Generalized source method for modeling nonlinear diffraction in planar periodic structures

Martin Weismann[a,b], Dominic F. G. Gallagher[b] and Nicolae C. Panoiu[a]

[a]Department of Electronic and Electrical Engineering, University College London, Torrington Place, London WC1E 7JE, UK;

[b]Photon Design Ltd, 34 Leopold Street, Oxford, OX4 1TW, UK

ABSTRACT

We present a new numerical method for the analysis of second-harmonic generation (SHG) in one- and two-dimensional (1D, 2D) diffraction gratings with arbitrary profile made of non-centrosymmetric optical materials. Our method extends the generalized source method (GSM), which is a highly efficient alternative to the conventional Fourier modal method, to quadratically nonlinear diffraction gratings. The proposed method consists of a two-stage algorithm. Initially, the electromagnetic field at the fundamental frequency is computed in order to obtain the second-harmonic polarization using the known second-order nonlinear susceptibility. Then the optical field at the second-harmonic frequency is computed using this polarization as an additional source term in the GSM. We show how to integrate this source term into the GSM framework without changing the structure of the basic algorithm. We use the proposed algorithm to investigate a doubly resonant mechanism that leads to strong enhancement of SHG in a nonlinear 2D circular GaAs grating mounted on top of a GaAs slab waveguide. We design this optical device such that slab waveguide modes at the fundamental and second-harmonic are simultaneously excited and phase matched by the grating. The numerically obtained resonance frequencies show good agreement with analytically computed resonance frequencies of the unperturbed slab waveguide.

Keywords: Diffraction gratings, Generalized source method, Fourier modal method (FMM), Rigorous coupled-wave analysis (RCWA), Nonlinear optics, Second harmonic generation, Iterative solver

1. INTRODUCTION

Diffraction gratings are ubiquitous devices used in a broad array of optics and photonics applications, including phase control,[1] frequency-selective surfaces,[2,3] local field enhancement for photovoltaic applications[4,5] or design of linear and nonlinear metamaterial surfaces.[6,7] One of the main reasons why diffraction gratings have such a multitude of applications and functionalities is that, due to the periodic nature of their structure, the optical modes supported by these gratings are markedly different from the modes of homogeneous optical media. In particular, the frequency dispersion of the Bloch modes of diffraction gratings is strongly dependent on the material and geometrical properties of the unit cell, which allows one to tailor the spectral response of the grating so as to suit a broad range of applications. This same design flexibility can also be used to engineer the optical near-field generated by diffraction gratings, a property that is particularly relevant in nonlinear optics applications. Nonlinear optical effects can be exploited to design devices for a variety of technological applications, including frequency conversion for all-optical signal processing,[8,9,10] nonlinear optical microscopy,[11,12,13,14] and nonlinear data storage.[15] While rewarding, this design flexibility poses major theoretical challenges as analytical solutions for the Bloch modes of diffraction gratings only exist in very few cases. This has generated continuous interest in the development of accurate and efficient numerical methods for modeling the linear and nonlinear optical response of diffraction gratings. This paper presents such a method for the computational study of second-harmonic generation (SHG) in 1D and 2D diffraction gratings.

The starting point of our work is the generalized source method (GSM), which was introduced in Ref. [16], substantially improved in Ref. [17], and which has proven to be an efficient alternative to the Fourier modal

Corresponding author: Martin Weismann, E-mail: martin.weismann.12@ucl.ac.uk

Optical Modelling and Design III, edited by Frank Wyrowski, John T. Sheridan, Jani Tervo, Youri Meuret, Proc. of SPIE Vol. 9131, 913115 · © 2014 SPIE
CCC code: 0277-786X/14/$18 · doi: 10.1117/12.2052247

method (FMM) also known as rigorous coupled-wave analysis (RCWA) for the exact analysis of one- and two-dimensional (1D, 2D) corrugated gratings, see Ref. [18]. In a simplified description, the GSM amounts to decomposing an arbitrary, complicated structure into a simple background structure with known analytical solution for any source-term and the remaining difference structure. Algorithmically, this difference structure gives rise to an artificial source term in the background structure, which in turn defines the electric field itself. This implicit description of the electromagnetic field results in a linear system of equations, which can be solved efficiently using an iterative solver. Despite its broad applicability, the current version of the GSM can be applied only to photonic structures that operate in the linear regime. Despite this, since the GSM is based on expressing the electric field in terms of a generalized, i.e. unphysical, electromagnetic source, it can naturally be extended to the computation of electromagnetic fields that are generated by an actual, i.e. physical, source embedded in a structure, as it is the case in second harmonic generation. The extension of the GSM to this important nonlinear optics phenomenon represents the main topic of this study.

The paper is organized as follows: Section 2 introduces the mathematical and physical model for SHG in periodic structures and the numerical algorithm used in our study. In Section 3 we illustrate how the nonlinear GSM can be applied to find the linear and nonlinear optical response of 1D and 2D diffraction gratings that contain quadratically nonlinear materials. In the last section the main results are summarized.

2. MATHEMATICAL AND PHYSICAL MODEL

This section is devoted to a brief outline of the main ideas and formalism on which the current generalized source method is based and present its extension to the nonlinear case of SHG in diffraction gratings. In addition, we investigate the convergence characteristics of this newly developed nonlinear GSM and validate its implementation by applying it to a generic test problem. To begin with, let us consider a quadratically nonlinear optical medium and a harmonic electric field, $\mathcal{E}(\mathbf{x},t)$, propagating in the medium at the fundamental frequency (FF), ω_1,

$$\mathcal{E}(\mathbf{x},t) = \mathbf{E}(\mathbf{x})e^{i\omega_1 t} + c.c.,$$

where $\mathbf{x}$ and t are the position vector and time, respectively. Through nonlinear interaction with the optical medium, this field gives rise to a nonlinear polarization, $\mathcal{P}(\mathbf{x},t) = \mathbf{P}^{\mathrm{NL}}(\mathbf{x})\exp(i\Omega t) + c.c.$, which oscillates at the second-harmonic (SH) frequency, $\Omega = 2\omega_1$. This polarization represents the source of the electromagnetic field generated at the SH and is related to the excitation field via a third-order nonlinear susceptibility tensor, $\chi^{(2)}$:

$$P_\gamma^{\mathrm{NL}}(\mathbf{x}) = \sum_{\alpha,\beta} \chi^{(2)}_{\gamma\alpha\beta} E_\alpha(\mathbf{x}) E_\beta(\mathbf{x}). \tag{1}$$

In this equation and what follows, Greek indices take the values of the Cartesian coordinates, x, y, z.

Since in this study we use the undepleted pump approximation, there is no influence of the SH field on the fundamental field. In other words, the nonlinear polarization generated by the SH field at the FF is neglected, which for most experimental setups is a valid approximation. Under these conditions, the solution process is straightforward: *i*) Calculate the field at the FF, ω_1; *ii*) evaluate the nonlinear polarization generated at the SH via Eq. (1); and *iii*) calculate the field at the SH. The first and last steps are performed using the linear and nonlinear (extended) versions of the GSM, respectively.

2.1 Description of the generalized source method

We will follow the derivation in Ref. [18] and show how to incorporate the additional source polarization (1) into the generic GSM. Thus, the GSM is a rigorous method for solving the electromagnetic wave equation

$$\nabla \times \nabla \times \mathbf{E} - \omega^2 \mu_0 \varepsilon \mathbf{E} = i\omega\mu_0 \mathbf{J}^{\mathrm{ext}} \tag{2}$$

with an external source current density, $\mathbf{J}^{\mathrm{ext}}$. To this end, one considers a simple background structure with permittivity distribution, $\varepsilon_b(\mathbf{x})$, for which the linear solution operator $\mathcal{N}_b$ to (2) with $\varepsilon = \varepsilon_b$ is known for any source $\mathbf{J}$, i.e.

$$\mathbf{E} = \mathcal{N}_b(\mathbf{J}). \tag{3}$$

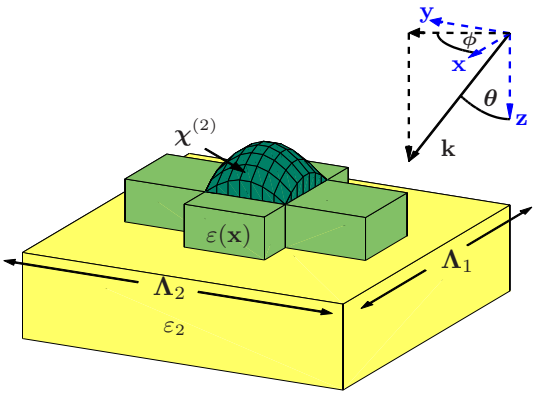

Figure 1: Setting for the GSM: Periodic cell of a 2D corrugated grating with linear and nonlinear materials and direction of incident wave $\mathbf{k}$.

This can be used to rewrite the solution of (2) as

$$\mathbf{E} = \mathbf{E}_{\text{ext}} + \mathcal{N}_b(\mathbf{J}^{\text{gen}}) = \mathcal{N}_b(\mathbf{J}^{\text{ext}}) + \mathcal{N}_b(-i\omega(\varepsilon - \varepsilon_b)\mathbf{E}), \tag{4}$$

where $\mathbf{E}_{\text{ext}} = \mathcal{N}_b(\mathbf{J}^{\text{ext}})$. The generalized source term $\mathbf{J}^{\text{gen}} = -i\omega(\varepsilon - \varepsilon_b)\mathbf{E}$ is an artificial and unphysical source term that arises due to the decomposition of the total permittivity ε. At the FF $\omega = \omega_1$, the external source term $\mathbf{J}^{\text{ext}}$ is defined such that the external field $\mathbf{E}^{\text{ext}}$ is an incident plane wave. At the SH, $\omega = \Omega$ in (2), the external source term is given in terms of the nonlinear polarization $\mathbf{P}^{\text{NL}}$ from Eq. (1): $\mathbf{J}^{\text{ext}} = -i\Omega\mathbf{P}^{\text{NL}}$. This implicit description of the solution $\mathbf{E}$ is the generic form of the GSM. The actual expression for the operator $\mathcal{N}_b$, the discretization procedure associated with the representation of the periodic fields, and an efficient algorithm for finding the discretized fields will be briefly described now. For a more detailed description of the linear GSM we refer the reader to Ref. [18].

We consider a two-dimensional corrugated grating, as depicted in Fig. 1. It consists of the grating region $0 < z \leq h$, with grating vectors $\mathbf{K}_1$ and $\mathbf{K}_2$ lying in the transverse (x, y) plane, and a given permittivity distribution, $\varepsilon(\mathbf{x})$, which is a Λ-periodic function of the x and y coordinates. This slab is sandwiched between the cover and the substrate, which consist of linear materials with permittivity ε_1 and ε_2, respectively. We assume that only the grating region contains nonlinear material, i.e. $\chi^{(2)} = 0$ if $0 < z$ or $z > h$. Due to its periodicity, the permittivity of the structure can be decomposed into a Fourier series

$$\varepsilon(x, y, z) = \sum_{n=-\infty}^{\infty} \varepsilon_n(z) e^{i[(n_1 K_{1x} + n_2 K_{2x})x + (n_1 K_{1y} + n_2 K_{2y})y]}, \tag{5}$$

where the sum over $n \equiv (n_1, n_2)$ is to be understood as a double infinite sum over n_1 and n_2, $n_{1,2} = -\infty, \ldots, \infty$. Because in the transverse plane $\varepsilon(\mathbf{x})$ is a periodic function, Bloch's theorem guarantees that the grating field can be expressed as a series of functions that are periodic on the transverse coordinates x and y,

$$\mathbf{E}(\mathbf{x}) = \sum_{n=-\infty}^{\infty} \mathbf{E}_n(z) e^{i(k_{nx}x + k_{ny}y)}, \tag{6}$$

where $k_{nx/y} = k_{0x/y} + n_1 K_{1x/y} + n_2 K_{2x/y}$ are the projections of the wavevector of the nth order diffraction mode. The principal direction of the central diffraction order, which corresponds to $n_1 = n_2 = 0$, is given by $\mathbf{k}_0 = k_0(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$, where $k_0 = \omega\sqrt{\varepsilon_0\mu_0}$ is the wavenumber in free space.

Since we do not include in our analysis nonlinear optical effects in the substrate and cover materials and we assume that the nonlinear polarization $\mathbf{P}^{\text{NL}}(\mathbf{x})$ arises due to a periodically distributed field (6) at the FF, the external current $\mathbf{J}^{\text{ext}} = -i\Omega\mathbf{P}^{\text{NL}}$ at the SH can also be expressed as a Fourier series

$$\mathbf{J}^{\text{ext}} = -i\Omega\mathbf{P}^{\text{NL}} = -i\Omega \sum_{n=-\infty}^{\infty} \mathbf{p}_n(z) e^{i(k_{nx}x + k_{ny}y)}.$$

Moreover, the generalized source term in (4) is given by the infinite sum

$$\mathbf{J}^{\mathrm{gen}} = -i\omega(\varepsilon - \varepsilon_b)\mathbf{E} = -i\omega \sum_{n=-\infty}^{\infty} \mathbf{j}_n(z) e^{i(k_{nx}x + k_{ny}y)}, \tag{7}$$

where the coefficients $\mathbf{j}_n(z)$ depend on all field modes, $\mathbf{E}_m$

$$\mathbf{j}_n(z) = [(\varepsilon(z) - \varepsilon_b(z))\mathbf{E}(z)]_n = \sum_{m=-\infty}^{\infty} [\varepsilon(z) - \varepsilon_b(z)]_{n-m} \mathbf{E}_m(z). \tag{8}$$

Since the work of Lalanne[19] and Li,[20,21] it has been well known that the factorization of products of periodic functions with a *finite* number of terms has to be performed with care, namely the correct approach is dictated by the continuity properties of the factors and the product: products of functions with non-simultaneous discontinuities (e.g., the tangential component of the electric displacement field $\mathbf{D}_{\parallel} = \varepsilon\mathbf{E}_{\parallel} + \mathbf{P}_{\parallel}^{\mathrm{NL}}$) can be factorized using Laurent's product rule, whereas continuous products of functions with simultaneous discontinuities (the normal component $\mathbf{D}_{\perp} = \varepsilon\mathbf{E}_{\perp} + \mathbf{P}_{\perp}^{\mathrm{NL}}$ is continuous at interfaces, whereas ε and $\mathbf{E}$ can be discontinuous) can be factorized using the inverse rule. In the context of numerical methods for periodic structures in the Fourier space, such as the FMM and GSM, the violation of the correct factorization rules causes slow convergence with respect to the number of retained diffraction orders. In this work, the correct factorization rules are applied in the linear part of the calculations, where $\mathbf{P}^{\mathrm{NL}} = 0$. For the nonlinear part, the same factorization rules for $\mathbf{D}$ are applied as in the linear part, which may lead to slow convergence. However, numerical results presented in Section 2.2 suggest that in our approach this is not a practical limitation. This potential shortcoming of the algorithm will be addressed in a future publication.

With this in mind, we define a local coordinate system that is used to decompose $\mathbf{D}$ and $\mathbf{E}$ into normal and tangential components at the discontinuities of ε (i.e., the interface of the grating structure) and define the block Toeplitz-matrices of its Fourier series components, $\Gamma_{\alpha\beta}$, $\alpha, \beta = x, y, z$. The correct factorization rules for $[\varepsilon(z) - \varepsilon_b(z)\mathbf{E}(z)]$ when using a finite number of diffraction orders $N = (N_1, N_2)$ in (8) can then be written as

$$j_{\alpha,n}(z) = \sum_{m=-N}^{N} \sum_{\beta} \hat{W}_{\alpha\beta,n-m}(z) E_{\beta,m}(z), \tag{9}$$

where the factorization matrix $\hat{W}_{\alpha\beta}(z) = \delta_{\alpha\beta}([\varepsilon(z)/\varepsilon_b] - [\varepsilon_b/\varepsilon(z)]^{-1})) - [1 - \varepsilon(z)/\varepsilon_b]\Gamma_{\alpha\beta}$ combines the normal field information contained in Γ with the structure geometry defined by ε.

The background solution operator $\mathcal{N}_b$ can be expressed as the convolution of the Green's function[16,22] $\mathbf{G}_n(z, z')$ corresponding to the background permittivity distribution $\varepsilon_b(z \geq h) = \varepsilon_1$, $\varepsilon_b(0 < z < h) = \varepsilon_b$, and $\varepsilon_b(z < 0) = \varepsilon_2$, with the periodic source distribution $\mathbf{j}_n(z)$. The solution at each layer z in this slab-background structure is described as a the upward and downward TE- and TM-amplitudes $\mathbf{a}_n(z) = (a_{en}^{+}(z), a_{en}^{-}(z), a_{hn}^{+}(z), a_{hn}^{-}(z))$ of a modified electric field

$$\mathbf{a}_n(z) = \bar{Q}_n \tilde{\mathbf{E}}_n(z) = \bar{Q}_n \left[\mathbf{E}_n(z) + i \frac{j_{nz}}{\omega\varepsilon_b} \mathbf{e}_z \right], \tag{10}$$

where the matrices $\bar{Q}_n$ and Q_n define a decomposition in Cartesian coordinates into TE- and TM amplitudes of upward and downward propagating waves and vice-versa, respectively. The modified field in Eq. (10) is introduced to eliminate the non-propagating term $j_{nz}(z)$ in the Green's function. The amplitudes of the solution operator can than be written as

$$\mathbf{a}_n(z) = \int_{z'=0}^{h} R_n(z, z') \bar{Q}_n \mathbf{j}_n(z') dz', \tag{11}$$

where $R_n(z, z')$ incorporates the propagation of TE- and TM-polarized plane waves inside the slab-background structure and their reflections at its top and bottom interfaces.

Combining equations (11), (9), and (4) we obtain a system of Fredholm integral equations of the second kind for the unknown amplitudes

$$\mathbf{a}_n(z) = \int_{z'=0}^{h} R_n(z,z')\bar{Q}_n \sum_{m=-N}^{N}\sum_{\beta} W_{\alpha\beta,n-m}(z')Q_m\mathbf{a}_m(z')dz' + \mathbf{a}_n^0(z), \quad \text{for } n=-N,\ldots,N, \tag{12}$$

where $\mathbf{a}_n^0(z)$ are either the amplitudes of the incident plane wave electric field in the background structure at the FF or the amplitudes of the external source field $\mathcal{N}_{bn}(J_{\text{ext}})$ at the SH, given by

$$\mathbf{a}_n^0(z) = -i\Omega \int_{z'=0}^{h} R_n(z,z')\bar{Q}_n\mathbf{p}_n(z')dz'. \tag{13}$$

The system (12) of integral equations is discretized at N_l points, $z_p = (p-0.5)h/N_l, p=1,\ldots,N_l$, using the midpoint rule, resulting in a system of linear equations

$$\mathbf{a}_{np} = \frac{h}{N_l}\sum_{q=1}^{N_l} R_n(z_p,q)\bar{Q}_n \sum_{m=-N}^{N}\sum_{\beta} W_{\alpha\beta,n-m}(z_q)Q_m\mathbf{a}_m(z_q) + \mathbf{a}_n^0(z), \text{ for } n=-N,\ldots,N, p=1,\ldots,N_l. \tag{14}$$

The total number of unknowns in (14) is $4N_oN_l$, with $N_o=(2N+1)^2$, and becomes prohibitively large for direct solvers to be used in practical cases.

In order to take full advantage of an iterative solver, the system (14) is reformulated[18] to

$$\mathbf{a} = \mathbf{a}^0 + RPU(M-QRPU)^{-1}Q\mathbf{a}^0 = \mathbf{a}^0 + RPUA^{-1}Q\mathbf{a}^0, \tag{15}$$

where $\mathbf{a}$ and $\mathbf{a}^0$ are the vectors corresponding to $\mathbf{a}_n(z_p)$ and $\mathbf{a}_n^0(z_p)$, respectively, and P and Q are block matrices corresponding to $\bar{Q}_n$ and Q_n, respectively. The matrix R corresponds to $R_n(z_p,z_q)$ and has Toeplitz-structure with respect to the layer indices p and q. The matrices U and M are defined so that $W=UM^{-1}$ but they do not contain inverted matrices themselves. Additionally, U and M are block-Toeplitz-Toeplitz-matrices with respect to the mode index $n=(n_1,n_2)$. Due to the Toeplitz-properties of its sub-matrices, multiplication of a vector $\mathbf{b}$ of size $3N_lN$ with the $3N_lN\times 3N_lN$ system matrix A can be performed in $\mathcal{O}(NN_l\log(N_l)\log(N))$ operations instead of $\mathcal{O}(N^2N_l^2)$ using standard matrix vector multiplication. Therefore, the use of an iterative solver based on matrix vector multiplications like GMRES is fully justified.

2.2 Numerical example and some remarks on convergence

In order to illustrate the convergence and runtime behavior of the proposed method, let us consider a basic two-dimensional grating structure, shown schematically in Fig. 1. It consists of a square periodic distribution of circular disks with height, $h=0.2\,\mu$m, and radius $r=\Lambda/3$, where $\Lambda=2\,\mu$m is the grating period and is assumed to be the same for both directions. The index of refraction of the substrate and the circular disks are chosen as $n_m=2.5$, whereas the cover region is air with $n_c=1$. We use an isotropic nonlinear susceptibility tensor $\chi^{(2)}_{jkl}=\bar{\chi}\delta_{jk}\delta_{kl}$, with $\bar{\chi}=10^{-8}\,\mathrm{m\,V^{-1}}$. The incident wave, at a wavelength of $\lambda=1.5\,\mu$m, is specified by the angles $\theta=30^\circ$ and $\phi=70^\circ$ and is 45° polarized with respect to the grating lattice vectors.

We perform simulations for increasing number of spatial harmonics, $N_o=(2N+1)^2$, $N\in\{3,5,\ldots,19\}$, and GSM-layers $N_l\in\{50,100,200,400\}$ and compute the total diffraction efficiencies for transmission ($T_{F/S}$) and reflection ($R_{F/S}$) at the FF (F) and SH (S), the number N_F^{iter} and N_S^{iter} of GMRES iterations necessary to solve (15) at the FF and SH, respectively, and the GSM runtime in minutes needed to complete the linear (t_F^{GSM}) and nonlinear (t_S^{GSM}) part of the computations. As a comparison for the linear results we use the diffraction efficiencies T_F^{FMM} calculated by using an implementation of the FMM. This implementation relies on a local normal vector field[23,24] to incorporate Li's rule for fast convergence, thus being particularly well suited for the simulation of gratings containing rounded-shape diffractive elements. The implementations of both methods in MATLAB were modestly optimized and all simulations were performed on a single 2.5GHz core for a fair comparison of runtimes.

Table 1: Convergence of linear and nonlinear diffraction efficiencies for a circular grating.

N	N_l	T_F	T_F^{FMM}	E_F	R_S	T_S	N_F^{iter}	t_F^{GSM}	t^{FMM}	N_S^{iter}	t_S^{GSM}
3	50	0.868117	0.867828	$7.55 \cdot 10^{-4}$	$5.28278 \cdot 10^{-19}$	$9.24272 \cdot 10^{-18}$	57	0.3	0.0	105	0.6
	100	0.868127		$7.55 \cdot 10^{-4}$	$5.28158 \cdot 10^{-19}$	$9.23861 \cdot 10^{-18}$	57	0.7		105	1.2
	200	0.868130		$7.55 \cdot 10^{-4}$	$5.28128 \cdot 10^{-19}$	$9.23759 \cdot 10^{-18}$	57	1.3		105	2.3
	400	0.868130		$7.55 \cdot 10^{-4}$	$5.28120 \cdot 10^{-19}$	$9.23733 \cdot 10^{-18}$	57	2.4		105	4.5
7	400	0.869809	0.869903	$8.68 \cdot 10^{-5}$	$7.00750 \cdot 10^{-19}$	$3.06343 \cdot 10^{-17}$	63	9.4	0.7	264	51.2
11	400	0.869979	0.870014	$2.85 \cdot 10^{-5}$	$6.41110 \cdot 10^{-19}$	$3.24153 \cdot 10^{-17}$	64	19.1	3.5	280	99.2
15	400	0.870036	0.870048	$5.54 \cdot 10^{-6}$	$6.38049 \cdot 10^{-19}$	$3.28962 \cdot 10^{-17}$	65	31.2	18.5	284	169.0
17	400	0.870026	0.870045	$7.54 \cdot 10^{-6}$	$6.41065 \cdot 10^{-19}$	$3.30973 \cdot 10^{-17}$	65	42.5	41.5	286	219.4
19	400	0.870038	0.870050	$9.31 \cdot 10^{-7}$	$6.42491 \cdot 10^{-19}$	$3.32217 \cdot 10^{-17}$	65	55.7	76.7	287	282.7
21	50	0.870024	0.870052	$1.48 \cdot 10^{-6}$	$6.45567 \cdot 10^{-19}$	$3.33974 \cdot 10^{-17}$	65	13.2	142.4	288	68.0
	100	0.870038		$1.49 \cdot 10^{-6}$	$6.43979 \cdot 10^{-19}$	$3.33348 \cdot 10^{-17}$	65	17.2		288	83.2
	200	0.870042		$1.49 \cdot 10^{-6}$	$6.43583 \cdot 10^{-19}$	$3.33191 \cdot 10^{-17}$	65	31.0		288	158.0
	400	0.870043		$1.49 \cdot 10^{-6}$	$6.43485 \cdot 10^{-19}$	$3.33152 \cdot 10^{-17}$	65	62.1		289	403.0

The results of these calculations are summarized in Table 1. The transmission T_F at the fundamental harmonic obtained by the GSM shows good agreement with the value predicted by the FMM, T_F^{FMM}. The energy balance between the incident, reflected, and transmitted field intensities is satisfied up to a relative error $E_F := |1 - T_F - R_F|$ of approximately 10^{-6} in all simulations with a sufficiently large number of unknowns. The convergence at the SH is slower than at the FF: reflected and transmitted intensities at the SH tend to values of $R_S = 6.434\,846 \times 10^{-19}\,\mathrm{W\,m^{-2}}$ and $T_S = 3.331\,522 \times 10^{-17}\,\mathrm{W\,m^{-2}}$, in the case of an incident wave with intensity of $1\,\mathrm{W\,m^{-2}}$; however, increasing the number of layers to $N_l = 400$ still changes the result by 0.6%. Yet for most practical purposes one considers this accuracy to be satisfactory, especially for the comparison with experimental results.

The GSM-runtimes $t_{F/S}^{\mathrm{GSM}}$ are dominated by the computation of the numerical solution of (15) and therefore by the number of iterations $N_{F/S}^{\mathrm{iter}}$, which is constant with respect to N_l and increases with N_o. The calculations of the SH field require a substantially higher runtime due to an approximately fourfold increase in the number of iterations, $N_S^{\mathrm{iter}} \approx 4N_F^{\mathrm{iter}}$. The simulation results show good agreement with the asymptotic runtime predicted in Section 2.1, i.e. $\mathcal{O}(N_{S/F}^{\mathrm{iter}} N_o \log(N_o) N_l \log(N_l))$. The FMM exhibits a runtime of $\mathcal{O}(N_o^3)$ since it requires to find the solution of an eigenvalue problem of size $2N_o$. In particular, our tests suggest that the linear GSM begins to outperform the FMM when $N \geq 15$, at $N_l = 100$. Moreover, due to the use of GMRES without restart as an iterative solver, which requires that all the past iterators perform a step, the memory consumption of the GSM is $\mathcal{O}(N_{F/S}^{\mathrm{iter}} N_o N_l)$.

3. APPLICATION TO NONLINEAR MODE MIXING IN PERIODICALLY TEXTURED SLAB WAVEGUIDES

In this section we illustrate how our method can be applied to a case of practical interest, namely to the generation of the SH via phase-matching techniques. To this end, we consider a slab waveguide made of a quadratically nonlinear optical material. We choose this material to be GaAs ($n_g = 3.4$), primarily because its large nonlinearity makes it suitable for a broad array of applications in nonlinear optics. Thus, GaAs crystallizes in the zincblende structure, so that it belongs to the crystal point group $\bar{4}3m$. As a result, the only non-vanishing components of its second-order susceptibility tensor are $\chi^{(2)}_{xyz} = \chi^{(2)}_{xyz} = \chi^{(2)}_{yxz} = \chi^{(2)}_{yzx} = \chi^{(2)}_{zxy} = \chi^{(2)}_{zyx} = 7.4 \times 10^{-10}\,\mathrm{m\,V^{-1}}$.[8] The slab is placed on a silica substrate ($n_s = 1.4$), the cover being air ($n_c = 1$). In this configuration, one can achieve efficient SH generation (SHG) when slab waveguide modes with frequencies ω_1 and $\Omega = 2\omega_1$ are simultaneously excited. Because of modal dispersion the propagation constants of these modes are not phase-matched, so that one customarily uses a diffraction grating to compensate for the residual wave vector mismatch. This diffractive element also insures that the incoming plane waves are coupled to and excite slab waveguide modes at the FF,

ω_1. A brief description of this grating-induced coupling mechanism as well as the results obtained by applying our numerical method to this photonic structure are presented in what follows.

3.1 Resonant grating-induced coupling of slab waveguide modes

In an lossless, asymmetric, dielectric slab waveguide of height h_w and refractive index n_1, placed in-between a cover and a substrate with refractive indices n_2 and n_3, respectively, an orthogonal set of solutions to the time-harmonic Maxwell's equations of the form $\mathbf{E} = \mathbf{E}_\nu(z)\exp(-i\beta_\nu x)$, $\nu \in \mathbb{N}$, are called the modes of the waveguide.[25] The propagation constant, β_ν, of the νth mode is determined from the dispersion relation

$$\kappa(\beta)h_w - \arctan\left[\frac{\gamma_2(\beta)}{\kappa(\beta)}\right] - \arctan\left[\frac{\gamma_3(\beta)}{\kappa(\beta)}\right] = \nu\pi, \tag{16}$$

where the reflection parameters $\kappa(\beta) = \sqrt{n_1^2k_0^2 - \beta^2}$ and $\gamma_{2,3}(\beta) = \sigma\sqrt{\beta^2 - n_{2,3}^2k_0^2}$ for TE ($\sigma=1$) and TM ($\sigma = n_1^2/n_{2,3}^2$) modes. By placing a grating, i.e. a Λ-periodic perturbation, on top of the waveguide the translational symmetry of the structure is broken and thus β_ν is constrained within the first Brillouin zone of the grating, $\beta_\nu \in \left(-\frac{K}{2}, \frac{K}{2}\right)$, where $K = 2\pi/\Lambda$ is the grating period in the reciprocal space. More specifically, the dispersion curve $\beta_\nu(\omega)$ is folded back into the first Brillouin zone

$$\beta_\nu^f(\omega) = \mathrm{mod}\left(\beta_\nu(\omega) + \frac{K}{2}, K\right) - \frac{K}{2}. \tag{17}$$

For normal incidence, at the wavelengths λ for which the folded (reduced) dispersion curve $\beta_\nu^f(\lambda)$ intersects the y-axis, the incoming plane wave excites a waveguide mode, because under these conditions the total tangent component of the wave vectors of the interacting waves adds up to a multiple of K. One expects a strong resonant effect at these wavelengths, which would manifest through steep changes of the reflection and transmission coefficients of the device and strong field enhancement in the slab waveguide.

In the context of SHG there are two kinds of resonances that can affect the nonlinear response of the considered device: First, *linear resonances* correspond to the grating-induced, resonant excitation of a waveguide mode at the FF by an incident wave with certain wavelength, λ_1. We expect then that the fundamental field is strongly enhanced and confined inside the waveguide, thus a stronger nonlinear polarization is created, which in turn excites a stronger field at the SH wavelength, $\lambda_2 = \lambda_1/2$. We will call this effect of the linear resonance at λ_1 on the nonlinear response an *inherited resonance* at λ_2. The second type of resonances are *intrinsic resonances*, which are observed when a waveguide mode with vanishing reduced wave vector is excited at a certain SH wavelength λ_2, namely $\beta_\nu^f(\lambda_2) = 0$. In this case the field at the SH is confined inside the waveguide and therefore shows remarkably high radiated SH intensity. Note that the intrinsic resonances are not necessarily grating coupled with the radiative continuum and therefore they can be viewed as nonlinear dark modes of the system. In particular, one expects that these resonances have large quality factors.

We call intrinsic and inherited resonances nonlinear modes, since they can only be observed at the SH. The SH wavelengths λ_2 for which both types of nonlinear modes are excited simultaneously are of particular interest because at these wavelengths, this doubly resonant mechanism of SHG can lead to a remarkably large nonlinear optical response of the device.

3.2 SHG in one-dimensional binary gratings

We first consider a 1D grating, which although has a simple structure provides us the necessary insights to completely understand the more complex 2D design. Thus, let us consider a 1D binary GaAs grating on top of a GaAs slab waveguide, placed on a silica substrate. The period of the grating with filling factor $\rho = 2/3$ is $\Lambda = 2\,\mu$m, the height of the grating region being $h_g = 50\,$nm. The height of the slab waveguide, h_w, is a free design parameter and varies from $0.3\,\mu$m to $0.8\,\mu$m. At the FF, we consider a normally incident plane wave with different polarizations (TE, TM, and 45°), with the wavelength λ_0 ranging from $1\,\mu$m to $6\,\mu$m. In our simulations we used $N = 5$ diffraction orders and $N_l = 35, \ldots, 85$ GSM-layers, each with height of 10 nm.

Consider the dispersion maps in Figure 2, which show the transmitted intensity I relative to the incident field intensity, determined for different wavelengths and device height, h_w. Blue (red) lines correspond to the

position of linear TE_ν (TM_ν) modes in the h_w-λ-space. Their mode numbers, $\nu = 0, 1, 2$, are encoded as solid, dashed and dotted lines, respectively. In the plots for the SH response, the inherited modes have been marked by red dots.

Figure 2a shows the transmitted intensity at λ_1 for a TM polarized incident plane wave with wavelength λ_1. The map is a smooth graph in the h_w-λ_1-plane, except for the resonances, which manifest themselves in a strong local change of the transmission. The position of these spikes agrees very nicely with the estimates from the analytical model. The slight differences can be explained by the fact that the analytical model considers a infinitely small perturbation of the waveguide whereas in the simulations, a grating of finite height, h_g, is mounted onto the waveguide. Decreasing h_g improves the agreement between the predictions of the theoretical model and simulations, yet as expected decreases the strength of the resonant effect. Note that for a 1D grating under incidence along the x-z or y-z-plane, the TE and TM components of the electric field decouple. Accordingly, only the linear TM modes are excited under TM-incidence.

Figure 2b shows the transmitted intensity at the SH wavelength $\lambda_2 = \lambda_1/2$ for the same incident fundamental field as in Figure 2a. This figure clearly shows an increase of the SH field due to the excitation of the inherited TM modes and nonlinear TE modes. No nonlinear TM modes are excited in this case. This is due to the particular anisotropic form of $\chi^{(2)}$: for a TM-incident fundamental field $\mathbf{E}(\mathbf{x}) = (E_x(\mathbf{x}), 0, E_z(\mathbf{x}))$ only the y-component (i.e. the TE component) of the nonlinear polarization is non-vanishing.

The SH transmitted intensity, I_S, at resonance is substantially higher than at off-resonance (e.g. $h_w = 0.6\,\mu$m and $\lambda_2 = 2.75\,\mu$m), which determines the reference value $I_S^{\mathrm{ref}} = I_S(0.6\,\mu\mathrm{m}, 2.75\,\mu\mathrm{m}) = 1.38 \times 10^{-18}\,\mathrm{W\,m^{-2}}$. Regarding the strength of the nonlinear resonances, several important conclusions can be drawn. *i*) The strongest intrinsic resonance (TE_1) at $h_w = 0.35\,\mu$m and $\lambda_2 = 1.245\,\mu$m has a relative intensity of $I_S^{\mathrm{nl}} = 1.66 \cdot 10^6 I_S^{\mathrm{ref}}$; *ii*) the strongest inherited resonance (TM_1) at $h_w = 0.68\,\mu$m and $\lambda_2 = 1.43\,\mu$m has a relative intensity of $I_S^{\mathrm{inh}} = 9.44 \cdot 10^9 I_S^{\mathrm{ref}}$; and *iii*) a *simultaneous* phase match is achieved at $h_w = 0.6925\,\mu$m and $\lambda_2 = 2.111\,\mu$m, which results in a relative intensity $I_S^{\mathrm{simult}} = 3.44 \cdot 10^{10} I_S^{\mathrm{ref}}$. The zoom-in of the phase matching region in Figure 2c shows the increase of intensity as an accumulation of the less pronounced effects of the intersecting branches of the intrinsic TE_0 mode (blue solid line) and inherited TM_0 mode (red dots) with relative intensity $1.26 \cdot 10^8 I_S^{\mathrm{ref}}$ and $1.96 \cdot 10^8 I_S^{\mathrm{ref}}$, respectively. As a general conclusion, our simulations indicate that the inherited resonances are stronger than the intrinsic ones for the considered device.

Figure 2d shows the transmitted intensity at the SH for an incident field at 45° polarization. All components of the fundamental electric field are non-vanishing, hence all components of $\mathbf{P}^{\mathrm{NL}}$ are non-vanishing as well and consequently all possible inherited and nonlinear TE- and TM-resonances can be observed. One can see a substantial increase of the SH intensity due to the excitation of inherited and intrinsic modes. The strongest SH is generated at the inherited TM resonance at $h_w = 0.68\,\mu$m and $\lambda_2 = 1.43\,\mu$m, with a relative intensity of $\tilde{I}_S^{\mathrm{inh}} = 2.36 \cdot 10^9 I_S^{\mathrm{ref}}$. Note that the maximum SHG-intensity at 45° polarization $I_S^{\mathrm{inh}} \approx 4\tilde{I}_S^{\mathrm{inh}}$. This is not surprising, since only half of the intensity of the incident field is due to the TM-polarization.

3.3 SHG in two-dimensional circular gratings

We investigate now a similar 2D device consisting of a circular grating with radius $5/6\,\mu$m and square unit cell of side-length $\Lambda = 2\,\mu$m. The height of the waveguide, h_w, is set to vary from $0.5\,\mu$m to $0.8\,\mu$m. The remaining

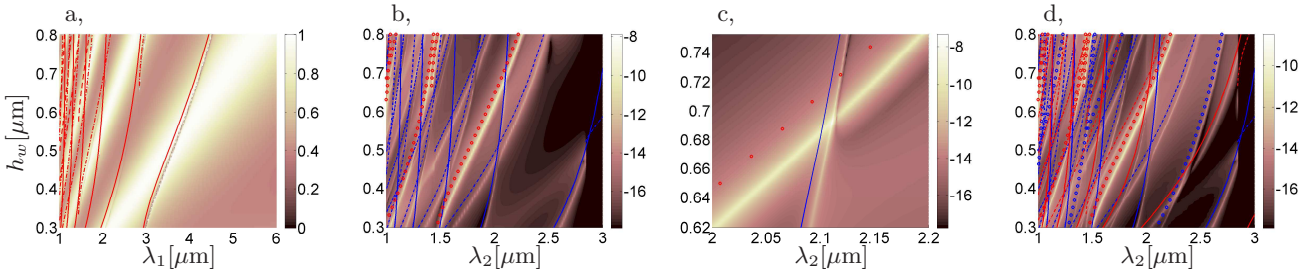

Figure 2: Wavelength scan of normalized intensity in transmission direction at fundamental λ_1 (a) and second harmonic wavelengths λ_2 (b,c,d; logarithmic scale) for different waveguide heights h_w for TM (a,b,c) and 45° (d) polarization and closeup of phase-matched region (c).

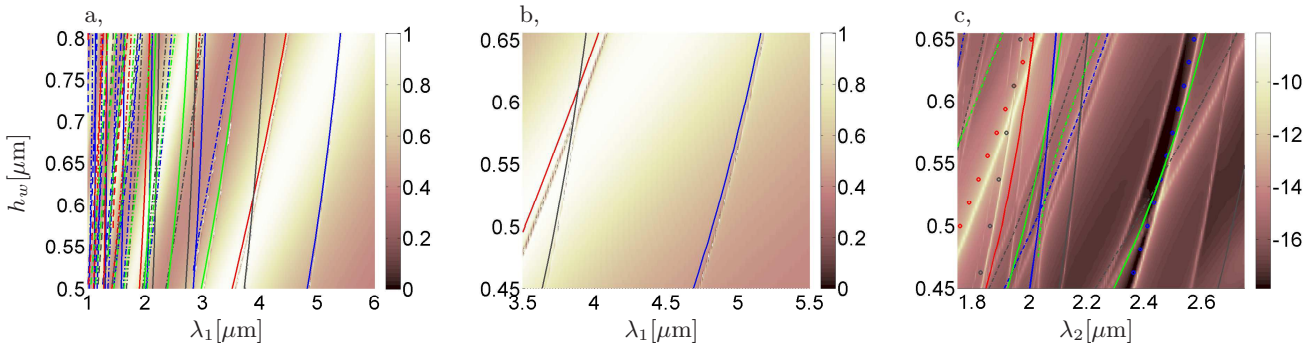

Figure 3: Wavelength scan of normalized intensity in transmission direction at fundamental λ_1 (a,b) and second harmonic wavelengths λ_2 (c; logarithmic scale) for different waveguide heights h_w and zoom-in of several phase-matched regions (b, c).

specification of the device is as in the 1D design. For the simulations, a total of $N_o = 81$ diffraction orders and $N_l = 50, \ldots, 80$ GSM-layers with height of 10 nm each are used. The results for normal incidence and TM polarization are summarized in Figure 3.

Figure 3a shows the transmitted relative intensity at the FF. One can identify deep drops in transmission near the analytically predicted values of the resonance wavelengths. Note that in this case linear TE and TM modes for period Λ can be excited. Since the structure is also periodic with periods $\tilde{\Lambda}_n = \sqrt{n}\Lambda$, $n = 2, 3, \ldots$, modes for $\tilde{\Lambda}_n$ can be excited, too. The analytically determined values of resonance wavelengths for $n = 2$ for linear TE (TM) modes are indicated by gray (green) lines and explain several of the peaks in the spectrum. Due to the very large number of waveguide modes that can be excited for smaller wavelengths of the incident wave, and the fact that the analytically predicted values do not exactly coincide with the simulation results, it is not always clear which mode is associated to a particular drop in the simulated transmission spectrum. A comparison of the electromagnetic field profile at each of the simulated resonances with the analytical profile of $\mathbf{E}_\nu(z)$ and their neighborhood can explain which mode is responsible for their excitation.

We focus the following discussion of the SH response to the region $h_w = 0.45\,\mu$m to $0.65\,\mu$m and $\lambda_1 = 3.5\,\mu$m to $5.5\,\mu$m around several phase-matched resonances. Figure 3b shows the relative transmitted intensity at the FF. We can clearly identify the excitation of the TM_0 (red) and TE_0 (blue) modes, similar to the 1D case (see Figure 2a). The average deviation from the analytical estimate for the TM_0 in 2D for $0.55\,\mu\text{m} \le h_w \le 0.65\,\mu$m is $\Delta\lambda_1^{2D} = 55$ nm whereas the average deviation in 1D is $\Delta\lambda_1^{1D} = 85$ nm. We speculate that the slightly better agreement in 2D is due to the fact that the perturbation of the waveguide by a circular grating with radius $\Lambda/3$, i.e. a filling factor of $\pi/9 = 0.3491$, in 2D is lower than the perturbation by a 1D binary grating with filling factor of $2/3$. Apart from the modes for $\Lambda = 2\,\mu$m, the simulation results show resonances near the TM_0 mode for period $\tilde{\Lambda}_2$, indicated by the gray line.

Figure 3c shows the relative transmitted intensity at the SH. We see good agreement with most of the analytical estimates. Again, we choose a non resonant point in the $h_w - \lambda$ space as reference $I_S^{\text{ref,2D}} = I_S^{\text{2D}}(0.52\,\mu\text{m}, 2.29\,\mu\text{m}) = 7.59 \times 10^{-18}\,\text{W m}^{-2}$. There are two types of simultaneously excited modes, which have no correspondent to the 1D-case: at $h_w = 0.59\,\mu$m and $\lambda_2 = 1.945\,\mu$m the inherited TM_0 mode for period Λ and the inherited TM_0 mode for period $\tilde{\Lambda}_2$ are simultaneously excited and result in radiated SH of $I_S = 3.35 \cdot 10^8 I_S^{\text{ref,2D}}$ relative intensity, and at $h_w = 0.52\,\mu$m and $\lambda_2 = 1.353\,\mu$m the inherited TE_0 mode for period Λ and the intrinsic TM_1 mode for period $\tilde{\Lambda}_2$ are excited simultaneously and result in radiated SH of $I_S^{\text{2D}} = 1.00 \cdot 10^9 I_S^{\text{ref,2D}}$. Similar to the 1D case, the inherited resonances result in a stronger generated SH than the intrinsic resonances and the strongest nonlinear response can be observed for simultaneous excitation of inherited and intrinsic modes.

4. CONCLUSION

We have extended the generalized source method to solve the problem of second harmonic generation in periodic structures with quadratically nonlinear optical materials and have discussed the main steps of the theoretical

derivation. In a numerical benchmark study we achieved similar convergence and favorable runtime as compared to a reference implementation of the FMM for the linear part of the calculations and sufficiently good convergence in the nonlinear part of the algorithm. The method was used to design a nonlinear photonic device for second-harmonic generation, consisting of a 1D or 2D grating textured slab waveguide. We revealed that simultaneous, resonant excitation of waveguide modes at the fundamental and second harmonic is possible in this device, which leads to order-of-magnitude enhancement of the generated second harmonic intensity. The simulation results show excellent agreement with analytically predicted values of the frequencies of the waveguide modes, which further validates our numerical method. We found several designs of the device which exhibit substantial second harmonic field generation up to 10 orders of magnitude stronger than a non-optimized reference design.

ACKNOWLEDGMENTS

The work of M. W. was supported through a UCL Impact Award graduate studentship funded by UCL and Photon Design Ltd.

REFERENCES

1. G. J. Tearney, B. E. Bouma, and J. G. Fujimoto, "High-speed phase- and group-delay scanning with a grating-based phase control delay line," *Opt. Lett.* **22**, pp. 1811–1813, 1997.
2. D. Sievenpiper, L. J. Zhang, R. F. J. Broas, N. G. Alexopolous, and E. Yablonovitch, "High-impedance electromagnetic surfaces with a forbidden frequency band," *IEEE Trans. Microwave Theory Tech.* **47**, pp. 2059–2074, 1999.
3. J. A. Porto, F. J. Garcia-Vidal, and J. B. Pendry, "Transmission resonances on metallic gratings with very narrow slits," *Phys. Rev. Lett.* **83**, pp. 2845–2848, 1999.
4. F. J. Garcia-Vidal, H. J. Lezec, T. W. Ebbesen, and L. Martin-Moreno, "Multiple paths to enhance optical transmission through a single subwavelength slit," *Phys. Rev. Lett.* **90**, 2003.
5. N. C. Panoiu and R. M. Osgood, "Enhanced optical absorption for photovoltaics via excitation of waveguide and plasmon-polariton modes," *Opt. Lett.* **32**, pp. 2825–2827, 2007.
6. J. Shen, P. Catrysse, and S. Fan, "Mechanism for designing metallic metamaterials with a high index of refraction," *Phys. Rev. Lett.* **94**, p. 197401, 2005.
7. W. Fan, S. Zhang, N. C. Panoiu, A. Abdenour, S. Krishna, R. M. Osgood, K. J. Malloy, and S. R. J. Brueck, "Second harmonic generation from a nanopatterned isotropic nonlinear material," *Nano Lett.* **6**, pp. 1027–1030, 2006.
8. R. W. Boyd, *Nonlinear Optics, 3rd ed.*, Academic Press, 2008.
9. Z. Sekkat and W. Knoll, *Photoreactive Organic Thin Films*, Academic Press, 2002.
10. A. Cowan and J. Young, "Mode matching for second-harmonic generation in photonic crystal waveguides," *Phys. Rev. B* **65**, 2002.
11. K. Kneipp, Y. Wang, H. Kneipp, I. Itzkan, R. R. Dasari, and M. S. Feld, "Population pumping of excited vibrational states by spontaneous surface-enhanced Raman scattering," *Phys. Rev. Lett.* **76**, pp. 2444–2447, 1996.
12. P. J. Campagnola and L. M. Loew, "Second-harmonic imaging microscopy for visualizing biomolecular arrays in cells, tissues and organisms," *Nature Biotechnol.* **21**, pp. 1356–1360, 2003.
13. M. Han, G. Giese, and J. Bille, "Second harmonic generation imaging of collagen fibrils in cornea and sclera," *Opt. Express* **13**, pp. 5791–5797, 2005.
14. V. K. Valev, B. De Clercq, C. G. Biris, X. Zheng, S. Vandendriessche, M. Hojeij, D. Denkova, Y. Jeyaram, N. C. Panoiu, Y. Ekinci, A. V. Silhanek, V. Volskiy, G. A. E. Vandenbosch, M. Ameloot, V. V. Moshchalkov, and T. Verbiest, "Distributing the optical near-field for efficient field-enhancements in nanostructures," *Adv. Mater.* **24**, pp. OP208–OP215, 2012.
15. S. Bidault, J. Gouya, S. Brasselet, and J. Zyss, "Encoding multipolar polarization patterns by optical poling in polymers: towards nonlinear optical memories," *Opt. Express* **13**, pp. 505–510, 2005.
16. A. V. Tishchenko, "A generalized source method for wave propagation," *Pure Appl. Opt.* **7**, pp. 1425–1449, 1998.

17. A. A. Shcherbakov and A. V. Tishchenko, "Fast numerical method for modelling one-dimensional diffraction gratings," *Quantum Electron.* **40**, pp. 538–544, 2010.
18. A. A. Shcherbakov and A. V. Tishchenko, "New fast and memory-sparing method for rigorous electromagnetic analysis of 2D periodic dielectric structures," *J. Quant. Spectrosc. Radiat. Transfer* **113**, pp. 158–171, 2012.
19. P. Lalanne and G. M. Morris, "Highly improved convergence of the coupled-wave method for TM polarization," *J. Opt. Soc. Am. A* **13**, pp. 779–784, 1996.
20. L. Li, "Use of Fourier series in the analysis of discontinuous periodic structures," *J. Opt. Soc. Am. A* **13**, pp. 1870–1876, 1996.
21. L. Li, "New formulation of the Fourier modal method for crossed surface-relief gratings," *J. Opt. Soc. Am. A* **14**, pp. 2758–2767, 1997.
22. J. E. Sipe, "New Green-function formalism for surface optics," *J. Opt. Soc. Am. B* **4**, pp. 481–489, 1987.
23. T. Schuster, J. Ruoff, N. Kerwien, S. Rafler, and W. Osten, "Normal vector method for convergence improvement using the RCWA for crossed gratings," *J. Opt. Soc. Am. A* **24**, pp. 2880–2890, 2007.
24. P. Götz, T. Schuster, K. Frenner, S. Rafler, and W. Osten, "Normal vector method for the RCWA with automated vector field generation," *Opt. Express* **16**, pp. 17295–17301, 2008.
25. R. R. A. Syms and J. R. Cozens, *Optical guided waves and devices*, McGraw-Hill, 1992.