Supplementary Material

This document contains additional information on the article "Grating couplers in Silicon-on-Insulator: The role of photonic guided resonances on lineshapes and bandwidth". In particular, we provide:

- Details on FDTD simulations;
- Summary of parameters of the presented structures;
- Details on fitting procedure;
- Details on the calculation with Rigorous Coupled Wave Analysis;
- Details on optimization procedure;
- Parameters of the structures composing the Pareto fronts;
- Tolerance analysis with respect to small variations in position, width and etching depth of the grooves;
- Comparison between 2D and full 3D FDTD simulations of the optimized structures with 4 µm MFD.

I. FINITE DIFFERENCE TIME DOMAIN SIMULATION

As the main numerical technique in this work, we employ the standard Finite Difference Time Domain method, implemented in the commercial software Lumerical FDTD Solutions. The idea of the method is to discretize the magnetic and electric fields in two interlaced grids in space and time and solve the discretized Maxwell equations in time domain. For the FDTD method to work properly, it is necessary to set up the environment of the simulation, specifying the discretization grid (mesh), the boundary conditions at the terminations of the simulation box, the sources and the monitors. In our case we decided to analyze the structures with 2D FDTD (an example of a simulated structure as directly extracted from the software can be seen in Supp. Fig. 1), a comparison with 3D FDTD simulations is presented in Suppl. Sec. VIII. We use the following settings:

- **mesh:** we opted for automatic non-uniform meshing (with maximum precision setting) for the whole simulation region, except in the waveguide region, where a vertical mesh in steps of 10 nm was forced to better reproduce the vertical geometry of the grating;
- **boundary condition:** we adopt the Perfectly Matched Layer (PML) formulation provided by the software, which absorbs almost all the radiation incident on the boundary;
- **source:** the source is chosen to inject a Gaussian beam on top of the grating structure with TE polarization and an axis inclined at 10°. This source injects a short broadband pulse with a sufficiently wide energy spectrum centered around the wavelength of 1.55 µm;
- **monitor:** we use a single monitor placed in the waveguide region at the end of the grating. This monitor calculates run-time the power flow through its surface, and after normalizing to the input power at the end of the simulation it directly provides the coupling efficiency spectrum.

The typical simulation time for one structure spans from a few second to a few minutes (depending on the simulation area, influenced by the MFD) on a standard quad-core desktop PC.

II. DETAILED PARAMETERS OF THE INVESTIGATED STRUCTURES

A grating coupler is characterized by a number of parameters, which define the vertical waveguide structure (silicon thickness and bottom-oxide thickness, etch depth), the grating area and the incident fiber mode (footprint FOOT, mode-field diameter MFD, fiber offet FIO, fiber distance FID, incidence angle θ), and the parameters defining the grating (period and duty cycle of the grooves in the case of uniform gratings, positions and widths of each groove in the case of apodized gratings). Most of these parameters are quoted in the article, however for the sake of clarity we summarize in Supp. Tab. I the complete set of parameters for all the structures pertaining to the data presented in the article.

III. FITTING OF SPECTRA

To better understand the physics at the origin of the coupling for uniform gratings we decided to fit the FDTD spectra to obtain their relevant parameters. For each spectrum we tried fitting with 3 different functions, namely Gaussian, Lorentzian and Voigt (a convolution between Lorentzian and Gaussian) profiles.

For the Gaussian and Lorentzian case we defined the fitting functions as standard (unnormalized) Gaussian and Lorentzian profiles, namely:

$$\begin{aligned} G(x; A, x_0, \sigma) &= A e^{-\frac{(x-x_0)^2}{2\sigma^2}} \\ L(x; A, x_0, \gamma) &= A \frac{\gamma^2}{(x-x_0)^2+\gamma^2}. \end{aligned} \tag{S1}$$

both with 3 fitting parameters. Note that in this unnormalized versions the coefficient A gives directly the value of the functions at their maximum (corresponding to the coupling efficiency).

The Voigt profile can be obtained as a convolution of Lorentzian and Gaussian, namely given the two normalized profiles:

$$G_n(x;\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} \quad L_n(x;\gamma) = \frac{1}{\pi}\frac{\gamma}{(x-x_0)^2+\gamma^2}. \quad \text{(S2)}$$

the normalized Voigt profile can be defined as:

$$V_n(x;\sigma,\gamma) = \int_{-\infty}^{+\infty} G_n(x';\sigma) L_n(x-x',\gamma) dx' \quad \text{(S3)}$$

For the practical implementation of the Voigt profile as fitting function it is worth noting that the Voigt profile can be related to the real part of the Faddeeva function $w(z)$ as:

$$V_n(x;\sigma,\gamma) = \frac{1}{\sigma\sqrt{2\pi}} \Re\left[w\left(\frac{x+i\gamma}{\sigma\sqrt{2}} \right) \right] \quad \text{(S4)}$$

where the Faddeeva function is a complex error function defined as:

$$w(z) = e^{-z^2}\left(1 + \frac{2i}{\sqrt{\pi}} \int_0^z e^{t^2} dt\right). \quad \text{(S5)}$$

Since the Feddeeva function is often provided as a module in standard scientific software, this provides a more practical implementation of the Voigt profile. As fitting function in the case of Voigt profile we define the function $V(x; A, x_0, \sigma, \gamma)$ with 4 fitting parameters as:

$$V(x; A, x_0, \sigma, \gamma) = A V_n(x - x_0; \sigma, \gamma) \quad \text{(S6)}$$

The fitting procedure was carried out within a Python framework, utilizing the standard non-linear least-square method provided by the Scipy module (`scipy.optimize.curve_fit`). The same module is also used to calculate the Faddeeva function (`scipy.special.wofz`).

To verify the goodness of our fits we decided to analyze the residual of the least-square minimization, namely the square root of the mean quadratic deviation:

$$R = \sqrt{\frac{\sum_{i=1}^{N} (f(x_i) - y_i)^2}{N}}. \quad \text{(S7)}$$

A small value of this parameter (with respect to the values of the fitted quantity) should indicate that the fit is good. We performed this analysis for each of the uniform structures analyzed in the article. We found that in almost all the cases, the Voigt profile provides a good fit with a small R (around 10^{-1} or smaller). Furthermore we found that the Gaussian profile provides a R value similar to the Voigt (and hence a good fit) only for small MFD and a bad fit for large MFD, while the opposite happens for the Lorentzian profile. This confirms the evolution of the spectra from Gaussian to Lorentzian on increasing the MFD. An example of the analysis of the residual is reported in Supp. Fig. 2.

IV. PARAMETER EXTRACTION FROM RCWA

As an analysis method for the infinite grating we opted for Rigorous Cuopling Wave Analysis. This method, based on an application of the Scattering Matrix, was developed to analyze multilayered structures in which each layer presents an independent 2D periodic patterning (with the only restriction of having the same 2D lattice in each layer). With the RCWA available to us it is possible to calculate reflection and transmission spectra of the infinite 1D grating under plane wave illumination at different angles and energies (and for both TE/TM polarization). From the analysis of the reflectance spectra it is possible to extract information on the quasi-guided mode that mediates the coupling, in particular the real part (resonance frequency) and the imaginary part (half the resonant linewidth) of the energy. To proceed with this analysis we have to model the reflection spectra for this kind of structures. The reflection in this case is composed of two contributions: a background contribution, given by the vertical structure of the multilayer, and a resonant one, provided by the 2D patterning. This kind of behavior is known in the literature as a Fano resonance, and it is characterized, in the reflection and transmission spectra, by a Fano lineshape, which can be parametrized as:

$$F(x; A, x_0, g, q) = A \frac{\left(q + \frac{2(x-x_0)}{g}\right)^2}{1 + \frac{4(x-x_0)^2}{g^2}} \quad \text{(S8)}$$

where x_0 is the position of the resonance, g its width, q is the ratio between the resonance and the background and A is a multiplicative constant introduced for fitting purposes. It is worth noting that in case of no background, the parameters q tends to infinity and the Fano lineshape tends to a Lorentzian with $\gamma = g$.

An example of a reflection spectrum calculated with RCWA can be seen in Supp. Fig. 3b. We could try to fit directly this curve to estimate the parameters, but there is still one subtlety. In order to obtain a Fano lineshape, a constant background (i.e., independent of energy) is usually assumed. This is not the case here since the background contribution is given by constructive/destructive interference in the vertical direction, and it is thus energy dependent. To remove this effect we tried to calculate the contribution of the background only, substituting the patterned region with a uniform slab with the averaged dielectric constant (this is also plotted in Supp. Fig. 3b). After rescaling the spectrum by eliminating the background, we fitted it with Eq. S8 to obtain the real part (x_0) and the imaginary part (g) of the resonance. An example of this fit is presented in Supp. Fig. 3c. Knowing the energy of the resonance, it is also possible to calculate the effective index of the patterned slab noting that the wavevectors of the incident and guided mode have to be related by momentum conservation:

$$k_g = k_i + \frac{2\pi}{\Lambda} \quad \text{(S9)}$$

where Λ is the period of the grating. Once this equation is rewritten as:

$$\frac{\omega}{c} n_{eff} = \frac{\omega}{c} n_{clad} \sin\theta_i + \frac{2\pi}{\Lambda} \tag{S10}$$

it can be used to calculate the effective index n_{eff} of the quasi-guided mode.

V. DETAILS ON OPTIMIZATION PROCEDURE

Since the MO-PSO algorithm is very time consuming from a computational point of view, we performed a series of preliminary studies for each value of MFD in order to get a valid starting point for the MO-PSO. In particular, we decided to include knowledge of the linear chirped grating in the optimization for each MFD, setting one of the agents of the swarm to start in this configuration. For the 10.4 µm MFD we directly included in the MO-PSO the knowledge of the apodized structure with maximum CE reported in the literature, so the following procedure was carried out only for the three smaller MFDs.

To find the chirped structure with maximum coupling efficiency we first conducted a sweep along the parameters of the uniform grating, namely duty cycle, etching depth, bottom oxide thickness and fiber offset. The period of each uniform grating was chosen to tune the maximum CE at 1550 nm.

After this first step we moved to linear chirped structures, namely structures in which the duty cycle is varied along the grating in a linear way from a value DC_{start} to a value DC_{stop} along a length L_c following:

$$DC(x) = \begin{cases} DC_{start} + (DC_{stop} - DC_{start})\frac{x}{L_c} & x < L_c \\ DC_{stop} & x > L_c \end{cases} \tag{S11}$$

We optimized the coupling efficiency with a single objective PSO with the free parameters DC_{start},DC_{stop},L_c, etching depth, bottom oxide thickness and fiber offset around the parameters obtained for the uniform grating (still modifying the period at each iteration to keep the maximum CE at 1550 nm). The results of this optimization can be seen in Supp. Tab. II. Then for the MO-PSO we fixed etching depth, bottom oxide thickness and fiber offset at the value provided by the chirped optimization, and optimized the position and width of each groove (but setting at least one agent to start in the chirped configuration already found).

VI. PARETO FRONTS' DATA

We include the detailed parameters of the structures obtained by the MO-PSO optimization. The data are presented in the attached files (one .dat file for each value of Mode Field Diameter). At the beginning of each file we report the parameters common to all structures in the Pareto front (corresponding to Fig. 1 in the main text). This is followed by tables with widths and positions of each groove for each specific structure, which is numbered with increasing Coupling Efficiency and decreasing Bandwidth. In the following, the links for the files containing these data are reported.

- supp_MFD_4.txt for MFD=4µm;
- supp_MFD_6.txt for MFD=6µm;
- supp_MFD_8.txt for MFD=8µm;
- supp_MFD_10.txt for MFD=10.4µm.

For some of the optimized structures (three for each value of MFD) we plot the grooves' widths versus their positions, to give a feeling of the effect of apodization. These plots can be seen in Supp. Fig. 5. The interpretation is discussed in the article.

VII. TOLERANCE ANALYSIS

The fabrication of grating couplers by lithograpy and etching is prone to introducing random variations of the parameters from the designed optimal structure. In the lack of an experimental validation of our designs, we decided to address this issue by performing a tolerance analysis of the results employing again FDTD simulations. We split the tolerance analysis in two parts: tolerance against variation of the etching depth of the grooves, and tolerance to variation in their width and position.

A. Etching depth

Controlling the etching depth in the fabrication of grating couplers is a critical and not easy task. In fact, the etching process has to be adapted and re-tuned every time the nominal etching depth is changed. Due to the difficulties in controlling an etching process to obtain the exact groove depth, the resulting fabricated structure may be slightly different from the designed one. This variation in etching depth, which should be no larger than ±10 nm, tends to be uniform among the different grooves of a given structure. Hence, we hereby study the tolerance with respect to the etching depth by considering a selection of the structures in the Pareto front, and we recalculate them at varying etching depth in the range ±20 nm from the optimal value, with a 1 nm step. In order to be sensitive to such small variations, the FDTD mesh along the vertical direction was refined down to 1 nm in the grating region. In Supp. Fig. 4 we report the behavior of CE and bandwidth for some of the structures in the Pareto front. As a general trend, we notice that a good tolerance to variations of the etching depth exists. In fact, all structures present a 10-15 nm range in which the CE is fairly constant, although it is not always symmetric around the starting value. This follows from the fact that the Pareto fronts were obtained at fixed etching depth, as explained in the main text. Around the

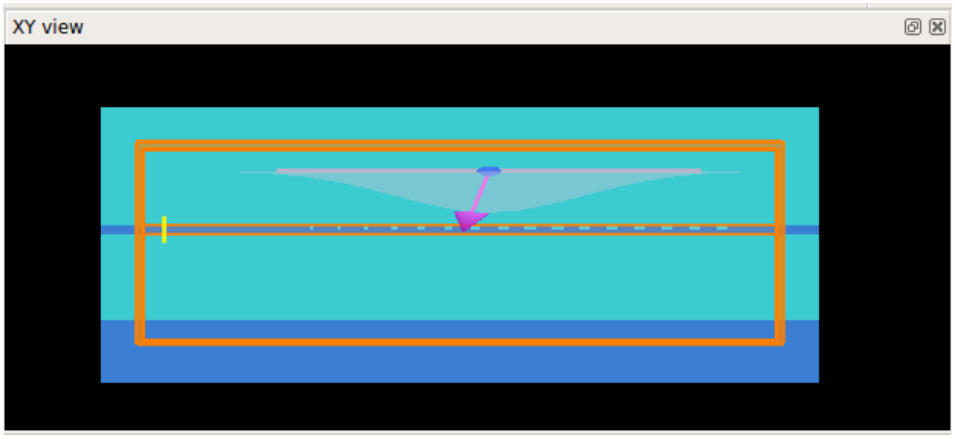
position of maximum CE the bandwidth tends to monotonically increase with etching depth, but the variations are limited to a few nm. Summarizing, tolerance against variations of the etching depth is quite good. In principle the results of the main text could be slightly improved by taking the etching depth as an optimization parameter when calculating the Pareto fronts, however this requires a much finer mesh along the vertical direction (1 nm instead of 10 nm) and would lead to a minor improvement in CE or bandwidth, of the same order of the tolerance against fabrication control.

B. Widths and positions

Another potential issue for the performances of fabricated structures is the random variation in the widths and positions of each groove, as induced by the fabrication process. The latter typically introduces a random and uncorrelated variation of the parameters in each groove of the designed structure. To test the tolerance of our designs against these variables, we run a Monte Carlo analysis limited to the structures in the Pareto fronts. For each structure a set of 20 new structures was generated and analysed, as obtained by varying the width and position of the grooves in the original structure by applying a uniform distribution of ± 10 nm broadening around the original one. In the case of 10.4 μm MFD we took the new configurations given by the tolerance analysis as starting points for an additional optimization run of the MO-PSO algorithm, which was found to improve convergence of the Pareto front. The results are reported in Supp. Fig. 6. We notice that the spread of performances is quite small (at least for the designs with the largest CE), and the general trend of the Pareto front is well reproduced. We also notice the presence of a few points lying slightly above the Pareto fronts, which can be explained by the uncertainty in convergence of the MO-PSO algorithm. Since most of the structures close to the optimal ones lie below the Pareto fronts, we conclude that convergence of the algorithm is comparable with the precision of fabrication, thus a better convergence should not improve the design of the fabricated structures.

VIII. COMPARISON WITH 3D SIMULATION

Since 1D grating couplers are uniform along the direction of the grooves, they are usually studied by using 2D simulation techniques. This is an approximation, which provides remarkably good results with considerably less computational effort as compared to full 3D simulations. In fact, the main source of approximation is due to the finite size of the excitation spot, which makes the simulated system non-uniform along the groove direction. The effect is larger for structures with smaller footprint. We hereby report a test of our 2D FDTD results for the 4 µm MFD Pareto front, by recalculating a few selected structures with this very small footprint. Results are shown in Supp. Fig. 7, where a very good agreement between 3D and 2D results is found. This analysis validates the use of 2D FDTD as a reliable tool for this problem.



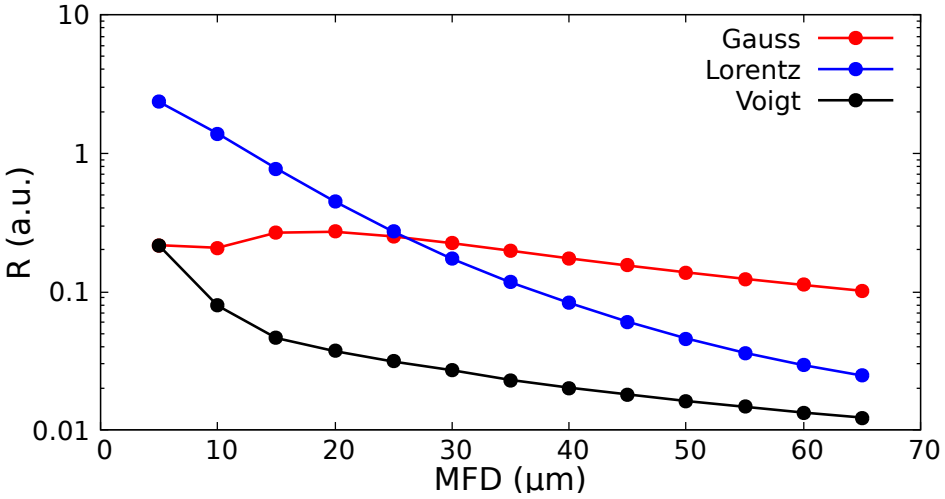
Supplementary Figure 1. CAD representation of the simulation area as directly extracted from Lumerical FDTD Solutions. The orange rectangle represents the boundary of the simulation region. The grey line and shaded region represent the source and its intensity profile, while the purple arrow represent the injection direction of the light and the blue arrow its polarization. The yellow line represent the monitor.

curve	DC (%) (%)	Etch (nm)	T_box (nm)	Period (nm)	MFD (µm)	FOOT (µm)	FIO (µm)	FID (µm)	θ deg
				Figure 2					
–	25	80	not def.	619	∞	∞	not def.	not def.	10
				Figure 3					
green	25	40	2000	603	varying	1.344*MFD	0.470*MFD	1.28	10
blue	25	60	2000	610	varying	1.344*MFD	0.470*MFD	1.28	10
red	25	80	2000	619	varying	1.344*MFD	0.470*MFD	1.28	10
black	25	120	2000	642	varying	1.344*MFD	0.470*MFD	1.28	10
				Figure 4					
–	25	80	2000	619	varying	1.344*MFD	0.470*MFD	1.28	10
				Figure 5					
–	25	80	2000	619	varying	1.344*MFD	0.470*MFD	1.28	10
				Figure 6					
red	not def.	150	2100	684 (av.)	4	5	2.6	1.28	10
blue	not def.	120	2000	662 (av.)	6	8	3.7	1.28	10
black	not def.	100	2000	644 (av.)	8	10	4.2	1.28	10
green	not def.	100	2000	634 (av.)	10.4	15	7.5	1.28	10

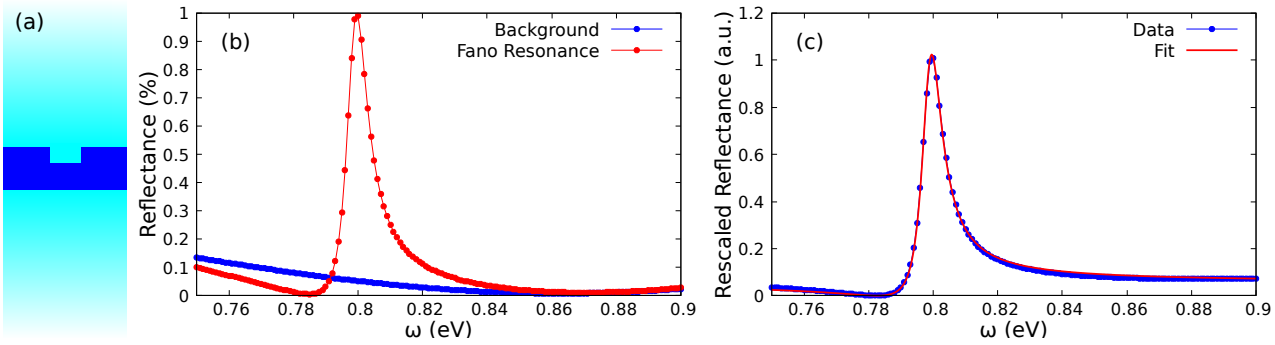
Supplementary Table I. Summary table of all parameters of the structures whose data are presented in the article.

MFD (µm)	CE (%)	Bandwidth (nm)	Start DC (%)	End DC (%)	Chirp Length (µm)	Period (nm)	Etch (nm)	T_BOX (nm)	FIO (µm)	FID (µm)	Footprint (µm)
4	56.7	89	10	35	2.13	684	150	2.10	2.48	1.28	9.0
6	60.1	67	10	40	4.97	662	120	2.02	3.68	1.28	10.6
8	61.6	54	10	40	4.98	644	100	2.00	4.19	1.28	12.9

Supplementary Table II. Parameter of the linear chirped configuration with maximum Coupling Efficiency.



Supplementary Figure 2. Representation of the residual of the fit from the same uniform structure as in Figs. 4 and 5 in the article (DC=25%, Etching = 80 nm, T_BOX= 2 µm). We can notice the goodness of the Voigt fit and the evolution of the spectrum from Gaussian to Lorentzian.



Supplementary Figure 3. How to extract data from RCWA calculations. (a) Unit cell for the calculation (the shaded areas are assumed to be semi-infinite). (b) Reflection spectra obtained from RCWA, with the pattern (red curve) and with the substitution of the patterned region with an uniform effective slab (blue curve). (c) Rescaled reflectance with no background variation and corresponding Fano fit for parameters.

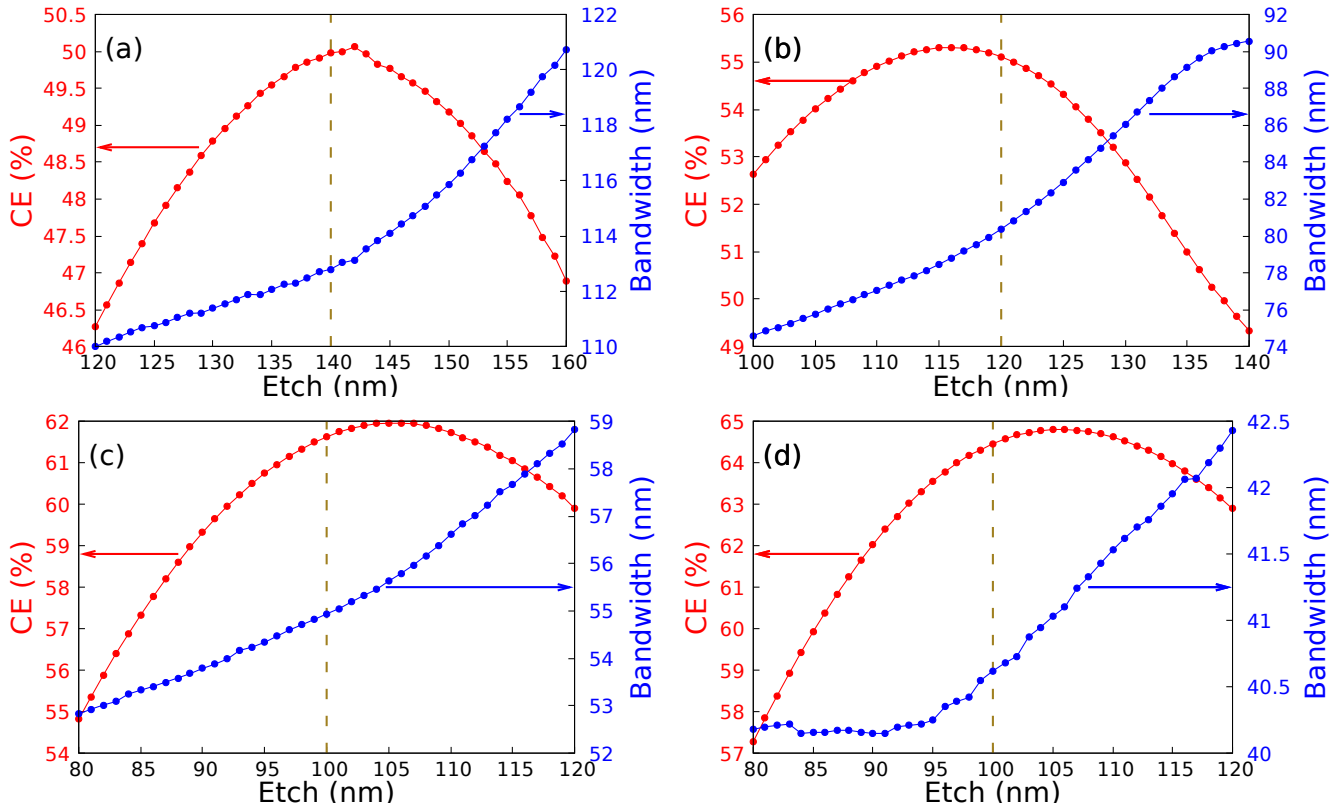
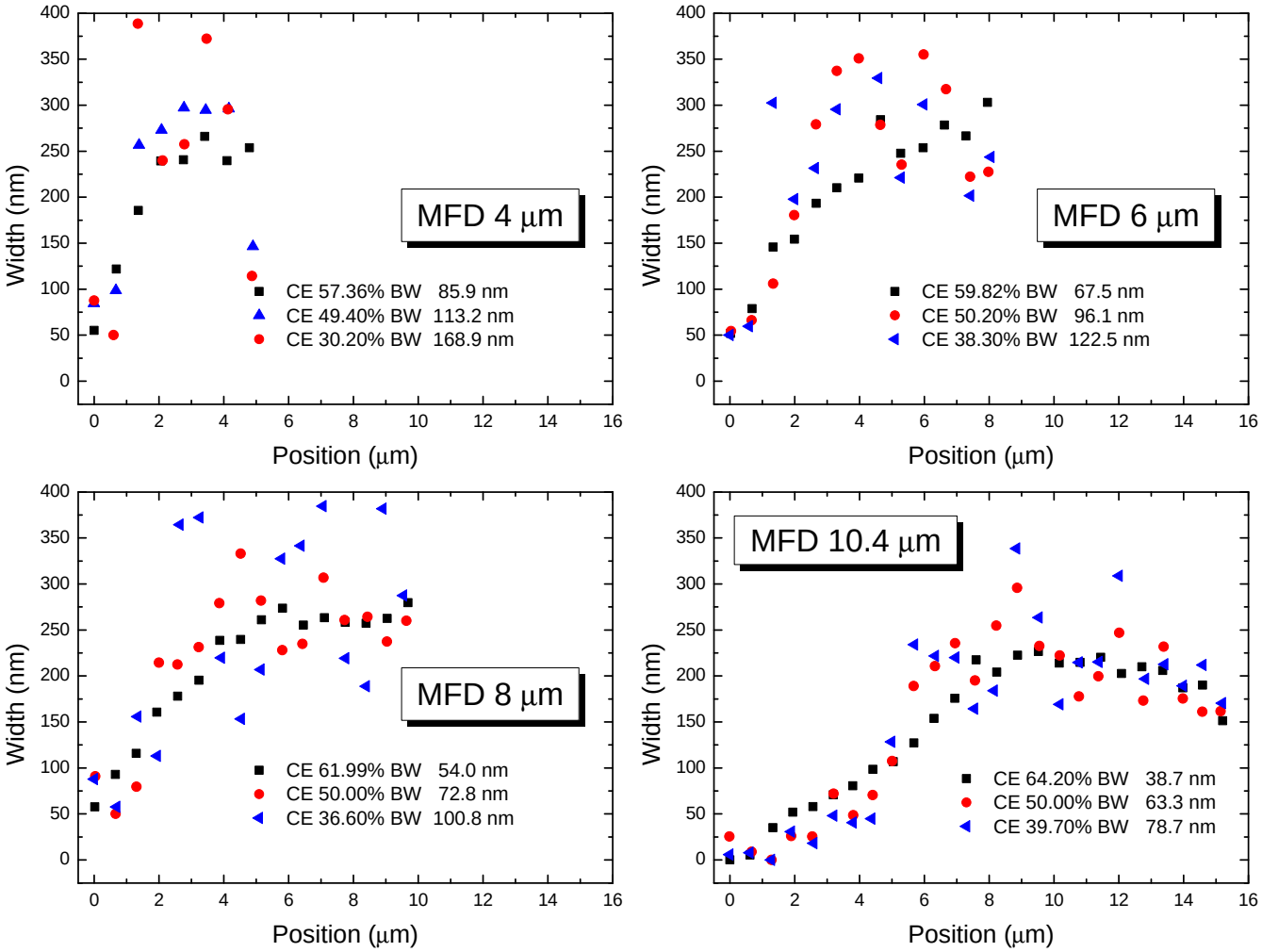
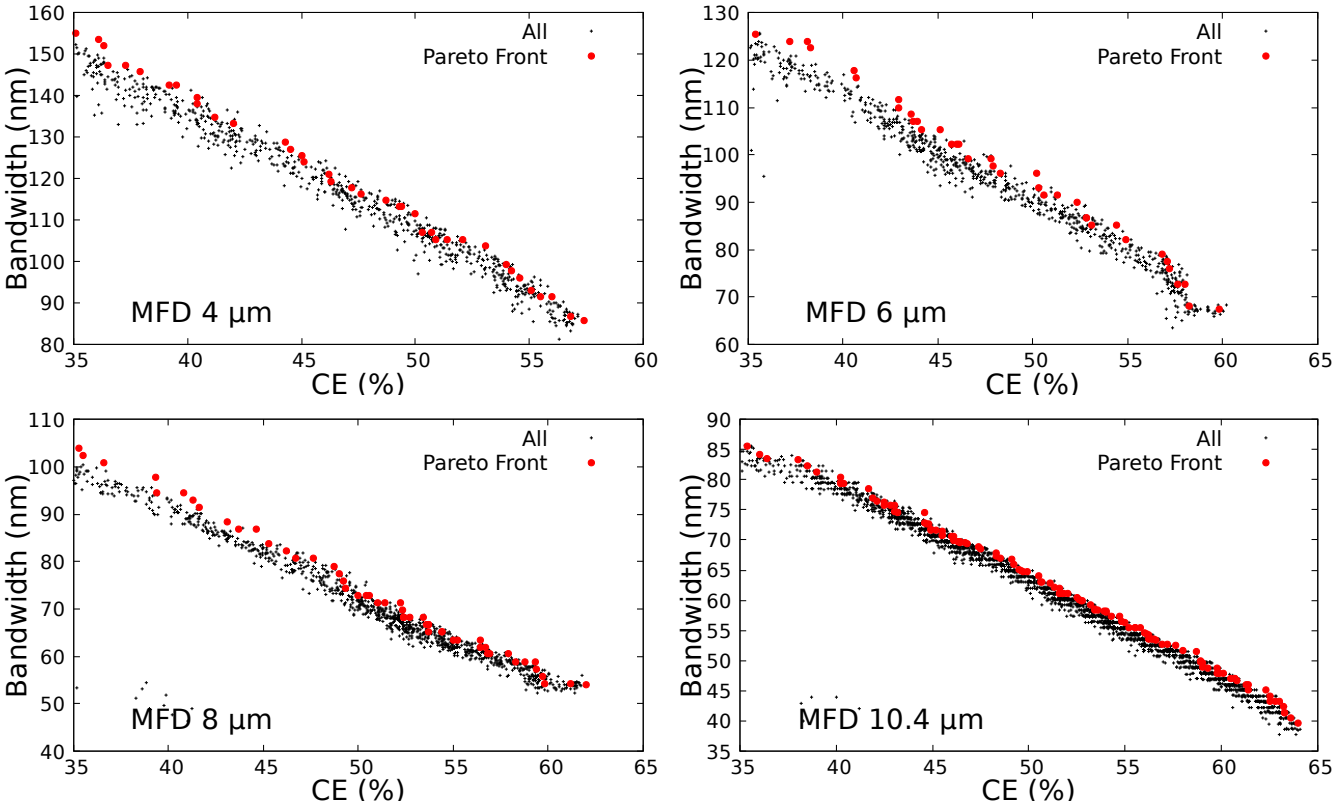


Figure	MFD (µm)	Structure n°
a	4	42
b	6	51
c	8	62
d	10.4	92

Supplementary Figure 4. Representation of CE and Bandwidth as a function of etching depth for four of the structures in the Pareto front. Each structure is identified in the table by its MFD and number, as reported in the .dat file containing the data of the structures. The dashed line in the graphs represents the etching depth of the original structure.



Supplementary Figure 5. Representation of some of the structures (groove's width vs groove's position) along each Pareto front. For each value of MFD we report 3 structures: the one with best coupling efficiency, a structure with high bandwidth and an intermediate one.



Supplementary Figure 6. Tolerance analysis with respect to position and width of the grooves. For each value of MFD the Pareto front (red circles) is reported, along with CE and Bandwidth of structures (20 for each of the points in the original Pareto front) obtained by random variation of the original structures (black cross points).

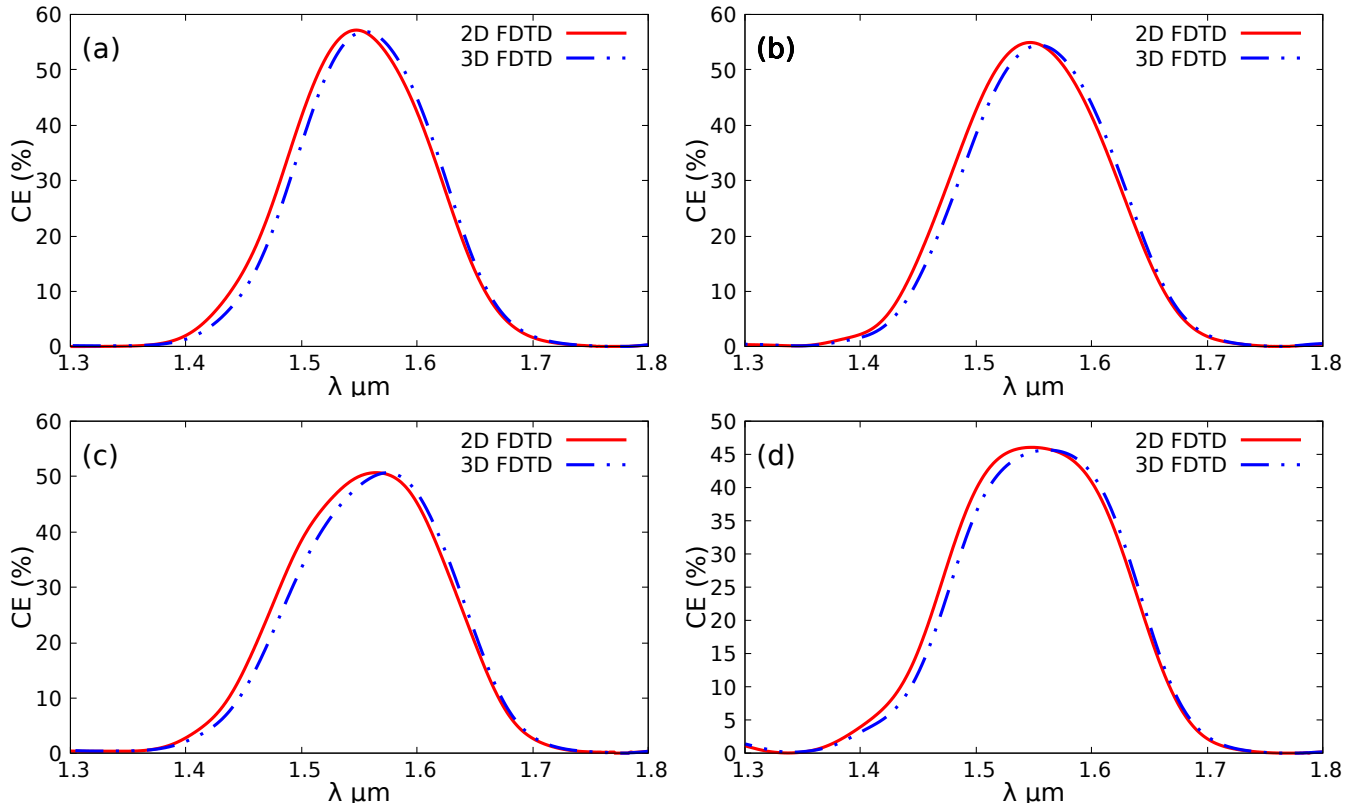


Figure	Structure n°	2D FDTD		3D FDTD	
		CE (%)	Bandwidth (nm)	CE (%)	Bandwidth (nm)
a	56	57.12	87.01	56.92	85.21
b	52	54.95	93.58	54.46	91.72
c	45	50.70	106.79	50.69	99.29
d	35	46.07	121.11	45.66	118.15

Supplementary Figure 7. Comparison between 2D and 3D FDTD simulations of four selected structures along the 4 µm Pareto front. A summary of the performances of these structures as calculated with the two methods is given in the table. The "Structure n°" column refers to the number of the structure as reported in supp_MFD_4.dat