

Efficient and spurious-free integral-equation-based optical waveguide mode solver

Amit Hochman and Yehuda Leviatan

*Department of Electrical Engineering
Technion - Israel Institute of Technology, Haifa 32000, Israel
leviatan@ee.technion.ac.il*

Abstract: Modal analysis of waveguides and resonators by integral-equation formulations can be hindered by the existence of spurious solutions. In this paper, spurious solutions are shown to be eliminated by introduction of a Rayleigh-quotient based matrix singularity measure. Once the spurious solutions are eliminated, the true modes may be determined efficiently and reliably, even in the presence of degeneracy, by an adaptive search algorithm. Analysis examples that demonstrate the efficacy of the method include an elliptical dielectric waveguide, two unequal touching dielectric cylinders, a plasmonic waveguide, and a realistic micro-structured optical fiber. A freely downloadable version of an optical waveguide mode solver based on this article is available.

© 2007 Optical Society of America

OCIS codes: (000.4430) Numerical approximation and analysis; (060.0060) Fiber optics and optical communications; (230.7370) Waveguides; (230.5750) Resonators; (240.6680) Surface Plasmons; (230.3990) Micro-optical devices;

References and links

1. K. Hayata, M. Koshiba, M. Eguchi, and M. Suzuki, "Vectorial Finite-Element Method Without Any Spurious Solutions for Dielectric Waveguiding Problems Using Transverse Magnetic-Field Component," *IEEE Trans. Microwave Theory Tech.* **34**, 1120–1124 (1986).
2. B. Rahman, F. Fernandez, and J. Davies, "Review of finite element methods for microwave and optical waveguides," *Proc. IEEE* **79**, 1442–1448 (1991).
3. J. T. Chen, T. W. Lin, K. H. Chen, and S. W. Chyuan, "True and spurious eigensolutions for the problems with the mixed-type boundary conditions using BEMs," *Finite Elem. Anal. Des.* **40**, 1521–1549 (2004).
4. C. Vassallo, "1993-1995 optical mode solvers," *Opt. Quantum Electron.* **29**, 95–114 (1997).
5. Y. Leviatan, A. Boag, and A. Boag, "Generalized formulations for electromagnetic scattering from perfectly conducting and homogeneous material bodies-theory and numerical solution," *IEEE Trans. on Antennas Propag.* **36**, 1722–1734 (1988).
6. A. Hochman and Y. Leviatan, "Analysis of strictly bound modes in photonic crystal fibers by use of a source-model technique," *J. Opt. Soc. Am. A* **21**, 1073–1081 (2004).
7. W. Schroeder and I. Wolff, "The origin of spurious modes in numerical solutions of electromagnetic field eigenvalue problems," *IEEE Trans. Microwave Theory Tech.* **42**, 644–653 (1994).
8. W. Schroeder and I. Wolff, "Full-wave analysis of the influence of conductor shape and structure details on losses in coplanar waveguide," *IEEE MTT S International Microwave Symposium Digest* **3**, 1273–1273 (1995).
9. O. Bucci, G. D'Elia, and M. Santojanni, "Non-redundant implementation of method of auxiliary sources for smooth 2D geometries," *Electron. Lett.* **41**, 22 (2005).
10. B. E. Spielman and R. F. Harrington, "Waveguides of Arbitrary Cross Section by Solution of a Nonlinear Integral Eigenvalue Equation," *IEEE Trans. Microwave Theory Tech.* **20**, 578–585 (1972).
11. V. Labay and J. Bornemann, "Matrix singular value decomposition for pole-free solutions of homogeneous matrix equations as applied to numerical modeling methods," *IEEE Microwave Guid. Wave Lett.* **2**, 49–51 (1992).

12. W. Schroeder and I. Wolff, "A reliable and efficient numerical method for indirect eigenvalue problems arising in waveguide and resonator analysis," *Int. J. Numer. Model.* **12**, 197–208 (1999).
13. N. Don, S. Germani, M. Bozzi, A. Kirilenko, and L. Perregri, "Determination of the mode spectrum of arbitrarily shaped waveguides using the eigenvalue-tracking method," *Microwave Opt. Technol. Lett.* **48**, 553 (2006).
14. The Source-Model Technique Package (SMTP) may be downloaded at:
<http://www.ee.technion.ac.il/leviatan/sntp/index.htm>.
15. Z. Altman, H. Cory, and Y. Leviatan, "Cutoff frequencies of dielectric waveguides using the multifilament current model," *Microwave Opt. Technol. Lett.* **3**, 294–295 (1990).
16. M. Golberg, *Solution Methods for Integral Equations: Theory and Applications* (Plenum Press, 1979).
17. C. Müller, *Foundations of the Mathematical Theory of Electromagnetic* (Springer-Verlag, 1969).
18. S. V. Boriskina, T. M. Benson, P. Sewell, and A. I. Nosich, "Highly efficient full-vectorial integral equation solution for the bound, leaky, and complex modes of dielectric waveguides," *IEEE J. Sel. Top. Quantum Electron.* **8**, 1225–1232 (1990).
19. H. Cheng, W. Y. Crutchfield, M. Doery, and L. Greengard, "Fast, accurate integral equation methods for the analysis of photonic crystal fibers," *Opt. Express.* **12**, 3791–3805 (2004).
20. D. L. Young, S. P. Hu, C. W. Chen, C. M. Fan, and K. Murugesan, "Analysis of elliptical waveguides by the method of fundamental solutions," *Microwave Opt. Technol. Lett.* **44**, 552–558 (2005).
21. R. Lehoucq, D. Sorensen, and C. Yang, *ARPACK Users' Guide: Solution of Large-Scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods* (SIAM Publications, Philadelphia, 1998).
22. C. F. Van Loan, "Generalizing the singular value decomposition," *SIAM J. Numer. Anal.* **13**, 76–83 (1976).
23. R. Brent, *Algorithms for Minimization Without Derivatives* (Dover Publications, 2002).
24. B. N. Datta, *Numerical Linear Algebra and Applications* (Brooks/Cole, Pacific Grove, CA, 1994).
25. C. G. Broyden, "The convergence of a class of double-rank minimization algorithms," *J. Inst. Math. Its Appl.* **6**, 76–90 (1970).
26. G. Golub, C. F. Van Loan, *Matrix Computations* (Johns Hopkins Univ Pr, 1996).
27. K. Lo, R. McPhedran, I. Bassett, and G. Milton, "An electromagnetic theory of dielectric waveguides with multiple embedded cylinders," *J. Lightwave Technol.* **12**, 396–410 (1994).
28. M. J. Steel and R. Osgood Jr, "Polarization and dispersive properties of elliptical-hole photoniccrystal fibers," *J. Lightwave Technol.* **19**, 495–503 (2001).
29. P. R. McIsaac, "Symmetry-induced modal characteristics of uniform waveguides-I: Summary of results," *IEEE Trans. Microwave Theory and Tech.* **23**, 421–429 (1975).
30. D. Marcuse, *Light Transmission Optics*, 2nd ed. (Rober E. Krieger Publishing Company, Malabar, Florida, 1989).
31. J. Kottmann, O. Martin, D. Smith, and S. Schultz, "Plasmon resonances of silver nanowires with a nonregular cross section," *Phys. Rev. B* **64**, 235,402 (2001).
32. J. Krenn and J. Weeber, "Surface plasmon polaritons in metal stripes and wires," *Philosophical Transactions: Mathematical, Physical and Engineering Sciences* **362**, 739–756 (2004).
33. H. Diltbacher, A. Hohenau, D. Wagner, U. Kreibig, M. Rogers, F. Hofer, F. Aussenegg, and J. Krenn, "Silver Nanowires as Surface Plasmon Resonators," *Phys. Rev. Lett.* **95**, 257,403 (2005).
34. L. Novotny and C. Hafner, "Light propagation in a cylindrical waveguide with a complex, metallic, dielectric function," *Phys. Rev. E* **50**, 4094–4106 (1994).
35. P. Johnson and R. Christy, "Optical Constants of the Noble Metals," *Phys. Rev. B* **6**, 4370–4379 (1972).
36. J. Burke, G. Stegeman, and T. Tamir, "Surface-polariton-like waves guided by thin, lossy metal films," *Phys. Rev. B* **33**, 5186–5201 (1986).
37. P. Berini, "Plasmon-polariton waves guided by thin lossy metal films of finite width: Bound modes of symmetric structures," *Phys. Rev. B* **61**, 10,484–10,503 (2000).
38. M. Szpulak, W. Urbanczyk, E. Serebryannikov, A. Zheltikov, A. Hochman, Y. Leviatan, R. Kotynski, and K. Panajotov, "Comparison of different methods for rigorous modeling of photonic crystal fibers," *Opt. Express* **14**, 5699–5714 (2006) <http://www.opticsinfobase.org/abstract.cfm?URI=oe-14-12-5699>.
39. A. Webb, F. Poletti, D. Richardson, J. Sahu, "Suspended-core holey fiber for evanescent-field sensing," *Opt. Eng.* **46**, 010,503 (2007).
40. A. Hochman and Y. Leviatan, "Calculation of confinement losses in photonic crystal fibers by use of a source-model technique," *J. Opt. Soc. Am. B* **22**, 474–480 (2005).

1. Introduction

Spurious solutions have been known to exist in various techniques for modal analysis of optical waveguides and resonators. Although more well-known in the context of the Finite-Element Method (FEM) [1,2], spurious solutions also appear when the method of moments (MoM) is applied to a surface formulation using local-domain basis functions, like in the Boundary-Element Method (BEM) [3], and also using entire-domain basis functions [4]. As shown in this paper,

spurious solutions also appear in methods like the Source-Model Technique (SMT), which are based on off-surface integral-equation formulations [5, 6]. Some spurious solutions are the result of the discretization of a continuous operator [7]. Others, however, can be traced to the continuous operator itself, which has, owing to its compactness, a continuum of eigenfunctions of arbitrarily small eigenvalue [8]. In this paper we focus on methods based on integral-equation formulations, such as the MoM and SMT, and show how spurious solutions may be eliminated by a modification of the matrix singularity measure used to determine the eigenvalues.

Methods based on surface and off-surface formulations are among the most successful electromagnetic modelling techniques. Compared with differential methods, the representation of the fields usually requires fewer unknowns, and curved boundaries and unbounded domains are easily accommodated. These advantages are largely a consequence of the basis functions used to approximate the fields. The basis functions are the Maxwellian fields of sources radiating in a homogeneous domain, located either on the media boundaries (as in the MoM), or slightly offset from them (as in the SMT). This choice exploits the tacit assumption that the material is piecewise homogeneous to render a very compact approximation of the fields. Indeed, in the context of scattering problems, it has been shown that the number of unknowns in the SMT is practically optimal and it approaches the number of degrees of freedom of the scattered fields [9].

Despite the advantages of integral-equation-based methods they are not as popular, in the context of mode determination, as the FEM and other differential methods. The main reason is probably that in contrast to the latter, mode determination in the MoM and the SMT leads to a *nonlinear* eigenvalue problem. The relevant impedance matrix depends nonlinearly on the frequency (and propagation constant, for waveguides), because the basis functions must obey Maxwell's equations for every potential eigen-frequency (and propagation constant). It turns out then, that the advantages of using Maxwellian fields as basis functions are offset by the cost of solving the resultant nonlinear eigenvalue problem.

Clearly, if the nonlinear eigenvalue problem could be solved efficiently and reliably, integral-equation methods would become more useful for modal analysis. A number of works have addressed this issue. The simplest approaches consist of sampling the determinant of the impedance matrix in ever increasing sampling rates to try and verify that no modes are missed [10]. The problem with this and similar approaches which search for minima of the smallest singular value [11] is that modes may be degenerate or almost degenerate, and hence, there is no *a priori* bound on the required sampling rate. To overcome this problem, more sophisticated methods that rely on the so-called *orthogonality in the local limit* [12] have been proposed [12, 13]. The idea is to take advantage of the (near) orthogonality of solution vectors that correspond to (nearly) degenerate modes by continuously correlating sets of candidate solution vectors at different frequencies or propagation constants. In [12], the minima of the smallest singular value are found by keeping track of singular values that correspond to all the candidate solution vectors on a given interval of frequency or propagation constant. Similarly, in [13], an eigenvalue-tracking method is used to track the smallest eigenvalues of the impedance matrix in the complex eigenvalue plane as the frequency is varied. The eigenmodes are determined by detecting real-line crossings of the complex eigenvalues. While these methods are more robust, their implementation is somewhat complicated and the overhead compared with a simple search cannot be entirely overlooked.

The search algorithm proposed in this paper makes use of the same quasi-orthogonality principle, though it does not require continuous tracking of many eigenvalues. Instead, the algorithm first tries to find the modes by simple (yet adaptive) sampling of the reciprocal of the matrix singularity measure introduced in this paper. This function is devoid of the irregularities caused by spurious solutions and typically resembles a sawtooth waveform whose minima

correspond to the eigenmodes. The adaptive sampling algorithm exploits this to sample the function sparsely. Only after a reasonably good image is obtained, the finer details are resolved by use of the orthogonality in the local limit.

Although the algorithm assumes that the modes have real frequency or propagation constant, moderately lossy or leaky modes can be found very efficiently by estimating the corresponding imaginary part and starting a search in the complex plane from this estimate. A dielectric waveguide solver based on this paper that includes many examples and a graphical-user interface is freely available for download [14].

The remainder of this paper is organized as follows: In Section 2, we discuss the source and character of the spurious solutions in integral-equation formulations. The difficulties caused by these solutions are described in Section 3. In Section 4, we present the proposed modification of the matrix singularity measure and show how it eliminates the spurious solutions. As explained in Section 5, once free from spurious solutions, the matrix singularity measure can be sampled adaptively to determine the modes. Numerical examples are given in Section 6, and the last section is a summary.

2. Spurious solutions of integral equations

Wolff and Schroeder have pointed out two different types of spurious solutions in electromagnetic modelling schemes. The more well-known type was discussed thoroughly in [7]. These solutions are due to the discretization process; they solve the discrete approximation of a continuous operator equation but do not correspond to any of its solutions. In effect, the spurious solutions solve the equation in a weighted average sense, but if the error of the solution is evaluated by use of slightly different weighting functions a much larger error is found. In the context of the SMT, these spurious solutions were also recognized in [15]. They can be avoided at a marginal extra cost by overdetermination, i.e., by using more weighting functions than basis functions (about twice as many).

The other type of spurious solutions was alluded to in [8], where the BEM was used to analyze coplanar waveguides. In contrast to the previous type of spurious solutions, these do not appear at discrete frequencies or propagation constants. Instead they form a continuum of spurious solutions which exist at each and every frequency or propagation constant. These spurious solutions do correspond to eigenfunctions of the continuous operator.

The integral equations that are discretized in the BEM and SMT are usually Fredholm equations of the first kind. For example, the integral equation for the TM modes of a perfectly conducting cylindrical waveguide oriented along the z direction is given by

$$\int_{C'} J_z(\mathbf{r}') H_0^{(2)}(k_\rho |\mathbf{r} - \mathbf{r}'|) dl' = 0. \quad (1)$$

Here, $H_0^{(2)}(\cdot)$ is the Hankel function of zero order and second kind, J_z is the unknown current distribution, and the radial wave number k_ρ obeys the separation equation

$$k_\rho = \sqrt{k_0^2 - \beta^2}, \quad (2)$$

where k_0 is the free-space wave number, equal to ω/c . The assumed and suppressed time and longitudinal variation are $\exp(j\omega t)$ and $\exp(-j\beta z)$, respectively, and c is the speed of light in vacuum. The radius vector \mathbf{r} points to the waveguide boundary contour, C , and the equation must be satisfied at all points $\mathbf{r} \in C$. The integration is carried out on the contour C' , and the radius vector that points to C' is \mathbf{r}' . The integration variable, dl' , is a differential line element. In conventional MoM surface formulations, both contours are the same, so $C \equiv C'$. This makes the integral equation kernel weakly singular. To avoid this, in the SMT the current distribution

is placed on C' which is a slightly dilated version of C . A non-trivial solution to Eq. (1) exists only for specific pairs of ω and β for which the integral operator has a zero eigenvalue.

In the context of integral equations, it is well-known that the compact integral operator,

$$\mathcal{L}(f) = \int K(\mathbf{r}, \mathbf{r}') f(\mathbf{r}') dl' \quad (3)$$

induced by a smooth or weakly-singular symmetric kernel, $K(\mathbf{r}, \mathbf{r}')$, has an infinite number of real eigenvalues, $\{\lambda_j\}$, which tend to zero [16, p. 38]. This is a direct consequence of the low-pass nature of the kernel function; rapidly varying eigenfunctions have, consequently, small eigenvalues. For example, in a circular geometry such an eigenfunction would be a cylindrical current sheet, of radius ρ_0 , carrying a unit-amplitude, circularly-harmonic current, $\exp(jn\phi)$. The longitudinal electric field, E_z , in the region $\rho \leq \rho_0$ due to this current is given by

$$E_z = -\frac{\pi\rho_0 k_\rho^2}{2\omega\epsilon_0} J_n(k_\rho\rho) H_n^{(2)}(k_\rho\rho_0) e^{jn\phi}; \quad \rho \leq \rho_0, \quad (4)$$

where J_n is the n th order Bessel function and ϵ_0 is the free-space permittivity. When $n \gg k_\rho\rho_0$, the spatial variation of the current is fast on a radial-wavelength scale, and E_z can be approximated by

$$E_z \approx -\frac{k_\rho^2\rho_0}{2j\omega\epsilon_0 n} \left(\frac{\rho}{\rho_0}\right)^n e^{jn\phi}; \quad \rho \leq \rho_0. \quad (5)$$

This shows that the field inside of a *unit amplitude* current sheet can be made arbitrarily small by using a sufficiently rapidly varying current. As n is increased, the field right on the current sheet decreases and it also decays more rapidly as the observation point moves away from the current sheet.

When the operator \mathcal{L} is discretized, the finite number of basis functions, N , limits the spatial variation of the solutions. Hence, when N is small, the error of the spurious solution is large, and assuming the true solutions have a smaller error they will be easily discernable. On the other hand, as N is increased the error of spurious solutions decreases and they may hide a true solution. In the SMT this problem is more severe owing to the factor $(\frac{\rho}{\rho_0})^n$ which appears in Eq. (5). As the current is displaced from the waveguide boundary, this factor strongly reduces the field there, thus reducing the error of the spurious solution.

Alternative surface formulations, such as Müller's [17], which lead to Fredholm equations of the second kind have also been proposed as a way of avoiding ill-conditioning (which is due to the spurious solutions). Applications of this approach to mode determination of optical waveguides can be found in [18, 19].

3. The effect of the continuum of spurious solutions on mode determination

Since the first type of spurious solutions can be handled quite easily by overdetermination, we focus now on the second type and show how it renders mode determination problematic. The difficulties caused by the spurious solutions are manifest even in the simplest analysis cases. This enables us to discuss these difficulties while avoiding undue complications, by determining the TM modes of a circular perfectly conducting waveguide. More challenging examples are given in Section 6. The method of analysis is the SMT, and we will look at the effect of varying the number of elementary sources and their distance from the boundary. It is assumed that the behavior of conventional MoM can be inferred from the behavior of the SMT in the limit that the number of sources increases and they approach the waveguide boundary.

In the SMT, TM modes are approximated by the fields of an array of N electrical current filaments, uniformly distributed on a circle outside the waveguide, concentric with the waveguide.

The filaments carry a longitudinally varying current, whose dependence on the longitudinal coordinate z , is $\exp(-j\beta z)$ (this dependence is suppressed). We denote the waveguide radius by R . The radius of the circle on which the sources are located is larger than R by a factor α ($\alpha > 1$). This geometry is depicted in Fig. 1. The amplitudes of the sources are arranged in an

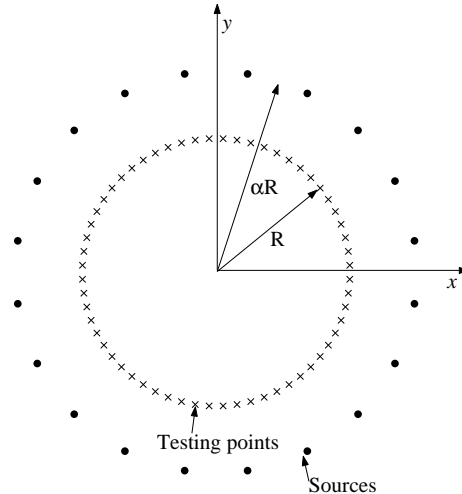


Fig. 1. Sources and testing points for a SMT analysis of a circular waveguide.

N -tuple column vector, denoted by \vec{I} . They are found by requiring that the longitudinal electric field, E_z , be zero at a set of M testing points, uniformly distributed on the waveguide boundary. This leads to a homogeneous matrix equation

$$[Z]\vec{I} = \vec{0}, \quad (6)$$

where $[Z]$ is the $M \times N$ impedance matrix. To avoid the first type of spurious solutions, more testing points than sources should be used. In this example, we used twice as many testing points as sources, so $M = 2N$.

Equation (6) is the discretized version of Eq. (1). It has nontrivial solutions if and only if $[Z]$ is singular, or as is the case in any approximate numerical solution, close to singular. To find the modes, a suitable measure of the singularity of $[Z]$ must be chosen and evaluated at the relevant values of the radial wave number, k_ρ . A common measure of singularity is the condition number of $[Z]$. Other measures that have been used are the closely related, smallest singular value of $[Z]$ [20], and the determinant of $[Z]$ [15], which, however, is not directly applicable to non-square matrices and is difficult to compute for even moderately large square matrices. Typically, the singularity measure is sampled on a sufficiently fine grid of $k_\rho R$, yielding a plot like the one shown in Fig. 2a, or Fig. 2 in [20]. The cut-off wave numbers, which are at the first zeros of the Bessel functions, can be clearly seen as peaks of the matrix condition number.

It is interesting to note that the matrix condition number is quite high *between* the peaks. The high condition number is of course due to the continuum of spurious solutions. In the discrete version, these solutions consist of sources of equal magnitude and alternating sign, that generate fields strongly confined to the vicinity of the sources. Because their E_z field is small on the waveguide boundary, they approximately solve Eq. (6), and this is reflected by the high condition number.

The difference between a true mode and a spurious one, is that while both have vanishingly small fields on the boundary, the fields of the spurious solution are also vanishingly small within

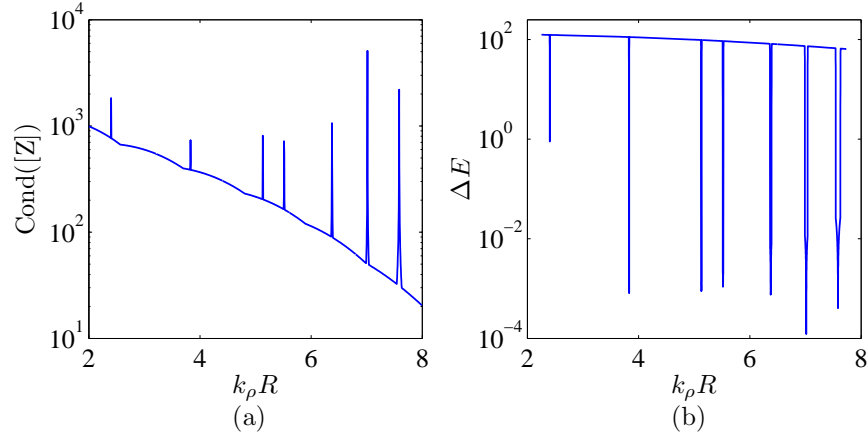


Fig. 2. Singularities at the zeros of the Bessel functions. For this graph, $N = 20$ and $\alpha = 1.5$. A very fine sampling grid of 1500 points was needed to reveal all the singularities.

the boundary. To distinguish between the two, we define a normalized error measure, ΔE , as

$$\Delta E = \frac{\text{rms}([\tilde{Z}]\vec{I})}{\text{rms}([\tilde{Z}]\vec{I})}, \quad (7)$$

where $[\tilde{Z}]$ is an impedance matrix that maps the amplitude vector \vec{I} to the values of E_z at a number of sampling points inside the waveguide. Also in Eq. (7), the 2-norm, or root-mean-square value, is denoted by $\text{rms}(\cdot)$. The sampling points can be arranged, for example, on a uniform grid of reasonable density. When analyzing dielectric structures, the sampling points can be the testing points already distributed on the media boundaries, since the modal fields are in general not zero there. This is a convenient choice because $[\tilde{Z}]$ can then be obtained directly from $[Z]$ after sign reversal of some of its entries. To evaluate ΔE , the vector that solves Eq. (6) in the least-squares sense is found by a singular value decomposition. The value of ΔE for \vec{I} found at every $k_\rho R$ of Fig. 2a, is shown in Fig. 2b, and it confirms that although the matrix is rather singular between the peaks, there are no modes there since their normalized error is high.

We now turn to take a closer look at the singularities shown in Fig 2. Zooming on the first of these singularities, we see in Fig. 3, that it is very sharp and discontinuous. This behavior has undesirable consequences. Because of the discontinuity, there can be no efficient search algorithm to find the singularities since there is no indication that a point near a singularity is in fact close, unless it is closer than the discontinuous edges. The only option is to sample the measure of singularity on a very fine grid. The width of the singularity determines the appropriate sampling resolution, and as shown in Fig. 3, the width depends on the number and location of the sources. Note that a small change in these parameters is enough to shrink the singularity considerably. If the sampling grid is kept constant while the number of sources is increased, the singularity may shrink to the point where it falls between grid points and goes undetected. As it is often essential to verify that the results are correct by increasing the number of sources or changing their positions slightly, this lack of reliability in the detection is a serious drawback.

The behavior seen in Fig. 3 is easily explained by existence of spurious solutions. If for a fixed α , the number of sources is increased, it would be like increasing n in Eq. (5), and if for a fixed N , α is increased, it would be like increasing ρ_0 in Eq. (5). In both cases, the absolute error of the spurious solution would decrease, contracting the singularity. To keep the singularities

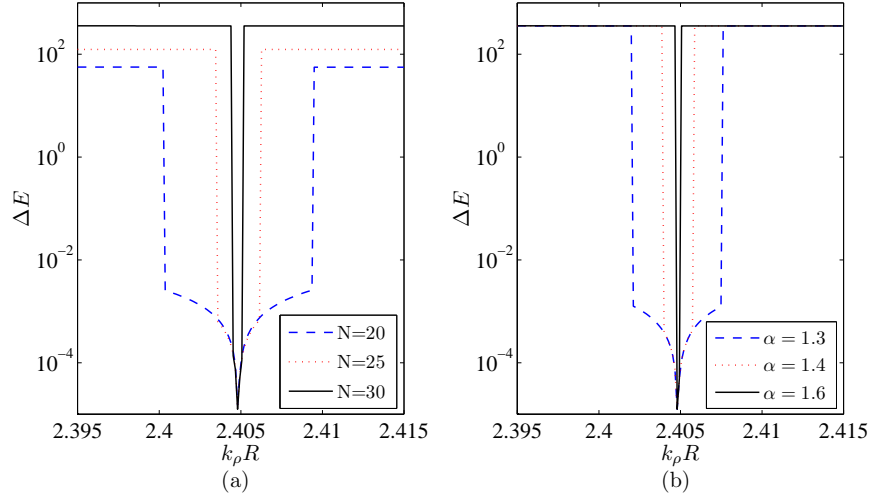


Fig. 3. Dependence of the shape of the singularity on (a) the number of sources and (b) the distance from the waveguide boundary. In (a), $\alpha = 1.5$, and in (b), $N = 30$.

wide enough, one should therefore use as few sources as possible and they should be placed as close as possible to the boundary. While this facilitates the detection of the singularities, the accuracy of the field approximation is bound to suffer. As we show in the next section, the need to balance these conflicting goals is obviated by the proposed method.

In a conventional surface-formulation, discretized by the MoM, the sources are placed on the boundary, i.e., $\alpha = 1$. Therefore, the singularities in the MoM are wider than in the SMT for an equal number of unknowns. Nevertheless, the problem will occur in the MoM as the number of basis functions is increased (Usually, a MoM solution will require more unknowns than an SMT solution [9]). Indeed, the continuum of spurious solutions was identified in [8] in the context of the BEM, which can be classified as a surface-formulation.

4. A modified singularity measure

In [8], a Tikhonov regularization was proposed to deal with the continuum of spurious solutions. The penalty function proposed there was proportional to the norm of the second derivative of the unknown fields at the interfaces between homogeneous regions. The method proposed here has the advantage that it does not involve derivatives of the fields and it also does not require calibration of the penalty function coefficient.

As explained in Section 3, it is the normalized error, ΔE , which takes into account field values inside the waveguide that reliably indicates the existence of a true mode. Hence, instead of evaluating ΔE for the least-squares solution of Eq. (6), it would be better to find the vector that minimizes ΔE for a given $k_\rho R$. The square of normalized error, $(\Delta E)^2$, is a ratio of two positive definite quadratic forms known as a *generalized Rayleigh quotient*. Direct differentiation shows that the stationary values of a generalized Rayleigh quotient are the generalized eigenvalues of the following generalized eigenvalue problem

$$[Z]^\dagger [Z] \vec{I} = \xi [\tilde{Z}]^\dagger [\tilde{Z}] \vec{I}. \quad (8)$$

The generalized eigenvector \vec{I}_{\min} , which corresponds to the minimum generalized eigenvalue, ξ_{\min} , yields the minimum ΔE , which is simply $\sqrt{\xi_{\min}}$. In essence, we have modified the matrix singularity measure; it can now be defined as the reciprocal of the square root of the small-

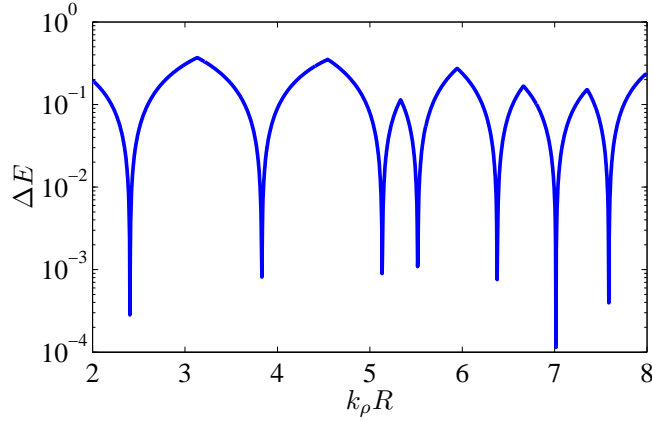


Fig. 4. Singularities at the zeros of the Bessel functions, calculated with the proposed matrix singularity measure. For this graph, $N = 20$ and $\alpha = 1.5$. The sampling grid is the same as that of Fig. 2, although a much coarser grid could have been used.

est generalized eigenvalue of Eq. (8). The main advantage of this scheme is that \vec{l}_{\min} and ξ_{\min} change continuously when moving from one singularity to the next, thus allowing reliable detection of the singularities.

The generalized eigenvalue decomposition can be carried out by a number of methods quite efficiently, i.e., without a significant increase in the computation time relative to any of the other common measures of singularity. Since we are only interested in the smallest eigenvalue, Arnoldi methods are most suitable [21]. They require, however, explicit multiplication of the matrices $[Z]$ and $[\tilde{Z}]$ by their Hermitian conjugates, and this operation is prone to round-off errors. In the vast majority of cases, however, this method yields excellent results and is used throughout this paper. A more robust, though slower, alternative would be to use Van Loan's algorithm for the generalized singular value decomposition (GSVD) [22], since the square root of the smallest generalized eigenvalue of Eq. (8) is the smallest singular value of the pair $([Z], [\tilde{Z}])$.

To demonstrate the effectiveness of the spurious-free formulation, the plots of Figs. 2 and 3 were recalculated with the same N and α using the proposed matrix singularity measure. The results are shown in Figs. 4 and 5, respectively. The slope from which the singularities protruded in Fig. 2 has disappeared in Fig. 4, and the minima of the curve can be found with far less iterations. This is the subject of the next section. As shown in Fig. 5, the high sensitivity of the width of the singularity to the location and number of sources has been completely eliminated.

5. Spurious-free mode determination

The first phase of the mode determination scheme consists of adaptive sampling of the singularity measure. The objective of this phase is to sample the singularity measure, ΔE , on a nonuniform sampling grid which reveals just enough detail for the next phase to detect all the minima in a prescribed interval of propagation constant or frequency. As the next phase utilizes the orthogonality in the local limit, closely spaced modes need not be resolved, but otherwise all the minima should be detected. In the description that follows it will be assumed that a waveguide is analyzed, and therefore the singularity measure depends both on the frequency, ω , and on the propagation constant, β . In the analysis of a resonator it would depend, of course, only on the frequency.

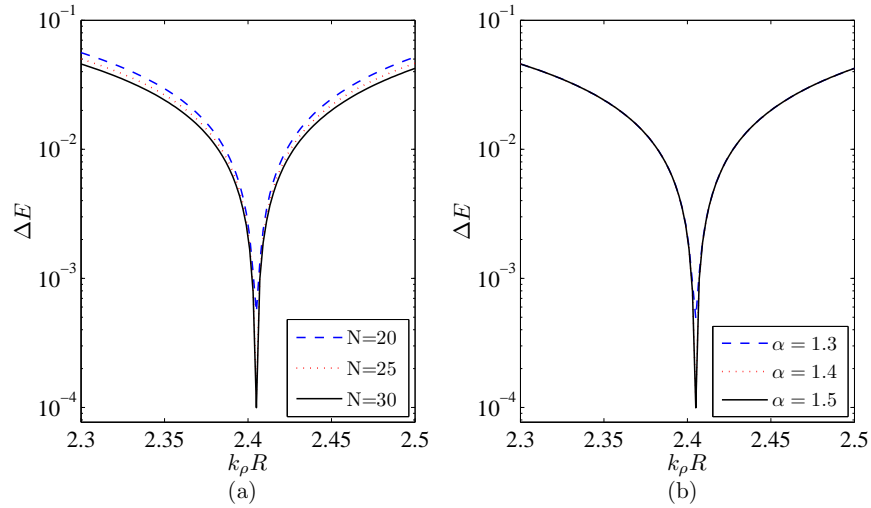


Fig. 5. The proposed matrix singularity measure. Dependence of the shape of the singularity on the number of sources (a) and the distance from the waveguide boundary (b). In (a), $\alpha = 1.5$, and in (b), $N = 30$. Note that the range of the abscissa is ten times that of Fig. 3.

5.1. Adaptive sampling algorithm

The adaptive sampling algorithm is based on the observation that $\Delta E(\beta)$ for a given frequency ω , resembles a sawtooth waveform. This can be readily observed in Fig. 6a, where $\Delta E(\beta)$ is plotted on a linear scale (as opposed to the logarithmic scale used in the previous figures). As shown in Fig. 6b, if β is kept fixed and the frequency is allowed to vary, a similar result is obtained for $\Delta E(\omega)$. The sawtooth-like shape is obtained when the proposed singularity measure is used. Of course, if Fig. 2 were plotted on a linear scale it would still be discontinuous and completely unsuitable for an efficient search algorithm.

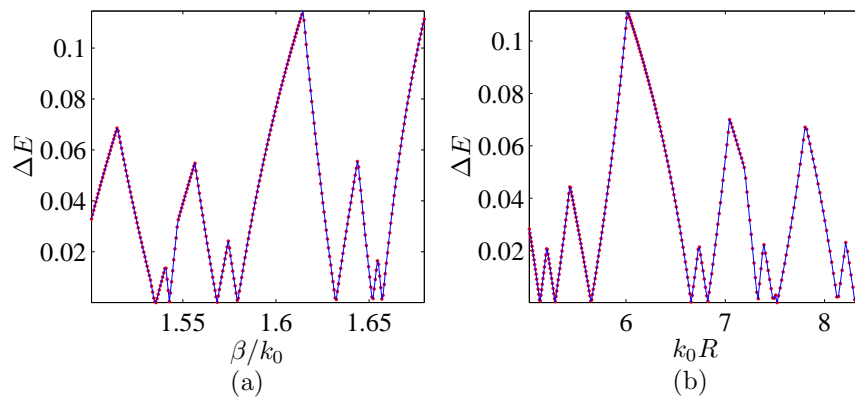


Fig. 6. Typical sawtooth-like dependence of the singularity measure, ΔE , on (a) the normalized propagation constant, or effective index, β/k_0 , and on (b) the normalized wave vector, $k_0 R$. The structure analyzed is a round step-index fiber. The relative permittivities are: $\epsilon_{rw} = 1.8^2$, and $\epsilon_{rc} = 1.5^2$, for the waveguide and cladding regions, respectively. In (a) the radius R is equal to the free-space wavelength, λ_0 . In (b) the normalized propagation constant, β/k_0 , is 1.6.

The algorithm, shown in Fig. 7, exploits the shape of ΔE by assuming that monotonic parts of the function are approximately straight lines. For a given interval of β , the algorithm attempts to ascertain as quickly as possible whether the function $\Delta E(\beta)$ is monotonic on the interval. As explained in the next paragraph, this is where the shape of $\Delta E(\beta)$ is utilized. If $\Delta E(\beta)$ is indeed monotonic, the algorithm ends with empty output. If, on the other hand, a minimum or a maximum is detected, its location is determined to higher accuracy by a standard golden section search [23]. The interval is then subdivided to intervals to the left and to the right of the extremum found, and the algorithm is called recursively with these subintervals as input. The output then is a concatenation of this extremum and those found in the two subintervals. At the end of the recursive process the output is a list of all the extrema encountered.

Algorithm 5.1: FINDEXTREMA($\beta_{\min}, \beta_{\max}$)

if $\Delta E(\beta)$ is monotonic on the interval ($\beta_{\min}, \beta_{\max}$)
then return ()
else $\left\{ \begin{array}{l} \text{Find an extremum } \beta_{\text{ext}} \in (\beta_{\min}, \beta_{\max}) \text{ by golden section search} \\ \text{return } (\beta_{\text{ext}}, \text{FINDEXTREMA}(\beta_{\min}, \beta_{\text{ext}}), \text{FINDEXTREMA}(\beta_{\text{ext}}, \beta_{\max})) \end{array} \right.$

Fig. 7. Adaptive search algorithm for finding the extrema of $\Delta E(\beta)$.

Ascertaining whether a function is monotonic on a given interval is of course a difficult problem in the general case. However, by exploiting the sawtooth form of the function, a simple and efficient algorithm may be proposed (see Fig. 8). The algorithm begins with three samples: at the interval endpoints and at its midpoint. It then adds samples as long as the sample series is monotonic and the number of samples, n , is smaller than a prescribed number, n_{\max} , which can be determined empirically as will be shown below. The samples are added as follows. Assuming the function has been sampled n times, a straight line is fit through every three consecutive samples by linear regression. For each fit, the correlation coefficient, r , is evaluated by

$$r^2 = 1 - \frac{\sum_{i=1}^3 [\Delta E(\beta_i) - \Delta E_f(\beta_i)]^2}{\sum_{i=1}^3 [\Delta E(\beta_i) - \langle \Delta E(\beta_i) \rangle]^2} \quad (9)$$

where β_i are the sample points, $\Delta E_f(\beta_i)$ is the value of the fit at these points, and the angle brackets denote the mean value. A correlation coefficient close to 1 indicates that the three points are approximately collinear. The algorithm finds the three consecutive points, β_{m-1} , β_m , and β_{m+1} , which correspond to the smallest value of r , i.e., the ones for which the straight line fit is of worst quality. This deviation from a straight line possibly indicates an extremum, and therefore, the following samples are taken to be $(\beta_{m-1} + \beta_m)/2$ and $(\beta_m + \beta_{m+1})/2$. If the sample series remains monotonic when the maximum number of samples is reached, it is concluded that $\Delta E(\beta)$ is monotonic. If an extremum is encountered during the sampling process, $\Delta E(\beta)$ is obviously not monotonic.

As illustrated in Fig. 9, the sampling points tend towards regions of high curvature where an extremum is likely to be found. However, this sampling scheme will have difficulty detecting both minima of a nearly degenerate mode pair. When the first minimum of the pair is found,

Algorithm 5.2: ISMONOTONIC($\beta_{\min}, \beta_{\max}, n_{\max}$)

comment: Begin by adding the endpoints and the midpoint to the sample series.

$$\beta_0 \leftarrow \beta_{\min}$$

$$\beta_1 \leftarrow (\beta_{\min} + \beta_{\max})/2$$

$$\beta_2 \leftarrow \beta_{\max}$$

$$n \leftarrow 3$$

comment: In the following, $\vec{\beta}$ is a shorthand for the entire sample series, $\{\beta_j\}_{j=0}^{n-1}$.

comment: Similarly, $\vec{\Delta E}$ is a shorthand for the error series, $\{\Delta E(\beta_j)\}_{j=0}^{n-1}$.

while $\vec{\Delta E}$ is monotonic **and** $n \leq n_{\max}$

do $\left\{ \begin{array}{l} \text{find the three less collinear points, } \beta_{m-1}, \beta_m, \text{ and } \beta_{m+1} \\ \beta_n \leftarrow (\beta_{m-1} + \beta_m)/2 \\ \beta_{n+1} \leftarrow (\beta_m + \beta_{m+1})/2 \\ n \leftarrow n + 2 \\ \text{sort } \vec{\beta} \end{array} \right.$

if $\vec{\Delta E}$ is monotonic **return** (**true**)

else return (**false**)

Fig. 8. Algorithm for attempting to determine whether $\Delta E(\beta)$ is monotonic by sampling it at most n_{\max} times.

the algorithm will be called with a new subinterval which includes the undetected mode near one of its endpoints. If the undetected mode is sufficiently close to the endpoint, the function will appear to be monotonic on the subinterval and the mode will be missed. This case can also be seen in Fig. 9, where a mode has been detected at the left endpoint in the previous call to the search algorithm. The mode just to the right of the left endpoint may go undetected, even after further subdivision of the intervals because the number of samples required for its detection could be more than n_{\max} . The probability of this happening may be reduced (at a relatively small expense) by choosing the first two samples very close to the endpoints. However, since minima may be arbitrarily close, a mode may still be missed. Such close modes are resolved by the next phase of the algorithm. The first phase is summarized in the animation shown in Fig. 10.

5.2. Resolving degenerate and nearly-degenerate modes

It is well-known that eigenmodes of waveguides and resonators obey certain orthogonality relations. In general, however, it is not true that the solution vectors of Eq. (6) that correspond to different modes obey an orthogonality relation. In fact, the number of modes is infinite whereas the number of linearly independent solution vectors is bounded by their length, N . However, the generalized eigenvectors of Eq. (8) corresponding to different eigenvalues (which are the con-

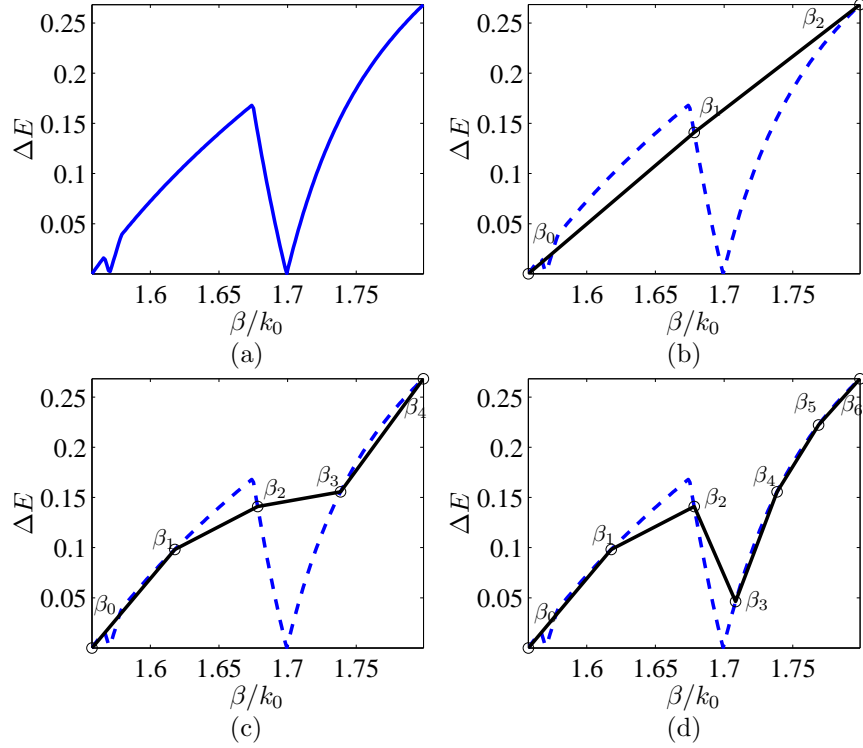


Fig. 9. Progress of the sampling scheme used to determine whether $\Delta E(\beta)$ is monotonic. A high resolution plot of ΔE is shown in (a) for the same step-index fiber of Fig. 6, but with $\lambda_0 = 2R$. In the first step of the sampling scheme (b), the endpoints and midpoint are sampled. Since the sample series is monotonic, two new samples are added at $(\beta_0 + \beta_1)/2$ and $(\beta_1 + \beta_2)/2$, as shown in (c). Since the sample series is still monotonic, the three less collinear consecutive points are found (β_3, β_4 , and β_5) and two points are added in between them (d). This last refinement reveals a minimum and a maximum.

tinuity condition errors) are orthogonal. So the solution vectors that correspond to degenerate modes are orthogonal. For a generalized eigenvalue problem,

$$[A]\vec{x} = \xi[B]\vec{x}. \quad (10)$$

where $[A]$ and $[B]$ are Hermitian and positive definite, the generalized eigenvectors \vec{x}_i are B-orthogonal [24, p.507], i.e.,

$$\vec{x}_i^\dagger [B] \vec{x}_j = 0, \quad i \neq j. \quad (11)$$

Similarly, if the modes are nearly degenerate, their solution vectors will be nearly orthogonal [12].

This quasi-orthogonality can be used to resolve degenerate and nearly degenerate modes. Suppose the previous phase of the algorithm has found one of the minima of two closely spaced minima. The solution vector at this minima is a by-product of the evaluation of the matrix singularity measure ΔE ; let this vector be denoted by \vec{I}_1 . Now, instead of $[Z]^\dagger [Z]$ in Eq. (8), substitute the matrix $[A]$ given by

$$[A] = [Z]^\dagger [Z] + \xi_0 \frac{[B] \vec{I}_1 \vec{I}_1^\dagger [B]^\dagger}{\vec{I}_1^\dagger [B]^\dagger \vec{I}_1}, \quad (12)$$

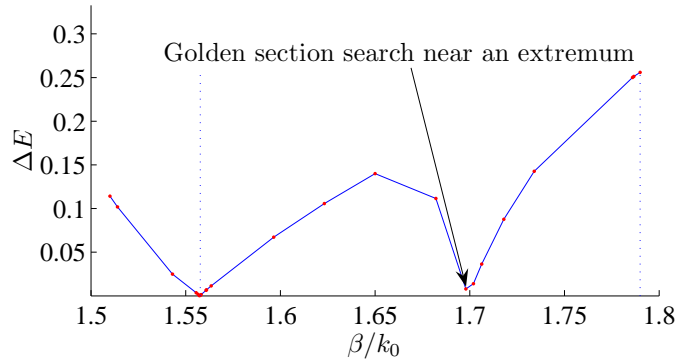


Fig. 10. An animation summarizing the sampling algorithm (AVI, 3.7MB), for the same step-index fiber shown in Fig. 9. The current search interval is marked by the dotted lines. The rate of the animation is much slower than the actual computation.

where the matrix $[B]$ is given by $[\tilde{Z}]^\dagger[\tilde{Z}]$. So constructed, the matrix pair $([A], [B])$ has the same eigenvectors as the original generalized eigenvalue problem, Eq. (6). The eigenvalues are also the same except for ξ_1 , the eigenvalue of \tilde{I}_1 , which is now $\xi_1 + \xi_0$ instead. By setting ξ_0 to a high enough value (say, $\xi_0 = 1$), the minimum which corresponds to ξ_1 can be removed from the $\Delta E(\beta)$ function which is defined as the smallest generalized eigenvalue at a given β . An important feature of this process is that it does not create any spurious minima. When the modes are only nearly-degenerate, the eigenvectors and eigenvalues of all the modes will be in general altered by the proposed substitution. Nevertheless, as can be seen in Fig. 11, this alteration is negligible when the modes are close enough.

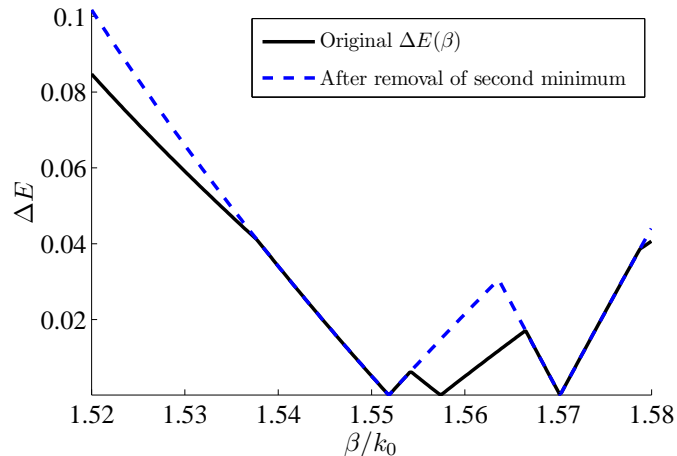


Fig. 11. The resolution of two close modes in the step-index fiber of Fig. 9. Assuming the sampling process missed the first minimum but found the second minimum, the neighborhood of the second minimum can be searched again, this time with the matrix $[A]$ of Eq. (12). The first minimum can then be easily detected.

The input of the second phase of the algorithm is the list of extrema found in the previous phase. The second phase has to check each minimum found for a nearby mode that was missed. To this end, the sampling algorithm could be run again for each interval defined by two adjacent

maxima (or a maximum and an endpoint), with the minimum between them removed. While this is probably the most robust possibility, a faster alternative is to remove the minimum and then check the value of $\Delta E(\beta)$ at the point of the removed minimum. If the value is small enough, it indicates that there is in fact another minimum nearby and that another search is warranted.

To determine whether another search is warranted the following criterion is used. The value of $\Delta E(\beta)$ with the minimum removed is checked against the values of $\Delta E(\beta)$ at the maxima on both sides of the minimum, with the minimum present. If it is smaller than the smallest of these values, another search between the two maxima is conducted. After the new minimum is found the criterion is tested again, this time with both minima removed. This process continues until it is concluded that there are no more minima in between the two maxima, and the search moves on to the next minimum found in the previous phase. The above criterion will be always correct when the modes are degenerate. As the modes move further apart, it may become incorrect and a mode could be missed. The probability of this happening is controlled by the resolution of the previous phase, which is determined by the parameter n_{\max} . By varying this parameter in numerous study-cases we have found, empirically, that $n_{\max} = 7$ leads to no modes being missed. However, the time complexity scales slower than linearly with n_{\max} , so, for added robustness, n_{\max} can be increased while incurring only a small increase in computation time.

5.3. Finding moderately lossy modes in the complex plane

It has been assumed until now that the modes are lossless and hence the minima are on the real β axis. For moderately lossy (or leaky) modes, the minima will be found in the complex plane, near the real axis. Usually, only the modes with the smallest losses are of interest, and these are very close to the real axis. The determination of these small losses can be quite challenging as the imaginary part of β , β_i , can be several orders of magnitude smaller than the real part, β_r . When this is the case, the effect of the losses is to blunt the minima of $\Delta E(\beta_r)$ when it is evaluated on the real line. It is then possible to estimate β_i from the shape of the blunted curve. Furthermore, this estimate can be used as the starting point for standard quasi-Newton methods (we used the MATLAB implementation of Broyden's algorithm [25]) which can converge very rapidly to the complex minimum given a good first estimate.

In the previous phases of the algorithm, and especially in the golden section searches, the function $\Delta E(\beta_r)$ is sampled many times in the neighborhood of a minimum. These samples can be used to estimate β_i . In analogy to the lossless case, it is assumed that $\Delta E(\beta)$ depends linearly on the distance from the minimum point, β_0 , i.e.,

$$\Delta E(\beta) \approx a|\beta - \beta_0|. \quad (13)$$

When evaluated on the real β axis, $[\Delta E(\beta_r)]^2$ is given by,

$$[\Delta E(\beta_r)]^2 \approx a^2[(\beta_r - \beta_{0r})^2 + \beta_{0i}^2], \quad (14)$$

Where $\beta_0 = \beta_{0r} + j\beta_{0i}$. This is a second degree polynomial in β_r with coefficients p_i , i.e.,

$$[\Delta E(\beta_r)]^2 \approx p_2\beta_r^2 + p_1\beta_r + p_0. \quad (15)$$

The polynomial which best fits the non-uniformly sampled data is easily obtained by the solving the corresponding Vandermonde matrix equation [26, p.119–124]. By comparing the coefficients in Eq. (13) and Eq. (15), the following estimate for β_0 can be obtained:

$$\beta_0 \approx -p_1/(2p_2) - j\sqrt{p_0/p_2 - [p_1/(2p_2)]^2}. \quad (16)$$

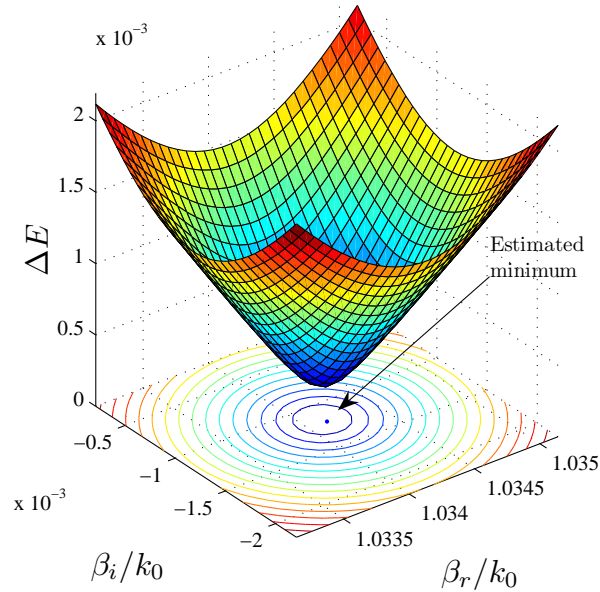


Fig. 12. The normalized error, $\Delta E(\beta)$, near a minimum in the complex β plane. Contours of equal ΔE are plotted at the bottom. The contours are equally spaced in ΔE .

An example of the use of this method is shown in Fig. 12. As shown in the figure, the estimate is very close to the real minimum, and a search starting from the estimate should converge quickly. The constant ΔE contours plotted in the bottom part of the figure are equally spaced in ΔE . Their nearly equal spacing in the complex β plane, as can be observed in the figure, indicates that Eq. (13) is a good approximation. To find degenerate modes in the complex plane, a minimum found can be eliminated, and the quasi-Newton method can be started again from the same estimate. It should then converge to a nearby minimum, if present.

6. Numerical results

To demonstrate the various aspects of the proposed algorithm, a few optical waveguide analysis examples are given in this section. Some of the results shown have been obtained previously by other methods; these are used to validate the code. New data, however, is also presented.

6.1. Two touching cylinders

The geometry analyzed consists of two touching dielectric cylinders of unequal radii R_1 and R_2 , surrounded by an air cladding, as shown in the inset of Fig. 13. This geometry has been analyzed previously in [27] by a Rayleigh formulation, and our results are validated against the six digit results for the effective index given there. We also give the dispersion curves (shown in Fig. 13) of the first modes of the two touching cylinders, which were not shown in [27]. The x coordinate of the figure is the normalized frequency, V , given by,

$$V = \frac{k_0 R_1}{\pi} \sqrt{\epsilon_{rw} - \epsilon_{rc}}, \quad (17)$$

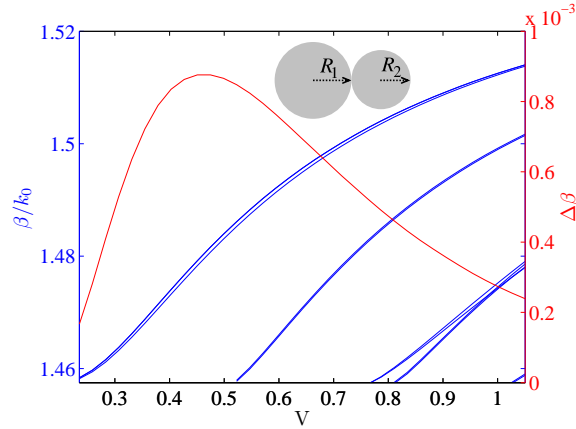


Fig. 13. Dispersion curves of the two touching cylinders analyzed in [27]. The material parameters are: $\epsilon_{rc} = 1.457^2$, and $\epsilon_{rw} = 1.457^2/0.9$, and the ratio of radii, R_1/R_2 , is 1.3.

where ϵ_{rc} is the relative permittivity of the cladding medium, and ϵ_{rw} is the relative permittivity of the two dielectric waveguides. In Fig. 13, the birefringence of the fundamental mode pair, given by,

$$\Delta\beta = n_x - n_y, \quad (18)$$

is also shown. Here, n_x and n_y are the effective indices of the x - and y -polarized modes. Similarly to an elliptical waveguide, the birefringence of the fundamental mode pair tends to zero at the high and low normalized frequency limits [28].

The convergence of the effective index with the number of sources used is shown in Table 1. To find hybrid modes, both electric and magnetic current filaments are used [15], and an equal

N	β/k_0	ΔE
10	1.459668	7.0×10^{-2}
20	1.461927	3.0×10^{-2}
30	1.462487	7.7×10^{-3}
40	1.463286	1.8×10^{-3}
50	1.463290	7.7×10^{-4}
60	1.463292	3.0×10^{-4}
70	1.463292	1.2×10^{-4}
Ref. [27]	1.463292	

Table 1. Convergence of the effective index. The normalized frequency, V , is 3.5. All the rest of the parameters are given in the caption of Fig. 13.

number of each type are placed inside and outside of each homogeneous region. In Table 1, and in all the tables and figures that follow, we use N to denote the number of sources of each type and on each side of a material boundary, so the total number of sources is actually $4N$. We find that our solver converges to the same value given in [27].

The real part of the z component of the complex Poynting vector, $\text{Re}(S_z)$, for the two lowest mode pairs, is shown in Fig. 14. As can be readily observed, the light concentrates in one of the

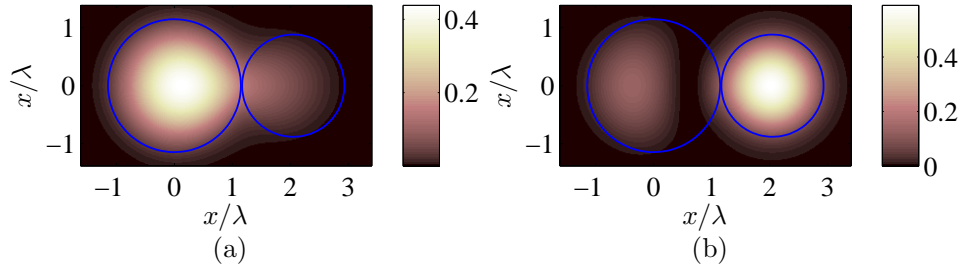


Fig. 14. Real part of the z component of the complex Poynting vector, $\text{Re}(S_z)$, for the x -polarized mode of the first (i.e. fundamental) mode pair (a), and for the x -polarized mode of second mode pair. For the y -polarized modes, the plots are almost identical to the x -polarized plots shown. The parameters are as in Fig. 13, and the effective indices of the modes shown are $n_x = 1.5155$ and $n_y = 1.5044$, for $V = 3.5$. The plots have been normalized to unit total power in the cross-section shown.

two cylinders, and penetrates quite moderately into the other cylinder. This is a consequence of the unequal radii, and it would result in only moderate coupling between the two cylinders, even though they are touching.

6.2. An almost-circular dielectric waveguide

The next example demonstrates the reliability of the method in finding numerous modes. An elliptical waveguide with an aspect ratio of 1.05 was chosen as a challenging example, as it is expected to have pairs of very closely spaced modes. The modes of the waveguide were found for three normalized frequencies $V = 1, 2, 3$. In this case, the normalized frequency is again given by Eq. (17), with R_1 the semiminor axis of the ellipse. The number of modes found at each normalized frequency as a function of the number of sources is shown in Fig. 15(a). As

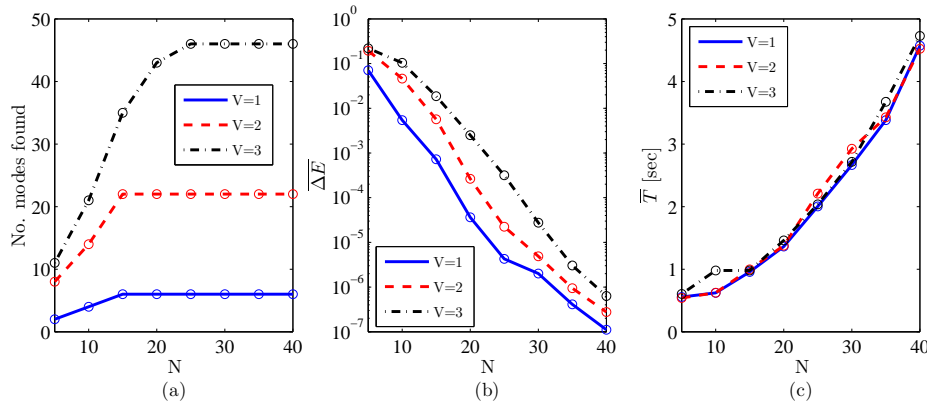


Fig. 15. Solver performance as function of the number of sources used. The number of modes found is shown in (a), the average error in continuity conditions per mode, $\overline{\Delta E}$, in (b), and the average computation time per mode on a 3.2 GHz Intel PC, \overline{T} , in (c). For the elliptical waveguide analyzed, $\epsilon_{rw} = 2$, and $\epsilon_{rc} = 1$.

can be readily observed in the figure, the number of modes found increases at first and then

saturates. Clearly, the saturation indicates that there are enough sources to simulate the field of the highest mode (i.e the mode with fastest spatial variation).

To verify that no modes are missed owing to possible deficiencies of the sampling algorithm (as opposed to insufficient number of sources), symmetry considerations [6, 29] were used, as follows. In an elliptical waveguide, modes can be classified into four different classes according to their symmetry. Since nearly-degenerate modes belong to different symmetry classes, it is easier to find the total number of modes by searching for each class separately. The total number of modes found in this way matched those found without exploiting symmetry. To verify the results further, the propagation constants found were compared with those of a circular dielectric waveguide, for which the characteristic equation (the solution of which yields the propagation constants) is well-known [30, p. 296]. This test also showed that the total number of modes found was correct, when enough sources were used. The interested reader is referred to the Appendix for the propagation constants of all the modes found.

The error in continuity conditions was averaged over all the modes found at a given frequency and number of sources. The average error, $\Delta\bar{E}$, is shown in Fig. 15(b). An exponential decrease of the average error with the number of basis functions, typical of integral equation methods, can be readily observed. Higher order modes require in general more sources owing to their faster spatial variation. This explains why the average error increases with frequency, for a given number of sources, or conversely, why higher frequencies require more sources to attain a prescribed average error.

In Fig. 15(c), the computation time per mode is shown for the three normalized frequencies studied. The three plots fall practically one on top of the other, even though, as the frequency increases, the average density of the modes increases, and their minimum separation decreases (for $V = 1$ it is 1.9×10^{-3} , for $V = 2$, it is 1.8×10^{-5} , and for $V = 3$, it is 4×10^{-8}). This is an appealing feature of the sampling algorithm. It implies that, for a given number of sources, the computation time would grow only linearly with the number of modes found, regardless of their average density and minimum spacing.

As usual in methods based on integral equations, the memory requirements are quite modest. Practically all the memory allocated is used up by the impedance matrix. For the calculations of Fig. 15, the number of testing points used was greater than the number of sources by a factor of 1.5, and each complex entry of the matrix required 16 bytes of storage. Therefore, the memory required, in kilobytes, is $0.375N^2$, which is 600 KBytes for the largest matrix used in the calculations of Fig. 15.

6.3. Circular and elliptical plasmonic nano-wire

The following example demonstrates complex eigenvalue determination. Plasmonic nano-wires have attracted considerable interest recently, owing to their unique wave-guiding (and light scattering) properties [31–33]. These lossy waveguides are made of a noble metal, such as silver, which is usually modelled by assuming a plasma-like frequency dependent permittivity. In some works (see [34] for example), a simple Drude model is assumed, whereas in others the permittivity is interpolated from measured values. We follow the latter practice and use the measured data given by Johnson and Christy [35].

The characteristic equation for a lossy circular nano-wire is the same as that of a lossless one, with a complex permittivity substituted for the real permittivity. The search for the effective indices, however, must be carried out in the complex plane, and the square root branch in Eq. (2) should be chosen according to the desired behavior at infinity. If, as in the following examples, proper modes are sought, the radial wave number outside the waveguide should have negative real and imaginary parts. For a circular silver nano-wire, the convergence of the effective index to the value obtained from numerical solution of the characteristic equation is shown in Table 2.

The initial estimate of the imaginary part, β_{0i}/k_0 , based on a few samples of ΔE on the real line

N	β_r/k_0	β_i/k_0	β_{0i}/k_0	Time [sec]
5	1.2541335	0.0066253	0.2746202	1
10	1.2188215	0.0054484	0.0468567	2
15	1.2180724	0.0055741	0.0073826	5
20	1.2180544	0.0055804	0.0056145	10
25	1.2180543	0.0055801	0.0055802	18
Exact	1.2180543	0.0055801		

Table 2. Convergence of the effective index and the estimate of the imaginary part, β_{0i}/k_0 , of the TM_0 mode of circular silver nano-wire. The free-space wavelength is 633nm, the radius of the wire is 50nm, and the relative permittivity of the wire (interpolated from [35]) is: $-18.3187697 - j0.5037517$.

also appears to converge to the exact value. It worth noting that when the number of sources is small, the estimate can err by a few orders of magnitude. However, the results improve greatly after the complex plane is searched, starting from this estimate. The reason for this is that when too few sources are used, the minimum of ΔE on the real line is blunted more by the numerical error than by the distance of the true minimum from the real line. Thus, using enough sources is important for ensuring rapid convergence to the minimum closest to the real line, especially if many minima are present.

The dispersion with semimajor axis, a , of the TM_0 mode of an elliptical nano-wire made of silver is shown in Fig. 16, for various aspect-ratios. In the circular case, very good agreement

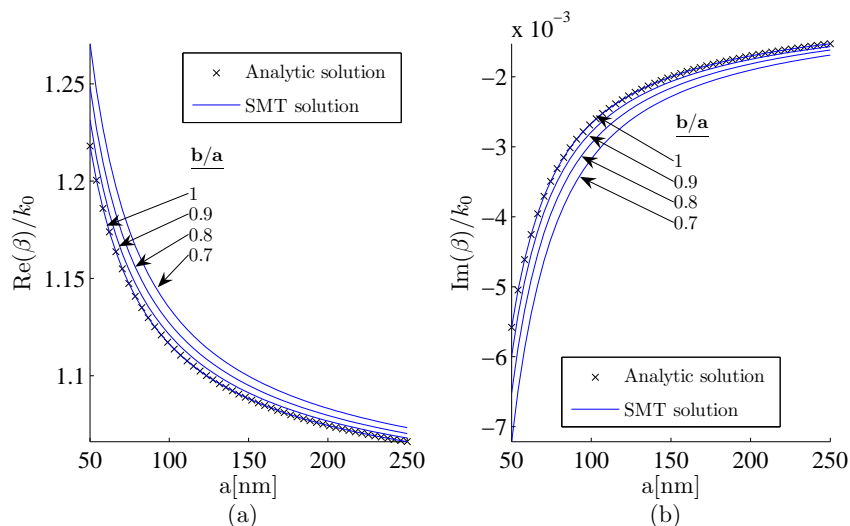


Fig. 16. Real (a) and imaginary (b) parts of the effective index of the TM_0 mode of circular and elliptical nano-wires made of silver. The free-space wavelength and permittivity are as in Table 2.

is obtained with the solution of the characteristic equation. The behavior of the complex prop-

agation constant is typical of symmetrical Surface Plasmon Polariton (SPP) modes [36, 37]. As the cross-sectional area of the nano-wire is decreased, the power flow density in the z direction concentrates in the metal and this is accompanied by more rapid dissipation and slower phase velocity.

7. A micro-structured optical fiber

Micro-structured optical fibers usually have quite complex cross-sections. To easily distribute the sources and testing points, it is most convenient to have an analytical parametric representation of the material boundaries. It is then easy to distribute the sources on contracted and dilated versions of the boundaries. The required parametric representation may be obtained, for example, by fitting spline curves to the material boundaries obtained from a thresholded SEM image. This approach was used in [38], where further details about it can be found.

The fiber analyzed in this example was proposed for gas sensing applications in [39]. The geometry of the fiber, superimposed on the calculated $\text{Re}(S_z)$ of the fundamental mode, is shown in Fig. 17. A parameter of interest in the design of the gas sensor is the fraction of power

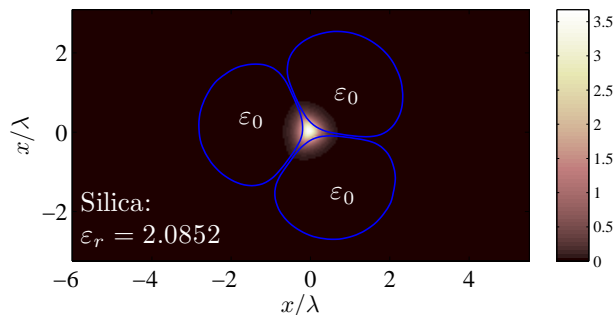


Fig. 17. Real part of the z -component of the complex Poynting vector, $\text{Re}(S_z)$, for the x polarized mode of the first (i.e. fundamental) mode pair ($\beta/k_0 = 1.1179$, $V = 0.38$). The plots have been normalized to unit total power in the cross-section shown.

propagating in the air holes. In Fig. 18 the results obtained with the SMT for this fraction are shown to be in very good agreement with those obtained in [39] by the FEM. As can be readily

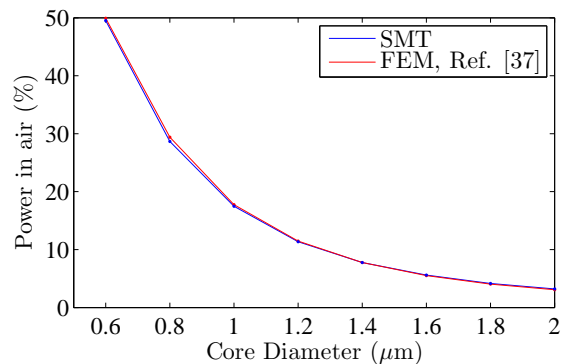


Fig. 18. Fraction of the power carried in air as function of core diameter, at $\lambda = 1.55\mu\text{m}$. The range of core diameters follows [39].

observed, for small core diameters, the fraction of power in air can reach 50%.

The dispersion curves and birefringence of the fundamental mode pair are shown in Fig. 19. The two modes are roughly polarized in the x and y directions, and the birefringence is again given by Eq. (18). It is interesting to note that, in contrast to the two-touching cylinders, the birefringence does not vanish in the low frequency limit. This occurs also in micro-structured fibers with elliptical veins [28]. As explained in [28], in contrast to a waveguide surrounded by air, the cladding of the micro-structured fiber is itself birefringent, and therefore, even though the field extends into the cladding in the low frequency limit, the birefringence does not vanish.

A common practice (and one we adopted) when analyzing micro-structured fibers, is to assume that the dielectric extends to infinity. This assumption implies that the modes are necessarily leaky, and their confinement losses can be determined by searching the complex β plane [40]. However, since the fiber has a solid core, the fundamental core mode is guided by total internal reflection. It is thus evanescent in air, and consequently, in a real fiber, any transverse radiation would be reflected back at the outer interface between the fiber and the surrounding air [6]. Confinement losses were therefore not calculated. On the other hand, the air jacket was not modelled either as its effect on the real part of the effective index is anticipated to be negligible away from cut-off.

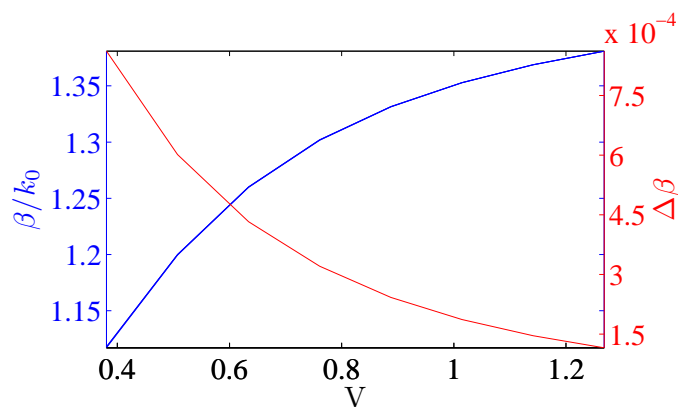


Fig. 19. Dispersion curves of the fundamental mode pair (plot on top of each other), and their difference, the birefringence. The range of V parallels the range of core diameters in Fig. 18.

8. Summary

We have expanded on the nature of spurious solutions in integral equation formulations, and described a simple method for their elimination. An adaptive sampling algorithm that takes advantage of the resulting form of the matrix singularity measure has been presented. Methods for resolving degenerate and nearly degenerate modes and for determining complex-valued eigenvalues have also been described. The various techniques have been combined in an efficient SMT-based waveguide solver, which has been demonstrated and characterized by a few numerical examples.

Acknowledgment

The authors would like to thank Mr. D. Peleg for useful discussions on generalized eigenvalue problems, and Dr. F. Poletti for sending us his data on the micro-structured fiber originally analyzed in [39]. This work was supported in part by the Israel Science Foundation.

Appendix

The effective indices of the modes of the almost-circular elliptical waveguide of Section 6.2 are given in this appendix.

Mode	β/k_0	Mode	β/k_0
1	1.28044536	12	1.18839178
2	1.27844665	13	1.13055864
3	1.11630744	14	1.12957045
4	1.08161112	15	1.10262569
5	1.06528023	16	1.10260786
6	1.06335567	17	1.08002675
		18	1.06404139
		19	1.05426579
		20	1.04860997
		21	1.00701462
		22	1.00691423

Table 3. Effective indices at $V = 1$.

Table 4. Effective indices at $V = 2$.

Mode	β/k_0	Mode	β/k_0	Mode	β/k_0	Mode	β/k_0
1	1.0389628	13	1.1334321	25	1.2207089	37	1.3249207
2	1.0403708	14	1.1334321	26	1.2207769	38	1.3250405
3	1.0403713	15	1.1432248	27	1.2362120	39	1.3267532
4	1.0453029	16	1.1512948	28	1.2428575	40	1.3288946
5	1.0554622	17	1.1515289	29	1.2468148	41	1.3627171
6	1.0626084	18	1.1515362	30	1.2543368	42	1.3648812
7	1.0706366	19	1.1644971	31	1.2736272	43	1.3658426
8	1.0706373	20	1.1659863	32	1.2736376	44	1.3681903
9	1.0712788	21	1.1715513	33	1.2790593	45	1.3950665
10	1.0715944	22	1.1800930	34	1.2795410	46	1.3952048
11	1.0875107	23	1.2103325	35	1.3086810		
12	1.0905844	24	1.2103332	36	1.3113530		

Table 5. Effective indices at $V = 3$.