Second order sensitivity analysis for a photonic crystal waveguide bend

Zhen Hu[1,*] and Ya Yan Lu[2]

[1] *Department of Mathematics, Hohai University, Nanjing, Jiangsu, China*

[2] *Department of Mathematics, City University of Hong Kong, Kowloon, Hong Kong*

compiled: November 5, 2014

Sensitivity analysis provides important information on the tolerance to fabrication imperfections for any designed devices. Standard sensitivity analysis relies on the first order partial derivatives of a response function of the device with respect to the design parameters. These derivatives are also useful in the device optimization process when a gradient based optimization method is used. The first order sensitivity analysis becomes inadequate if the first order derivatives are small or zero. In this paper, an efficient method for computing the second order partial derivatives is developed for idealized two-dimensional photonic crystal devices with circular cylinders, where the response function is the normalized power transmission coefficient in an output waveguide and the design parameters are the radii of the cylinders. Based on that, a second order sensitivity analysis is performed for a 90° photonic crystal waveguide bend at the frequency where the transmission coefficient is close to 1.

OCIS codes:

http://dx.doi.org/10.1364/XX.99.099999

1. Introduction

For practical applications of photonic crystal (PhC) devices [1–3], it is important to study the sensitivity of device properties with respect to design parameters [4]. Such a sensitivity analysis gives valuable information about fabrication tolerance and is also useful in the optimal design process [5–7]. For PhC devices, a typical device property could be the transmission coefficient for normalized power in an output waveguide, and the design parameters could be the radii, centers, refractive indices, shape variations of the dielectric rods or air-holes, etc.

The standard sensitivity analysis [4, 8] is based on the first order partial derivatives of a response function (such as the transmission coefficient) with respect to the design parameters. For design parameters in a small neighborhood of the design point, the response function can be approximated by its first order Taylor expansion. It is possible to perform a sensitivity analysis by computing the response function with each design parameter (alone) slightly varied away from the design point [5], and approximating the partial derivatives by difference formulae. If many design parameters are involved, such a direct approach is not efficient, since the response function must be repeatedly calculated when each design parameter is varied. The adjoint variable method [4, 9–11] is much more efficient, since it computes all the partial derivatives using only two field computations.

The first order sensitivity analysis becomes inadequate if the first order partial derivatives are zero or small. For example, if the transmission coefficient of a linear device is close to 1 at some frequency, all its partial derivatives with respect to the design parameters must be nearly zero, since 1 is the absolute maximum. Notice that even when the first order partial derivatives are zero, the device may still be quite sensitive to the design parameters. This happens in devices that exhibit Fano resonances, where the transmission spectrum can change rapidly between a maximum and a minimum. In any case, the response function can be more accurately approximated if second order partial derivatives are available. Second order derivatives can also be used in nonlinear programming algorithms [12] for optimal design of photonic devices.

In this paper, we develop an efficient method for second order sensitivity analysis of idealized two-dimensional (2D) PhC devices with circular inclusions (dielectric rods or air-holes) in square or triangular lattices. To simplify the presentation, we concentrate on a specific example, namely, the 90° PhC waveguide bend proposed by Mekis *et al.* [3]. Our method is based on a frequency-domain computational method (the so-called DtN-map method) that is particularly efficient for analyzing 2D PhC devices with circular inclusions [13, 14]. The DtN-map method relies on the Dirichlet-to-Neumann (DtN) maps of the unit cells to establish a small system equation on the edges of the unit cells only. For PhC devices with circular inclusions, the DtN maps

* Corresponding author: huzhen1230@gmail.com

are approximated by small matrices based on cylindrical wave expansions. Our sensitivity analysis takes advantage of the analytic solutions, i.e. the cylindrical waves.

2. DtN-map method

In order to present our method for second order sensitivity analysis, we include a brief summary for the DtN-map method below and consider the 90° PhC waveguide bend proposed by Mekis *et al.* [3] and shown in Fig. 1. The background PhC is an infinite square lattice of di-

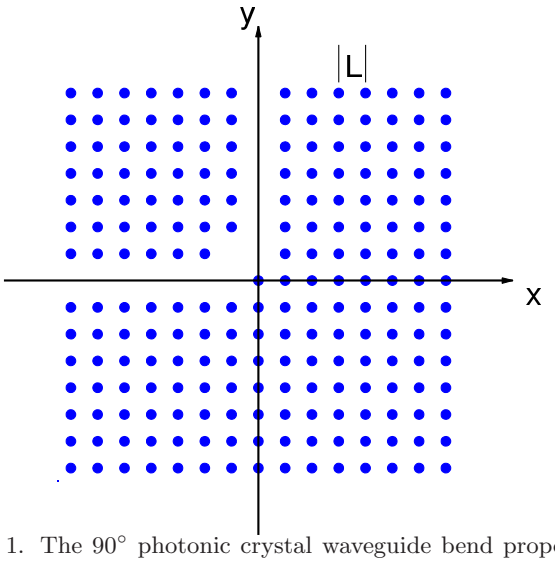

Fig. 1. The 90° photonic crystal waveguide bend proposed by Mekis *et al.* [3].

electric rods surrounded by air, where the lattice constant is L, the radius and refractive index of the rods are $a = 0.18L$ and $n = 3.4$, respectively. For the E polarization, the electric field is parallel to the axes of the rods, and it satisfies the following Helmholtz equation

$$\partial_x^2 u + \partial_y^2 u + k_0^2 n^2(x,y)\, u = 0, \tag{1}$$

where k_0 is the free space wavenumber, $n(x,y)$ is the refractive index function, $\{x, y, z\}$ is a Cartesian coordinate system, z is the axis parallel to the dielectric rods, and u is the z component of the electric field. For this polarization, the PhC has a bandgap given by $0.302 < \omega L/(2\pi c) < 0.443$ [3], where ω is the angular frequency and c is the speed of light in vacuum. As shown in Fig. 1, the 90° bend connects two semi-infinite PhC waveguides along the negative x (horizontal) and positive y (vertical) directions. The PhC waveguide is formed by removing a row of rods, and it has a single propagating mode for $0.312 < \omega L/(2\pi c) < 0.443$ [3].

In the horizontal waveguide, an incident wave is specified. It is a propagating mode propagating towards the bend, and it can be written as

$$u^{(i)}(x,y) = \psi_1(x,y) e^{\mathbf{i}\beta_1 x}, \tag{2}$$

where ψ_1 is periodic in x with period L and decays to zero as $y \to \pm\infty$. The incident wave gives rise to a reflected wave and a transmitted wave in the horizontal and vertical waveguides, respectively. The transmitted wave can be expanded in the Bloch modes of the PhC waveguide as

$$u^{(t)}(x,y) = \sum_{j=1}^{\infty} C_j \psi_j(y,-x) e^{\mathbf{i}\beta_j y}, \tag{3}$$

where C_j is the coefficient of the jth Bloch mode. Notice that we use $\psi_j(x,y)$ to denote a Bloch mode in the horizontal waveguide, then $\psi_j(y,-x)$ is a Bloch mode in the vertical waveguides. Since the PhC waveguide has only one propagating mode (i.e., β_1 is real and $\mathrm{Im}(\beta_j) > 0$ for all $j > 1$), the transmission coefficient (the power carried by the transmitted wave, normalized by the power of the incident wave) is simply $T = |C_1|^2$.

Using the DtN-map method [13], the original Helmholtz equation is replaced by the following linear system

$$\mathbf{A}\boldsymbol{u} = \boldsymbol{f}, \tag{4}$$

where $\boldsymbol{u}$ is a column vector for u on the edges of the unit cells in a truncated domain. The truncated domain is a square with $M \times M$ unit cells around the bend. The case for $M = 11$ is shown in Fig. 2. Each edge of a unit cell is

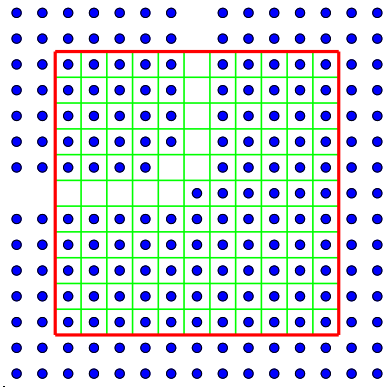

Fig. 2. Truncated domain and square unit cells for analyzing the 90° PhC waveguide bend by the DtN-map method.

discretized by N points, where a typical value is $N = 5$. The boundary condition on the lower and right sides of the truncated domain is simply $u = 0$, since these sides are sufficiently far away from the waveguide core. The boundary conditions on the top and left sides are rigorous nonlocal conditions established based on expanding the transmitted and reflected waves in Bloch modes of the PhC waveguides [13]. It is also possible to use approximate and simpler boundary conditions on the top and left sides [15], but the truncated domain must be slightly increased. The boundary condition on the left side is inhomogeneous, since there is an incident wave in the horizontal waveguide, and it gives rise to the vector $\boldsymbol{f}$ in Eq. (4). Notice that $\boldsymbol{u}$ is a vector of length $2NM^2$, since it involves only $2N$ points for each unit cell.

The DtN-map method relies on the DtN maps of the unit cells. Let Ω_i be a square unit cell with a circular cylinder of radius a_i at its center, its DtN map is an operator $\mathbf{\Lambda}_i$ satisfying

$$\mathbf{\Lambda}_i\, u = \partial_{\boldsymbol{\nu}} u \quad \text{on} \quad \partial\Omega_i \tag{5}$$

for u satisfying Eq. (1) in Ω_i, where $\partial\Omega_i$ denotes the boundary of Ω_i and $\partial_{\boldsymbol{\nu}}$ denotes the normal derivative operator on $\partial\Omega_i$ (we take it to be ∂_x and ∂_y on the vertical or horizontal edges of Ω_i, respectively). If each edge of Ω_i is discretized by N points, then $\mathbf{\Lambda}_i$ is approximated by a $(4N)\times(4N)$ matrix. Equation (4) is established by matching $\partial_{\boldsymbol{\nu}} u$ on each edge of the unit cells in the truncated domain. If Γ_{ij} is the common edge of two neighboring unit cells Ω_i and Ω_j, we can evaluate $\partial_{\boldsymbol{\nu}} u$ on Γ_{ij} by $\mathbf{\Lambda}_i$ and $\mathbf{\Lambda}_j$. If an edge of a unit cell Ω_i is on the left or top sides of the truncated domain, we can evaluate $\partial_{\boldsymbol{\nu}} u$ on that edge by $\mathbf{\Lambda}_i$ and by the nonlocal boundary condition. In any case, $\partial_{\boldsymbol{\nu}} u$ evaluated in the two different approaches must equal to each other.

The DtN-map method can be implemented semi-analytically or fully numerically, depending on how the DtN maps are approximated by matrices. For unit cells with a general index profile, the finite element method and the boundary integral equation method may be used [16]. But for simple square or hexagon unit cells with circular inclusions, the analytic method based on cylindrical wave expansions is the most efficient [17, 18]. In that case, the solution of Eq. (1) in Ω_i is approximated by a linear combination of $4N$ analytic solutions:

$$u(\boldsymbol{r}) \approx \sum_{l=1}^{4N} c_l \Phi_l(\boldsymbol{r}) \quad \text{in} \quad \Omega_i, \tag{6}$$

where $\boldsymbol{r}=(x,y)$, Φ_l is an exact solution (a cylindrical wave) given in Appendix, and c_l is an arbitrary coefficient. Let $\boldsymbol{r}_j$ for $1\le j\le 4N$, be $4N$ discretization points on $\partial\Omega_i$ (with N points on each edge equally spaced), then

$$u(\boldsymbol{r}_j) \approx \sum_{l=1}^{4N} c_l \Phi_l(\boldsymbol{r}_j), \tag{7}$$

$$\frac{\partial u}{\partial \boldsymbol{\nu}}(\boldsymbol{r}_j) \approx \sum_{l=1}^{4N} c_l \boldsymbol{\nu}(\boldsymbol{r}_j)\cdot\nabla\Phi_l(\boldsymbol{r}_j), \tag{8}$$

for $1\le j\le 4N$. If we introduce two matrices $\mathbf{D}$ and $\mathbf{N}$ with their (j,l) entries given by $\Phi_l(\boldsymbol{r}_j)$ and $\boldsymbol{\nu}(\boldsymbol{r}_j)\cdot\nabla\Phi_l(\boldsymbol{r}_j)$, respectively, then $\vec{u}\approx\mathbf{D}\,\vec{c}$, $\partial_{\boldsymbol{\nu}}\vec{u}\approx\mathbf{N}\,\vec{c}$, and

$$\partial_{\boldsymbol{\nu}}\vec{u} \approx \mathbf{N}\mathbf{D}^{-1}\vec{u}, \tag{9}$$

where $\vec{u}$ and $\partial_{\boldsymbol{\nu}}\vec{u}$ are column vectors of $u(\boldsymbol{r}_j)$ and $\partial_{\boldsymbol{\nu}}u(\boldsymbol{r}_j)$ for $1\le j\le 4N$, and $\vec{c}$ is a column vector of c_l for $1\le l\le 4N$. The matrix $\mathbf{N}\mathbf{D}^{-1}$ is an approximation to the DtN map $\mathbf{\Lambda}_i$.

Once $\boldsymbol{u}$ is calculated, we can use its values on the top side of the truncated domain to obtain a numerical version of the Bloch mode expansion (3). This leads to a vector $\boldsymbol{d}$ with only NM nonzero elements corresponding to the discretization points on the top side of the truncated domain, such that

$$C_1 = \boldsymbol{d}^{\mathsf{T}}\boldsymbol{u}, \tag{10}$$

where the superscript T denotes the transpose operation.

Using the DtN-map method, the transmission and reflection spectra of the 90° bend are calculated and shown

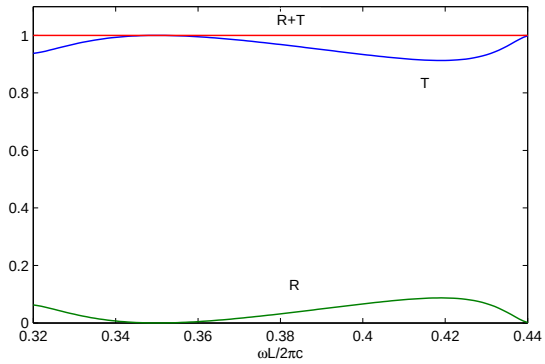

Fig. 3. Transmission (T) and reflection (R) spectra of the 90° PhC waveguide bend proposed by Mekis *et al.* [3].

in Fig. 3. These results are obtained with $M=11$ and $N=5$.

3. First order derivatives

The normalized power transmission coefficient T depends on the frequency ω, as well as other physical and geometric parameters. A sensitivity analysis is to analyze how T varies with design parameters, such as the refractive indices, radii, centers of the cylinders in each unit cell. In the following, we assume the refractive indices of the cylinders and the surrounding medium are fixed, the cylinders are strictly circular and exactly centered on lattice points of a square lattice, and consider only the sensitivity with respect to the radii of the cylinders. While the original 90° bend involves either an empty unit cell with no rod or a regular unit cell with a circular rod of radius a, where a is fixed at $0.18L$, we now consider a more general structure for which the ith unit cell Ω_i contains a circular rod of radius a_i. Consequently, we have $T=T(\omega;a_1,a_2,a_3,...)$, and our objective is to calculate the first order partial derivatives $\partial_{a_i}T$.

Since $T=|C_1|^2$, we have $\partial_{a_i}T = 2\mathrm{Re}(\overline{C}_1\,\partial_{a_i}C_1)$, where $\overline{C}_1$ is the complex conjugate of C_1. Therefore, we can concentrate on computing the partial derivatives of C_1. In Eq. (10), the vector $\boldsymbol{d}$ is only related to the Bloch modes of the PhC waveguide (it is independent of a_i), thus

$$\frac{\partial C_1}{\partial a_i} = \boldsymbol{d}^{\mathsf{T}}\frac{\partial \boldsymbol{u}}{\partial a_i}. \tag{11}$$

Similarly, the vector $\boldsymbol{f}$ in Eq. (4) is independent of a_i.

Therefore,

$$\mathbf{A}\frac{\partial \boldsymbol{u}}{\partial a_i} = -\frac{\partial \mathbf{A}}{\partial a_i}\boldsymbol{u}. \tag{12}$$

In principle, we can solve $\partial_{a_i}\boldsymbol{u}$ from Eq. (12) and then evaluate $\partial_{a_i}C_1$ using Eq. (11). If there are many design parameters, this approach could be expensive. The key observation of the adjoint variable method [4] is that if $\boldsymbol{g}$ satisfies

$$\mathbf{A}^{\mathsf{T}}\boldsymbol{g} = \boldsymbol{d}, \tag{13}$$

then

$$\frac{\partial C_1}{\partial a_i} = -\boldsymbol{g}^{\mathsf{T}}\frac{\partial \mathbf{A}}{\partial a_i}\boldsymbol{u}. \tag{14}$$

Consequently, it is only necessary to solve one equation for $\mathbf{g}$, then all partial derivatives are easily evaluated from (14).

Since the system matrix $\mathbf{A}$ is constructed from the DtN maps of the unit cells and the boundary conditions for terminating the waveguides, $\partial_{a_i}\mathbf{A}$ depends on $\partial_{a_i}\mathbf{\Lambda}_i$. Since $\mathbf{\Lambda}_i = \mathbf{N}\mathbf{D}^{-1}$, we take the partial derivative with respect to a_i for both sides $\mathbf{\Lambda}_i\mathbf{D} = \mathbf{N}$, and get

$$\frac{\partial \mathbf{\Lambda}_i}{\partial a_i} = \left(\frac{\partial \mathbf{N}}{\partial a_i} - \mathbf{\Lambda}_i\frac{\partial \mathbf{D}}{\partial a_i}\right)\mathbf{D}^{-1}. \tag{15}$$

The entries of $\mathbf{D}$ and $\mathbf{N}$ are related to the special solutions Φ_l $(1 \le l \le 4N)$ of the Helmholtz equation (1) in the unit cell Ω_i. Clearly, Φ_l depends on the radius a_i of the rod. If we write Φ_l as $\Phi_l(\boldsymbol{r}; a_i)$, then the (j,l) entries of $\mathbf{D}$ and $\mathbf{N}$ are

$$\mathbf{D}_{jl} = \Phi_l(\boldsymbol{r}_j; a_i), \quad \mathbf{N}_{jl} = \boldsymbol{\nu}(\boldsymbol{r}_j)\cdot\nabla\Phi_l(\boldsymbol{r}_j; a_i), \tag{16}$$

where $\boldsymbol{r}_j$ (for $1 \le j \le 4N$) are the discretization points on four edges of Ω_i, $\boldsymbol{\nu}(\boldsymbol{r}_j)$ is a unit normal vector at $\boldsymbol{r}_j$. The partial derivatives of $\mathbf{D}_{jl}$ and $\mathbf{N}_{jl}$ with respect to a_i can be explicitly evaluated and they are given in Appendix.

4. Second order derivatives

For second order sensitivity analysis, we need to calculate the second order partial derivatives of T with respect to a_i and a_j, where a_i and a_j are the radii of rods in unit cells Ω_i and Ω_j. From $T = |C_1|^2$, we have

$$\frac{\partial^2 T}{\partial a_i \partial a_j} = 2\,\mathrm{Re}\left[\overline{C}_1\frac{\partial^2 C_1}{\partial a_i \partial a_j} + \frac{\partial C_1}{\partial a_i}\overline{\frac{\partial C_1}{\partial a_j}}\right]. \tag{17}$$

Since we already know how to evaluate the first order derivatives, we concentrate on $\partial^2_{a_i,a_j}C_1$ which can be written as

$$\frac{\partial^2 C_1}{\partial a_i \partial a_j} = \boldsymbol{d}^{\mathsf{T}}\frac{\partial^2 \boldsymbol{u}}{\partial a_i \partial a_j}. \tag{18}$$

Taking the partial derivative with respect to a_j for Eq. (12), we obtain

$$\mathbf{A}\frac{\partial^2 \boldsymbol{u}}{\partial a_i \partial a_j} = -\frac{\partial^2 \mathbf{A}}{\partial a_i \partial a_j}\boldsymbol{u} - \frac{\partial \mathbf{A}}{\partial a_i}\frac{\partial \boldsymbol{u}}{\partial a_j} - \frac{\partial \mathbf{A}}{\partial a_j}\frac{\partial \boldsymbol{u}}{\partial a_i}. \tag{19}$$

Multiplying $\mathbf{A}^{-1}$ and $\boldsymbol{d}^{\mathsf{T}}$ to the above, we have

$$\frac{\partial^2 C_1}{\partial a_i \partial a_j} = -\boldsymbol{g}^{\mathsf{T}}\frac{\partial^2 \mathbf{A}}{\partial a_i \partial a_j}\boldsymbol{u} - \boldsymbol{g}^{\mathsf{T}}\frac{\partial \mathbf{A}}{\partial a_i}\frac{\partial \boldsymbol{u}}{\partial a_j} - \boldsymbol{g}^{\mathsf{T}}\frac{\partial \mathbf{A}}{\partial a_j}\frac{\partial \boldsymbol{u}}{\partial a_i}. \tag{20}$$

Unlike the adjoint variable method for first order sensitivity analysis, it appears impossible to avoid solving the linear systems for $\partial_{a_i}\boldsymbol{u}$. We use the following steps to evaluate the first and second order partial derivatives of C_1.

1. Solve the original system (4) and the adjoint system (13) for $\boldsymbol{u}$ and $\boldsymbol{g}$.
2. For each i, evaluate $\partial_{a_i}\mathbf{A}$, and
$$\boldsymbol{h}_i^{\mathsf{T}} = \boldsymbol{g}^{\mathsf{T}}\frac{\partial \mathbf{A}}{\partial a_i} \quad \boldsymbol{p}_i = \frac{\partial \mathbf{A}}{\partial a_i}\boldsymbol{u}, \quad \frac{\partial C_1}{\partial a_i} = -\boldsymbol{h}_i^{\mathsf{T}}\boldsymbol{u}.$$
3. Solve $\partial_{a_i}\boldsymbol{u}$ for all i from the following system with multi-column right hand side
$$\mathbf{A}\left[\frac{\partial \boldsymbol{u}}{\partial a_1}, \frac{\partial \boldsymbol{u}}{\partial a_2}, ...\right] = -[\boldsymbol{p}_1, \boldsymbol{p}_2, ...].$$
4. For each pair (i,j), evaluate $\partial^2_{a_i,a_j}\mathbf{A}$, and
$$\frac{\partial^2 C_1}{\partial a_i \partial a_j} = -\boldsymbol{g}^{\mathsf{T}}\frac{\partial^2 \mathbf{A}}{\partial a_i \partial a_j}\boldsymbol{u} - \boldsymbol{h}_i^{\mathsf{T}}\frac{\partial \boldsymbol{u}}{\partial a_j} - \boldsymbol{h}_j^{\mathsf{T}}\frac{\partial \boldsymbol{u}}{\partial a_i}.$$

The second order partial derivatives of the system matrix $\mathbf{A}$ appear in the above equations. For the case $i \ne j$, if the two unit cells Ω_i and Ω_j are neighbors, $\partial^2_{a_i,a_j}\mathbf{A}$ is directly related to $\partial_{a_i}\mathbf{\Lambda}_i$ and $\partial_{a_j}\mathbf{\Lambda}_j$, otherwise $\partial^2_{a_i,a_j}\mathbf{A}$ is the zero matrix, because $\mathbf{A}$ is constructed by matching the normal derivative of u on common edges of neighboring unit cells. If $i = j$, $\partial^2_{a_i}\mathbf{A}$ is related to $\partial^2_{a_i}\mathbf{\Lambda}_i$. Taking the second order derivatives for the equation $\mathbf{\Lambda}_i\mathbf{D} = \mathbf{N}$, we get

$$\frac{\partial^2 \mathbf{\Lambda}_i}{\partial a_i^2} = \left(\frac{\partial^2 \mathbf{N}}{\partial a_i^2} - \mathbf{\Lambda}_i\frac{\partial^2 \mathbf{D}}{\partial a_i^2} - 2\frac{\partial \mathbf{\Lambda}_i}{\partial a_i}\frac{\partial \mathbf{D}}{\partial a_i}\right)\mathbf{D}^{-1}. \tag{21}$$

Therefore, we need to evaluate the matrices $\partial^2_{a_i}\mathbf{D}$ and $\partial^2_{a_i}\mathbf{N}$. Since the entries of $\mathbf{D}$ and $\mathbf{N}$ are known explicitly, the second order derivatives can be evaluated explicitly and they are given in Appendix.

5. The 90° bend

In this section, we perform a second order sensitivity analysis for the 90° PhC waveguide bend described in Section 2. From the transmission spectrum shown in Fig. 3, it is clear that $T \approx 1$ for $\omega L/(2\pi c) = 0.35$. For this frequency, the first order derivatives are small and the second order derivatives are important.

Instead of analyzing the sensitivity to all rods, we consider ten rods near the center of the waveguide bend as shown in Fig. 4, since they are the most important ones. Let a_i for $1 \le i \le 10$, be the radii of these

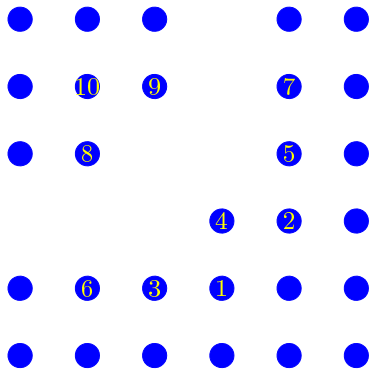

Fig. 4. Ten rods near the center of the 90° PhC waveguide bend.

Rod No. i	$\partial_{a_i} T$ in L^{-1}	$\partial^2_{a_i} T$ in L^{-2}
1	0.000031	-0.0150
2	0.000030	-0.0150
3	0.005782	-7.2801
4	-0.003039	-1.7304
5	0.005800	-7.2793
6	0.000271	-10.489
7	0.000275	-10.489
8	0.004195	-3.6595
9	0.004200	-3.6592
10	-0.000860	-0.1422

Table 1. The first and second order partial derivatives of transmission coefficient T with respect to the radii of ten selected rods at $\omega L/(2\pi c) = 0.35$.

ten rods, we calculate the first and second order partial derivatives $\partial_{a_i} T$ and $\partial^2_{a_i} T$ for the normalized frequency $\omega L/(2\pi c) = 0.35$. The results at $a_i = a = 0.18L$ (i.e. the design point) are listed in Table 1. The first and second order derivatives are given in scaled units $1/L$ and $1/L^2$, respectively. It can be seen that the first order derivatives are close to zero. Thus, the second order derivatives are needed in the sensitivity analysis. Notice that T is particularly sensitive to the 3rd and 5th rods, since the first order derivatives with respect to a_3 and a_5 have the largest magnitudes. The second order mixed derivatives are also available, but not listed.

Based on the partial derivatives above, we can predict the transmission coefficient $T(\omega; a_1, a_2, ..., a_{10})$ near the design point by Taylor expansions. We consider the special case where the radius of only one rod (the 3rd rod in Fig. 4) is varied. In Table 2, we compare the "exact" value of $T(\omega; a, a, a_3, a, ..., a)$ with its Taylor expansions truncated to first and second order derivatives. Notice that the Taylor approximation based only on the first order derivative gives some T values larger than 1 and it is physically impossible. On the other hand, the Taylor approximation based on both first and second

$(a_3 - a)/L$	Taylor I	Taylor II	Exact value
0.002	1.000009	0.999994	0.999995
0.005	1.000026	0.999935	0.999940
0.008	1.000044	0.999811	0.999831
0.010	1.000055	0.999691	0.999729
0.050	1.00029	0.99118	0.99250

Table 2. Transmission coefficients obtained by first and second order Taylor expansions and the DtN-map method, for a few values of a_3.

order derivatives gives accurate results when $|a_3 - a|$ is small. Hence, this is an additional good reason for using a second order sensitivity analysis. The "exact" value is calculated using the DtN-map method for the given a_3 directly.

Based on the first and second order partial derivatives of T, we can find the approximate probability distribution of T if the radii are regarded as random variables. For simplicity, we only consider the effect of varying the radius of one rod a_i. To simplify the notations, we denote the transmission coefficient as $T(a_i)$ (i.e., hide its dependence on all other parameters), and denote $\partial_{a_i} T$ and $\partial^2_{a_i} T$ by T' and T'', respectively. If a_i is a random variable with mean a and standard deviation σ, then the second order Taylor expansion gives

$$\begin{aligned} T(a_i) &\approx T(a) + T'(a)X + \frac{1}{2}T''(a)X^2 \\ &= \frac{T''(a)}{2}\left[X + \frac{T'(a)}{T''(a)}\right]^2 + T(a) - \frac{[T'(a)]^2}{2T''(a)}, \end{aligned}$$

where $X = a_i - a \sim N(0, \sigma^2)$ is a random variable following the normal distribution with zero mean and standard deviation σ. Clearly, $Y = [X + T'(a)/T''(a)]/\sigma$ is a normal distribution with mean $T'(a)/[\sigma T''(a)]$ and unit standard derivation. According to the definition of non-central χ^2 distribution, Y^2 is a non-central χ^2 distribution with degrees of freedom 1 and non-centrality $\lambda = \{T'(a)/[\sigma T''(a)]\}^2$, i.e., $\chi^2(1, \lambda)$. Therefore,

$$T(a_i) \sim \frac{\sigma^2 T''(a)}{2}\chi^2(1, \lambda) + T(a) - \frac{[T'(a)]^2}{2T''(a)}. \qquad (22)$$

The probability density function of $\chi^2(1, \lambda)$ is

$$f(z; 1, \lambda) = \frac{e^{-(z+\lambda)/2}}{2}\left(\frac{z}{\lambda}\right)^{-1/4} I_{-1/2}(\sqrt{\lambda z}), \qquad (23)$$

where $I_{-1/2}$ is a modified Bessel function of the first kind.

To test the validity of Eq. (22), we consider the 90° waveguide bend with a random perturbation on the radius of the 3rd rod. For the normalized frequency $\omega L/(2\pi c) = 0.35$, we have $T(a) = 0.99999741$, $T'(a) = 0.005782L^{-1}$ and $T''(a) = -7.2801L^{-2}$, where T' and T'' are derivatives with respect to a_3 and evaluated at $a_3 = a = 0.18L$. Assuming a_3 follows a normal distribution with mean a and standard deviation $\sigma = 0.006L$,

then the non-centrality parameter is $\lambda = 0.01752$. Notice that σ corresponds to a 3.3% relative error in the rod radius. If $L = 542.5\,\text{nm}$, then the wavelength is $1.55\,\mu\text{m}$, the radius a is 97.7 nm, and $\sigma = 3.26\,\text{nm}$. We take 20,000 random samples of a_3, calculate the transmission coefficient $T(a_3)$ by the DtN-map method, and estimate the distribution of

$$Z = \frac{2}{\sigma^2 T''(a)} \left\{ T(a_3) - T(a) + \frac{[T'(a)]^2}{2T''(a)} \right\}, \qquad (24)$$

and compare it with the probability density function of $\chi^2(1,\lambda)$ in Fig. 5. For each simulation, the main work is

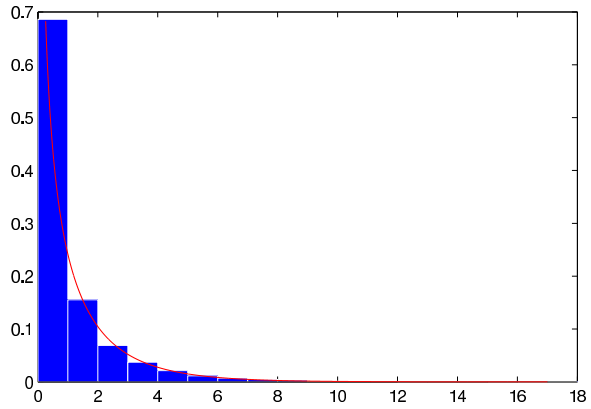

Fig. 5. Red curve: the probability density function of $\chi^2(1,\lambda)$ for $\lambda = 0.01752$, i.e., $f(z;1,\lambda)$ given in Eq. (23). Blue bars: estimated distribution (i.e., normalized histogram) of Z given in Eq. (24) based on 20,000 samples.

to solve the linear system (4), where $\mathbf{A}$ is a 1210×1210 matrix and it is somewhat sparse. To solve 20,000 such linear systems, the time required on a current iMac is about 45 minutes. However, this can be significantly improved. Let $\mathbf{A}_0$ be the matrix of the unperturbed structure, and $\mathbf{A}$ be the matrix where one unit cell is changed, then $\mathbf{A} = \mathbf{A}_0 + \mathbf{E}$, where $\mathbf{E}$ is a low rank matrix. After $\mathbf{A}_0^{-1}$ is calculated, we can use the Sherman-Morrison-Woodbury formula to quickly evaluate $\mathbf{A}^{-1}\boldsymbol{f}$. The red curve in Fig. 5 is just the probability density function $f(z;1,\lambda)$ for λ given above. The blue bars show the normalized histogram of 20,000 samples of Z. We observe that excellent agreement is obtained. This indicates that the distribution of $T(a_i)$ given by Eq. (22) is correct.

6. Conclusion

In this paper, an efficient second order sensitivity analysis technique for 2D PhC devices is developed and demonstrated using a 90° PhC waveguide bend. A second order sensitivity analysis requires the computation of first and second order partial derivatives of a response function with respect to design parameters. For the waveguide bend, the response function is the normalized power transmission coefficient, and the design parameters are chosen as the radii of dielectric rods around the center of the bend. Standard sensitivity analysis involving only the first order derivatives may have limitations near extreme points of the response function where the first order derivatives are nearly zero. The second order theory produces more accurate predictions when the structure is perturbed, avoids unphysical results such as $T > 1$ (transmitted power is larger than the incident power), and gives rise to well-known probability distributions for the response function if the design parameters are assumed to be random variables. While the standard first order sensitivity analysis is sufficient in most cases, the second order sensitivity analysis should be highly relevant in resonant structures where the extreme points of the response function are more sensitive to structural variations.

The method presented in this paper is based on the DtN-map method for 2D PhC devices [13]. The partial derivatives of the response function are related to the partial derivatives of the system matrix $\mathbf{A}$, and are further related to the partial derivatives of the DtN maps of the unit cells. For PhC devices with circular inclusions, the DtN maps are usually constructed from cylindrical wave expansions, and their derivatives can be calculated analytically. For PhC devices with arbitrary unit cells, the general procedure is still applicable, but a fully numerical method for computing the derivatives of the DtN maps is needed. In this work, we assume the radii of rods are the design parameters, but our method can be extended to study the sensitivity of PhC devices with respect to other parameters, such as the centers of the rods and distortions of the circular shape.

Acknowledgment

This work was partially supported by the National Natural Science Foundation of China under Project 11101122, and City University of Hong Kong Project 7004068.

Appendix

Consider a square unit cell Ω_i containing a circular rod of radius a_i with the center of the rod located at the origin. The unit cell Ω_i is given by $-L/2 < x, y < L/2$, where L is its side length. The rod is given by $r < a_i$, where (r,θ) is the polar coordinate system. In Eq. (6), a general solution of Helmholtz equation (1) is approximated by a sum of $4N$ special solutions $\Phi_l(\boldsymbol{r})$ for $1 \le l \le 4N$, where $\boldsymbol{r} = (x,y)$ and N is the number of discretization points on each edge of Ω_i. The solution $\Phi_l(\boldsymbol{r})$ depends on the radius a_i and is a cylindrical wave given by

$$\Phi_l(\boldsymbol{r}) = \phi_m(r,a_i)e^{\mathbf{i}m\theta}, \quad m = l - 2N - 1,$$

where

$$\phi_m(r,a_i) = \begin{cases} J_m(k_0 n_1 r), & r < a_i, \\ AJ_m(k_0 n_2 r) + BY_m(k_0 n_2 r), & r > a_i. \end{cases}$$

The coefficients A and B depend on a_i, and they satisfy

$$J_m(\alpha_2)A + Y_m(\alpha_2)B = J_m(\alpha_1)$$
$$J'_m(\alpha_2)A + Y'_m(\alpha_2)B = \rho J'_m(\alpha_1),$$

where $\rho = n_1/n_2$, $\alpha_1 = k_0 n_1 a_i$, $\alpha_2 = k_0 n_2 a_i$, J_m and Y_m are Bessel functions, J'_m and Y'_m are their derivatives. The explicit formulas for A and B are

$$A = \frac{\rho s' q - s q'}{D}, \quad B = \frac{s p' - \rho p s'}{D}, \tag{25}$$

where

$$\begin{aligned} &D = qp' - pq', \\ &p = J_m(\alpha_2), \\ &p' = [J_{m-1}(\alpha_2) - J_{m+1}(\alpha_2)]/2, \\ &p'' = [J_{m-2}(\alpha_2) - 2J_m(\alpha_2) + J_{m+2}(\alpha_2)]/4, \\ &p''' = [J_{m-3}(\alpha_2) - 3J_{m-1}(\alpha_2) + 3J_{m+1}(\alpha_2) \\ &\quad -J_{m+3}(\alpha_2)]/8, \end{aligned}$$

q, q', q'' and q''' are defined as above with J_m replaced by Y_m, and s, s', s'' and s''' are defined with J_m but with α_2 replaced by α_1. The terms with double and triple primes will be used below.

The DtN map $\mathbf{\Lambda}_i$ of Ω_i is related to two matrices $\mathbf{D}$ and $\mathbf{N}$. The entries of $\mathbf{D}$ are just Φ_l evaluated at $4N$ points on the boundary of Ω_i, and the entries of $\mathbf{N}$ are related to $\nabla\Phi_l(\mathbf{r})$. We have

$$\nabla\Phi_l = \left[\phi'_m \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} + \frac{\mathbf{i}m\phi_m}{r} \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}\right] e^{\mathbf{i}m\theta},$$

where ϕ'_m denotes the derivative with respect to r.

The derivative of the DtN map $\mathbf{\Lambda}_i$ is related to the derivatives of $\mathbf{D}$ and $\mathbf{N}$ with respect to a_i. Clearly, they are further related to the derivatives of ϕ_m and ϕ'_m with respect to a_i. Since these functions are evaluated at the boundary of Ω_i, we use the expression of $\phi_m(r, a_i)$ for $r > a_i$. Therefore, $\partial_{a_i}\phi_m(r, a_i)$ is reduced to partial derivatives of A and B with respect to a_i. We have

$$\frac{\partial A}{\partial a_i} = \frac{f}{g}, \quad \frac{\partial B}{\partial a_i} = \frac{h}{g} \tag{26}$$

where

$$\begin{aligned} f &= k_0(\rho n_1 s'' q - n_2 s q'')D - \\ &\quad k_0 n_2(\rho s' q - s q')(qp'' - pq''), \\ h &= k_0(n_2 s p'' - n_1 p s'')D - \\ &\quad k_0 n_2(s p' - \rho p s')(qp'' - pq''), \\ g &= D^2. \end{aligned}$$

Notice that for $r > a_i$

$$\phi'_m(r, a_i) = k_0 n_2 \left[A J'_m(k_0 n_2 r) + B Y'_m(k_0 n_2 r)\right].$$

Therefore $\partial\phi'_m(r, a_i)/\partial a_i$ is also related to the first order derivatives of A and B.

Similarly, the process of taking the second order derivative of the DtN map $\mathbf{\Lambda}_i$ is passed to the matrices $\mathbf{D}$ and $\mathbf{N}$, to functions $\phi_m(r, a_i)$ and $\phi'_m(r, a_i)$, and finally to the two coefficients A and B. We have

$$\frac{\partial^2 A}{\partial a_i^2} = \frac{f'g - fg'}{g^2}, \quad \frac{\partial^2 B}{\partial a_i^2} = \frac{h'g - hg'}{g^2}, \tag{27}$$

where

$$\begin{aligned} f' &= k_0^2\left(\rho n_1^2 s''' q + n_1^2 s'' q' - n_1 n_2 s' q'' - n_2^2 s q'''\right) D \\ &\quad -k_0^2 n_2^2\left(\rho s' q - s q'\right)\cdot\left(q' p'' + q p''' - p' q'' - p q'''\right), \\ h' &= k_0^2\left(n_2^2 p''' s - n_1 n_2 p'' s' + n_1^2 p' s'' - \rho n_1^2 p s'''\right) D \\ &\quad -k_0^2 n_2^2\left(s p' - \rho p s'\right)\cdot\left(q' p'' + q p''' - p' q'' - p q'''\right), \\ g' &= 2k_0^2 n_2^2\left(q p'' - p q''\right) D. \end{aligned}$$

References

[1] D. Joannopoulos, S. G. Johnson, J. N. Winn, and R. D. Meade, *Photonic Crystals: Molding the Flow of Light,* 2nd ed. (Princeton University Press, Princeton, NJ, 2008).

[2] D. W. Prather, S. Shi, A. Sharkawy, and G. J. Schneider, Photonic Crystals: Theory, Applications, and Fabrication, (John Wiley & Sons, Hoboken, NJ 2009).

[3] A. Mekis, J. C. Chen, I. Kurland, S. H. Fan, P. R. Villeneuve and J. D. Joannopoulos, "High transmission through sharp bends in photonic crystal waveguides," Phys. Rev. Lett. **77**, 3787-3790 (1996).

[4] G. Veronis, R. W. Dutton, and S. Fan, "Method for sensitivity analysis of photonic crystal devices," Opt. Lett. **29**, 2288-2290 (2004).

[5] J. Smajic, C. Hafner, and D. Erni, "Design and optimization of an achromatic photonic crystal bend," Opt. Express **11**, 1378-1384 (2003).

[6] Z. Hu and Y. Y. Lu, "Improved bends for two-dimensional photonic crystal waveguides," Opt. Commun. **284**, 2812-2816 (2011).

[7] V. Liu, Y. Jiao, D. A. B. Miller, and S. Fan, "Design methodology for compact photonic crystal based wavelength division multiplexers," Opt. Lett. **36**, 591-593 (2011).

[8] K. K. Choi, "Shape design sensitivity analysis and optimal design of structural systems," in S. A. Mota Soares (ed.), *Computer Aided Optimal Design: Structural and Mechanical Systems*, Springer-Verlag, Berlin, pp. 439-492 (1987).

[9] N. K. Georgieva, S. Glavic, M. H. Bakr, and J. W. Bandler, "Feasible adjoint sensitivity technique for EM design optimization," IEEE Trans. Microwave Theory Tech. **50**, 2751-2758 (2002).

[10] C. M. Lalau-Keraly, S. Bhargava, O. D. Miller, and E. Yablonovitch, "Adjoint shape optimization applied to electromagnetic design," Opt. Express **21**, 21693-21701 (2013).

[11] A. C. R. Niederberger, D. A. Fattal, N. R. Gauger, S. Fan, and R. G. Beausoleil, "Sensitivity analysis and optimization of sub-wavelength optical gratings using adjoints," Opt. Express **22**, 12971-12981 (2014).

[12] K. Svanberg, "The method of moving asymptotes – a new method for structural optimization," International Journal for Numerical Methods in Engineering **24**(2), 359-373 (1987).

[13] Z. Hu and Y. Y. Lu, "Efficient analysis of photonic crystal devices by Dirichlet-to-Neumann maps," Opt. Express **16**, 17383-17399 (2008).

[14] Z. Hu and Y. Y. Lu, "Improved Dirichlet-to-Neumann map method for modeling extended photonic crystal devices," Opt. Quant. Electron. **40**, 921-932 (2008).

[15] Z. Hu and Y. Y. Lu, "A simple boundary condition for terminating photonic crystal waveguides," J. Opt. Soc. Am. B **29**(6), 1356-1360 (2012).

[16] J. Yuan, Y. Y. Lu and X. Antoine, "Modeling photonic crystals by boundary integral equations and Dirichlet-to-Neumann maps," J. Comput. Phys. **227**, 4617-3629 (2008).

[17] Y. Huang and Y. Y. Lu, "Scattering from periodic arrays of cylinders by Dirichlet-to-Neumann maps," J. Lightw. Technol. **24**, 3448-3453 (2006).

[18] J. Yuan and Y. Y. Lu, "Photonic bandgap calculations using Dirichlet-to-Neumann maps," J. Opt. Soc. Am. A **23**, 3217-3222 (2006).