

constants being much higher than that of the incident wave (we normalized all wave vectors to the wave vector of the incident light, meaning that the square shouldn't be more than a couple orders of magnitude larger than a unit square), or (ii) propagate the modes with substantial attenuation (i.e., with large imaginary parts of their propagation constants).

Next, the rectangular domain is split in half (in a real situation, any ratio may be used) over vertical and horizontal dimensions and use (24) again in order to locate “non-empty” (containing zeros) quadrants. Such iterations (splitting of zero-containing quadrants in four parts) are repeated until the resulting quadrants are small enough - meaning that zeros have been located close enough to ensure convergence of a given root-polishing algorithm. This process is schematically represented in Fig. 2, where crosses correspond to the values of k_x (eigenvalues) of a binary periodic structure.

To test our approach, we consider a layer consisting of only two elemental materials (see Fig. 1(b)). This binary structure (Material 1) is composed of a gain-doped silica layer (with thickness $\delta_1 = 45$ nm and complex dielectric permittivity $\epsilon_1 = 2.7224 - i0.029615$), and a layer of silver with $\delta_2 = 5$ nm and $\epsilon_2 = -26.079 + i0.882$. The incident p -polarized light of a wavelength $\lambda = 740$ nm enters at the incident angle, $\theta = \pi/3$, as shown in Fig. 1.

For this problem, the nonlinear equation to be solved is given by expression (21). We seek to get all k_x in a rectangle with the lower left corner being at $-100 - i100$, and the upper right corner at $100 + i100$.

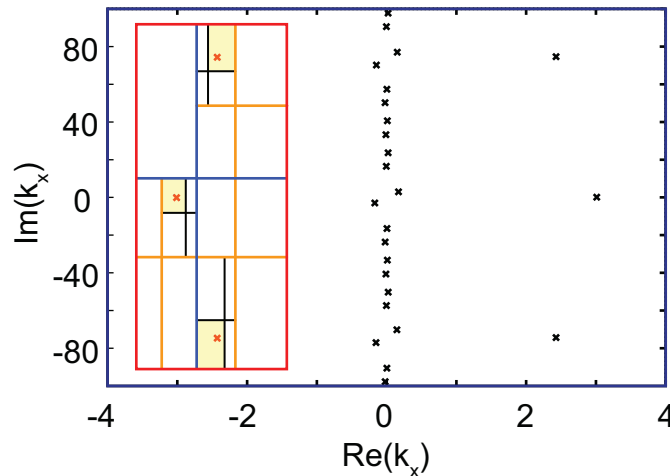


Fig. 2. Schematic representation of the Lehmer-Schur algorithm for Material 1. The red crosses (\times) correspond to the values of k_x detected through several consecutive splittings; the black crosses (\times) indicate the remaining values of k_x yet to be found.

Several consecutive splittings of the initial search area (a red rectangle) are shown in Fig. 2 by blue, orange, and black lines. The final zero-containing quadrants are shown as filled yellow rectangles. Overall, this is simply a generalization of the dichotomy method for the complex plane. *As calculations for each quadrant are completely independent from each other, these calculations can be done in parallel.*

It is convenient to use (25) in the following formulation:

$$\frac{1}{2\pi} \Delta_{\Gamma} \arg(f(z)) = Z - P,$$

here $\Delta_{\Gamma} \arg(f(z))$ is a change of argument of function $f(z)$ along the closed contour Γ . In this case, it is not required to calculate the derivatives of the function. However, troubles arise because available functions for calculating arguments of a complex number are well defined only on one of the branches (usually the principal one). In order to calculate $\Delta_{\Gamma} \arg(f(z))$, Γ is split into relatively small steps, and is checked whether arguments on every step change by a small value with respect to 2π . If this change is greater than some threshold value, let's call it argument smoothness tolerance (say, 0.1), then we split our small interval in half and repeat calculation of the argument for both parts. Iterations are continued until this branch point is located, or it is understood that there is no branch point on this interval. At this point, $\arg(f(z))$ has a jump of 2π , so we can correct our calculations of the argument function, which follows the principal branch only, by adding this jump to the sum of argument changes. It should be noted that instead of the argument function, the complex logarithm function can be used.

During application of the Lehmer-Schur algorithm, we should watch closely for the conservation of the total number of zeros. If the total number of zeros is not conserved, it usually means that there is a zero on one of the boundaries and this boundary must be moved. After the zeros are separated, the hybrid Powell algorithm (a modification of 2D Newton's algorithm) [36] is used in order to obtain every zero with a desired accuracy. Currently, an implementation provided by the GNU Scientific Library [37] is used.

3.1. Accuracy of the proposed algorithm

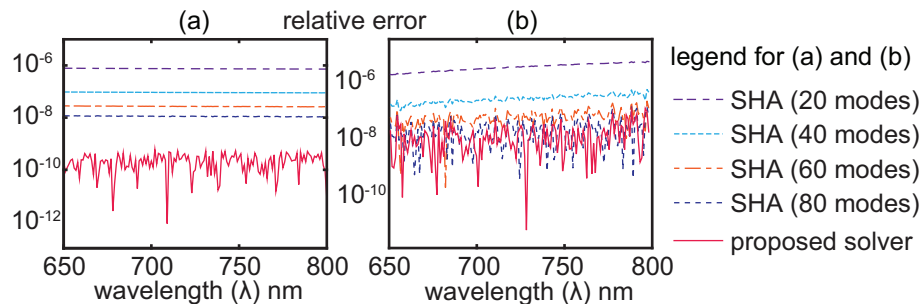


Fig. 3. Relative errors in real (a) and imaginary (b) parts of the propagating constant for the lowest eigenmode as a function of the incident wavelength λ obtained for test Material 2.

As a reference point, we have used the simulation results from 2D spatial harmonics analysis (SHA), proposed in [28] and validated in [27, 38] (2D-case) and in [27] (3D-case). The accuracy of the SHA method depends on the total number of Fourier modes. In the reference runs, 300 modes were taken. This number of modes is, with confidence, beyond the limit of convergence of the algorithm. This number is so high because the usual configuration of current nanostructures involves a complex piecewise constant function for the dielectric permittivity, which results in a slow convergence of the Fourier series.

We compared the simulation results of the lowest propagating mode (k_x has the smallest imaginary value) as a function of the incident light's wavelength. Results are shown in Fig. 3(a) and 3(b), where the binary periodic structure (Material 2) had the following configuration: layer of silica doped with gain inclusions of thickness $\delta_1 = 20$ nm, and a layer of silver of thickness $\delta_2 = 20$ nm, incident angle $\theta = 0$. The optical constants of silver and doped silica were kept the same as in Material 1.

We have used the following parameters of the algorithm: zeros tuning was started after the size of all the quadrants in the Lehmer-Schur algorithm smaller than 10^{-3} was achieved, ac-

curacy of the zero tuning was 10^{-16} (the absolute value of the function at the zero-point). It can be seen that the same (up to the round-off error) results for 300 harmonics in SHA have been obtained. These results are very promising, taking into account the performance tests discussed later. Here, it should also be stressed that the proposed algorithm takes into account the piecewise character of the structure by using the natural proper functions. This means that the exact eigenvalues are obtained up to the accuracy of calculated zeros, which is limited from below by a round-off error of a given floating-point representation.

In Fig. 4, we demonstrate that the propagation constants (k_x) from the MEP-SHA method converge to the results of the proposed method after increasing the number of harmonics. In this case (Material 3) the binary structure had the following composition: silica layer of thickness $\delta_1 = 20$ nm and $\epsilon_1 = 2.723$, and a layer of silver of thickness $\delta_2 = 20$ nm and $\epsilon_2 = -25.274 + i0.85436$, and the wavelength of normally incident light is 730 nm. Even in our quite simple example, only after taking into account at least 300 Fourier modes the MEP-SHA was capable of computing almost the same numerically exact (up to a floating-point representation accuracy) eigenvalues as the proposed method.

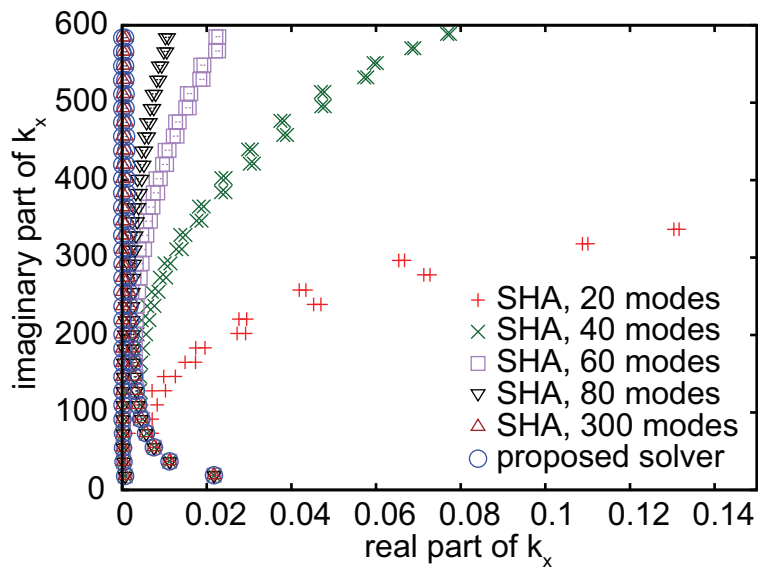


Fig. 4. Convergence of the SHA method (non-circular markers of different color) with increase of the total number of modes to the results of the new method (blue circles) simulated for test Material 3. The results for SHA with 300 modes (brown triangles) and for proposed solver (blue circles) are indiscernible.

3.2. Parallelization of the algorithm

After each iteration of the Lehmer-Schur process, a set of zero-containing quadrants is found. Every such quadrant is processed independently. Thus, we have a natural, parallel part of the algorithm. Another important part is the final tuning of zeros with Powell's algorithm, where each zero is independent. However, as we know that this part of the entire algorithm takes a small fraction of the total computation time, it is much more beneficial to optimize the calculation of $\arg(f(x))$ in the Lehmer-Schur process.

Currently, only the simplest Lehmer-Schur parallelization described above has been implemented. The shared memory model with the Linux implementation of POSIX threads as a parallelization mechanism is used.

For our current requirements (simulation of metasurfaces), we are satisfied with the present performance, which is more than one order of magnitude faster than the traditional, MEP-based SHA. The comparison of the performances of SHA against the proposed algorithm is shown in Fig. 5. As one can see from Fig. 5(a), scalability of the proposed algorithm is better than in the case of SHA, but still far from linear. The total performance is much better, which is demonstrated in Fig. 5(b). Moreover, in order to obtain comparable accuracy in MEP-based SHA (even for the lowest evanescent modes), one must use hundreds of modes.

For the proposed approach, we have used a freely available gcc compiler and standard libraries, while for the SHA code we have used a highly optimized Intel compiler together with a specialized Intel Math Kernel Library. So we still have a room for further acceleration of the new algorithm implementation. Hence, the current speed up numbers at Fig. 5(b) can be considered as minimal values. All the performance tests have been performed on Intel Xeon 5450 CPU operating at 3.0 GHz under CentOS 5.6 Linux.

4. Conclusion

We propose an algorithm for simulating periodic metasurfaces and cascaded metamaterial structures. The current formalism is developed for 2D-geometry, but can be generalized for 3D-case without any difficulties. We decided to limit ourselves to the 2D-case in order to keep all derivations relatively simple and clear for understanding. The proposed algorithm was implemented, and its accuracy and performance have been thoroughly tested. The preliminary results are truly encouraging - both in terms of accuracy and speed. Even the proof-of-concept, pilot version of the proposed algorithm is significantly (at least one order of magnitude) faster than the traditional, MEP-based implementation of SHA.

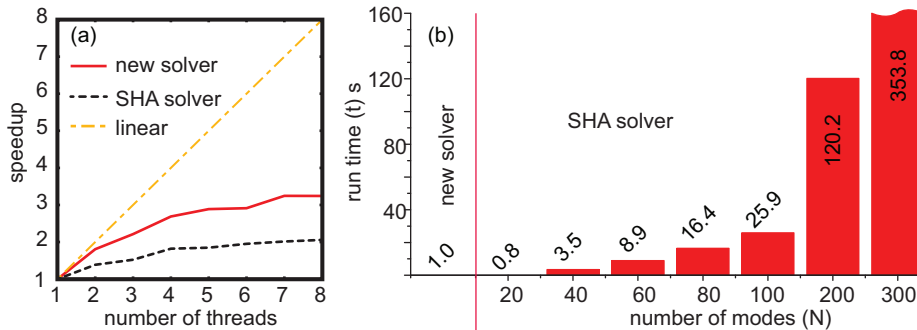


Fig. 5. (a) Scaling of the current implementation of the algorithm. (b) Computation time of the new solver and the MEP-based SHA solver vs. number of modes (N).

We must point out that even the simplest case of a slit in a metal film requires using at least 80 modes in the MEP-based SHA (or similar algorithms) in order to get any reasonable correspondence with experimental results [39]. In this case, our method offers more than one order-of-magnitude less time. At the same time, more involved configurations may require even more modes, as the main problem of the Fourier basis in MEP-based approaches is poor representation of piecewise constant distribution of dielectric constants in the media, which leads to slow convergence of higher frequency modes. This means that the further from the origin we go in the spatial spectrum, the more significant is the error (see Fig. 4). In other words, higher frequency modes obtained by SHA have a tendency to deviate from their exact values stronger than lower frequency modes. In contrast with the MEP-based approaches, the proposed method includes naturally the sharp boundaries. As a result, all calculated propagation constants are exact up to a desired level of accuracy, which is limited from below by a round-off error of floating-point representation.

One may speculate that the non-propagating modes are not that important. However, current metasurfaces and metamaterials operate mostly due to near-field coupling effects. In the

case of cascaded layers (which is natural for construction of bulk metamaterials) the near-field interaction becomes crucial. As it has already been demonstrated in [28], the significance of lower evanescent modes in cascaded structures increases dramatically with respect to single-layer case, and is still important even in thin symmetric structures on a substrate [40]. As a result, accurate calculation of the evanescent modes becomes absolutely necessary. This leads to an increasing number of modes in MEP-based methods, with N reaching several hundred modes. In this case, usage of our method is a must, because of several orders of magnitude advantage in performance.

As the number of required operations is dramatically decreased, the proposed solver demonstrates somewhat weak scalability, which is still better than the scalability of MEP-based methods. However, due to the exceptional speed enhancement of our new solver, even its current, non-optimal implementation exhibits very good performance. The solver already allows for fast global optimization of modern periodic metamaterials and metasurfaces by using a conventional, four-core desktop computer.

We truly believe that the proposed solver is a key to both fast computation of periodic nano-structures on conventional personal computers and feasible application of the optimization algorithms, which requires many repeating calculations of the given structure. Future work will include the implementation of the proposed algorithm into our free, in-the-cloud simulation tool [38], which currently employs a conventional, MEP-based version of 2D SHA. Speed-up techniques for spectral and parametric sweeps will also receive our special attention.

Acknowledgments

A.V.K. gratefully cites the support from the Gordon and Betty Moore Foundation Grant GBMF3385 and ARO Grant W911NF-13-1-0226. He also thankfully acknowledges motivating discussions with Prof. E. E. Narimanov. A.O.K. acknowledges useful discussions on numerical methods with Prof. K. S. Turitsyn, MIT, along with partial support by the University of New Mexico RAC grant #11-19. The authors thank developers of FFTW [41] and the whole GNU project (especially GNU Scientific Library [37]) for creating, developing, and supporting this free and useful software.