MITSUBISHI ELECTRIC RESEARCH LABORATORIES
http://www.merl.com

Application of numerical optimization to the design of InP-based wavelength combiners

Ozbayat, S.; Kojima, K.; Koike-Akino, T.; Wang, B.; Parsons, K.; Singh, S.; Nishikawa, S.; Yagyu, E.

TR2014-019 July 2014

Abstract

We applied a numerical device optimization scheme, where tens of parameters can be optimized simultaneously with multiple target performance criteria that are given. The key items of the design scheme are the selection of the best optimization algorithm, metric, and consideration for fabrication errors. This method was then applied to design an MMI beam combiner with rectangular effective refractive steps with up to 75 parameters, and we obtained a simulated insertion loss of 0.8 dB for a 1.4 mm-long 2x1 wavelength combiner, and a simulated insertion loss of 4.2 dB for a 1.9 mm-long 4x1 wavelength combiner, both with 20 nm wavelength spacing. This methodology could also be applied to other types of optical devices.

Optics Communications

Application of Numerical Optimization to the Design of InP-based Wavelength Combiners

Selman Özbayat[a,b], Keisuke Kojima[a,*], Toshiaki Koike-Akino[a], Bingnan Wang[a], Kieran Parsons[a], Siddharth Singh[a,c], Satoshi Nishikawa[d], Eiji Yagyu[d]

[a] *Mitsubishi Electric Research Laboratories, 201 Broadway, Cambridge, MA, 02139*
[b] *Electrical and Computer Engineering Department, University of Massachusetts Amherst, Amherst, MA, 01003*
[c] *Department of Electrical and Computer Engineering, Stony Brook University, NY, 11794*
[d] *Advanced Technology R&D Center, Mitsubishi Electric Corporation, 8-1-1 Tsukaguchi-Honmachi, Amagasaki, Hyogo, Japan*

Abstract

We applied a numerical device optimization scheme, where tens of parameters can be optimized simultaneously with multiple target performance criteria are given. The key items of the design scheme are the selection of the best optimization algorithm, metric, and consideration for fabrication errors. This method was then applied to design an MMI beam combiner with rectangular effective refractive steps with up to 75 parameters, and we obtained a simulated insertion loss of 0.8 dB for an 1.4 mm-long 2×1 wavelength combiner, and a simulated insertion loss of 4.2 dB for an 1.9 mm-long 4×1 wavelength combiner, both with 20 nm wavelength spacing. This methodology could also be applied to other types of optical devices.

Keywords: Multimode interference combiner, optimization, wavelength combiner, evolution strategy

*Corresponding author
Email address: kojima@merl.com (Keisuke Kojima)

1. Introduction

Recently, there have been increasing interest in designing optical devices using computer algorithms to optimize many parameters. Spühler *et al.* used evolutionary strategy (ES) and finite difference beam propagation method (FD-BPM) to optimize segmented patterns of a spot-size converter represented by 27 integers on a silicon planar waveguide [1]. Felici and Gallagher optimized taper waveguide structure represented by 9 parameters using Quasi-Newton method, and maximized the coupling efficiency [2]. Liu *et al.* used combinatory search algorithm to optimize the photonic crystal structure for spatial mode conversion [3]. In these cases, device structures were optimized by using non-intuitive structures to outperform conventional straightforward structures.

Use of optimization algorithms for Multi-mode interference (MMI) couplers has also been reported. Wang *et al.* used genetic algorithm (GA) to optimize 3 parameters of an $N \times N$ MMI couplers [4]. West and Honkanen applied GA for a weak guiding MMI [5]. Wang *et al.* used Particle Swarm Optimization (PSO) to design ultra compact 2×2 MMI [6].

There has been an interest in InP-based photonic integrated circuits with multiple lasers, modulators, and beam combiners for realizing high performance and very compact transmitters for WDM (Wavelength Division Multiplexing) optical communication systems. One of the applications is a 40 Gb/s and 100 Gb/s Optical Ethernet, where multiple wavelengths are combined into a single fiber [7].

One way to achieve it monolithically is to use an MMI power combiner [8]. Its design and fabrication processes are well established, however, there is an inherent $3N$ dB insertion loss for 2^N:1 coupling, where $N = 1, 2, ...$ Another method for coupling is to use wavelength combiners/splitters. For example, 960 μm-long MMI-based wavelength splitters have been reported, but the wavelength separation needs to be very large (1.3/1.55 μm) [9], or device length needs to be as long as several millimeters for narrow wavelength spacing. A 120 μm-long MMI wavelength splitter using slot waveguide was proposed for 1.3/1.55 μm [10], however, slot waveguides cannot be applied to InGaAsP/InP material systems.

Alternatively, an InP-based compact 4×4 arrayed waveguide grating (AWG) was fabricated using deep reactive ion etching process, and the insertion loss around 5 dB has been reported [11]. An InP-based Mach-Zehnder Interferometer is also an alternative for wavelength coupler/splitter [12]. In

order to create optical path difference between two arms, typically of the order of tens of microns, the device occupies a long and wide area. A compact wavelength combiner/splitter has been demonstrated based on MMI-interferometer on silicon-on-insulator (SOI) platform [13]. However, this is not directly applicable to InGaAsP/InP material systems, in which very sharp bending is not possible.

In this paper, we propose an optimization scheme of many parameters to design device structures using the latest variant of evolution strategy [14], including consideration for fabrication tolerance. As examples, we apply it to MMI-based 2×1 and 4×1 wavelength combiners for 20 nm wavelength spacing, to achieve maximum transmittance. The MMI has multiple rectangular patches of low refractive index, whose sizes and positions are optimized to minimize the insertion loss. Up to 75 parameters were simultaneously optimized. The simulation results show that an insertion loss of 0.8 dB can be achieved with 1.4 mm total device length for a 2×1 combiner, and an insertion loss of 4.2 dB with 1.9 mm device length for a 4×1 combiner. Since it is straight and narrow ($< 10\mu$m), multiple devices can be placed densely, for various applications.

2. Physical design and numerical method

The base MMI device structure consists of an $\mathrm{In}_{1-x}\mathrm{Ga}_x\mathrm{As}_{1-y}\mathrm{P}_y$ ($y = 0.4$) core layer, sandwiched between the InP substrate and the InP upper cladding layer.

If we solely rely on MMI's basic wavelength selectivity, then the following relationship holds,

$$L = M \times L_\pi^{\lambda_1} = (M+1) \times L_\pi^{\lambda_2}, \tag{1}$$

where $L_\pi^{\lambda_1}$ and $L_\pi^{\lambda_2}$ are the beat lengths at wavelengths λ_1 and λ_2, respectively, and M is an integer. For an MMI device of lateral width W at wavelength λ, $L_\pi \propto W^2/\lambda$ [15], and the required L for a typical 1.27/1.29 μm wavelength splitter with $W = 8$ μm would be much longer than 5 mm.

The wavelength selectivity in this work is instead achieved by creating non-uniform refractive index distribution within an MMI, as shown in Fig. 1(a). One possible way is to etch the InGaAsP core layer by a small constant thickness T_g at pre-determined rectangular shapes and filling the grooves with an InP regrown layer as shown in Fig. 1(b). The etched depth T_g

determines the beam propagation inside the MMI. We used the effective refractive index method [16], which will allow a 2-dimensional simulation of the device in the course of optimization. Once the effective refractive indices for etched region n_{low} and the non-etched region n_{high} are fixed, then the 2-dimensional refractive index distribution (e.g. the top-view in Fig. 1(a)) of the MMI based device is sufficient for an approximate determination of the beam propagation in the device. We chose $T_g = 0.1$ μm, creating effective refractive index step ($\Delta n = n_{high} - n_{low}$) of 0.0167 at 1.27 μm, and 0.0162 at 1.29 μm.

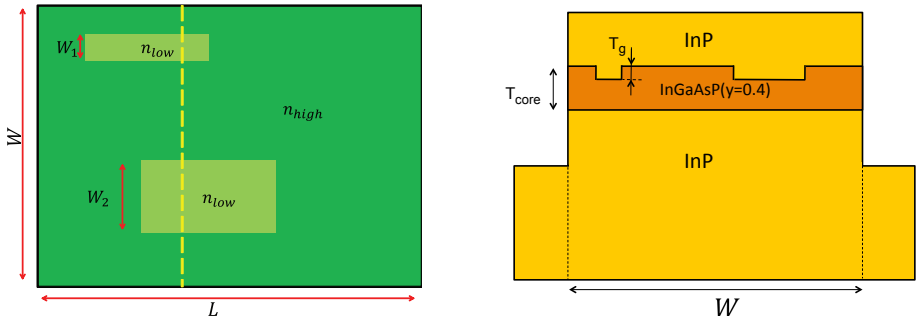

Figure 1: (a) Top view of the MMI with two rectangles with low effective refractive index, and (b) cross-sectional view along the yellow line of the top view, showing two regions of thinner core layer

Since our main target is an integrated wavelength combiner for a transmitter, we focus on transverse electric (TE) modes. We use the 2D finite difference beam propagation method (FD-BPM) [17] for fast computation. As a validation, we compared our results with those obtained by a commercial simulator FIMMWAVE [2] using a 3D EigenMode Expansion (EME) method for an MMI with two rectangles fixed devices, and they produced transmittance peak intensities within $\pm 1.5\%$.

3. Optimization strategy

It is not straightforward to derive the best configuration of low refractive index (n_{low}) patches buried into a high refractive index (n_{high}) comprising of a fixed MMI-based device, which would result in desired wavelength selectivity in a short device. Therefore, we determine a set of parameters

for randomizing the distribution in a flexible way, whereas an efficient optimization algorithm is employed to converge to the best realization of each parameter. The covariance matrix adaptation evolutionary strategy (CMA-ES) optimization is a rather recent and powerful algorithm (Refs. [14], [18]) which requires only a single input a-priori, thus has the desirable feature of self-adaptation. This is a major advantage for optimization problems like ours, where the underlying theory is less explanatory and it is hard to choose set of optimizer parameters.

The MMI lateral width W is fixed, whereas the total device length L along propagation axis is kept a variable. The refractive index distribution is randomized by employing a fixed number of rectangular patches N_p, each identified by 4 variables: patch lateral width W_p, patch offset O_p, patch position P_p and patch length L_p. For the 2×1 wavelength combiner, two input waveguides that couples their fundamental TE modes to the MMI device is assumed along with a similar waveguide at the output of the MMI device. The positions of these three waveguides are identified by lateral axis offsets O_{in1}, O_{in2} and O_{out}, whereas the widths of all three are an identical variable, W_{port}. An example of 5 patches positioned randomly, along with the input and output waveguides are depicted in Fig. 2. Each patch's 4 variables are assigned low and high supports in such a way that they can overlap, reside partially inside the MMI body, or completely end up outside the MMI body. Therefore, it is important to assign reasonable supports for W_p's, O_p's, P_p's and L_p's, as well as L. With this convention, the total number of variables in the above optimization problem is $N_{var} = 4 \times N_p + 5$, where N_p is the number of patches initially set.

Like other evolutionary type strategies, CMA-ES searches for the global optimum in the function space based on several particles (also called agents), and the particles' history of function evaluation distributed on evolving hyper ellipse determines the direction of the algorithm in the next iteration.

Another aspect of the optimization process is the metric function, i.e. the value returned by each particle at each iteration, that is used to evaluate the optimizer behavior at a given iteration. The wavelength combiner is designed to couple the beam at wavelength λ_1 and λ_2 to the output port, the metric should be chosen according to the functionalities of the device. Among several metric functions we have considered, we chose the one with which the average convergence of CMA-ES is maximized:

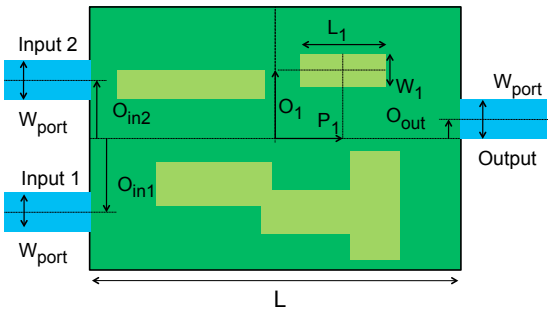

Figure 2: An example of the device geometry with multiple rectangular patches, each represented by 4 variables.

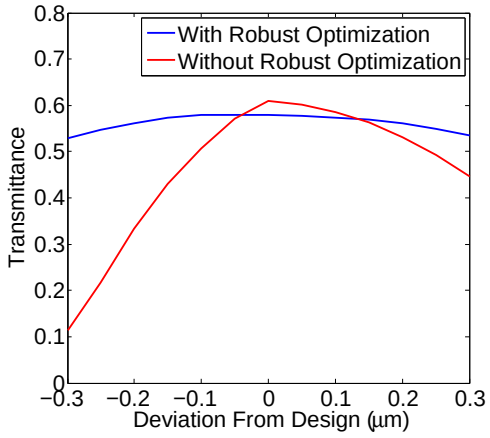

Figure 3: Transmittance of the 2×1 wavelength combiner with 4 rectangular patches vs patch width deviation from design. Two devices, with and without consideration of robust optimization, are shown.

$$Metric = \min(P_1^{\lambda_1}, P_2^{\lambda_2}), \tag{2}$$

where $P_1^{\lambda_1}$ and $P_2^{\lambda_2}$ are the transmittances from the 1st and 2nd input waveguide at wavelength λ_n and λ_2, respectively. Transmittance is the ratio of the fundamental output mode power to the input fundamental mode power, calculated from the overlap integral including phase. This choice of metric means we try to maximize the worst of the transmittance at two wavelengths.

When the device is optimized naively just to maximize the metric, they tend to be very sensitive the fabrication variations, such as the patch width. In order to optimize the design to be less sensitive to this variation, we chose to use a simplified version of robust optimization [19], where metric was calculated at three points: one with nominal dimensions, one with each patches 0.2 μm narrower, and one with each patch 0.2 μm wider than nominal values. The worst metric among the three is passed to the optimizer, assuring optimization of the worst combination that accounts for a ± 0.2 μm fabrication error.

Figure 3 compares the result with and without robust optimization techniques. It is clearly shown that optimized device with robust optimization shows less sensitivity for fabrication variations. For the rest of the paper, we only show results when robust optimization is used.

We have also compared three widely-used global optimization methods as shown in Fig. 4. PSO optimizes a problem by having a population of candidate solutions, here dubbed particles, and moving these particles around in the search-space according to simple mathematical formulae over the particle's position and velocity [20]. Continuous ant colony optimization (CACO) is an optimization algorithm, whose original idea came from the behavior of ants seeking a path between their colony and a source of food [21]. Even though CMA-ES did not converge very quickly, it almost always gave best performance at the end. It should also be stated that these optimizers are not guaranteed to converge to the same optimized device over two independent identical simulations. In fact, many of the runs resulted in trivial solutions where MMI is marginally better than a power combiner. Thus we run about 100 simulations in parallel in a computer cluster, out of which the best runs are displayed in the figure.

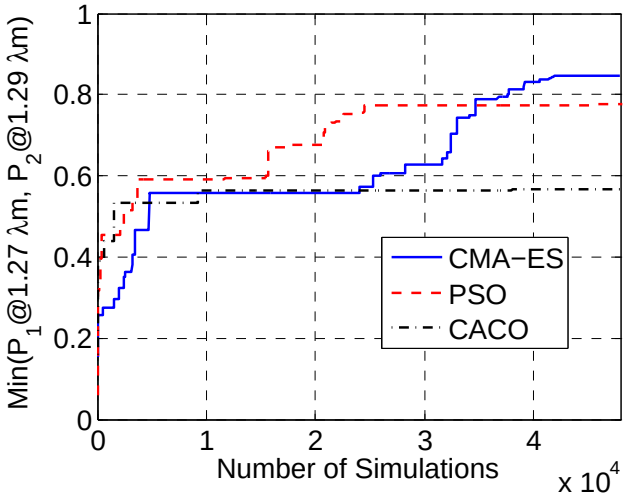

Figure 4: Metric (worst of the transmittance at two wavelengths) as a function of number of simulations, for three widely-used global optimization algorithms, for 2×1 wavelength combiners with 14 patches.

Variable Name	Lowest Value(μm)	Highest Value(μm)
L	1000	1500
W_{port}	2.5	3
O_{in1}	-2.4	-1.6
O_{in2}	1.6	2.4
O_{out}	-2.4	2.4
$W_{pn},\ n = 1, 2, .., N_p$	0.35	2.0
$O_{pn},\ n = 1, 2, .., N_p$	-4.9	4.9
$P_{pn},\ n = 1, 2, .., N_p$	-750	750
$L_{pn},\ n = 1, 2, .., N_p$	50	500

Table 1: Lowest and highest values for each of each variable.

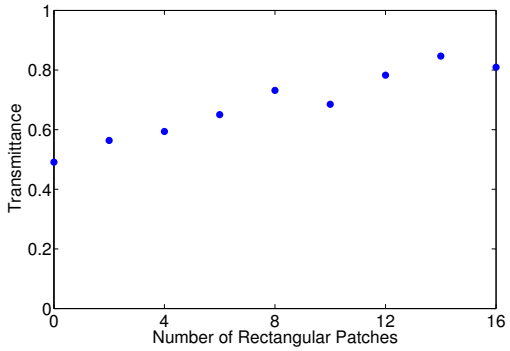

Figure 5: Transmittance (worst of the 2 wavelengths) of 2×1 wavelength combiners as a function of the number of patches.

4. Two Beam Combiner Results

We first used the above methodology to design a 2×1 beam combiner. We used a fixed device width (along lateral x axis) of $W = 8$ μm. $\lambda_1 = 1.27$ μm and $\lambda_2 = 1.29$ μm were used as they are commonly used for data communication. The number of patches to parameterize N_p was varied from 2 to 16, , i.e. the CMA-ES optimizer will maximize a metric function up to $N_{var} = 69$ degrees of freedom. We kept the maximum iteration number for the optimizer as 800 and number of particles was set to 20, hence the total number of accesses to the FD-BPM solver is 48,000 ($800 \times 20 \times 3$) during the course of one optimization run. The factor of 3 comes from the fact that the metric is evaluated at three positions, at nominal patch width, and ± 0.2 μm from nominal values. We typically conducted about 100 optimization runs for each N_p, on a computer cluster. Grid sizes were $\Delta x = 0.005$ μm, and $\Delta z = 0.2$ μm, along lateral and propagation axes, respectively. Note that the minimum value for the number of particles N_{part} was given by [14]

$$N_{part} = 4 + [3 \times \ln(N_{var})], \tag{3}$$

where [ ] is a floor function. This gives $N_{part} = 20$ for $N_{var} = 69$ We did a small comparison study between $N_{part} = 20$ and 50 with a fixed number of simulations, and no obvious difference was observed.

Table 1 tabulates each of the variable's support range used here.

The refractive index distribution of the optimized device and the electric field propagation patterns are shown in Fig.6 and Fig.7, respectively.

The refractive index distribution of the optimized device is shown in Fig. 6. The MMI length was 1396 μm. Propagation of electric fields at 1.27 μm and 1.29 μm are shown in Figs. 7 (a) and (b), respectively. Figure 8 shows the transmittance from input ports 1 and 2 as a function of wavelength. The loss of 0.77 dB and extinction ratio of 10.8 dB are obtained for this device.

Figure 9 shows the transmittance with 1.27 μm into Input Port 1 and 1.29 μm into Input Port 2, as a function of patch width deviation from the nominal values. The transmittance degrades only 4% when the width changed by ± 0.2 μm, which is achieved by the consideration of fabrication errors as explained in the previous section.

Figure 10 shows the transmittance with 1.27 μm into Input Port 1 and 1.29 μm into Input Port 2, with W_{MMI} varied by ± 0.05 μm from the nominal 8 μm. In this case, the maximum transmittance degraded by 8%. Even

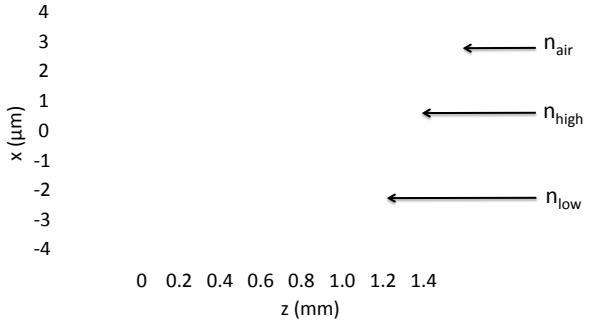

Figure 6: The optimized 2×1 wavelength combiner using 14 rectangular patches.

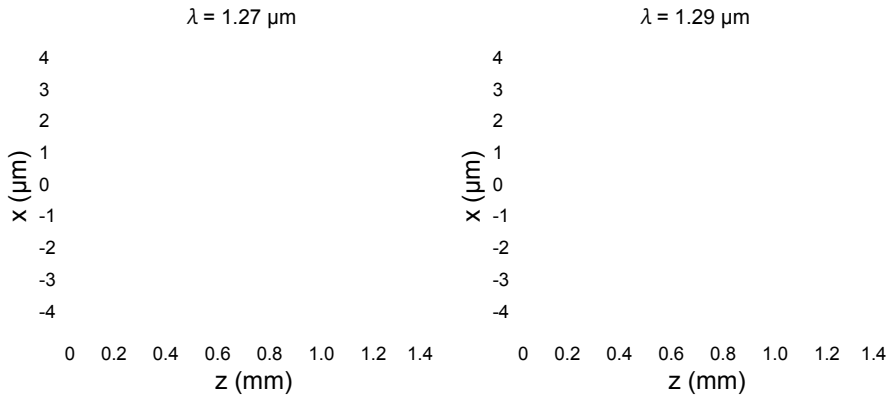

Figure 7: Propagation pattern of the TE mode with (a) 1.27 μm light into the lower port, and (b) 1.29 μm light into the upper port.

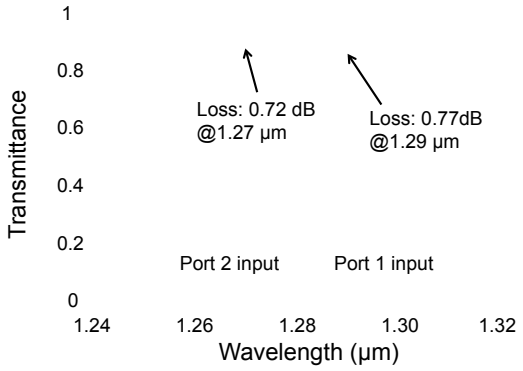

Figure 8: Transmittance for 2 input ports as a function of wavelength of the optimized 2×1 wavelength combiner.

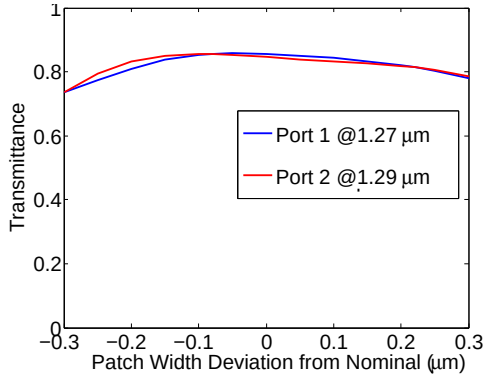

Figure 9: Transmittance of the 2×1 wavelength combiner for 1.27 μm light input Port 1 and 1.29 μm light input Port 2 as a function of patch width deviation from the nominal value.

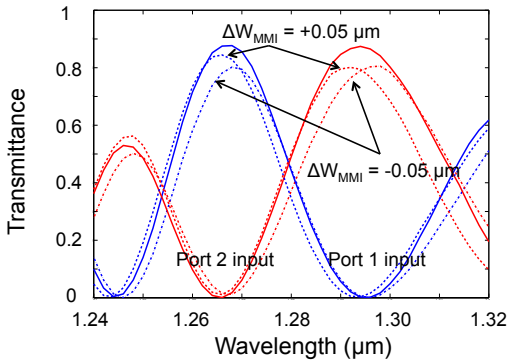

Figure 10: Transmittance for 2 input ports with W_{MMI} deviation from the nominal value.

though ± 0.05 μm is within the state-of-the-art processing capability, further design improvement would be desirable.

5. Four Beam Combiner Results

We further extended this scheme for a 4×1 wavelength combiner for 1.271 μm, 1.291 μm, 1.311 μm, and 1.331 μm, using 16 patches. In this case, $N_{var} = 4 \times N_p + 11$, where $N_p = 16$. The additional 11 parameters include the width and offset of all the input and output ports, and the total MMI length. Therefore, a total of 75 dimensions had to be optimized. For this device, after using a global optimizer CMA-ES, we used the Nieder-Mead local optimizer [22] to fine tune the variables to gain extra 3% in transmittance. Fig. 11 shows the optimized refractive index patterns, and Fig. 12 shows the transmittance from one of the 4 ports to the output port, as a function of wavelength. The MMI length was 1901 μm. The worst case insertion loss of 4.2 dB is achieved by simulation, which is a major improvement compared to the inherent 6.0 dB loss of a conventional MMI-based power combiner.

6. Discussions

At this moment, the the number of parameters to optimize are limited by the available computer resources. By using more CPUs, we may be able to obtain even better results by optimizing more design parameters.

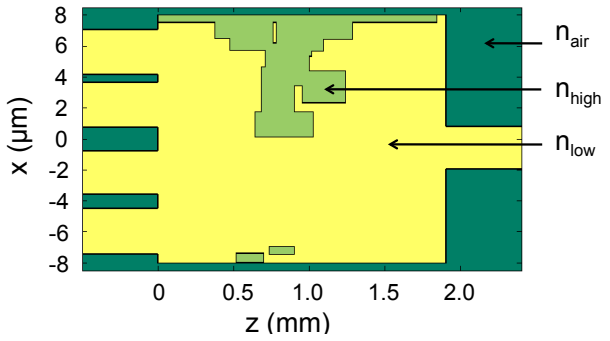

Figure 11: Effective refractive indes distribution of the optimized 4×1 wavelength combiner with 16 patches.

This design optimization process does not necessarily guarantee the absolute best performance physically possible. However, sufficiently increasing the number of runs will bring the results very close to the optimum structure.

The device structure introduced here only uses simple rectangular patches. If we introduce tapered shape structures for smooth transition, we may expect further improvements.

Even though all the simulations were conducted for InP-based materials, this design methodology can be readily applied to other material systems, including silicon on insulator (SOI).

7. Conclusion

We propose a method of designing and optimizing optical devices, where device structure is represented by tens of parameters. The CMA-ES global optimizer takes advantage of multiple independent cases simulated simultaneously. We also proposed a method of robust design considering fabrication errors. This was applied to the design of an InP-based 2×1 wavelength combiner, resulting in a simulated insertion loss of 0.8 dB. In addition, a 4×1 wavelength combiner was designed with a simulated insertion loss of 4.2 dB with as many as 75 optimized parameters. Compared to AWGs or MZIs, our proposed device is straight and narrow, and multiple devices can be placed close to each other, which may be beneficial in various applications. This design methodology is applicable to broad range of optical devices that had not been realized with conventional simpler structures.

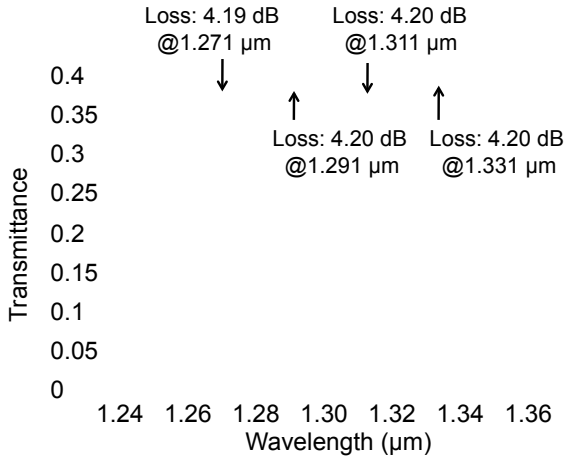

Figure 12: Transmittance for 4 input ports as a function of wavelength of the optimized 4×1 wavelength combiner.

References

[1] M. Spühler, B. J. Offrein, G.-L Bona, R. Germann, I. Massarek, D. Erni, J. Lightwave Technol. 16 (1988) 1680-1685.

[2] T. P. Felici, D. F. G. Gallagher, Proc. SPIE, 4986 (2003) 48-58.

[3] V. Liu, D.A.B. Miller, S. Fan, Optics Expr. 20 (2012) 28388-28397.

[4] Q. Wang, J. Lu, S. He, Opt. Commun. 209 (2002) 131-136.

[5] B. R. West, S. Honkanen, Opt. Expr. 12 (2004) 2716-2722.

[6] Q. Wang, S.-T. Ho, J. Lightwave Technol. 28 (2010) 1298-1304.

[7] IEEE Standard 802.3ba.

[8] S. Kanazawa, T. Fujisawa, A. Ohki, H. Ishii, N. Nunoya, Y. Kawaguchi, N. Fujiwara, K. Takahata, R. Iga, F. Kano, and H. Oohashi, Proceeding of Semiconductor Laser Conf. (2010) 70-71.

[9] C. Yao, H.G. Bach, R. Zhang, G. Zhou, J.H. Choi, C. Jiang, R. Kunkel, Optics Expr. 20 (2012) 18248-18253.

[10] J. Xiao, X. Liu, X. Sun, Optics Expr. 15 (2007) 8300-8308.

[11] Y. Barbarin, X. J. M. Leijtens, E. A. J. M. Bente, C. M. Louzao, J. R. Kooiman, M. K. Smit, IEEE Photonics Technol. Lett. 16 (2004) 2478-2480.

[12] K. Watanabe, N. Koshobu, K. Hirabayashi, J. Kobayashi, H. Ishii, Electron. Lett. 47 (2011) 1245-1246.

[13] F. Xia, M. O'Boyle, L. Sekaric, Y.A. Vlasov, Proceedings of InP and Related Materials Conf. (2006) 429-430.

[14] M.D. Gregory, Z. Bayraktar, D.H. Werner, IEEE Trans. Antennas Propagat. 59 (2011) 1275-1285.

[15] L.B. Soldano, E.C.M. Pennings, J. Lightwave Technol. 13 (1995) 615-627.

[16] R.G.Hunsperger, Integrated Optics, 6th ed. Springer, 2009.

[17] D. Schulz, C. Glingener, E. Voges, IEEE J. Quantum Elect. 30 (1994) 1132-1140

[18] N. Hansen, A. Ostermeier, Evol. Comput. 9 (2001) 159-195.

[19] D. Bertsima, M. Sim, Operations Res. 52 (2004) 35-53.

[20] J. Robinson, Y. R. Samii, IEEE Trans. Antennas Propag. 52 (2004) 397-407.

[21] K. Socha, M. Dorigo, European J. Operational Res. 185 (2008) 1155-1173.

[22] J. A. Nelder, R. Mead, Computer J., 7 (1965) 308-313.