Accurate near-field calculation in the rigorous coupled-wave analysis method

Accurate near-field calculation in the rigorous coupled-wave analysis method

Martin Weismann[1,2], Dominic F G Gallagher[2] and Nicolae C Panoiu[1]

[1] Department of Electronic and Electrical Engineering, University College London, Torrington Place, WC1E 7JE, London, UK
[2] Photon Design Ltd, 34 Leopold Street, OX4 1TW, Oxford, UK

E-mail: martin.weismann.12@ucl.ac.uk

CrossMark

Abstract

The rigorous coupled-wave analysis (RCWA) is one of the most successful and widely used methods for modeling periodic optical structures. It yields fast convergence of the electromagnetic far-field and has been adapted to model various optical devices and wave configurations. In this article, we investigate the accuracy with which the electromagnetic near-field can be calculated by using RCWA and explain the observed slow convergence and numerical artifacts from which it suffers, namely unphysical oscillations at material boundaries due to the Gibbs phenomenon. In order to alleviate these shortcomings, we also introduce a mathematical formulation for accurate near-field calculation in RCWA, for one- and two-dimensional straight and slanted diffraction gratings. This accurate near-field computational approach is tested and evaluated for several representative test-structures and configurations in order to illustrate the advantages provided by the proposed modified formulation of the RCWA.

1. Introduction

The study of the interaction of electromagnetic waves with matter has spawned a large variety of methods to analytically or numerically solve the Maxwell equations (MEs). The results obtained from those methods are invaluable for understanding, validating, predicting, and guiding experimental efforts and for the design process of electromagnetic and optical devices. Amongst the numerical methods used in computational electromagnetics, one can fundamentally distinguish between time-domain methods, which directly incorporate the transient behavior of electromagnetic waves, such as the finite-difference time-domain method [1], and frequency-domain methods, like the finite-element frequency-domain method [2], which directly determine time-harmonic solutions to MEs. The former methods are general purpose methods and are capable of simulating virtually any

electromagnetic structure comprising metallic or dielectric objects of arbitrary size and shape. However, for certain structures and applications, other methods can be superior in terms of runtime and accuracy. For example, the multiple scattering method [3] is a series expansion method in the frequency domain that is tailored to calculate interaction of light and clusters of spherical particles [4] or cylindrical rods [5], and as such it is superior to other methods when applied to these particular geometries.

An important field of applications, where several specialized numerical algorithms exist [6], is the study of periodic optical structures, including diffraction gratings [7] and periodic metamaterials [8]. Amongst the numerical methods for periodic structures, such as the differential theory of gratings [9], the $\mathcal{C}$-method [10], integral methods [6, 11], and the generalized source method [12, 13], the most widely used method is the rigorous coupled-wave analysis (RCWA) [14–16]. It is a modal method in the frequency domain and is based on the decomposition of the periodic structure and the pseudo-periodic solution of MEs in terms of their Fourier series (FS) expansion, hence the periodicity is naturally incorporated into the numerical method. RCWA was initially
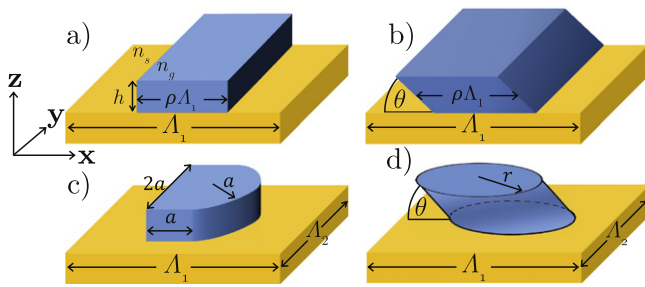
Figure 1. Grating structures mounted on a homogeneous substrate (refractive index, n_s) considered in this study: (a) 1D binary grating, filling factor, ρ; (b) 1D binary grating, slanted by $\theta = \pi/4$ w.r.t. the x-axis. (c) 2D grating with rectangle-semi-circular cross-section. (d) 2D cylindrical grating, slanted by $\theta = \pi/4$ w.r.t. the x-axis. The height of all gratings is denoted with h, their period is Λ_1 (and Λ_2 for 2D-periodic gratings), and their refractive index is denoted by n_g. The cover medium is vacuum with $n_c = 1$.

developed for modeling one-dimensional (1D) diffraction gratings, but with the introduction of fast converging Fourier factorization rules for modeling 1D [17, 18] and two-dimensional (2D) [19–21] structures, its formulation for isotropic and anisotropic materials [22] as well as multilayered [23] and oblique structures [24], RCWA has evolved to describe arbitrary, 2D-periodic structures. The method has been successfully applied to model diffraction gratings, diffractive optical elements, surface coatings, spectroscopic applications, photonic crystals, and periodic metamaterials. It should be noted, that in most cases the functionality of these applications of periodic structures depends on the electromagnetic far-field, i.e. the propagating diffraction orders, and it is known that the RCWA delivers fast converging and accurate far-field results, as reported in many works [14–24].

There is, however, a range of novel applications, which rely on the optical near-field of a periodic structure, especially at their surface, such as surface-enhanced Raman spectroscopy [25], surface second-harmonic generation [26–28], and near-field sensing [29–32]. These applications require a suitably designed near-field distribution, usually optimized for maximum field enhancement within specific spatial domains. Although there have been significant advances in experimental optical near-field measurement techniques [33], these techniques are still in their development state and not readily available to accurately characterize complex photonic nanostructures. These applications and experimental shortcomings lead to a critical demand for numerical methods for periodic structures that can facilitate an accurate calculation of electromagnetic near-fields, and more importantly, the design of gratings with optimized near-field patterns. With very few exceptions [20, 34, 35], a thorough investigation of numerically calculated near-fields in the RCWA has been largely neglected during the development of the method, as it is often merely considered a post-processing step. Moreover, additional reasons for the scarcity of reports on the convergence and the accuracy of the numerically computed electromagnetic near-field in RCWA, are the slow near-field

convergence and spurious oscillations displayed by these fields [35].

In this paper we address the issue of inaccurate near-field calculations in the RCWA in several ways. First, we use some generic cases of diffraction gratings to illustrate the slow convergence of RCWA for near-field calculations and reveal the reasons for this behavior. Based on this analysis and the continuity properties of the electromagnetic fields, a new formulation of the field evaluation is proposed, which yields faster convergence of the electromagnetic near-fields and explicitly fulfills the continuity properties of the electromagnetic field at material interfaces. This improved formulation of the numerical evaluation of the near-fields is then benchmarked against the current formulation, with the aim of making RCWA an effective numerical method for modeling modern nanophotonic applications that rely on highly accurate near-field calculations.

The remainder of the paper is organized as follows: section 2 will introduce the reader to the problem of inaccurate/spurious near-fields in RCWA for 1D structures and explain the overall strategy for the accurate field evaluation. Section 3 will extend the ideas gained from the analysis of 1D structures to arbitrary, straight or obliquely etched 2D gratings and will present the underlying mathematical formalism. Then in section 4, computational results for 2D gratings will be presented and discussed. Section 5 will investigate near- and far-field convergence of the modified method for slanted gratings, before final conclusions about the capability of the improved RCWA for modeling electromagnetic near-fields will be drawn in section 6.

2. Accurate near-field evaluation for 1D-periodic structures

Although the last major roadblock that was precluding RCWA from becoming a highly effective method for modeling 1D periodic structures has been removed when the correct Fourier factorization rules for transverse magnetic (TM) polarization were introduced [17, 18], a few topics have continued to attract attention, namely the convergence for slanted 1D-periodic gratings [24] and the numerical instabilities associated to highly conductive gratings [36]. These improvements and refinements of the method are concerned with the accuracy of the far-field calculation and they achieve excellent convergence results for the coefficients of and power carried by the diffraction orders in the cover and substrate regions of a grating, namely the far-field. This picture changes if the near-field is considered. Thus, in this section we investigate the numerically calculated near-field inside the periodic grating region via RCWA, and explain its slow convergence and the origin of the observed high spatial frequency oscillations. Our approach is based on a numerical method introduced in [34], for improved near-field accuracy and reduced numerical artifacts for straight 1D-periodic gratings. This numerical approach is thoroughly investigated here and will form the basis of the general formalism for accurate near-field evaluation in arbitrary, straight or slanted,
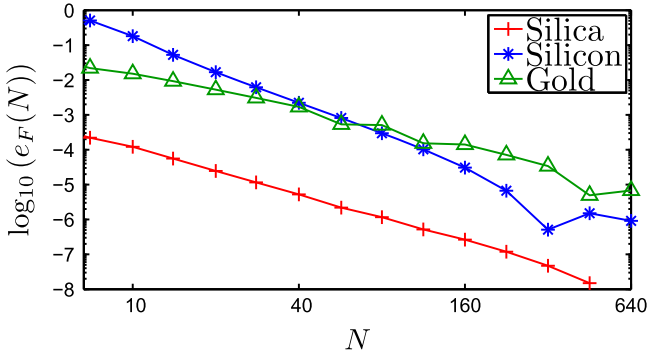
Figure 2. Far-field relative error versus the number of harmonics, determined for three different materials.

1D- and 2D-periodic gratings proposed in the subsequent sections.

The ideas developed in this paper are applicable to any periodic grating structure with sharp boundaries, but in order to conduct a comprehensive assessment of the accuracy of the near-field calculations achievable with the improved RCWA introduced in this work, a number of generic test structures (1D- and 2D-periodic, straight and slanted) have been chosen, as per figure 1. To widen the spectrum of test configurations, three different materials are considered for the grating: silica (SiO$_2$) as a dielectric with low index of refraction ($n_g = n_{SiO_2} = 1.45$), silicon (Si) as a high refractive index dielectric ($n_g = n_{Si} = 3.4$) and a metal, gold (Au), with index of refraction $n_g = n_{Au} = 0.97 + 1.87i$ [37]. In all examples the substrate is SiO$_2$ ($n_s = n_{SiO_2}$). Normal incidence and TM-polarization is considered in all of the sections, i.e. the incident electric field amplitude is oriented along the x-axis: $\mathbf{E}_{inc}(\mathbf{x}) = (1, \ 0, \ 0)^T E_{inc}(\mathbf{x})$.

The example considered in this section is depicted in figure 1(a). It consists of a binary grating with period, $\Lambda_1 = 1 \, \mu m$, height, $h = 0.25 \, \mu m$, and filling factor, $\rho = 0.5$. The grating is illuminated by a normally incident plane wave with wavelength, $\lambda = 0.51 \, \mu m$.

As the first step of our investigation, we used RCWA to calculate the far-field. In order to quantify the convergence of the RCWA method, the relative error of the far-field is defined as:

$$e_F(N) = \frac{\sqrt{\left| T^{(N)} - T^{ref} \right|^2 + \left| R^{(N)} - R^{ref} \right|^2}}{\sqrt{\left| T^{ref} \right|^2 + \left| R^{ref} \right|^2}}, \quad (1)$$

where $T^{(N)}$ ($R^{(N)}$) denotes the relative transmitted (reflected) power corresponding to a discretization with $2N + 1$ complex FS coefficients. The discretization parameter, N, represents the number of harmonics retained for each dimension. Moreover, $T^{ref}(R^{ref})$ is a reference value that is considered to be the exact solution or a sufficiently good approximation of the exact solution. Due to the absence of the exact solution of the diffraction grating problem, the reference values are chosen to be numerical values obtained by high-N simulations; in our case we chose $T^{ref} = T^{(905)}$ and $R^{ref} = R^{(905)}$.

The far-field relative error, $e_F(N)$, for increasing number of harmonics, $N = 5, \ldots, 640$, is depicted in figure 2. As this figure illustrates, RCWA converges quickly for all three materials. Specifically, in order to achieve a self-error of $e_F(N) < 1\%$ (as a generic criterion adopted here for an accurate far-field calculation), for the three materials, silica (red crosses), silicon (blue stars), and gold (green triangles), a relatively small number of harmonics is necessary, namely $N > 5$, $N > 25$, and $N > 13$, respectively.

As mentioned in the introductory section, the far-field of optical gratings and periodic structures has been the physical quantity of most interest from experimental point of view, hence the characterization of a numerical method by means of the far-field convergence has usually been the adopted strategy. This approach, however, largely neglects the electromagnetic near-field predicted by a specific method. On the other hand, the near-field is of fundamental importance for modeling plasmonic effects or optical nonlinear phenomena in devices with size comparable to or smaller than the operating wavelength, effects whose description relies on accurate calculations of the electromagnetic near-field.

In order to characterize the numerically obtained near-fields, we define the grating norm, $\| \cdot \|_G$, of a scalar or vectorial function, f, in the grating region as follows:

$$\| f \|_G = \left(\int_0^h \int_{-\Lambda_1/2}^{\Lambda_1/2} | f(x, z) |^2 \, dx \, dz \right)^{\frac{1}{2}}, \quad (2)$$

where the z-integration extends over the bulk of the periodic region. The grating norm is used to define the near-field error, $e\{E_\alpha^{(N)}\}$, of the scalar field components $E_\alpha^{(N)}$, $\alpha = x, y, z$, of a near-field, $\mathbf{E}^N$, numerically obtained using N harmonics:

$$e\left\{ E_\alpha^{(N)} \right\} = \left\| E_\alpha^{(N)} - E_\alpha^{ref} \right\|_G \Big/ \left\| E_\alpha^{ref} \right\|_G. \quad (3)$$

Here, $E_\alpha^{(N)}$ denotes the FS reconstruction given by the $2N + 1$ central Fourier coefficients, $E_{\alpha n}$, which are calculated by RCWA:

$$E_\alpha^{(N)}(x) = \sum_{n=-N}^{N} E_{\alpha n} \exp\left(in\frac{2\pi}{\Lambda_1}x \right).$$

Similarly to the far-field calculations, $\mathbf{E}^{ref} = \mathbf{E}^{(905)}$ is obtained by a high-resolution RCWA simulation with $N = 905$. The near-field self-convergence for the tangential component, $E_t = E_z$, and the normal component, $E_n = E_x$, of the electric field is depicted in figure 3. For all three materials, the self-convergence of E_t (blue circles) is fast and comparable to the far-field self-convergence (see figure 2). The normal component, E_n (red crosses), however, exhibits much slower convergence and even at the highest numerical resolution of $N = 640$ a relative error of $e\{E_x^{(640)}\} > 0.9\%$ still remains, whereas the error of the tangential field, $e\{E_z^{(640)}\} < 0.08\%$, for all materials.

One intriguing question raised by the data plotted in figures 2 and 3 is why the normal component of the near-field converges much more slowly than the far-field power and the tangential component of the near-field. Or put it the other way around: How can the far-field converge quickly when the
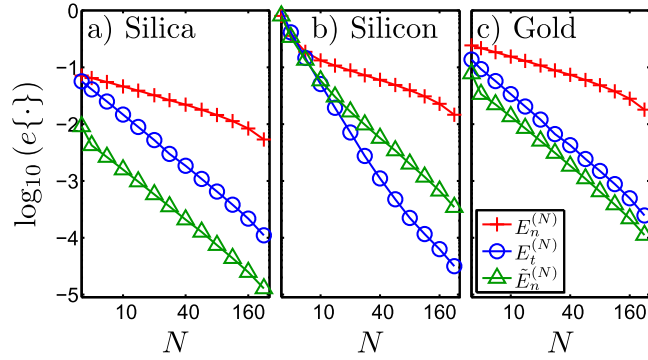
Figure 3. Self convergence of the tangential and normal electric field components E_z and E_x and $\tilde{E}_x$ inside the grating structure described by their self-errors $\Delta E_t(N)$ (green triangles), $\Delta E_n(N)$ (red crosses) and $\Delta \tilde{E}_n(N)$ (blue circles), respectively for the three benchmark structures made of silica (a), silicon (b) and gold (c).

near-fields have not converged yet? There are two factors that explain this behavior: (i) the far-field consists of a super-position of a small number of propagating diffraction orders, i.e., plane waves. Hence, the far-field requires a small number of FS components to be reconstructed and does not suffer from the Gibbs phenomenon. (ii) RCWA does not depend on the representation of the discontinuous normal field E_n. Instead, the method relies on the correct Fourier factorization of the continuous normal component of the displacement field, D_n, which can be accurately described by a FS, i.e. without spurious oscillations. The second observation is at the core of the accurate near-field calculation that is introduced in this work. Specifically, instead of reconstructing a discontinuous physical quantity, i.e. $E_n(\mathbf{x})$, directly from its FS coefficients, it is more effective to reconstruct a continuous quantity, the normal component of the displacement field, $D_n(\mathbf{x})$, and then divide it by a discontinuous quantity, the electric permittivity, $\varepsilon_0 \varepsilon(\mathbf{x})$, which is known in the space-domain where one seeks to solve the diffraction problem.

For 1D-periodic gratings, we define the modified normal component of the electric field:

$$\tilde{E}_n^{(N)}(\mathbf{x}) = \varepsilon_0^{-1} D_n^{(N)}(\mathbf{x}) / \varepsilon(\mathbf{x}), \qquad (4)$$

where ε_0 is the vacuum permittivity and $\varepsilon(\mathbf{x})$ the relative permittivity at position, $\mathbf{x}$. Note that $E_n^{(N)}(\mathbf{x})$ and $\tilde{E}_n^{(N)}(\mathbf{x})$ represent the same physical quantity, namely the normal component of the electric field. However, whereas the former is found by using RCWA to solve directly for the electric field, the latter one is determined by first calculating the displacement field and then the electric field via equation (4).

The error, $e\{\tilde{E}_n^{(N)}\}$, of the modified normal component, $\tilde{E}_n$, is shown in figure 3 and encoded in green triangle. It is found to converge as fast as the fast-convergent tangential component, E_t, and as fast as the power in the far-field shown in figure 2. Even at the highest considered resolution, $N = 640$, the conventional formulation of the RCWA yields a self error of $e\{E_n^{(640)}\} > 9 \times 10^{-3}$ for all three materials. By contrast, the same self error $e\{\tilde{E}_n^{(N)}\} < 9 \times 10^{-3}$ of the modified normal field $\tilde{E}_n$ can be achieved by using as few as $N = 70$ harmonics.
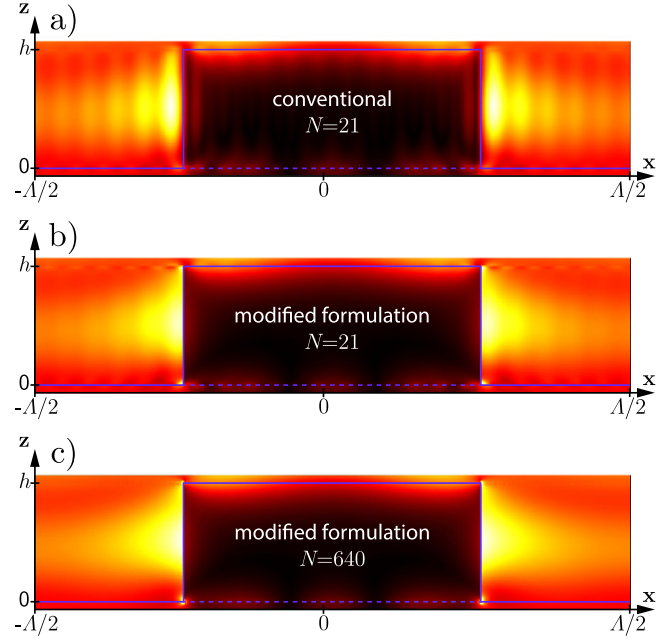

Figure 4. Near-field distribution of the normal component of the electric field, $|E_n| = |E_x|$, in and around the Au grating. (a) Calculated with $N = 21$ harmonics using the conventional RCWA method (unphysical field oscillations can be observed). (b) Calculated using the improved formulation, $\tilde{E}_n = D_n / \varepsilon$, and $N = 21$ harmonics. (c) Calculated using the improved formulation and $N = 640$ harmonics.

The spatial profile of the electric field in the grating region illustrates the full benefits of the modified field calculation. Figure 4(a) depicts the conventional normal field component $|E_n^{(21)}| = |E_x^{(21)}|$ in the gold grating for a moderately coarse discretization of $N = 21$ harmonics. It can be seen that $|E_n^{(21)}|$ exhibits unphysical oscillations with a spatial frequency equal to the period of the smallest spatial frequency component in the FS expansion of the solution. This is the well known Gibbs phenomenon, which occurs when describing a discontinuous function with a truncated FS. On the other hand, the modified normal field, $\tilde{E}_n$, does not suffer from such spurious oscillations at the interface. In particular, even for a small number of harmonics, $N = 21$, $\tilde{E}_n^{(21)}$ is smooth, as per figure 4(b). At very large number of harmonics, $N = 640$, the modified normal field is free of any numerical artifacts, as can be seen in figure 4(c).

The improved formulation of RCWA exhibits another benefit, namely $\tilde{E}_n$ is by construction discontinuous and exactly fulfills the corresponding boundary condition

$$\varepsilon^{(\text{in})} \mathbf{E}^{(\text{in})}(\mathbf{x}_s) \cdot \mathbf{n} = \varepsilon^{(\text{out})} \mathbf{E}^{(\text{out})}(\mathbf{x}_s) \cdot \mathbf{n}, \qquad (5)$$

at surface points of the grating, $\mathbf{x}_s$, where $\mathbf{E}^{\text{in}}$ and $\varepsilon^{(\text{in})}$ ($\mathbf{E}^{\text{out}}$ and $\varepsilon^{(\text{out})}$) denote the electric field and permittivity inside (outside) the grating, respectively, and $\mathbf{n}$ is the unit vector normal to the surface. In the conventional formulation of the RCWA, the field E_n^N does not satisfy equation (5), because $E_n^{(n)}$ is—as a FS containing a *finite* number of terms—inherently continuous. The closeup of the interfacial field around $x = 0.25\Lambda$ in figure 5 emphasizes these ideas. Thus,
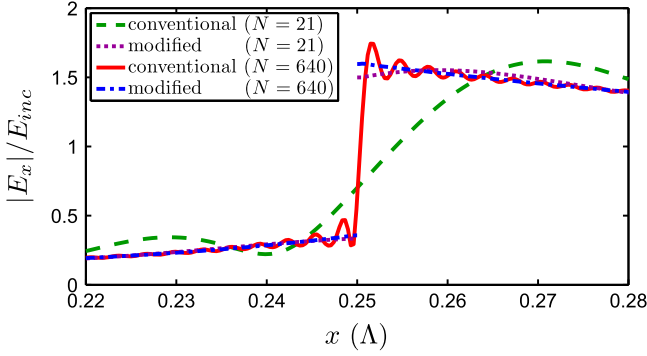
Figure 5. Closeup of the normal component of the electric field, $|E_n|$ and $|\tilde{E}_n|$, near the metal–air interface at $z = h/2$, computed using the conventional and modified RCWA methods, respectively.

the normal component of the field calculated using the conventional formulation of RCWA shows spurious oscillations for both small $N = 21$ (dashed green) and high $N = 640$ (solid red) number of harmonics, whereas the modified formulation is free of such unphysical oscillations and discontinuous for any number of harmonics, N (see the dotted purple and dashed–dotted blue lines corresponding to $N = 21$ and $N = 640$, respectively).

It should be clear now that the modified normal-field evaluation in 1D by means of the displacement field represents an improvement of the conventional evaluation of E_n in two ways: first, it exhibits optimal self-convergence in the sense that it is as accurate as the far-field and the continuous tangential field component. Second, it explicitly fulfills the boundary condition (5) at the material interfaces. Equally important, the improved near-field calculation is achieved with minimal additional computational cost and does not alter the mathematical framework of the core RCWA algorithm in 1D.

3. Formulation of accurate near-field evaluation for 2D-periodic structures

The ideas of the previous section are straightforward to implement because in 1D-periodic structures there is a trivial and unambiguous distinction between the tangential and normal components of **E**. However, this situation becomes more intricate in the case of 2D-periodic diffraction gratings.

As it will be shown now, accurate near-field evaluation is strongly related to correct Fourier factorization in 2D-periodic structures. Fourier factorization means in this context the decomposition of a product of periodic functions, $f = g \cdot h$, into its periodic factors g and h. Depending on the continuity properties of the factors and the product, different rules must be applied in order to obtain fast convergence when increasing the number of FS terms According to these rules, if f and h possess simultaneous discontinuities, but g is continuous at those locations, the product rule yields fast convergence with respect to the number of FS terms, namely one uses $f = g \cdot h$. Moreover, if g and h are simultaneously discontinuous, but f is continuous, the inverse rule should be

used, i.e., f must be factorized as $f = (1/g)^{-1} \cdot h$. A rigorous explanation of these rules and the solution to the 1D Fourier factorization problem is given in [18].

Because of the trivial distinction between the continuous (tangential) and discontinuous (normal) components of the electric field in 1D, Fourier factorization is straightforward in that case. But for 2D-periodicity, three different approaches to achieve the correct Fourier factorization have been proposed: (i) approximate the material boundaries by a coordinates-aligned staircase-contour [19]. (ii) Devise a coordinate system in which a given grating geometry is coordinate system aligned and use approach (i) to obtain the correct Fourier factorization [20, 38]. (iii) Construct a normal vector field (NVF) to decompose **E** into its normal and tangential components and then apply the corresponding correct factorization rules to them [21].

Since the accurate near-field evaluation relies on the decomposition of the electric field into normal and tangential components, it is natural to use the factorization approach (iii), the NVF approach. It is out of the scope of this paper to derive the full formulation of 2D-RCWA, so that only the crucial and unconventional steps will be given here. Thus, the normal and tangential components of $\mathbf{D} = \varepsilon \mathbf{E}$ have to be decomposed using the product- and inverse rule, respectively. This leads to the following relations for the α-component of the displacement field [21, 24, 39]:

$$[D_\alpha] = \varepsilon_0 \sum_{\beta=1}^{3} \left(\delta_{\alpha,\beta} [\![\varepsilon]\!] - \Delta_{\alpha\beta} \right) [E_\beta], \qquad (6)$$

where $\delta_{\alpha\beta}$ is the Kronecker delta. Here, $[f]$ denotes the vector of FS coefficients of a scalar function f, $[\![g]\!]$ is the Toeplitz matrix of FS coefficients of g, and the matrix $\Delta_{\alpha\beta}$ is given by $\Delta_{\alpha\beta} = \frac{1}{2} \left(\Delta [\![N_\alpha N_\beta]\!] + [\![N_\alpha N_\beta]\!] \Delta \right)$ with $\Delta = [\![\varepsilon]\!] - [\![1/\varepsilon]\!]^{-1}$ and N_α is the α-component of the NVF, $\mathbf{N} = (N_1, N_2, N_3)^T$, of the material boundary. The matrix $\delta_{\alpha,\beta} [\![\varepsilon]\!] - \Delta_{\alpha\beta}$ implements the three steps of the Fourier factorization: the decomposition of **E** into normal and tangential components (N_β in $[\![N_\alpha N_\beta]\!]$), factorization using either the inverse rule ($[\![1/\varepsilon]\!]^{-1}$) or the product rule ($[\![\varepsilon]\!]$), and back-projection to Cartesian coordinates (N_α in $[\![N_\alpha N_\beta]\!]$).

Note that one can also use asymptotically equivalent definitions of $\Delta_{\alpha\beta}$. However, our investigations have shown that despite the fact that the choices $\Delta_{\alpha\beta} = \Delta [\![N_\alpha N_\beta]\!]$ and $\Delta_{\alpha\beta} = [\![N_\alpha N_\beta]\!] \Delta$ produce similar convergence speed, they do not conserve the power for lossless structures, whereas the choice $\Delta_{\alpha\beta} = [\![N_\alpha]\!] \Delta [\![N_\beta]\!]$ yields power conservation but at the price of slower convergence. All three formulations are asymptotically equivalent with respect to the number of terms in the FS expansion, due to the commutation of Toeplitz operators [39].

Normal vector fields can be constructed analytically for a variety of structures and automated algorithms to obtain a NVF for arbitrary grating geometries have been developed [40]. It should be noted that this formulation allows inclined NVFs, i.e. NVFs with simultaneously non-vanishing *x*, *y*, and *z* components. Such a NVF is required to accurately model obliquely etched structures (see figure 1(d)) and hence can be

viewed as a generalization of the methods presented in [24] ($N_y = 0$, figure 1(b)) and [21] ($N_z = 0$, figure 1(c)).

Given the constitutive relation (6), one can derive the RCWA eigenvalue-problem

$$
\mathsf{M}
\begin{pmatrix}
[S_x] \\
[S_y] \\
[U_x] \\
[U_y]
\end{pmatrix}
= \beta
\begin{pmatrix}
[S_x] \\
[S_y] \\
[U_x] \\
[U_y]
\end{pmatrix}
\tag{7}
$$

for the electromagnetic modes described by $[S]$ and $[U]$ and propagation constant, β. The most general formulation of the system-matrix M reads:

$$
\mathsf{M} =
\begin{pmatrix}
\mathsf{K}_x \mathsf{B} \Delta_{zx} & \mathsf{K}_x \mathsf{B} \Delta_{zy} \\
\mathsf{K}_y \mathsf{B} \Delta_{zx} & \mathsf{K}_y \mathsf{B} \Delta_{zy} \\
\Delta_{yx} + \Delta_{yz} \mathsf{B} \Delta_{zx} - \mathsf{K}_x \mathsf{K}_y & \mathsf{K}_x \mathsf{K}_x + \Delta_{yz} \mathsf{B} \Delta_{zy} - \mathsf{C}_y \\
\mathsf{C}_x - \Delta_{xz} \mathsf{B} \Delta_{zx} - \mathsf{K}_y \mathsf{K}_y & \mathsf{K}_y \mathsf{K}_x - \Delta_{xz} \mathsf{B} \Delta_{zy} - \Delta_{xy}
\end{pmatrix}
$$

$$
\begin{pmatrix}
\mathsf{K}_x \mathsf{B} \mathsf{K}_y \mathsf{I} - \mathsf{K}_x \mathsf{B} \mathsf{K}_x \\
\mathsf{K}_y \mathsf{B} \mathsf{K}_y - \mathsf{I} - \mathsf{K}_y \mathsf{K}_x \\
\Delta_{yz} \mathsf{B} \mathsf{K}_y - \Delta_{yz} \mathsf{B} \mathsf{K}_x \\
-\Delta_{xz} \mathsf{B} \mathsf{K}_y \Delta_{xz} \mathsf{B} \mathsf{K}_x
\end{pmatrix},
$$

with $\mathsf{B} = (\llbracket \varepsilon \rrbracket - \Delta_{zz})^{-1}$ and $\mathsf{C}_\alpha = \llbracket \varepsilon \rrbracket - \Delta_{\alpha\alpha}$. The matrices $\mathsf{K}_\alpha = \mathrm{diag}(k_{\alpha n})$ are diagonal matrices of the in-plane propagation constants, k_{xn} for $\alpha = x$ and k_{yn} for $\alpha = y$, of the diffraction orders and I is the identity matrix. The size of matrices B, C_α, K_α, and I is $N_0 \times N_0$, whereas the size of M is $4N_0 \times 4N_0$, where $N_0 = (2N + 1)^2$ is the total number of FS coefficients in the calculation and the same number of FS terms for the truncation of the FS in the *x*- and *y*-direction is assumed.

The mode amplitudes $[S]$ and $[U]$ are determined by matching the tangential components of the electromagnetic field at the top and bottom of the grating region, whereas multilayered structures are described by the staircase approximation in conjunction with the numerically stable $\mathcal{S}$-matrix algorithm. This yields the FS coefficients, $[\mathbf{E}]$, of the electric field everywhere in the grating region, in the cover, and the substrate.

Let us now denote by $\mathcal{R}([f])$ the FS reconstruction of a Fourier coefficient vector $[f]$,

$$
\mathcal{R}([f])(x, y) = \sum_{n=-N}^{N} \sum_{m=-N}^{N} f_{nm} \exp\left(\mathrm{i}m\frac{2\pi}{\Lambda_1}x + \mathrm{i}n\frac{2\pi}{\Lambda_2}y \right),
$$

where $f_{nm} = [f]_{(n+N)(2N+1)+m+N+1}$ are the N_0 FS coefficients of f. Within the conventional RCWA framework, each component of $\mathbf{E}$ is evaluated as:

$$
E_\alpha^{(N)} = \mathcal{R}([E_\alpha]).
\tag{8}
$$

This relation does not take into account the continuity properties of the different components of $\mathbf{E}$ and hence will lead to spurious oscillations and slow convergence of the near-field as it was seen in section 2.

The modified field evaluation for 2D-periodic diffraction gratings requires one to define the continuous normal

component of $\mathbf{D}$ and the tangential component of $\mathbf{E}$. Their FS coefficient vectors are given by:

$$
[\mathbf{D}_{\mathrm{n}}] = \frac{1}{2}\varepsilon_0 \big(\llbracket \mathbf{N}\mathbf{N}^T \rrbracket \llbracket 1/\varepsilon \rrbracket^{-1} + \llbracket 1/\varepsilon \rrbracket^{-1} \llbracket \mathbf{N}\mathbf{N}^T \rrbracket \big)[\mathbf{E}], \tag{9a}
$$

$$
[\mathbf{E}_{\mathrm{t}}] = \llbracket 1 - \mathbf{N}\mathbf{N}^T \rrbracket [\mathbf{E}], \tag{9b}
$$

where 1 denotes the 3×3 identity matrix and $\mathbf{N}\mathbf{N}^T$ is the 3×3 projection matrix defined by the NVF at any point in space, $\mathbf{x}$. Note that by construction, $[\mathbf{D}_{\mathrm{n}}]$ and $[\mathbf{E}_{\mathrm{t}}]$ are FS coefficient vectors of vector fields that are continuous at material interfaces. Hence their reconstructions, $\mathcal{R}([\mathbf{D}_{\mathrm{n}}])$ and $\mathcal{R}([\mathbf{E}_{\mathrm{t}}])$, do not suffer from the Gibbs phenomenon at the interface. With this observation in mind, the electric field in the improved RCWA method for 2D-periodic structures at a point, $\mathbf{x}$, is given by:

$$
\tilde{\mathbf{E}}^{(N)}(\mathbf{x}) = \varepsilon_0^{-1}\mathcal{R}([\mathbf{D}_{\mathrm{n}}])(\mathbf{x}) \big/ \varepsilon(\mathbf{x}) + \mathcal{R}([\mathbf{E}_{\mathrm{t}}])(\mathbf{x}) \tag{10}
$$

and is expected to yield fast near-field convergence and non-oscillatory spatial field profiles. To investigate the validity of these predictions, the accurate field evaluation was implemented in a commercially available RCWA computer software, OmniSim/RCWA [41].

It should be noted that the other electromagnetic fields can be easily calculated with our improved method, too. Specifically, the displacement field, $\mathbf{D}$, can be evaluated using the modified electric field $\tilde{E}$, namely $\mathbf{D}^{(N)}(\mathbf{x}) = \varepsilon_0 \varepsilon(\mathbf{x}) \tilde{\mathbf{E}}^{(N)}(\mathbf{x})$, and hence will have the same convergence properties as $\tilde{\mathbf{E}}$. The magnetic induction $\mathbf{B}$ and the magnetic field $\mathbf{H} = \mu \mathbf{B}$ do not require special attention, because they are continuous in non-magnetic materials and hence behave similar to the continuous tangential component of the electric field.

4. Quantification of the accurate near-field evaluation in 2D-periodic structures

In this section, the improved formulation for accurate near-field calculations in 2D-periodic structures is assessed using two test structures under different configurations. To this end, we first extend the definition of the grating norm (2) of a scalar or vector function, f, to 2D-periodic structures in the following straightforward way:

$$
\| f \|_G = \left(\int_0^h \int_{-\Lambda_2/2}^{\Lambda_2/2} \int_{-\Lambda_1/2}^{\Lambda_1/2} |f(x, y, z)|^2 \; \mathrm{d}x \; \mathrm{d}y \; \mathrm{d}z \right)^{\frac{1}{2}}, \tag{11}
$$

where the integral is evaluated over the three-dimensional grating region.

4.1. Analysis of a 1D-periodic grating using 2D-RCWA

The first 2D-periodic diffraction grating under consideration is a 1D binary grating, as shown in figure 1(a), rotated by $\pi/4$ in the *x*–*y*-plane. It should be obvious that it can be modeled as a double-periodic 2D grating with periods $\Lambda_1 = \Lambda_2 = \sqrt{2}\,\tilde{\Lambda} = \sqrt{2}\,\mu$m, where $\tilde{\Lambda} = \Lambda_1$ is the period of the grating when it is viewed as a 1D-periodic structure. For clarity, the
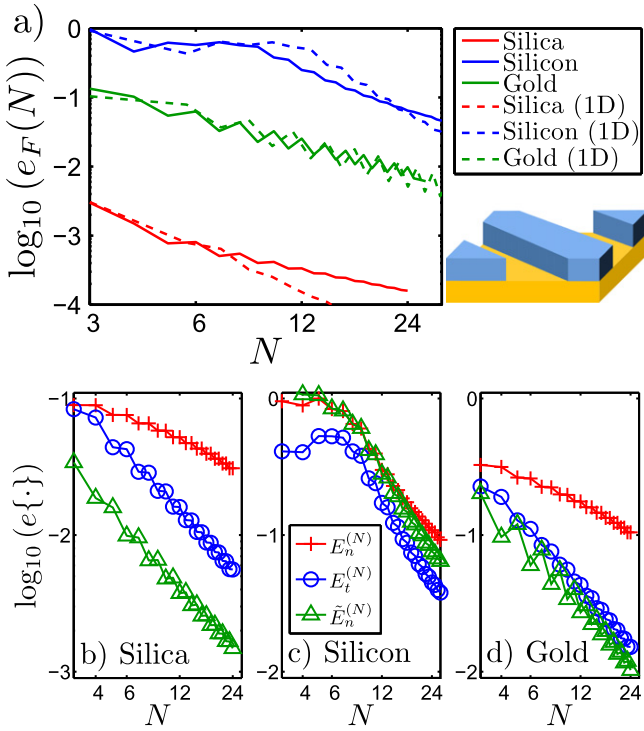
Figure 6. Computational results for a rotated binary 1D grating. (a) Error of calculated far-field versus N, determined for three gratings made of different materials. The decrease in the error follows that of the far-field (dashed lines) of the 1D simulations (see also section 2). (b)–(d) Near-field error corresponding to the silica, silicon, and gold grating, respectively.

primary unit cell of the grating is depicted in the inset of figure 6(a). With this choice, the reference quantities R^{ref}, T^{ref}, and $\mathbf{E}^{\text{ref}}$ can be calculated using 1D simulations, an approach inspired by an example in [21].

The error in the calculation of the far-field, e_{F} from equation (1), when 2D simulations are employed is depicted in figure 6(a). The convergence of the calculations for the silicon and gold gratings closely follows the convergence trends observed when 1D simulations are performed and, as expected, it shows somewhat worse, yet still good, agreement for the silica structure. This difference is explained by the fact that the NVF introduced in the 2D-RCWA formulation is discontinuous away from material interfaces and hence it can degrade the convergence rate, especially for the low-index of refraction contrast case.

The conclusions of our analysis of the convergence of the near-field are summarized in figures 6(b)–(d). Thus, for both the silica and gold diffraction gratings, the normal component of the modified electric field, $\tilde{\mathbf{E}}$, exhibits faster convergence than this same field component calculated using the conventional form of RCWA. In both cases, $e\{\tilde{E}_{\text{n}}^{(20)}\}$ is one order of magnitude smaller than $e\{E_{\text{n}}^{(20)}\}$. For the silicon grating, only marginal differences between the two formulations can be observed. This is in agreement with the results obtained in the 1D case, as per figure 3(b). For small N, $\tilde{\mathbf{E}}^{(N)}$ and $\mathbf{E}^{(N)}$ are determined with a comparable degree of accuracy, but for the larger number of harmonics considered in figure 6(c), i.e.
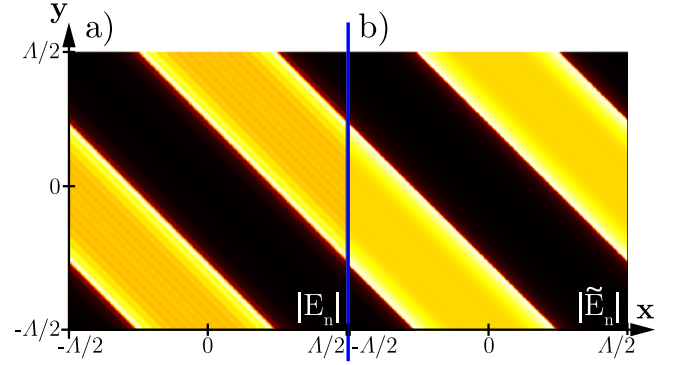

Figure 7. (a) Normal component of the electric field, $|E_{\text{n}}|$, in the grating region at $z = h/2$, determined by using the conventional RCWA and $N = 27$. (b) Normal electric field component, $|\tilde{E}_{\text{n}}|$, determined for the same grating parameters as in (a) but using the improved algorithm. The blue vertical line was added for clarity and merely separates the two plots.

$N = 30$, a higher accuracy of our improved formulation of the RCWA can clearly be observed.

These conclusions are further validated by the profile of the electric field, as presented in figure 7. This figure shows the spatial distribution of the normal component of the electric fields, $\mathbf{E}^{(27)}$ and $\tilde{\mathbf{E}}^{(27)}$, calculated in the median plane of the grating. A simple examination of these field profiles confirms that the spurious oscillations of the field $\tilde{\mathbf{E}}^{(27)}$ near the surface have much smaller amplitude as compared to that of the variations of $\mathbf{E}^{(27)}$. Moreover, a closer inspection of the surface-fields shows that the boundary condition (5) is fulfilled by $\tilde{\mathbf{E}}^{(27)}$ only. This first test-case already reveals that the improved near-field evaluation is more accurate in the case of 2D-periodic structures, too.

4.2. Near-field calculations for an intrinsically 2D-periodic grating

In order to thoroughly test the near-field evaluation for 2D-periodic structures using the improved RCWA presented in this article, in what follows we consider the challenging test structure depicted in figure 1(c). The grating region consists of a coordinate system aligned parallelepiped with the length of the sides aligned to the x-, y-, and z-axis being $a = 0.5\Lambda_1$, $2a$, and h, respectively, placed adjacently to a semicircular cylinder with radius a and height h, with $\Lambda = \Lambda_1 = \Lambda_2 = 0.25\,\mu\text{m}$. The structure is illuminated normally by a x-polarized plane wave with wavelength $\lambda = 0.5\,\mu\text{m}$.

Since it is generally computationally time consuming to obtain high-accuracy solutions in the case of 2D-periodic structures and due to the fact that the higher the ratio Λ/λ and the refractive index $|n|$, the more harmonics are necessary to achieve convergence [35], a relatively small period-to-wavelength ratio of $\Lambda/\lambda = 0.25\,\mu\text{m}/0.5\,\mu\text{m} = 0.5$ is chosen for this example.

As reference values in the definition of the far-field error, e_{F} from equation (1), $T^{\text{ref}} = T^{31}$ and $R^{\text{ref}} = R^{31}$, namely results obtained from simulations with $N = 31$ are chosen.
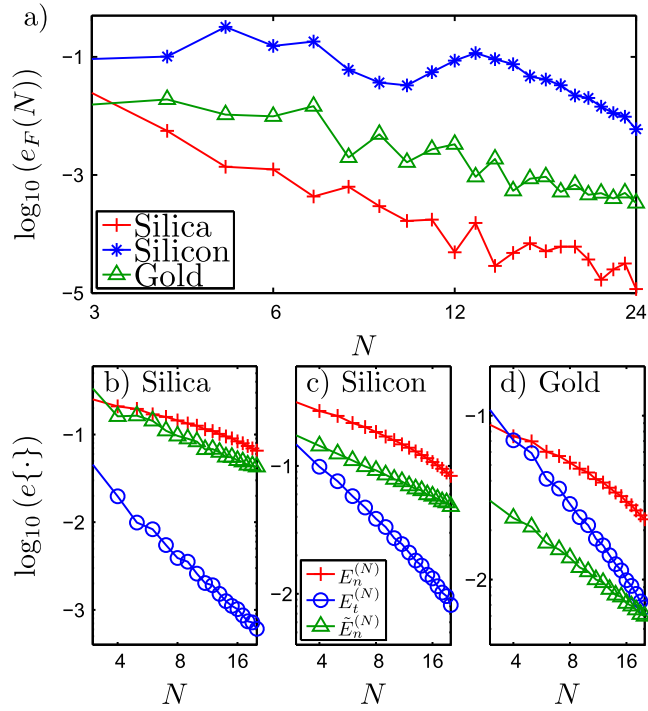
Figure 8. Computational results for the 2D grating shown in figure 1(c). (a) Error of calculated far-field versus N, determined for three gratings made of different materials. (b)–(d) Near-field error corresponding to the silica, silicon, and gold grating, respectively.
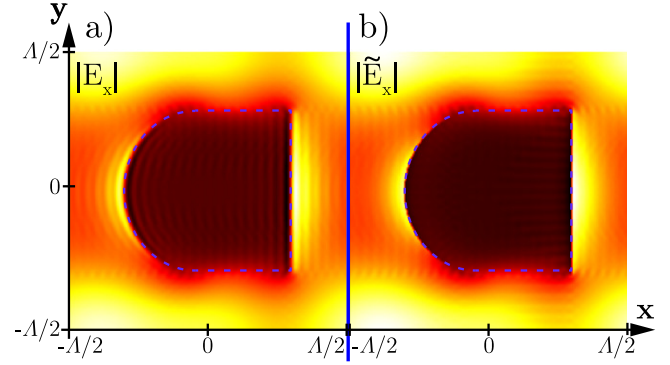

Figure 9. (a) Spatial distribution of the dominant component of the electric field, $|E_x|$, in the 2D grating region at $z = h/2$, determined by using the conventional RCWA and $N = 27$. (b) Spatial distribution of the dominant component of the electric near-field, $|\tilde{E}_x|$, determined for the same grating parameters as in (a) but using the improved algorithm. The blue vertical line was added for clarity and merely separates the two plots.

The convergence of the far-field, $e_F(N)$, is shown in figure 8(a), where the far-field physical quantity considered are the transmission and reflection coefficients, T and R, respectively. As expected, the fastest convergence can be observed for the silica grating, because it has a low refractive index. The numerical error obtained for the gratings made of gold and silicon are one and two orders of magnitude larger than in the case of the silica grating, respectively. This behavior is similar to that seen in the 1D case for a small number of harmonics, $N < 30$ (see figure 2).

Figures 8(b)–(d) contain the dependence of the near-field self-error determined for the three material configurations, silica, silicon, and gold, respectively. The three physical quantities plotted in each case are the z-component of $\mathbf{E}$ and the in-plane components, $\mathbf{E}_{xy} := (E_x, E_y, 0)^T$ and $\tilde{\mathbf{E}}_{xy} := (\tilde{E}_x, \tilde{E}_y, 0)^T$, obtained by using the conventional and improved versions of RCWA, respectively. Note that the in-plane component contains the discontinuous normal component, $\mathbf{N} \cdot \mathbf{E}$ ($\mathbf{N} \cdot \tilde{\mathbf{E}}$), and the continuous tangential component, $(\mathbb{1} - \mathbf{N}\mathbf{N}^T)\mathbf{E}_{xy}$ ($(\mathbb{1} - \mathbf{N}\mathbf{N}^T)\tilde{\mathbf{E}}_{xy}$).

It can be seen that in all cases, the z-component of $\mathbf{E}$, which is continuous at vertical surfaces inside the grating region, converges much faster than the in-plane component, in both the conventional and modified formulations. In addition, the modified formulation leads to a somewhat smaller error than the conventional formulation. Specifically, it was found that $e\{\tilde{\mathbf{E}}_{xy}^{(N)}\} \approx 0.5\,e\{\mathbf{E}_{xy}^{(N)}\}$ for the silica grating and $e\{\tilde{\mathbf{E}}_{xy}^{(N)}\} \approx 0.9\,e\{\mathbf{E}_{xy}^{(N)}\}$ for the silicon and gold gratings. However, the corresponding convergence speed, i.e. the slope

of $e\{\mathbf{E}_{xy}^{(N)}\}$ and $e\{\tilde{\mathbf{E}}_{xy}^{(N)}\}$, is the same. Moreover, in the modified formulation of the RCWA the convergence speed of the tangential component of the near-field is larger than that of the in-plane component. Three factors contribute to this behavior: (i) the decomposition of the near-field in a normal and tangential component in 2D-periodic structures relies on the specific definition of the NVF, $\mathbf{N}$ and it is not directly performed in Cartesian coordinates as in the 1D case. Hence, the inexact field decomposition by $\mathbf{N}$ introduces an additional error. (ii) The field of normal vectors characterizing the structure is only uniquely defined at the interfaces defining the grating, except at the corners, and, more importantly, away from the grating surface. This ambiguity can lead to choices of NVFs which are not optimal for the convergence and accurate calculation of the near-field. (iii) The NVF itself has discontinuities, which can cause additional oscillations in the spatial profile of the electromagnetic field.

It is also worthwhile to investigate the spatial profile of the near-field. The dominant x-component of the electric field in a horizontal cross-section through the grating region at $z = h/2$ is depicted in figure 9. As in the 1D case, these maps show that the field $\tilde{E}_x$ exhibits spatial oscillations with smaller amplitude as compared to the variations of the field E_x, especially at y-aligned interfaces (outlined with blue dashed lines in figure 9).

5. Out-of-plane normal vector fields for oblique diffraction gratings

Oblique diffraction gratings as those shown in figures 1(b) and (d) are modeled in the RCWA within the staircase approximation along a direction perpendicular onto the grating plane. Each computational slice is assumed to be a z-independent structure, for which the eigenmodes can be found numerically by solving equation (7). The field in each computational layer is then found by using the boundary conditions at the top and bottom interfaces of the grating,

employing a numerically stable $\mathcal{S}$-matrix formulation. The validity of this staircase approximation for 1D-periodic gratings under TE-polarization has been proven in [35]. In the context of the NVF formulation of RCWA for oblique 1D-periodic structures, it has been found that the use of an out-of-plane NVF, i.e. $\mathbf{N} = (N_x, 0, N_z)^T$, is beneficial for the far-field convergence speed [24, 35]. In this section we will study the relation between the accurate field formulation (10) and the near-field convergence speed for slanted 1D-periodic structures.

The situation for oblique 2D-periodic structures has not yet been explored in the context of RCWA. However, based on the results presented so far, one can conjecture that both the near- and far-field convergence of RCWA can be improved by using an out-of-plane 3D NVF, $\mathbf{N} = \left(N_x, N_y, N_z\right)^T \neq 0$, and that the near-field evaluation would be more accurate as well. The validity of this supposition will be explored in the second part of this section.

5.1. Analysis of slanted 1D-periodic binary diffraction gratings

In order to analyze oblique 1D-periodic structures, we consider the slanted binary grating depicted in figure 1(b). The period of the grating is $\Lambda = 1\ \mu$m, the filling factor, $\rho = 0.5$, the height, $h = 0.25\ \mu$m, and the slanting angle is $\theta = \pi/4$. Only the gold grating is considered in this section as this would be the most challenging case. If the unit cell is assumed to extend from $x = -\Lambda/2$ to $x = \Lambda/2$ and the center of the binary grating is set to be $x = 0, z = h/2$, a suitable out-of-plane NVF is given by $\tilde{\mathbf{N}}(x, y, z) = \text{sign}(x - z)\sqrt{2}\ (1,0,1)^T$.

The two formulations of the RCWA compared in this section are the in-plane NVF, which is used in conventional RCWA together with the conventional field evaluation (8), and the out-of-plane NVF, $\tilde{\mathbf{N}}$, combined with the improved field evaluation formulation (10). In contrast to the results presented in the previous sections, the two formulations yield different results for both the near- and far-field quantities. Moreover, for the sake of the clarity of the presentation, all physical quantities corresponding to the out-of-plane NVF formulation are denoted with a tilde symbol.

Numerical results for increasing number of harmonics, $N = 2, ..., 320$, and number of computational layers, $M = 2, ..., 256$, are presented in figure 10. It can be inferred from this figure that the in-plane formulation requires both a high number of FS coefficients, N, and layers, M, to achieve convergence to a result of $R^{\text{ref}} = 0.287\,88$, whereas the out-of-plane formulation yields fast convergence to $\tilde{R}^{\text{ref}} = R^{(320,256)} = 0.288\,37$ with respect to N, as per figures 10(a) and (b), respectively. This behavior is in agreement with the findings reported in [24]. The convergence of the calculated near-field, illustrated in figures 10(c) and (d), exhibits similar features. Specifically, the in-plane NVF formulation requires both high N and M to achieve a small self-error of $e\{\mathbf{E}^{(226,256)}\} = 4.7 \times 10^{-2}$, whereas this self-error can already be achieved with $N = 10$, $M = 256$ in the out-of-plane formulation. This clearly demonstrates a drastically improved efficiency to the calculation of the near-field of oblique diffraction gratings of the
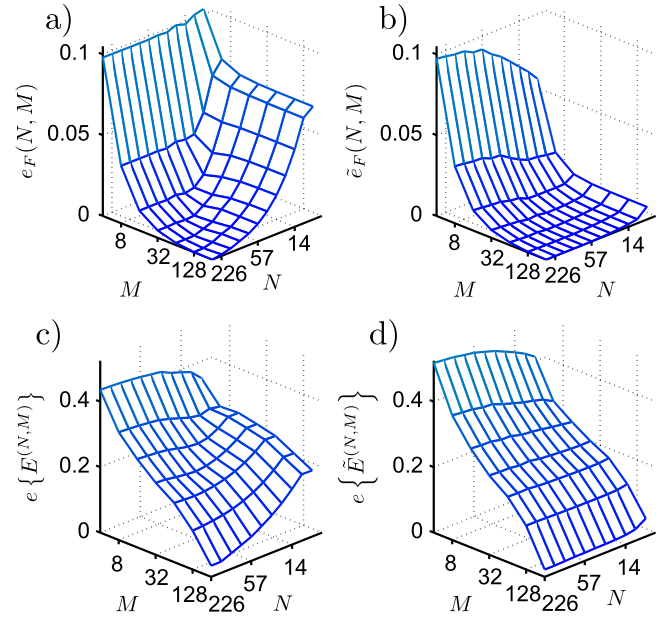

Figure 10. Computational results for the slanted 1D binary grating shown in figure 1(b). (a) Far-field self-error $e_{\text{F}}(N, M)$ versus N and M corresponding to the in-plane NVF formulation. (b) Far-field self-error $\tilde{e}_{\text{F}}(N, M)$ corresponding to the out-of-plane NVF formulation. (c) Near-field self convergence $e\{\mathbf{E}^{(N,M)}\}$ corresponding to in-plane NVF. (d) Near-field self convergence $e\{\tilde{\mathbf{E}}^{(N,M)}\}$ corresponding to out-of-plane NVF.
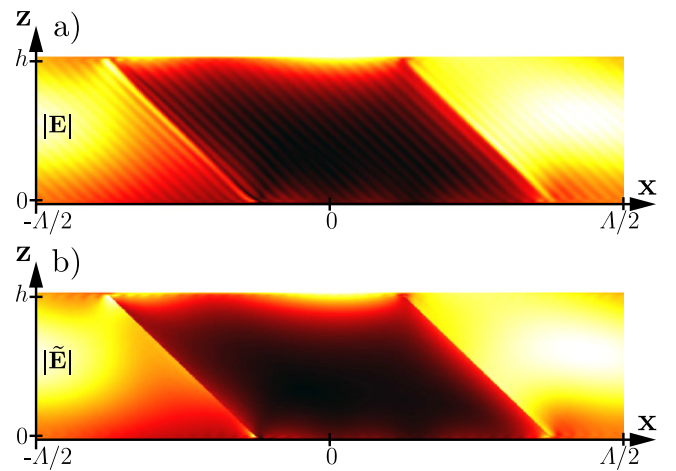

Figure 11. (a) Spatial distribution of the electric near-field, $|E_x^{(40,256)}|$, determined using the conventional in-plane NVF. (b) Spatial distribution of the electric near-field, $|\tilde{E}_x^{(40,256)}|$, determined for the same grating parameters as in (a) but using the modified out-of-plane NVF formulation.

approach based on the combination of out-of-plane NVF and the accurate near-field formulation. The highly improved near-field profile is illustrated in figure 11(b), which exhibits no unphysical oscillations near the gold–vacuum interface. This is in sharp contrast to the conventional field evaluation of the in-plane formulation (see figure 10(a)), which clearly suffers from spurious oscillations.
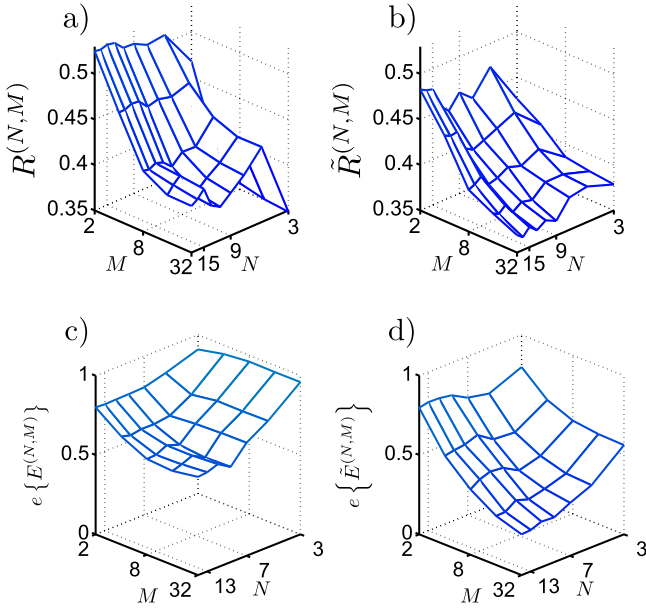
Figure 12. Computational results for the slanted 1D binary grating shown in figure 1(d). (a) Reflection coefficient R versus N and M, determined using the in-plane NVF formulation. (b) Reflection coefficient $\tilde{R}$ versus N and M, determined using the out-of-plane NVF formulation. (c), (d) Near-field convergence of conventional and modified formulation, respectively.

5.2. Analysis of slanted 2D-periodic cylindrical diffraction gratings

In this section we investigate the efficiency of using the accurate field evaluation and the out-of-plane NVF formulation to model a challenging, slanted 2D-periodic diffraction grating. The grating with periods, $\Lambda_1 = \Lambda_2 = \Lambda = 1\,\mu$m, is schematically depicted in figure 1(d) and consists of a cylindrical rod with radius $r = 0.3\,\mu$m and height $h = 0.125\,\mu$m, which is slanted by $\theta = \pi/4$ along the x-axis. Again, only the gold grating is considered in this section. Finally, the incident plane wave is impinging onto the grating along the normal direction, is polarized along the x-axis, and has a wavelength of $\lambda = 2\,\mu$m.

The reflection coefficient calculated for $N = 3, \ldots, 19$ harmonics and $M = 2, \ldots, 32$ layers is shown in figures 12(a) and (b) and was determined by using the conventional in-plane NVF and the out-of-plane NVF formulation, respectively. It can be seen that both approaches converge rather slowly, neither one achieves convergence even for the highest considered values of $M = 32$ and $N = 19$. Moreover, in order to characterize the error of the near-field calculations in the two formulations, the self-error with respect to the reference solutions obtained with $N = 19$ and $M = 32$ is presented in figures 12(c) and (d). The in-plane formulation achieves a relative self-error of $e\{\mathbf{E}^{(13,32)}\} = 0.625$, whereas for the same values of N and M, the modified field evaluation in conjunction with the out-of-plane NVF achieves a substantially lower self-error of $e\{\mathbf{E}^{(13,32)}\} = 0.261$. It has to be stressed, that the necessary accuracy for full convergence could not be achieved in our simulations. The evolution of the computational results for the shown values of $N \leqslant 19$ and

$M \leqslant 32$ however can be interpreted in favor of the out-of-plane NVF formulation, due to the lower near-field self-error $e\{\mathbf{E}^{(13,32)}\} < e\{\mathbf{E}^{(13,32)}\}$. It can be supposed that future simulations with finer discretization will reveal the practical benefit of the out-of-plane NVF formulation in conjunction with the modified field formulation for 2D-periodic slanted structures, similar to the case of 1D-periodic slanted diffraction gratings.

6. Conclusions

To summarize, we have analyzed the numerical near-field calculated using the RCWA and identified the Gibbs phenomenon as the main reason for their slow convergence and the spurious oscillations from which they suffer. As a solution to these deficiencies of RCWA, we proposed an improvement of this method that can be applied for modeling arbitrary diffraction gratings and, more generally, periodic optical structures. The modified formulation significantly improves the accuracy of 1D-RCWA calculations, for both straight and slanted gratings, where it speeds up convergence and removes the numerical artifacts from the calculated near-fields. The accuracy of 2D-periodic grating simulations can be enhanced, however, to a lesser extent than in the 1D case. The reduced performance in 2D can be attributed to the discontinuity and non-exactness of the numerical NVF, which is at the core of the modified formulation. Therefore, it might be fruitful to investigate more elaborate NVF-formulations and their suitability for near-field calculations, such as a complex valued NVF [42], which unlike the NVF used here is continuous everywhere in the grating region.

We expect that the proposed modification of the RCWA method will greatly advance its computational capabilities, especially for 1D periodic optical structures. In particular, this improved method could prove instrumental to accurate modeling of periodic plasmonic structures, diffraction gratings, and surface-nonlinear devices, namely to simulation of physical systems whose functionality rely on the electromagnetic near-field at interfaces.

References

[1] Taflove A and Hagness S C 2005 *Computational Electrodynamics* (Boston, MA: Artech House)
[2] Jin J M 2014 *The Finite Element Method in Electromagnetics* (New York: Wiley)

[3] Martin P A 2006 *Multiple Scattering: Interaction of Time-Harmonic Waves with N Obstacles* (Cambridge: Cambridge University Press)

[4] Mishchenko M I, Travis L D and Mackowski D W 1996 *J. Quant. Spectrosc. Radiat. Transfer* **55** 535–75

[5] Felbacq D, Tayeb G and Maystre D 1994 *J. Opt. Soc. Am.* A **11** 2526–38

[6] Popov E (ed) 2012 *Gratings: Theory and Numeric Applications* (Marseille: Institut Fresnel) (www.fresnel.fr/files/gratings/GratingsTheoryandNumericApplications.pdf)

[7] Palmer C A and Loewen E G (ed) 2005 *Diffraction Grating Handbook* (Springfield, Ohio: Newport Corporation)

[8] Cui T J, Smith D and Liu R 2009 *Metamaterials: Theory Design, and Applications* (Berlin: Springer)

[9] Neviere M and Popov E 2002 *Light Propagation in Periodic Media: Differential Theory and Design* (London: Taylor and Francis)

[10] Chandezon J, Dupuis M T, Cornet G and Maystre D 1982 *J. Opt. Soc. Am.* **72** 839–46

[11] Maystre D 1978 *J. Opt. Soc. Am.* **68** 490–5

[12] Shcherbakov A A and Tishchenko A V 2012 *J. Quant. Spectrosc. Radiat. Transfer* **113** 158–71

[13] Weismann M, Gallagher D F G and Panoiu N C 2015 *J. Opt. Soc. Am.* B **32** 523–33

[14] Moharam M G and Gaylord T K 1981 *J. Opt. Soc. Am.* **71** 811–8

[15] Moharam M G and Gaylord T K 1986 *J. Opt. Soc. Am.* A **3** 1780–7

[16] Moharam M G, Gaylord T K, Grann E B and Pommet D A 1995 *J. Opt. Soc. Am.* A **12** 1068–76

[17] Lalanne P and Morris G M 1996 *J. Opt. Soc. Am.* A **13** 779–84

[18] Li L 1996 *J. Opt. Soc. Am.* A **13** 1870–6

[19] Li L 1997 *J. Opt. Soc. Am.* A **14** 2758–67

[20] Essig S and Busch K 2010 *Opt. Express* **18** 23258–74

[21] Schuster T, Ruoff J, Kerwien N, Rafler S and Osten W 2007 *J. Opt. Soc. Am.* A **24** 2880–90

[22] Li L 2003 *J. Opt. A: Pure Appl. Opt.* **5** 345

[23] Mohram M G, Pommet D A, Grann E B and Gaylord T K 1995 *J. Opt. Soc. Am.* A **12** 1077–86

[24] Gushchin I and Tishchenko A V 2010 *J. Opt. Soc. Am.* A **27** 1575–83

[25] Kneipp K, Wang Y, Kneipp H, Perelman L T, Itzkan I, Dasari R R and Feld M S 1997 *Phys. Rev. Lett.* **78** 1667

[26] van Nieuwstadt J A H, Sandtke M, Harmsen R H, Segerink F B, Prangsma J C, Enoch S and Kuipers L 2006 *Phys. Rev. Lett.* **97** 146102

[27] Cao L, Panoiu N C and Osgood R M 2007 *Phys. Rev.* B **75** 205401

[28] Biris C G and Panoiu N C 2010 *Phys. Rev.* B **81** 195102

[29] Homola J, Yee S S and Gauglitz G 1999 *Sensors Actuators* B **54** 3–15

[30] Kabashin A V, Evans P, Pastkovsky S, Hendren W, Wurtz G A, Atkinson R, Pollard R, Podolskiy V A and Zayats A 2009 *Nat. Mater.* **8** 867–71

[31] Liu N, Mesch M, Weiss T, Hentschel M and Giessen H 2010 *Nano Lett.* **10** 2342–8

[32] Biris C G and Panoiu N C 2011 *Nanotechnology* **22** 235502

[33] Adam A J L, Brok J M, Seo M A, Ahn K J, Kim D S, Kang J H, Park Q H, Nagel M and Planken P C M 2008 *Opt. Express* **16** 7407–17

[34] Lalanne P and Jurek M P 1998 *J. Mod. Opt.* **45** 1357–74

[35] Popov E, Neviere M, Gralak B and Tayeb G 2002 *J. Opt. Soc. Am.* A **19** 33–42

[36] Lyndin N M, Parriaux O and Tishchenko A V 2007 *J. Opt. Soc. Am.* A **24** 3781–8

[37] Johnson P B and Christy R W 1972 *Phys. Rev.* B **6** 4370–9

[38] Weiss T, Granet G, Gippius N A, Tikhodeev S G and Giessen H 2009 *Opt. Express* **17** 8051–61

[39] Popov E and Neviere M 2001 *J. Opt. Soc. Am.* A **18** 2886–94

[40] Götz P, Schuster T, Frenner K, Rafler S and Osten W 2008 *Opt. Express* **16** 17295–301

[41] Photon Design Ltd *OmniSim/RCWA* http://photond.com/products/omnisim.htm

[42] Antos R and Veis M 2010 *Opt. Express* **18** 27511–24