Design methodology for compact photonic-crystal-based wavelength division multiplexers

Victor Liu,[1,*] Yang Jiao,[1,2] David A. B. Miller,[1] and Shanhui Fan[1]

[1]*Department of Electrical Engineering, Stanford University, Stanford, California, USA*

[2]*Currently with The D. E. Shaw Group, 120 West Forty-Fifth Street, 39th Floor, New York, New York 10036, USA*

**Corresponding author: vkl@stanford.edu*

Received December 7, 2010; revised January 16, 2011; accepted January 18, 2011;
posted January 26, 2011 (Doc. ID 139309); published February 15, 2011

We present an extremely compact wavelength division multiplexer design, as well as a general framework for designing and optimizing frequency selective devices embedded in photonic crystals satisfying arbitrary design constraints. Our method is based on the Dirichlet-to-Neumman simulation method and uses low rank updates to the system to efficiently scan through many device designs.

OCIS codes: 000.3860, 130.3120, 130.7408, 230.7408, 260.2110.

Photonic crystal (PhC)-based circuits provide many attractive features for on-chip manipulation of light. A key ingredient to realizing planar integrated PhC circuits is a compact wavelength division multiplexer (WDM) used for separating different wavelengths of light. Existing PhC WDM designs are based on resonators [1–5], waveguide couplers [6–8], or superprism effects [9–11]. Devices based on waveguide couplers or superprism effects tend to have large areas. Resonator-based devices are compact; however, the key issue is to suppress reflection [1].

We present here a systematic methodology for designing ultracompact WDM structures in PhCs. As an illustration, we construct a three-frequency WDM for the TM polarization with an out-of-plane electric field, in a PhC consisting of a square lattice of silicon rods with refractive index 3.4 and radius $0.18a$, where a is the lattice constant ($a = 600\,\mathrm{nm}$ at $\lambda = 1.55\,\mu\mathrm{m}$). To achieve the filter function, the structure consists of three output waveguides, each coupled to a high-Q resonator cavity tuned to a different specified frequency. These cavities are coupled to a common coupler region, which is also connected to the input waveguide. All these waveguides are single moded.

With temporal coupled mode theory, it was shown in [12] that, for this class of structure shown in Fig. 1, in principle, the common coupler region can be designed to act as an impedance transformer to eliminate reflection. The detailed design of such a coupler, however, is nontrivial. An optimized design, in which the coupler region consists of an aperiodic array of rods [highlighted yellow region, Fig. 1(a)], indeed functions as a WDM. At the three target frequencies of $\omega = 0.3857$, 0.3863, and $0.3868 \times 2\pi c/a$, incident light from the input waveguide is routed into the three output waveguides. At these three frequencies there is very little reflection in the input waveguide, as well as very little cross talk, defined to be the sum of outgoing power in any output waveguide other than the one specified. In contrast, the unoptimized design [Fig. 1(b)] does not work.

The optimized design in Fig. 1(a) is obtained with a combination of a fast numerical simulation technique based on the Dirichlet-to-Neumann (DtN) method [13,14] and systematic optimizations. We first briefly summarize the DtN method. The DtN method takes advantage of the many identical cells of a PhC system. For each unique type of unit cell, an operator Λ, called the DtN map, is computed, mapping the fields at the edge of the cell to their normal derivative. In practice, the fields are discretized with n points along the cell edges, so that, for a square unit cell, Λ is a $4n \times 4n$ matrix. For the results in this Letter, we used $n = 5$; simulating the same structure with $n = 7$ did not qualitatively change the results.

The DtN method stores only the fields at the edges of all unit cells. The system matrix $A(\omega)$, which describes the coupling of these fields, is constructed from the operators Λ by applying continuity of the normal derivatives over cell edges inside the computational region, and boundary conditions for the edges that lie at the boundary of the computational region. We use perfect electrical conductor (PEC) boundary conditions for edges along

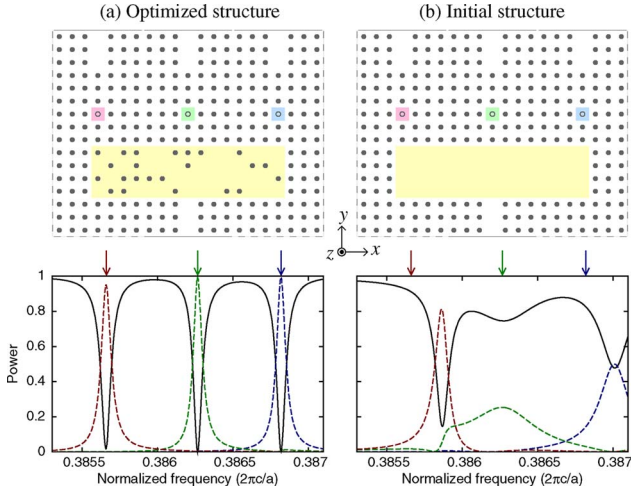

Fig. 1. (Color online) (a) Optimized WDM structure. Filled circles indicate dielectric rods. The coupler region is highlighted in yellow. The cavities are highlighted in different colors with hollow circles. The solid horizontal lines indicate the positions where the waveguide power flux is measured, while the dashed lines indicate PEC boundary conditions. (b) Initial unoptimized structure. The response spectra for each structure are shown below the respective structure. The solid black curves are the reflection spectra in the input waveguide. The colored dashed curves are the transmission spectra for each output waveguide, with the colors matching the color code of the cavity. The target frequencies are indicated by arrows.

the dashed lines in Fig. 1(a), where the field is expected to vanish. For boundary edges terminating a waveguide along the solid lines in Fig. 1(a), we impose an outgoing-wave boundary condition obtained by a decomposition into the complete outgoing waveguide mode basis [15]. The dimension N of the system matrix is determined by the total number of unknowns on all cell edges (approximately $2n$ unknowns per unit cell).

To calculate the transmission properties of the structure, we launch a waveguide mode in the input waveguide. The field distribution x is then determined by solving a linear system:

$$A(\omega)x(\omega) = b(\omega), \tag{1}$$

where $b(\omega)$ describes the input. The ω dependence is explicitly stated above to make clear that this is a single-frequency computation, and will be implied in the remainder of this Letter. We must solve one system for each output frequency of interest for the purposes of WDM design.

From the field distribution x in the computational domain, the output power in each waveguide is obtained by performing a decomposition of the field distribution in each waveguide into propagating waveguide modes and can be summarized for the ith waveguide as

$$p_i = x^H B_i x, \tag{2}$$

where x^H is the conjugate transpose of x and B_i is a positive definite Hermitian matrix computed from the complete outgoing mode basis vectors for the ith waveguide.

In the DtN method, the dimension of the matrix A is small enough to be inverted directly, even for structures with a few thousand unit cells. The ability to store A^{-1} itself is of crucial importance for optimization purposes. Such an ability enables efficient calculation of device sensitivity with respect to design parameters, which is required for gradient-based local optimization strategies. Within the DtN framework, the derivative A' of the system matrix A with respect to various parameters, such as the radius, position, and index of a rod, as well as frequency, may be obtained by direct differentiation of the formulas used to compute DtN maps. The derivative of the output power from the ith waveguide with respect to perturbations is then

$$\begin{aligned} p_i' &= x'^H B_i x + x^H B_i x' = 2\mathrm{Re}(x^H B_i x') \\ &= -2\mathrm{Re}(x^H B_i A^{-1} A' x), \end{aligned} \tag{3}$$

which is obtained by differentiating Eqs. (1) and (2).

The ability to store A^{-1} itself also leads to efficient solution of structures that are locally perturbed from a structure already solved. Because of the local nature of the DtN system matrix, a single cell perturbation results in a low rank (rank $4n$, in this case 20) perturbation of A, because only the coupling between the four edges adjacent to the perturbed cell is changed. The updated matrix equation becomes

$$(A + \delta A)x = b, \tag{4}$$

where δA is the perturbation, and can always be expressed as $\delta A = U \Delta V^T$ with Δ a rank $4n$ matrix whose elements are taken from the difference between the DtN maps of the perturbed and original cells. The matrices U and V are of dimension $N \times 4n$, and are zero except for a single $n \times n$ identity matrix in each block column determined by the location of the perturbed cell. By the matrix inversion lemma [16], the updated inverse is

$$(A + U\Delta V^T)^{-1} = A^{-1} - A^{-1}U(\Delta^{-1} + V^T A^{-1} U)^{-1} V^T A^{-1}, \tag{5}$$

requiring only two inversions of a small matrix. Using this formula is much more efficient than solving the entire system from scratch and is, therefore, particularly useful for optimization purposes.

We use simulated annealing to optimize the coupler region, considering only binary rod flips between empty unit cells and cells identical to the background PhC unit cell. The error metric is taken to be the sum of squares of reflection and cross talk for all frequencies. The optimization aims to achieve a low error metric by repeating the following steps:

First, we select a rod in the coupler region to flip by randomly choosing a rod in the coupler region, with preferential weighting toward rod flips that will reduce the error metric. We compute the expected change in the error metric with respect to the rod's index. If the sign of the change is favorable, we increase the probability of choosing a rod depending on the magnitude of the derivative, up to twice the probability of an unfavorable rod.

Second, the updated system is solved using Eq. (5) to update the stored full inverse matrix by a low rank adjustment. After perturbation of the coupler, the cavity center frequencies shift enough to detune the transmission peaks significantly from the target frequencies. In order to lock the frequencies of the transmission peak in each output waveguide to the target frequencies, we optimize the cavities to adjust their resonance frequencies. We use an adaptive step-size gradient descent method to simultaneously adjust the three-parameter space of rod indexes of each cavity. The solution of the updated system is again computed using the low rank adjustment and the gradients were computed using the analytic output power derivatives described previously. A maximum of 32 iterations were allowed because this step formed the bulk of the runtime due to the serial nature of the gradient descent algorithm.

Finally, we use a simulated annealing acceptance criterion to decide whether to keep the updated structure [17]. For modifications that reduce the error metric relative to the previous design, the updated structure was kept unconditionally. Otherwise, the acceptance probability was exponentially reduced throughout the optimization run.

In our design process, the three cavities were initially tuned to the output frequencies by adjusting the rod indexes (approximately 1, 1.02, and 1.04), and the coupler was initialized to an empty region without any rods. The transmission spectrum of the initial structure is shown in Fig. 1(b). The simulated annealing optimization was run on a 2.7 GHz AMD Opteron, with each frequency

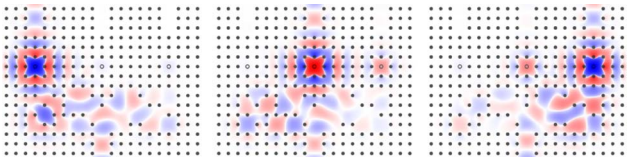

Fig. 2. (Color online) E_z field patterns at each of the three target frequencies.

simulation running on a separate core. After 48 h of runtime in which 73,000 candidate structures were considered, the optimal coupler region rod pattern is shown in Fig. 1(a). The error metric of the optimal structure is 2.6×10^{-4}, with the maximum error being 3.0. The transmission spectra of the device are shown in Fig. 1 with a maximum reflected power of 5%. Representative field patterns at each target frequency are shown in Fig. 2. The final design may be fine-tuned further by either enlarging the optimization region or using a gradient descent method considering additional design parameters, such as rod indexes or positions using the analytic derivative information.

Although the rod indexes of the cavities in the final design did not differ appreciably from the initial values, attempting to run the optimization without tuning the cavities along with the rod flipping did not converge to any working structure. It is clear that these adjustments are crucial to systematic WDM design.

In the simulated annealing process, complete convergence is, in fact, not necessary. We already see several performing structures with very low error metrics before full convergence in simulated annealing is reached. Repeated runs of the optimization procedure starting from the same initial blank configuration reliably yielded several working structures after 48 h; the design shown is simply a representative structure. The DtN method is similar to previous methods using a Wannier basis in that the field is expanded in a basis localized to a cell so that cell perturbations are low rank [18,19]. However, using the DtN method results in a more sparse system matrix A. Also, unlike the Wannier function method, here no additional difficulty arises in simulating an out-of-plane magnetic field. Comparing the DtN method to traditional full-field methods, the DtN method allows simulation of a single structure in approximately 1 s, while a comparable simulation using the finite-difference frequency-domain method would take several minutes.

We have presented an extremely compact design for a planar PhC WDM filter. We have also described a robust algorithm for optimizing such devices. Our method is generally applicable to any frequency selective device due to the ability to lock the spectral response of the device to specified frequencies. This method makes possible the automated design and optimization of large and complex photonic devices.

This work is supported in part by U.S. Air Force Office of Scientific Research (USAFOSR) grant No. FA9550-09-1-0704 and National Science Foundation (NSF) grant No. DMS-0968809. Victor Liu is supported by a Stanford Graduate Fellowship.

References

1. S. Fan, P. R. Villeneuve, J. D. Joannopoulos, and H. A. Haus, Phys. Rev. Lett. **80**, 960 (1998).
2. S. Kim, I. Park, H. Lim, and C.-S. Kee, Opt. Express **12**, 5518 (2004).
3. A. Sharkawy, S. Shi, and D. W. Prather, Appl. Opt. **40**, 2247 (2001).
4. B.-S. Song, T. Asano, Y. Akahane, Y. Tanaka, and S. Noda, J. Lightwave Technol. **23**, 1449 (2005).
5. C.-W. Kuo, C.-F. Chang, M.-H. Chen, S.-Y. Chen, and Y.-D. Wu, Opt. Express **15**, 198 (2007).
6. E. Centeno, B. Guizal, and D. Felbacq, J. Opt. A **1**, L10 (1999).
7. M. Koshiba, J. Lightwave Technol. **19**, 1970 (2001).
8. T. Niemi, L. H. Frandsen, K. K. Hede, A. Harpoth, P. I. Borel, and M. Kristensen, IEEE Photon. Technol. Lett. **18**, 226 (2006).
9. K. B. Chung and S. W. Hong, Appl. Phys. Lett. **81**, 1549 (2002).
10. M. Gerken and D. A. B. Miller, IEEE Photon. Technol. Lett. **15**, 1097 (2003).
11. A. Jugessur, L. Wu, A. Bakhtazad, A. Kirk, T. Krauss, and R. De La Rue, Opt. Express **14**, 1632 (2006).
12. C. Jin, S. Fan, S. Han, and D. Zhang, IEEE J. Quantum Electron. **39**, 160 (2003).
13. Y. Huang and Y. Y. Lu, J. Lightwave Technol. **24**, 3448 (2006).
14. Z. Hu and Y. Y. Lu, Opt. Express **16**, 17383 (2008).
15. Y. Huang, Y. Y. Lu, and S. Li, J. Opt. Soc. Am. B **24**, 2860 (2007).
16. R. A. Horn and C. R. Johnson, *Matrix Analysis* (Cambridge University, 1990).
17. S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, Science **220**, 671 (1983).
18. K. Busch, S. F. Mingaleev, A. Garcia-Martin, M. Schillinger, and D. Hermann, J. Phys. Condens. Matter **15**, R1233 (2003).
19. Y. Jiao, S. Fan, and D. A. B. Miller, Opt. Lett. **30**, 141 (2005).