left\langle (f_i - \bar{f})^2 \mid_{i=1,2,\dots,n+1;i \neq h} \right\rangle \right)^{1/2}$  (descent halting criterion)
7       $j := j + 1$  // increment
8      assign  $\{r, e, c, s\}_j \in \{[0.5, 1], [2, 4], [0.25, 0.5], [0.3, 0.7]\}$  // random, uniformly distributed
9      call
10      $[\mathbf{S}_j, \mathbf{F}_j, \text{step}, v_c] = \text{ESPNM\_Step}(\mathbf{P}_h, \mathbf{P}_l, \bar{\mathbf{P}}, \mathbf{\Omega}, f_h, f_l, r, e, c, s, \mathbf{S}_j, \mathbf{F}_j, \mathbf{P}_1, \mathbf{\Xi}_{\text{ESPNM}}, M, n, k, v_c, v_m)$ 
11     call  $[f_h, f_l, \mathbf{P}_h, \mathbf{P}_l, \bar{f}, \bar{\mathbf{P}}] = \text{SmxAssessm}(\mathbf{S}, \mathbf{F})$  // simplex assessment
12 endwhile // end of iteration loop
13 return  $\mathbf{P}_l, f_l, \sigma_j$  // output.
```

The concept behind  $\mathbf{\Xi}_{\text{ESPNM}}$  is that it stores a leftmost set of direction vectors that propagate the perturbation envelope starting from the first element in  $\mathbf{P}_1$  (1st column of  $\mathbf{\Xi}_{\text{ESPNM}}$ ) and stopping when the opposite end of the envelope reaches the last element ( $[n - (k - 1)]$ -th column). In this way there remain  $k - 1$  unfilled columns in  $\mathbf{\Xi}_{\text{ESPNM}}$  ( $[n - (k - 2)]$ -th to  $n$ -th column) that are assigned as shown. The later, in conjunction with subalgorithm  $\text{ESPNM\_SmxObjGener}$ , will allow the selection of the best vertex so far ( $\mathbf{P}_{\text{opt}}$ ) and its subsequent perturbation with emphasis to its most influential elements (importance sampling). In this case, the aforementioned influential elements are the first two chosen on the basis that they control the offset of the active core where the pump photons absorption takes place. Subalgorithm  $\text{ESPNM\_SmxObjGener}$  describes the stochastic assignment of each perturbation propagation instance along  $\mathbf{P}_1$  (line 4). In addition to the simplex matrix it also returns the objective matrix since the simplex is generated dynamically based on the feedback from the function evaluations. Then it evaluates the objective function at the perturbed vertices and selects the fittest ( $\mathbf{P}_{\text{opt}}$ ) amongst them (line 8). Its final operation is to randomly scramble the core offset along positive directions within the optimal cross section, represented by the decoded  $\mathbf{P}_{\text{opt}}$ , in search for objective improving coordinates (lines 9-13). The initial polytope generated in this way is again a numerically non-degenerate structure.

**Subalgorithm ESPNM\_PattGener.** ESPNM pattern generation:

$$[\Xi_{\text{ESPNM}}] = \text{ESPNM\_PattGener}(n, \sigma_N, k)$$

*Input:* (number of variables, maximum standard deviation of the normal distribution and number of perturbed variables (odd positive integer)). *Output:* [ESPNM pattern matrix].

```

1   $\varepsilon := (k-1)/2$  // number of variables in either branch of the normal distribution
2  assign  $\{\sigma_{N,i} \mid i=1,2,\dots,n-2\varepsilon\} \in [\sigma_N/2, \sigma_N]$  // uniformly distributed random values
3  for each ESPNM pattern matrix column vector in the set  $\{\xi_{i-\varepsilon} \mid i=\varepsilon+1,\dots,n-\varepsilon\}$ 
4       $\mathbf{N}_{i-\varepsilon} := (\nu_1, \nu_2, \dots, \nu_k)^T \mid_{k=2\varepsilon+1; \tau \in \mathbb{Z}^*}$  //  $\{\nu_q \mid q=1,2,\dots,k\} = \mathcal{N}(\mu, \sigma_{N,i-\varepsilon}^2) \in (0, 1/\sigma_{N,i-\varepsilon}\sqrt{2\pi})$ 
5       $\xi_{i-\varepsilon} := (\xi_1, \xi_2, \dots, \xi_n)^T \equiv (0, 0, \dots, 0)^T$  // additive identity ( $n \times 1$  zero vector)
6       $\left[ \left( \xi_{i-\varepsilon}, \xi_{i-(\varepsilon-1)}, \dots, \xi_{i-1}, \xi_i, \xi_{i+1}, \dots, \xi_{i+(\varepsilon-1)}, \xi_{i+\varepsilon} \right)^T \right]_{k \times 1} := \mathbf{N}_{i-\varepsilon}$  // bell shaped perturbation
7  endfor
8  for each ESPNM pattern matrix columns in the set  $\{\xi_i \mid i=n-(k-2), n-(k-3), \dots, n\}$  //  $k-1$ 
    vectors
9       $\xi_i := (\xi_1, \xi_2, \dots, \xi_n)^T = [(1, 1, 0, 0, \dots, 0)]^T$  // preferential perturbation pattern
    vectors
10 endfor
11  $\Xi_{\text{ESPNM}} := (\xi_1, \xi_2, \dots, \xi_n)_{n \times n}$  // ESPNM pattern matrix
12 return  $\Xi_{\text{ESPNM}}$  // output.
```

Using this technique in high dimensions means that the initial simplex is formed by a search process with an extra element of intelligence which is the selective collection of information within a subset of dimensions offering higher probability to deliver substantially optimized objective function values and/or second order information. In other words, a subset of simplex vertices record a certain space of higher interest, while keeping the coordinates in the rest of the dimensions frozen, adding an element of exploratory search right from the start of the process. The aforementioned assignment process of the initial simplex is graphically illustrated in figure 8 for a small number of perturbed elements  $k=5$  selected to assist the demonstration of the selective randomization concept. It is also assumed there that  $\mathbf{P}_3$  performed optimally amongst the vertices from  $\mathbf{P}_1$  to  $\mathbf{P}_{n+1-(k-1)}$  and as shown it is vertex  $\mathbf{P}_3$  that is further processed and used as the basis to assign the remaining  $k-1$  vertices of the initial simplex ( $\mathbf{S}_0$ ). The top two elements of each of the vertices  $\mathbf{P}_{n+1-(k-2)}$  to  $\mathbf{P}_{n+1}$  in figure 8 show the way the represented core centre coordinates are randomly altered to capture further and better focused objective function information in the sub-dimensions of higher probability to capture optimal objective function values. A schematic visualization of the above process is given in figure 7(b) where the cross sections shown

**Subalgorithm ESPNM\_SmxObjGener.** ESPNM simplex, objective matrices generation:

$$[\mathbf{S}, \mathbf{F}] = \text{ESPNM\_SmxObjGener}(\mathbf{P}_1, \Xi_{\text{ESPNM}}, \mathbf{\Omega}, M, n, k)$$

*Input:* (start point  $\mathbf{P}_1$ , ESPNM pattern, optimization domain, mesh size, length of  $\mathbf{P}_1$  and number of perturbed variables). *Output:* [simplex and objective matrices].

```

1  assign  $\{a_i \mid i=1,2,\dots,n-(k-1)\} \in [-M, M]$  // uniformly distributed random bell-amplitude values
2  assign  $\{m_i \mid i=1,2,\dots,2(k-1)\} \in [-M, M]$  // uniformly distributed random mesh size values
3  for each simplex vertex in the set  $\{\mathbf{P}_i \mid i=2,3,\dots,(n+1)-(k-1)\}$  // perturbation propagation loop
4       $\mathbf{P}_i := \mathbf{P}_1 + a_i (1/\xi_{\max,i-1}) \xi_{i-1}$  // where,  $\xi_{\max,i-1} = \max\{\xi_{w,i-1} \mid w=1,2,\dots,n\}$ 
5      call  $[f_i] = \text{FuncEval}(\mathbf{P}_i, \mathbf{\Omega})$  // function evaluation at the simplex vertices
6  endfor
7  call  $[f_1] = \text{FuncEval}(\mathbf{P}_1, \mathbf{\Omega})$  // function evaluation at the start point
8   $f_{\text{opt}} := \min\{f_i \mid i=1,2,\dots,(n+1)-(k-1)\}$ ; assign  $\mathbf{P}_{\text{opt}} \mid_{f_{\text{opt}} = f(\mathbf{P}_{\text{opt}})}$  // optimal vertex selection
9  for each simplex vertex in the set  $\{\mathbf{P}_i \mid i=(n+1)-(k-2), (n+1)-(k-3), \dots, n+1\}$ 
10      $\xi_{1,i-1} := \xi_{1,i-1} m_{2(i-n+k-2)-1}$ ;  $\xi_{2,i-1} := \xi_{2,i-1} m_{2(i-n+k-2)}$ 
11      $\mathbf{P}_i := \mathbf{P}_{\text{opt}} + \xi_{i-1}$  // stochastic perturbation of core centre coordinates in selected  $\mathbf{P}_{\text{opt}}$ 
12     call  $[f_i] = \text{FuncEval}(\mathbf{P}_i, \mathbf{\Omega})$  // function evaluation at the perturbed  $\mathbf{P}_{\text{opt}}$ 
13 endfor
14  $\mathbf{S} := [\mathbf{P}_1 \mathbf{P}_2 \dots \mathbf{P}_{n+1}]_{n \times (n+1)} \mid_{\text{vol}(\mathbf{S}) > 0}$ ;  $\mathbf{F} := [f_1 f_2 \dots f_{n+1}]_{1 \times (n+1)}$  // simplex; objective matrices
15 return  $\mathbf{S}, \mathbf{F}$  // output.
```

along the first row are instances of the stochastic bell shape propagation while the second row shows the importance sampling process [46] which is practically a uniformly random search for improved core offsets in the vicinity of  $\mathbf{P}_{\text{opt}}$ . The aforementioned process is invoked once by algorithm ESPNM during the initial simplex ( $\mathbf{S}_0$ ) generation at line 3 and then recursively during the line search process (simplex descent) at line 21 of subalgorithm ESPNM\_Step. The later is executed conditionally in the vicinity of the currently best vertex ( $\mathbf{P}_l$ ) when subsequent shrinkages are recorded indicating descent on a problematic landscape (noisy, discontinuous, nonconvex with many narrow and deep basins). It is also executed adaptively by halving the mesh size prior each new simplex generation around the preserved  $\mathbf{P}_l$  in order to accelerate convergence (ESPNM\_Step line 20). This process resembles the mesh size contraction in GPS and MADS and places ESPNM in the class of methods that optimize a function by iterative processes executed on a tower of meshes [29]. An important aspect of the initial simplex generation at line 3 of algorithm ESPNM is to choose appropriate values for  $M$  and  $\sigma_N$  parameters such that the initial simplex spans an



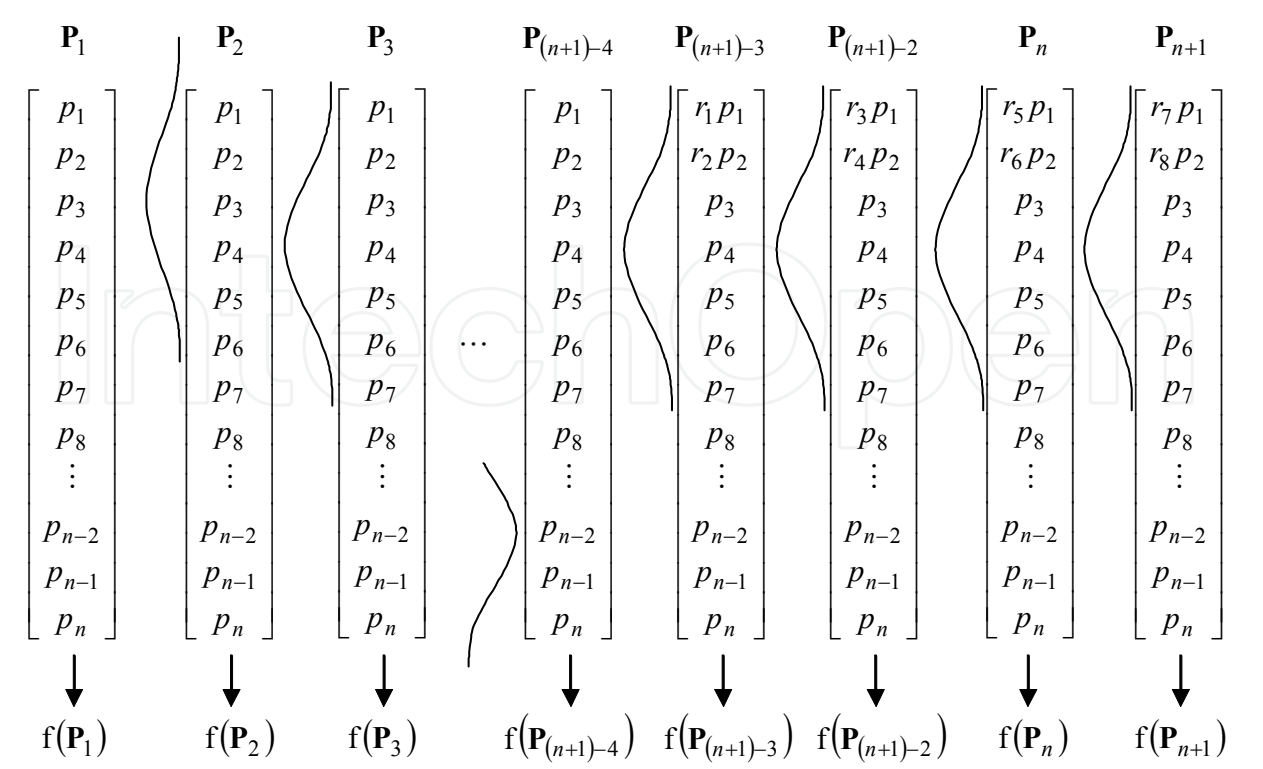


Fig. 8. Illustration of the selectively randomized initial simplex generation scheme for perturbed vertex elements number  $k = 5$ . The last four ( $k - 1$ ) simplex vertices are versions of the vertex ( $\mathbf{P}_3$ ) that was the optimal point found, containing the core centre coordinates altered by the set of normally distributed pseudorandom coefficients  $\{r_1, r_2, \dots, r_8\}$ .

area that includes many valleys (nonconvex objective function) as opposed to forming a small initial simplex with all its vertices located inside a single valley. The latter will almost certainly result in local convergence.

Subalgorithm ESPNM\_Step implements a line search operation that guides the simplex when descending in  $\Re^n$ . The aforementioned subalgorithm NM\_Step is a subset of ESPNM\_Step formed by removing the if-then-else-endif module (lines 19-25) after keeping lines 23 and 24. It includes a stronger expansion condition (line 4) and strict inequalities (lines 11 and 12). Also, the seven input arguments are removed as well as the last output argument. In previous work [21], the weaker expansion condition was used in NM\_step( $f_e < f_l$  as in the original algorithm [20]).

**Subalgorithm ESPNM\_Step.** Interpretation of the ESPNM step operation:

$[S_j, F_j, step, v_c] = \text{ESPNM\_Step}(\mathbf{P}_h, \mathbf{P}_l, \bar{\mathbf{P}}, \mathbf{\Omega}, f_h, f_l, r, e, c, s, S_j, \mathbf{F}_j, \mathbf{P}_1, \mathbf{\Xi}_{\text{ESPNM}}, M, n, k, v_c, v_m)$

*Input:* (worse, best points, centre of polytope excluding  $\mathbf{P}_h$ , bounds, highest, lowest function values, reflection, expansion, contraction, shrinkage coefficients, current simplex, objective matrices, start point  $\mathbf{P}_1$ , ESPNM pattern, mesh size, length of  $\mathbf{P}_1$ , number of perturbed variables, consecutive and max consecutive shrinkages). *Output:* [current simplex, objective matrices, operation step, consecutive shrinkages].

```

1   $\mathbf{P}_r := (1+r)\bar{\mathbf{P}} - r\mathbf{P}_h$ ; call  $[f_r] = \text{FuncEval}(\mathbf{P}_r, \mathbf{\Omega})$  // calculate; evaluate reflection point
2  if  $(f_r < f_l)$  then
3       $\mathbf{P}_e := e\mathbf{P}_r + (1-e)\bar{\mathbf{P}}$ ; call  $[f_e] = \text{FuncEval}(\mathbf{P}_e, \mathbf{\Omega})$  // calculate; evaluate expansion point
4      if  $[(f_e < f_l) \text{ AND } (f_e < f_r)]$  then // stronger expansion condition (modern NM)
5           $\mathbf{P}_h := \mathbf{P}_e$  in  $\mathbf{S}_j$ ;  $f_h := f_e$  in  $\mathbf{F}_j$ ; step := 'expansion' // expansion operation
6      else
7           $\mathbf{P}_h := \mathbf{P}_r$  in  $\mathbf{S}_j$ ;  $f_h := f_r$  in  $\mathbf{F}_j$ ; step := 'reflection' // reflection
8      endif
9  else
10      $f_m := \max\{f_i \mid i=1,2,\dots,n+1; i \neq h\}$ 
11     if  $(f_r \geq f_m)$  then
12         if  $(f_r < f_h)$  then
13              $\mathbf{P}_h := \mathbf{P}_r$  // improved  $\mathbf{P}_h$  to be used in line 16
14              $\mathbf{P}_h := \mathbf{P}_r$  in  $\mathbf{S}_j$ ;  $f_h := f_r$  in  $\mathbf{F}_j$ ; step := 'reflection' // reflection
15         endif
16          $\mathbf{P}_c := c\mathbf{P}_h + (1-c)\bar{\mathbf{P}}$ ; call  $[f_c] = \text{FuncEval}(\mathbf{P}_c, \mathbf{\Omega})$  // contraction point
17         if  $(f_c > f_h)$  then
18              $\{\mathbf{P}_i := c(\mathbf{P}_i + \mathbf{P}_l) \mid i=1,2,\dots,n+1; i \neq l\}$ ; step := 'shrinkage';  $v_c := v_c + 1$  // shrinkage
19             if  $(v_{\text{cons}} = v_{\text{max}})$  then
20                  $\mathbf{P}_1 := \mathbf{P}_l$ ;  $M := M/2$ ;  $v_c := 0$  // preservation; adaptation; reset
21                 call  $[\mathbf{S}_j, \mathbf{F}_j] = \text{ESPNNM\_SmxObjGener}(\mathbf{P}_1, \mathbf{\Xi}_{\text{ESPNNM}}, \mathbf{\Omega}, M, n, k)$  //
22                 new smx
23             else
24                 for each  $\{\mathbf{P}_i \mid i=1,2,\dots,n+1; i \neq l\}$  call  $[f_i] = \text{FuncEval}(\mathbf{P}_i, \mathbf{\Omega})$  endfor
25                  $\mathbf{F}_j := [f_1 f_2 \dots f_{n+1}]_{1 \times (n+1)}$  // evaluation of shrunk simplex
26             endif
27         else
28              $\mathbf{P}_h := \mathbf{P}_c$  in  $\mathbf{S}_j$ ;  $f_h := f_c$  in  $\mathbf{F}_j$ ; step := 'contraction' // contraction
29         endif
30     else
31          $\mathbf{P}_h := \mathbf{P}_r$  in  $\mathbf{S}_j$ ;  $f_h := f_r$  in  $\mathbf{F}_j$ ; step := 'reflection' // reflection
32     endif
33 return  $\mathbf{S}_j, \mathbf{F}_j, \text{step}, v_c$  // output.

```

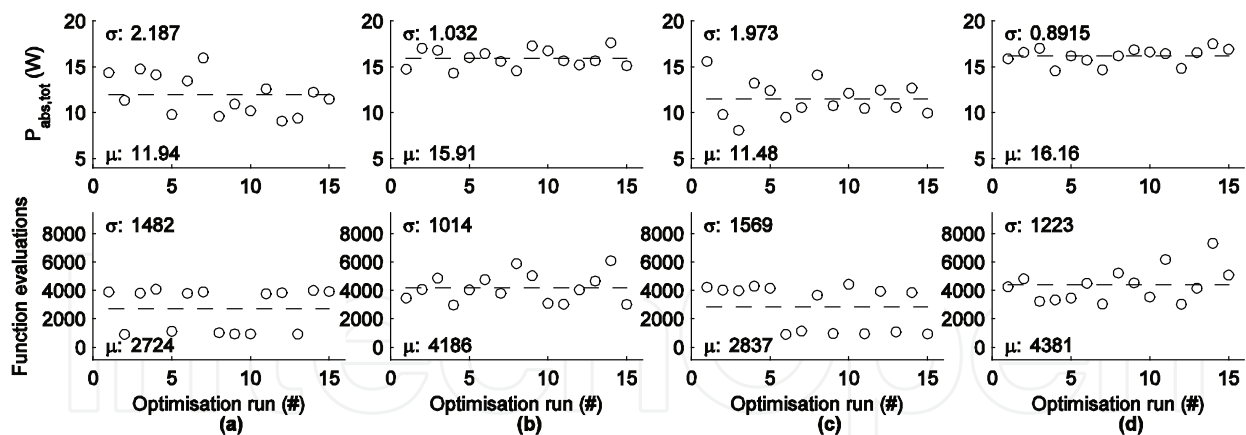


Fig. 9. Four groups of 15 optimizations in  $\mathcal{R}^{182}$  from the same starting point and driven by different algorithms: (a) SPNM{D,D}  $\equiv$  PNM. (b) ESPNM{S,D}. (c) SPNM{D,S}. (d) ESPNM{S,S}.

Figure 9 presents a comparison of the algorithms proposed in this section. The  $\{*, *\}$  notation denotes {simplex generation, descent coefficients} pairs that can be either deterministically (D) or stochastically (S) assigned. The corresponding start point was a circular non-hole inner cladding with a centred core which absorbed 5.6W of pump power. The reported results indicate that the best performing algorithm is ESPNM{S, S} because it delivered, on average, the optimal function values exhibiting at the same time the lowest spread around their mean value. It demonstrated a 152% improvement of the mean  $P_{abs,tot}$  compared to the 113% offered by PNM for a 61% increase in computation cost over PNM.

#### 4. Optimization results

The inner cladding of a conventional DCF has a numerical aperture (NA) of 0.48 while the core NA is 0.175. The core doping density is 20,000ppm-by-volume, the launched pump power is 100W and the fibre length is 10cm for all the optimization results presented in this section. The pump light has a random modal content, its energy is propagated via 288 rays in 10 time steps and the absorption computation grid of the active core is comprised of 100 volume elements. The pump light wavelength is  $\lambda_p = 975\text{nm}$  at which the  $\text{Yb}^{+3}$  ( $\text{Er}^{+3}$ - $\text{Yb}^{+3}$  ion system) absorption cross section is  $2.1 \times 10^{-24} \text{m}^2$ . The simulated fibres are single-end pumped by a 600 $\mu\text{m}$  diameter pure silica core (standard fibre bundled pump delivery fibre) and NA of 0.48 when pumping a fibre with polymer outer clad or it is assumed to be surrounded by an air outer cladding when pumping a double-clad fibre which also has an air outer cladding. This work focuses on a set of fibre topologies that are thoroughly optimized and computationally compared on a common basis that avoids confusion and develops intuition into their absorption trends. Although space restrictions did not allow comprehensive parametric optimization, a sample of parametric optimization results in  $\mathcal{R}^{10}$  is presented in figure 10 which shows a set of fairly similar optimizers exhibiting almost identical absorption characteristics. Algorithm NM converged to the reported shapes for different pairs of fibre length and pump power values correspondingly. The optimization process started from the same initial cross section and run under the same settings. Figure 10 demonstrates the generality of the optimization results reported in tables 1-4 which are approximately valid within the ranges [0.1, 1]W and [0.1, 1]m of pump power and fibre length correspondingly.

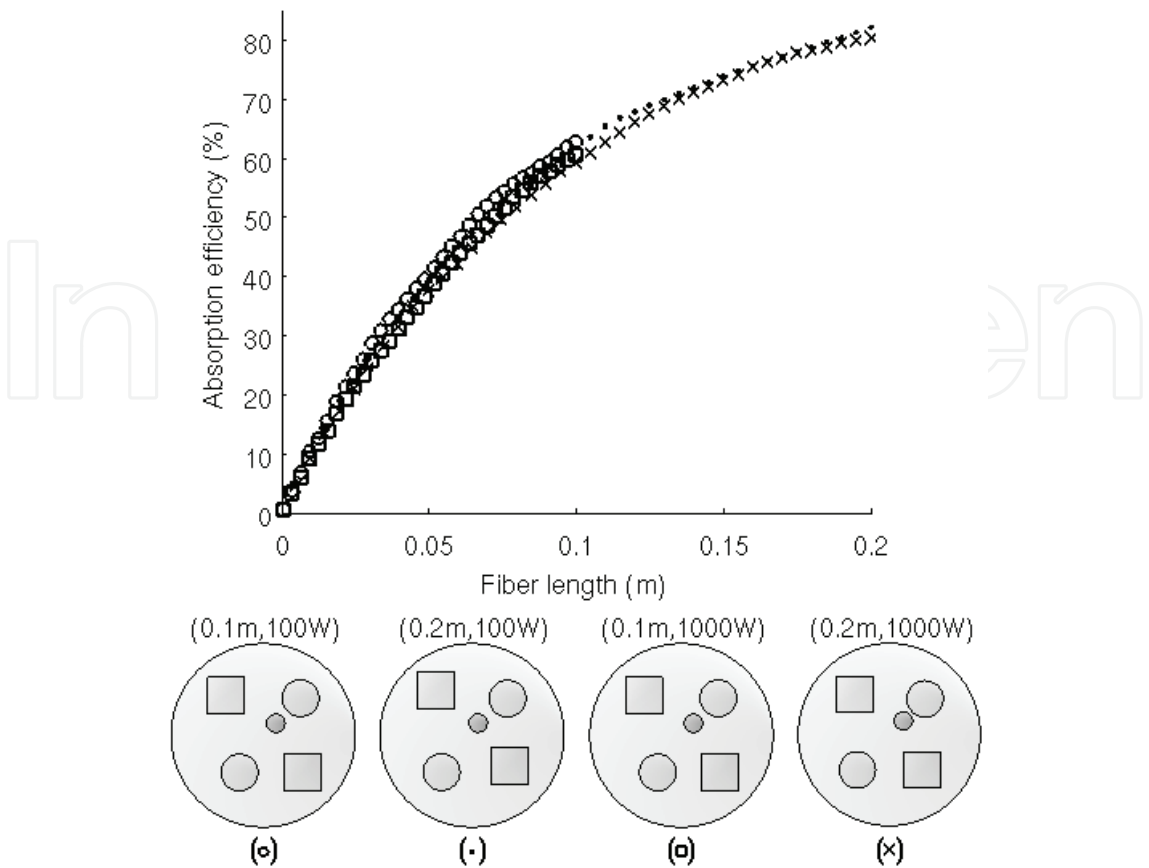


Fig. 10. Absorption performance of four convergence points resulting from optimization runs under different fibre length and pump power values.

The computing platform used for the optimizations reported in this chapter, is the same as the platform described in reference [19]. The CPU time consumed for the objective function evaluation at each start point is shown in the tables of this chapter for a more informative presentation. The strongest influence on the recorded CPU times originates from the total number of scattering operations which fluctuates slightly during an optimization. The computational efficiency of the 3-D fibre simulation method used was compared in [19] with other methods reported in the literature.

All Monte Carlo algorithms proposed in this chapter made use of the built-in MATLAB random number generator to produce the required sequences of uniformly distributed pseudorandom numbers. The built in function is based on the random number generator by Marsaglia and Zaman [47] which was specifically designed to produce floating point values and uses a lagged Fibonacci generator with a cache of 32 floating point numbers between 0 and 1 combined with a separate, independent random integer generator based on bitwise logical operations. As a result, MATLAB's built-in generator has a period of  $2^{1492}$  (number of values produced before the sequence begins to repeat itself) and can theoretically generate all numbers between  $2^{-53}$  and  $1-2^{-53}$ , all with equal probability to occur.

Figure 11 demonstrates the effort to optimize the offset of the core inside a circular (1st row figures) and a square (2nd row figures) inner cladding. The CPU time required for a single function evaluation for the circular fibre was approximately 28s on the MATLAB platform. Figures 11(a) and 11(e) show the corresponding pump power absorption surfaces generated by sampling the total absorbed pump power ( $P_{abs,tot}$ ) calculated at 49 nodes (by moving

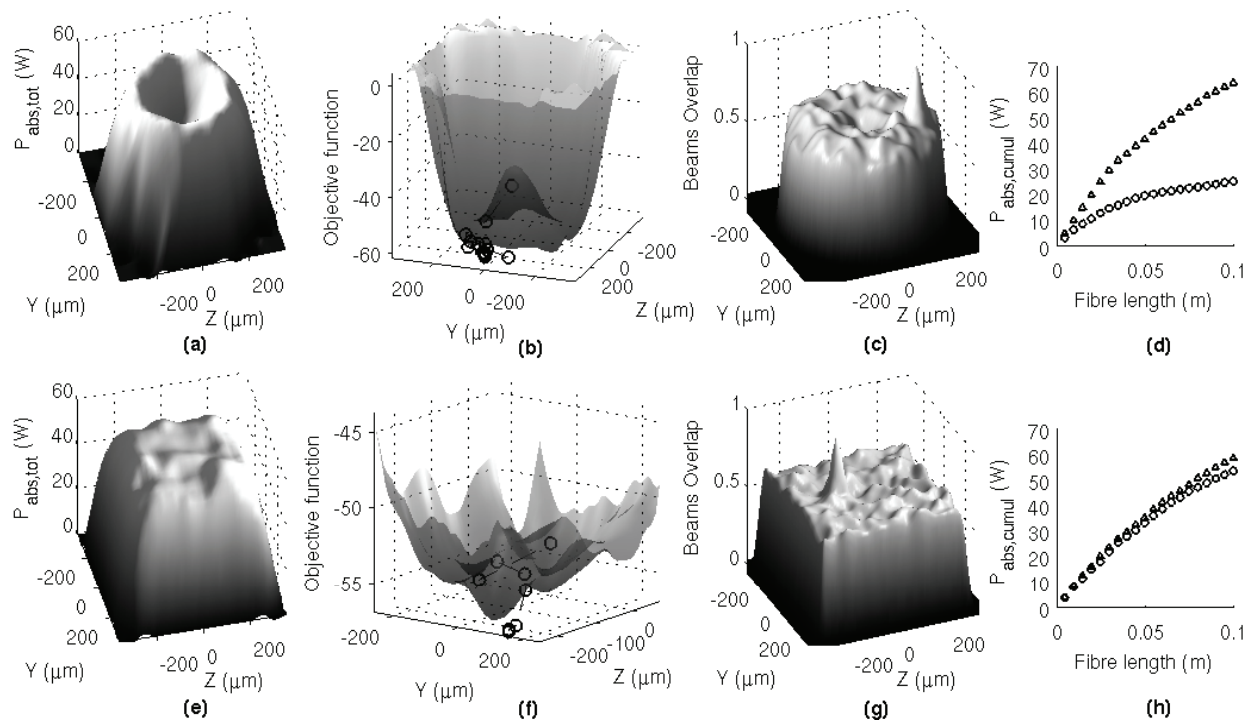


Fig. 11. Core offset optimization (in  $\mathbb{R}^2$ ) inside a circular (1<sup>st</sup> row) and a square inner-clad (2<sup>nd</sup> row). (a),(e) Transverse distribution of total absorbed power. (b),(f) Interpolated objective function surface and simplex descent path on the actual surface. (c),(g) Beam overlap images. (d),(h) Cumulative absorption of the initial guess (circles) and the convergence point (triangles).

the core on each one) of a Cartesian grid covering a square area  $4900\mu\text{m}^2$  and interpolating the values on a 784 nodes grid covering the same area. This information is plotted here in order to observe the behaviour of the referred to as the modern interpretation of the Nelder-Mead algorithm adopted in this work. For the circular inner cladding, the  $P_{abs,tot}$  values exhibit the well known symmetrical distribution around the centre of the cross section with the peak appearing near the inner-to-outer cladding interface. Figure 11(b) shows the surface that plots the corresponding values ( $-P_{abs,tot}$ ) of the objective function and the path followed by the lowest vertex of the simplex (which is a triangle here in  $\mathbb{R}^2$ ). The descent started from the region of the initial guess which was the centre of the cross section  $(y_{c,init}, z_{c,init}) = (0, 0)\mu\text{m}$  and the algorithm converged at the point  $(y_{c,opt}, z_{c,opt}) = (-38, -203)\mu\text{m}$  denoting that the optimum offset of the core from the centre is approximately at a distance of 69% of its radius for the considered operation point.

The corresponding path for the square DCF is shown in figure 11(f) on a fragment of the objective function surface. Here the simplex started again from the cross section centre and converged this time to the point  $(y_{c,opt}, z_{c,opt}) = (-24, 126)\mu\text{m}$  where the core is situated at a distance from the centre that is approximately 21% of the inner cladding side length. In both figures it is apparent that the direct search method achieved better landscape resolutions and at lower computation cost than those achieved through the initial evaluation of  $P_{abs,tot}$  at the grid nodes. Furthermore, figure 11(b) suggests graphically that first-order convergence from an arbitrary starting point (global convergence) is achieved at a point



$\mathbf{P}_{optim} \in \mathfrak{R}^2$  very close to an optimizer  $x_*$  that is a stationary point of the objective function satisfying the second-order sufficiency condition ( $\exists x_* : \nabla^2 f(x_*) > 0$  for a differentiable function). The spatial distribution of  $P_{abs,tot}$  across the cross section plane of the circular DCF is also clearly followed by the lowest order standing wave that developed in the beam overlap image in figure 11(c). The corresponding surface for the square DCF shows the improved scrambling of the modes achieved by this cross section. The peak standing high above the rest on each surface denotes the location of the core within the inner cladding. Figure 11(d) shows the dramatic improvement of absorption in the offset core of the circular DCF which is the direct result of the simplex descent to a deep valley while figure 11(h) demonstrates that there is comparatively little room for improvement when offsetting the core within a square DCF.

Table 1 presents the results from the simultaneous optimization of the cross section and refractive index performed mostly by the stochastic variants of NM at relatively low dimensions. The listed schemes (column 5) optimized the offset, size, shape and refractive index of an encompassed lamina while the shape of the inner cladding remained constant. These results represent a telescopic view into the considered optimization domains facilitated by the parsimonious nature of NM and SNM methods. All dielectric holes shown are assumed to be made of CBYA alloy-glass [40] apart from the row 3 optimizer representing an attempt to search for improved refractive index values. The increased CPU times recorded for the most complicated and absorbent topologies is due to the correspondingly larger number of scatterings occurring inside them. The optimal cross section in table 1 is the row 8 optimizer, found by stochastic search in  $\mathfrak{R}^{18}$  where the offset as well as ellipticity and size of four large area holes were allowed to vary independently. The single hole designs demonstrated high potential to achieving optimal absorption while when square shapes for the inner cladding or embedded holes were used, the absorption dropped considerably. The same was the case when air holes or hexagonal CBYA holes of variable offset and size were optimized (not shown). As far as the preliminary results in table 1 are considered, the cross sections worth to invest on in terms of computational expense for further optimization by the MADS method appear to be the:

- Four elliptical holes scheme (row 8)
- Circular hole topologies because they are easier to manufacture and showed improved absorption potential (row 6) after initiating a second optimization from a previous optimizer (row 5)
- Single large-hole cross section due to its simplicity and good performance.

The most promising topologies from table 1 are taken to the next level for optimization by MADS which promises to deliver global optimizers with higher probability but a significant increase in computational cost is expected. Prior to discussing the results in table 2 it is useful to describe the algorithmic settings of GA, MADS and GPS methods because they had an impact on all corresponding results. The optimizations executed by GA in section 2 started with  $n+1$  members in the initial population (to match the number of vertices maintained by a simplex), the elite population size was set to the nearest integer of  $(n+1)/10$ , the cross over factor was 0.8, the migration factor was 0.2 and the migration interval was set to 20. With the need to make the GPS and MADS as computationally efficient as possible with a minimum negative impact on their global convergence properties, they were set up as follows. Neither complete search nor complete poll




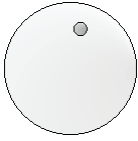
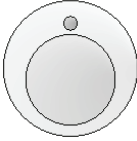
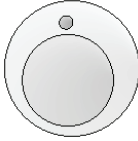


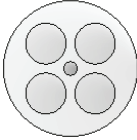
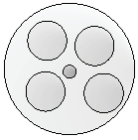
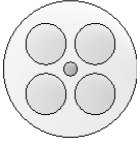
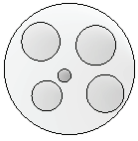
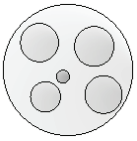
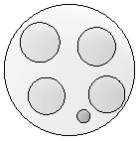
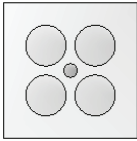
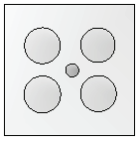
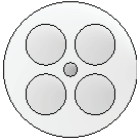
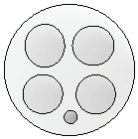
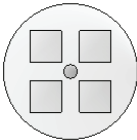
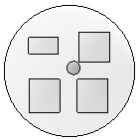
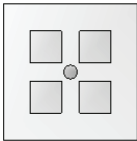
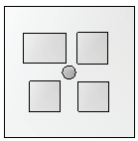
Start point	Opti-mizer	Start point $P_{abs,tot}$ (W)	Opti-mizer $P_{abs,tot}$ (W)	Encoding scheme	Algorithm	Optimi- zation space	Func Evals (#)	Start point $t_{CPU}$ (s)
		25.3	63.8	Offset	NM	$\Re^2$	78	27.6
		63.6	69.1	Offset-Diameter	SNM {S, S}	$\Re^5$	335	65.8
		63.6	69.5	Offset-Diameter - index	NM	$\Re^6$	328	65.8
		57.9	64.6	Offset	SNM {S, S}	$\Re^{10}$	227	109.0
		57.9	67.0	Offset-Diameter	SNM {S, D}	$\Re^{14}$	304	109.0
		67.0	69.8	Offset-Diameter	SNM {S, D}	$\Re^{14}$	425	101.7
		54.4	58.6	Offset-Diameter	SNM {S, D}	$\Re^{14}$	306	84.7
		57.9	70.7	Offset-Major-Minor	SNM {S, D}	$\Re^{18}$	558	109.0
		56.2	65.0	Offset-Major-Minor	SNM {S, D}	$\Re^{18}$	412	91.3
		54.7	59.3	Offset-Major-Minor	SNM {S, D}	$\Re^{18}$	355	76.4

Table 1. Optimization results for polymer outer-clad and holey inner-clad with NM variants.



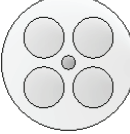
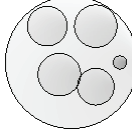
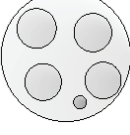
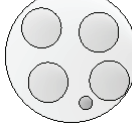
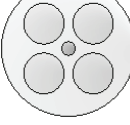
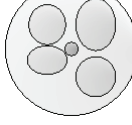
Start point	Opti-mizer	Start point $P_{abs,tot}$ (W)	Opti-mizer $P_{abs,tot}$ (W)	Encoding scheme	Algorithm	Optimi- zation space	Func Evals (#)	Start point $t_{CPU}$ (s)
		63.6	71.1	Offset-Major-Minor-Index	MADS {Np1, 2N}	$\Re^7$	112	65.8
		57.9	71.0	Offset-Diameter	MADS {Np1, 2N}	$\Re^{14}$	441	109.0
		69.8	69.8	Offset-Diameter	MADS {Np1, 2N}	$\Re^{14}$	235	97.3
		57.9	65.3	Offset-Major-Minor	MADS {Np1, 2N}	$\Re^{18}$	452	109.5

Table 2. MADS optimization results for polymer outer-clad and holey inner-clad.

operations were allowed resulting in an opportunistic style of direct search iteration that stops as soon as a better point has been found. Also, the first direction of search after a successful poll or search step is set to be the one that was successful in the previous iteration (exploratory search tactic). A so called tabu list that records the already visited points was maintained so that the expense of unnecessary function re-evaluations would be avoided. This added a tabu search metaheuristic element to MADS and GPS that was found to offer up to approximately 40% reduction in function evaluations. Tabu search is not recommended for stochastic functions but in this case the stochastic noise was suppressed. One other setting that can reduce the computation expense is to accelerate the rate at which the mesh size is adapted after a non-successful iteration. This setting was not enabled in this work because it was found to significantly reduce the probability to discover a global optimizer (at a benefit of 20% reduction in function evaluations). The last setting, shared by all optimization methods used here is the minimization halting criterion. In order to achieve an equally economical minimization that avoids unnecessary function evaluations at the vicinity of an already well approximated optimizer, all halting criterions were set to stop the minimization when saturation in the improvement of the lowest recorded objective function value as a function of the number of iterations was observed. Regarding MATLAB’s ‘Genetic Algorithm and Direct Search Toolbox’ used to implement the GA, GPS, and MADS optimizations, it was found via observation that the above halting condition was satisfied for ‘Function Tolerance’ (a parameter compared against the cumulative change in the best function value over a number of iterations) values of  $10^{-6}$ ,  $10^{-7}$  and  $10^{-4}$  correspondingly. In algorithm NM and all its forms proposed in sections 2 and 3, the saturation of the fittest function value was observed for

$$\sigma_{halt} \cong (1/10)\sigma_0$$

(17)

where  $\sigma_0$  is the standard deviation of the initial objective matrix elements excluding the highest value. The success of (17) depends on the standard deviation of the function values, stored in the initial objective matrix ( $F_0$ ), not being too large so that the simplex will reach the neighbourhood of an optimizer before the condition  $\sigma_j \geq \sigma_{halt}$  is satisfied at the end of the  $j$ -th iteration. When the aforementioned criterion fails to halt the simplex after acceptably approximating an optimizer, then the descent halts after a relatively small number of iterations and a large improvement in the objective (row 1 in table 3, row 5 in table 4). Then the process is restarted using the discovered point as a new start point (row 2 in table 3, row 6 in table 4). In this way, the inherent tendency of NM (and proposed NM-based methods) to perform unnecessary iterations after having adequately approximated an optimizer was avoided.


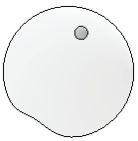
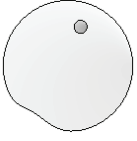
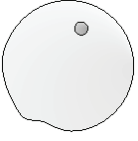
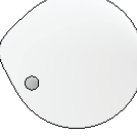


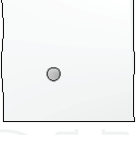


Start point	Opti-mizer	Start point $P_{abs,tot}$ (W)	Opti-mizer $P_{abs,tot}$ (W)	Encoding scheme	Algorithm	Optimi-zation space	Func Evals (#)	Start point $t_{CPU}$ (s)
		25.3	64.8	Offset-Perimeter	ESPNM {S, S}	$\Re^{182}$	727	27.6
		64.8	68.7	Offset-Perimeter	ESPNM {S, S}	$\Re^{182}$	9196	27.2
		61.5	67.2	Offset-Perimeter	ESPNM {S, S}	$\Re^{182}$	11662	26.1
		53.5	61.3	Offset-Perimeter	ESPNM {S, S}	$\Re^{186}$	10146	28.8
		63.6	69.7	Offset-Perimeter	ESPNM {S, S}	$\Re^{362}$	1238	65.8

Table 3. Optimization results for polymer outer cladding with algorithm ESPNM.

After the direct comparison of several algorithms in section 2, the MADS method was chosen as the most successful at lower dimensions in terms of probability to find global optimizers. The most distinctive topologies listed in table 1 are re-optimized in table 2 under MADS. The 1st row of table 2 shows the results from an attempt to optimize the same start point as in row 3 of table1 but this time with an added dimension. The discovered optimizer outperformed all optimizers from table 1 showing that an offset core topology with a single large hole of optimal ellipticity delivers the strongest pump absorption. Although the refractive index was independently varied during the optimization, the MADS algorithm

converged to an optimizer with the exactly the same hole refractive index, a manifestation of the discrete nature of pattern search. One other aspect of the MADS algorithm is that it demonstrates an inherent tendency to preferentially search along those directions that exhibit the stronger influence on the objective function values. The results in row 4 disappointed because although the start point was the same as in row 8 of table 1, the MADS algorithm converged to an optimizer in  $\mathfrak{R}^{18}$  that was strongly outperformed by the SNM found optimizer (for a higher cost though this time). This observation suggests that a surprisingly low number of function evaluations is a sign of local convergence. However, the discovered optimizer indicates that if a centred core topology is sought after then the design parameters of the holes can be optimized for improved absorption strength. Row 2 in table 2 shows a successful optimization in  $\mathfrak{R}^{14}$  that improved over the later suggesting that the optimization of the hole-ellipticity may not be justified if it is significantly more difficult to manufacture. An interesting result is reported in row 3 of table 2 where MADS converged to the start point after about 1/3 of the expected number of function evaluations. This behaviour of MADS was observed several times and showed that its success depends strongly on starting the process from a point far away from an optimizer, a property which is not shared by SNM as suggested by the results in row 6 of table 1.

Remaining in the class of topologies with polymer outer cladding, table 3 presents the optimization of inner cladding and hole perimeters along with the core offset at high dimensions. The two-stage optimization (rows 1,2) of a circular inner cladding with centred core resulted in a cross section with a minor spiral deformation to the inner cladding perimeter and an offset core. Row 3 adopted a start point resembling the spiral fibre proposed by Kouznetsov and Moloney [48] and converged to an inner cladding shape that is a perturbed spiral shape with the core located closer to the centre. The optimizer in row 3 suggests that a spiral cross section can be further improved. Row 4 demonstrates that a square fibre has limited prospects for competitive improvement while row 5 shows a case of local convergence in  $\mathfrak{R}^{362}$  where a global optimization is potentially very expensive due to the high dimensions. Finally, table 4 presents a set of optimization attempts for double-clad topologies with air outer cladding. The CPU times recorded here are much higher than in the polymer outer cladding case because the air-clad designs support higher order modes (rays of higher transmission angles under the absorption model in [19]) resulting in significantly increased number of scattering operations on the dielectric interfaces. An interesting finding was that an optimized polymer hole (row 2) can be very efficient in decoupling the pump light from its volume. In this way the pump modes are forced to propagate inside the significantly reduced inner cladding volume with a dramatic effect on the increase of the pump photons overlap with the active core volume. This design can be used with moderate pump power levels though due to the low damage threshold of a polymer. However, it has been demonstrated that high glass-transition temperature thermoplastic polymers can be thermally co-drawn into micro-sized structures without cracking or delamination [49]. A direct comparison between MADS and ESPNM is provided by the results in rows 3 and 4 where a dodecagon shaped inner cladding with offset core is optimized. The two algorithms converged to optimizers of the same absorption performance but ESPNM did so at a significantly lower cost. The dodecagon shape was chosen due to the small number of perimetric sampling points involved which did not allow MADS to generate trial points without physical meaning (or lacking manufacturability), contrary to the cases in figure 5. Furthermore, an air outer cladding may be easier to fabricate around a polygonic inner cladding by means of a comb of suitably shaped air holes.


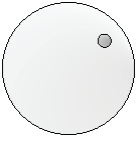
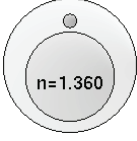
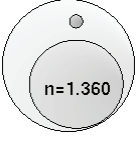

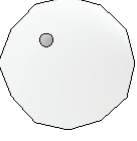

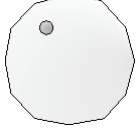




Start point	Opti-mizer	Start point $P_{abs,tot}$ (W)	Opti-mizer $P_{abs,tot}$ (W)	Encoding scheme	Algorithm	Optimi-zation space	Func Evals (#)	Start point $t_{CPU}$ (s)
		28.1	66.7	Offset	NM	$\Re^2$	68	58.9
		71.3	76.3	Offset-Major-Minor-Index	MADS {Np1, 2N}	$\Re^7$	126	155.7
		71.3	73.4	Offset-Perimeter	ESPNM {S, S}	$\Re^{26}$	629	55.3
		71.3	73.4	Offset-Perimeter	MADS {Np1, 2N}	$\Re^{26}$	898	55.3
		28.1	71.2	Offset-Perimeter	ESPNM {S, S}	$\Re^{182}$	669	58.9
		71.2	73.9	Offset-Perimeter	ESPNM {S, S}	$\Re^{182}$	11199	60.0

Table 4. Optimization results for air outer cladding.

The predictions reported here may be compared to the 35% pump absorption enhancement reported by Baek *et al* [14] and to the 18% improvement measured by Jeong *et al* [15] for a circular fibre with centred core. Based on the current results, in the case of polymer coated DCFs, it is predicted that the optimizer in row 1 of table 2 can offer an approximate enhancement of 180% compared to a conventional circular DCF with centred core. Against a conventional circular DCF with optimally offset core (table 1, row 1 optimizer), an enhancement of 11% is predicted. For the air outer cladding case, assuming high power operation (no polymer holes), a 160% improvement (table 4, row 6 optimizer) is predicted against a centred circular DCF and 10% enhancement compared to the circular optimizer with optimally offset core.

5. Summary

Several stochastic algorithms based on the deterministic Nelder-Mead method were proposed and benchmarked against pattern search methods and a genetic algorithm. In low dimensions, the proposed Monte Carlo NM variants offered improved computational efficiency via a simple sampling approach. Implicitly constrained search combined with



importance sampling offered efficient global convergence in high dimensions. Smoothly perturbed patterns were proposed that may find theoretical support for constrained optimization. The fittest algorithms were applied to the cross section geometry and corresponding refractive index profile optimization. The identified advantages of the aforementioned pump absorption enhancement concept were:

- In the case of the holey DCFs the size of the inner cladding can be scaled to accept more pump power without the need to increase the core size. The solid state holes can be correspondingly scaled to retain their pump light tapering effect into the core volume.
- The proposed holey cross sections are compatible with the helical core concept and most side pumping schemes. Multi-core ribbon lasers [12] may also benefit from optimized solid-state holes
- No fibre machining is needed while also compatibility with standard fibre manufacturing is maintained

The main limitation may be the low fabrication tolerance implied by the complexity of most proposed topologies. On the front of correctly predicting their relative absorption performance, limitations are imposed by error levels induced by stochastic and numerical noise during optimization as well as simulation inaccuracies induced during function evaluations.

## 6. References

- [1] Jeong Y, Sahu J, Payne D and Nilsson J 2004 Ytterbium-doped large-core fiber laser with 1.36 kW continuous-wave output power *Optics Express* 12 6088-92
- [2] Yahel E, Hess O and Hardy A A 2006 Modeling and Optimization of High-Power Nd<sup>3+</sup>-Yb<sup>3+</sup> Codoped Fiber Lasers *IEEE J. Lightwave Technol.* 24 1601-9
- [3] Yahel E and Hardy A 2003 Modeling High-Power Er<sup>3+</sup>-Yb<sup>3+</sup> Codoped Fibre Lasers *IEEE J. Lightwave Technol.* 21 2044-52
- [4] Vienne G G, Caplen E J, Dong L, Minelly D J, Nilson J and Payne N D 1998 Fabrication and Characterization of Yb<sup>3+</sup>:Er<sup>3+</sup> Phosphosilicate Fibres for Lasers *IEEE J. Lightwave Technol.* 16 1990-2001
- [5] Federighi M and Di Pasquale F 1995 The Effect of Pair-Induced Energy Transfer on the Performance of Silica waveguide Amplifiers with High Er<sup>3+</sup>/Yb<sup>3+</sup> Concentrations *IEEE Photonics Tech. Lett.* 7 303-5
- [6] Lassila E, Hernberg R and Alahautala T 2006 Axially symmetric fiber side pumping *Optics Express* 14 8638-43
- [7] Yan P, Gong M, Li C, Ou P, Xu A 2005 Distributed pumping multifiber series fiber laser *Optics Express* 13 2699-706
- [8] Polynkin P, Temyanko V, Mansuripur M and Peyghambarian N 2004 Efficient and Scalable Side Pumping Scheme for Short High-Power Optical Fiber Lasers and Amplifiers *IEEE Photonics Tech. Lett.* 16 2024-26
- [9] Koplow J P, Moore S W and Kliner D A V 2003 A New method for Side Pumping of Double-Clad Fiber Sources *J. Quantum Electr.* 39 529-40
- [10] Koplow J P, Goldberg L, and Dahv A. V. Kliner D A V 1998 Compact 1-W Yb-Doped Double-Cladding Fiber Amplifier Using V-Groove Side-Pumping *IEEE Photonics Tech. Lett.* 10 793-5



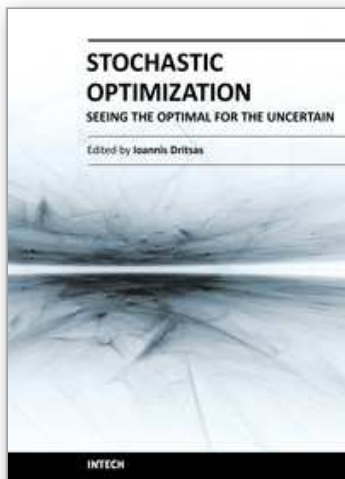
- [11] Wang P, Cooper L J, Sahu J K and Clarkson 2006 Efficient single-mode operation of a cladding-pumped ytterbium-doped helical-core fiber laser *Optics Letters* 31 226-8
- [12] Wang P, Clarkson W A, Shen D Y, Copper L J and Sahu J K 2006 Novel concepts for high-power fibre lasers *Solid state lasers and amplifiers II: Proc. SPIE* (Strasbourg, France, 5-6 April 2006) Vol. 6190, 61900I (Apr. 17, 2006) ed A. Sennaroglu *et al* pp 1-12
- [13] Jiang Z and Marciante J R 2006 Mode-area scaling of helical-core, dual-clad fiber lasers and amplifiers using an improved bend-loss model *J. Opt. Soc. Am. B* 23 2051-8
- [14] Baek S, Roh S, Jeong Y and Lee B 2006 Experimental Demonstration of Enhancing Pump Absorption Rate in Cladding-Pumped Ytterbium-Doped Fiber Lasers Using Pump-Coupling Long-Period Fiber Gratings *IEEE Photonics Tech. Lett.* 18 700-2
- [15] Jeong Y, Baek S, Nilsson J and Lee B 2006 Simple and compact, all-fibre retro-reflector for cladding-pumped fibre lasers *Electronics Letters* 42 15-6
- [16] Kouznetsov D and Moloney J V 2003 Highly Efficient, High-Gain, Short-Length, and Power-Scalable Incoherent Diode Slab-Pumped Fiber Amplifier/Laser *J. Quantum Electr.* 39 1452-61
- [17] Kouznetsov D and Moloney J V 2004 Slab Delivery of Incoherent Pump Light to Double-Clad Fiber Amplifiers: An Analytic Approach *J. Quantum Electr.* 40 378-83
- [18] Peterka P, Kašík I, Matejec V, Kubeek V and Dvořák P 2006 Experimental demonstration of novel end-pumping method for double-clad fiber devices *Optics Letters* 31 3240-2
- [19] Dritsas I, Sun T and Grattan K T V 2006 Numerical simulation based optimization of the absorption efficiency in double-clad fibres *IoP J. Opt. A: Pure Appl. Opt.* 8 49-61
- [20] Nelder J A and Mead R 1965 A simplex method for function minimization *Computer Journal* 7 308-13
- [21] Dritsas I, Sun T and Grattan K T V 2006 Double-clad fibre numerical optimization with a simplex method *Solid state lasers and amplifiers II: Proc. SPIE* (Strasbourg, France, 5-6 April 2006) Vol. 6190, 61900L (Apr. 17, 2006) ed A. Sennaroglu *et al* pp 1-12
- [22] Conn A R, Gould N I M and Toint P L 1991 A Globally Convergent Augmented Lagrangian Algorithm for Optimization with General Constraints and Simple Bounds *SIAM J. Numer. Analysis* 28 545-72
- [23] Audet C and Dennis J E, jr. 2003 Analysis of generalized pattern searches *SIAM J. Optim.*, 13 889-903
- [24] Lewis R M and Torczon V 2002 A globally convergent augmented lagrangian pattern search algorithm for optimization with general constraints and simple bounds *SIAM J. Optim.*, 12 1075-89
- [25] Lewis R M and Torczon V 2000 Pattern Search Methods for Linearly Constrained Minimization *SIAM J. Optim.*, 10 917-41
- [26] Lewis R M and Torczon V 1999 Pattern Search Algorithms for Bound Constrained Minimization *SIAM J. Optim.*, 9 1082-99
- [27] Torczon V 1997 On the Convergence of Pattern Search Algorithms *SIAM J. Optim.*, 7 1-25
- [28] Audet C and Dennis J E, jr. 2006 Mesh adaptive direct search algorithms for constrained optimization *SIAM J. Optim.*, 17 188-217

- [29] Abramson M A, Audet C and Dennis J E, jr. 2006 Nonlinear Programming by Mesh Adaptive Direct searches *SIAG/OPT Views-and-news* vol 17 no 1 pp 2-11
- [30] Lagarias J C, Reeds J A, Wright M H and Wright P E 1998 Convergence Properties of the Nelder-Mead Simplex Method in Low Dimensions *SIAM J. Optim.*, 9 112-47
- [31] Kolda T G, Lewis R M and Torczon V 2004 Optimization by Direct Search: New Perspectives on Some Classical and Modern Methods *SIAM Review* 45 385-482
- [32] Torczon V 1991 On the convergence of the multidirectional search algorithm *SIAM J. Optim.*, 1 123-45
- [33] Teng C-H, Chen Y-S and Hsu W-H 2006 Camera self-calibration method suitable for variant camera constraints *Applied Optics* 45 688-96
- [34] Lhommé F, Caucheteur C, Chah K, Blondel M and Mégret P 2005 Synthesis of fiber Bragg grating parameters from experimental reflectivity: a simplex approach and its application to the determination of temperature-dependent properties *Applied Optics* 44 493-7
- [35] McKinnon K I M 1998 Convergence of the Nelder-Mead simplex method to a nonstationary point *SIAM J. Optim.* 9 148-58
- [36] Kelley C T 1999 Detection and remediation of stagnation in the Nelder-Mead algorithm using a sufficient decrease condition *SIAM J. Optim.*, 10 43-55
- [37] Tseng P 1999 Fortified-descent simplicial search method: A general approach *SIAM J. Optim.*, 10 269-88
- [38] Zhang R and Shi F G 2004 A Novel Algorithm for Fiber-Optic Alignment Automation *IEEE Trans. Advance. Packaging* 27 173-8
- [39] Lin J Wu Y Huang T S 2004 Articulate Hand Motion Capturing Based on a Monte Carlo Nelder-Mead Simplex Tracker *Proceedings of the 17th International Conference on Pattern Recognition (ICPR'04-23-26 Aug.)* 4 pp 975-978
- [40] Zhang L, Gan F and Wang P 1994 Evaluation of refractive-index and material dispersion in fluoride glasses *Applied Optics* 33 50-6
- [41] Amar J G 2006 The Monte Carlo Method in Science and Engineering *IEEE Computing in Science & Engineering* vol 8 no 2 pp 9-19
- [42] Bandler J W, Koziel S and Madsen K 2006 Space Mapping for Engineering Optimization *SIAG/OPT Views-and-news* vol 17 no 1 pp 19-26
- [43] Renders J-M and Flasse S P 1996 Hybrid methods using genetic algorithms for global optimization *IEEE Trans. Systems, Man and Cybernetics, Part B* 26 243-58
- [44] Wessel S, Trebst S and Troyer M 2005 A renormalization approach to simulations of quantum effects in nanoscale magnetic systems *SIAM J. Multiscale model simul.* 4 237-49
- [45] Wen M and Yao J 2006 Birefringent filter design by use of a modified genetic algorithm *Applied Optics* 45 3940-50
- [46] Luijten E 2006 Fluid Simulation with the Geometric Cluster Monte Carlo Algorithm *IEEE Computing in Science & Engineering* vol 8 no 2 pp 20-9
- [47] Marsaglia G and Zaman A 1991 A new class of random number generators *Ann. Appl. Probab.* 3 462-80
- [48] Kouznetsov D and Moloney J 2002 Efficiency of pump absorption in double-clad fibre amplifiers. II. Broken circular symmetry *J. Opt. Soc. Am. B* 19 1259-63

- [49] Temelkuran B, Hart S D, Benoit G, Joannopoulos J D and Fink Y 2002 Wavelength-scalable hollow optical fibres with large photonic bandgaps for CO<sub>2</sub> laser transmission *Nature* 420 650-3

IntechOpen

IntechOpen



## **Stochastic Optimization - Seeing the Optimal for the Uncertain**

Edited by Dr. Ioannis Dritsas

ISBN 978-953-307-829-8

Hard cover, 476 pages

**Publisher** InTech

**Published online** 28, February, 2011

**Published in print edition** February, 2011

Stochastic Optimization Algorithms have become essential tools in solving a wide range of difficult and critical optimization problems. Such methods are able to find the optimum solution of a problem with uncertain elements or to algorithmically incorporate uncertainty to solve a deterministic problem. They even succeed in “fighting uncertainty with uncertainty”. This book discusses theoretical aspects of many such algorithms and covers their application in various scientific fields.

### **How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Ioannis Dritsas, Tong Sun and Ken Grattan (2011). Global Optimization of Conventional and Holey Double-Clad Fibres by Stochastic Search, *Stochastic Optimization - Seeing the Optimal for the Uncertain*, Dr. Ioannis Dritsas (Ed.), ISBN: 978-953-307-829-8, InTech, Available from: <http://www.intechopen.com/books/stochastic-optimization-seeing-the-optimal-for-the-uncertain/global-optimization-of-conventional-and-holey-double-clad-fibres-by-stochastic-search>

**INTECH**  
open science | open minds

### **InTech Europe**

University Campus STeP Ri  
Slavka Krautzeka 83/A  
51000 Rijeka, Croatia  
Phone: +385 (51) 770 447  
Fax: +385 (51) 686 166  
[www.intechopen.com](http://www.intechopen.com)

### **InTech China**

Unit 405, Office Block, Hotel Equatorial Shanghai  
No.65, Yan An Road (West), Shanghai, 200040, China  
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元  
Phone: +86-21-62489820  
Fax: +86-21-62489821

© 2011 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](https://creativecommons.org/licenses/by-nc-sa/3.0/), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen