Rectangular dipoles in the discrete dipole approximation

Dmitry A. Smuneva, Patrick C. Chaumetb, Maxim A. Yurkinc,d,*

Belarusian State University, Nezavisimosti av. 4, 220030 Minsk, Belarus
Aix Marseille University, CNRS, Centrale Marseille, Institut Fresnel, UMR 7249, 13013 Marseille, France
Institute of Chemical Kinetics and Combustion SB RAS, Institutskaya Str. 3, 630090 Novosibirsk, Russia
Novosibirsk State University, Pirogova Str. 2, 630090 Novosibirsk, Russia

ABSTRACT

We performed a comprehensive analysis of the extension of the discrete dipole approximation (DDA) to a rectangular cuboid lattice of dipoles. The theoretical analysis of two different approaches, based either on the point–dipole interaction or on the integration of Green’s tensor (IGT), was performed starting with the rigorous integral equation for the electric field. We showed that the expressions for polarizability and interaction terms must strictly conform to each other, which resolves the existing controversy in the literature. Moreover, there are large differences between the spectra of the interaction matrix in the static limit for those DDA formulations. In particular, the point–dipole formulation leads to unphysical edges of the spectrum that deteriorate the convergence of the iterative solver with increasing refractive index. This severely limits the applicability of point–dipole DDA formulations with rectangular dipoles in contrast to the case of cubic dipoles. We implemented both above formulations in the open-source code ADDA and illustrated their performance on a number of test cases. In particular, we considered a graphene sheet, with thickness much smaller than the wavelength. The use of rectangular dipoles (with IGT) resulted in up to 100-times decrease of both simulation time and memory requirements, keeping the satisfactory accuracy. Similar improvements are expected for any strongly oblate or prolate particles in which the smallest dimension is much smaller than the wavelength.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

The discrete dipole approximation (DDA) is a widely used method to simulate scattering and absorption of electromagnetic waves by particles of arbitrary shape and internal structure [1,2]. Initially the DDA was proposed by Purcell and Pennypacker [3] on the basis of the physical picture of the point dipoles set. This approach was further advocated by Draine and coworkers [1,4,5]. However, the DDA can also be rigorously derived by discretization of the volume integral equation for the electric field [2,6–8]. In most cases the DDA is used with a cubic lattice of dipoles, because it permits the use of the fast Fourier transform (FFT) to greatly accelerate computations [9]. For such cubic lattice both DDA variants are mathematically equivalent, leading to the same equations. In principle, many DDA formulations exist, modifying either the interaction between dipoles or a formula to calculate dipole polarizability from the scatterer’s refractive index [2]. However, they all aim at rather small corrections, usually of relative order \((kd)^2\), where \(k\) is the free-space wave number and \(d\) is the dipole size (lattice spacing). In particular, all those
formulation converge to the exact solution, when number of dipoles goes to infinity for a fixed scattering problem [10]. Due to the above equivalence the term “dipole” is used interchangeingly both for the point dipole in a lattice (empirical DDA picture) and for volume discretization element (derivation from the integral equation) in the literature. We will further refer to this standard case as “cubic dipoles”.

Certain DDA applications are concerned with very prolate or oblate particles [11–16], where cubic dipoles are definitely inefficient. Consider, for example, a graphene nanoplate with sizes 9 μm × 9 μm × 20 nm [11]. Dipole size has to be small to fit into the nanoplate thickness, which implies redundantly large number of dipoles to cover the width. The natural solution to this problem is use of “rectangular dipoles”, more precisely, point dipoles or volumetric elements in the form of rectangular parallelepipeds placed on a rectangular cuboid lattice, which can be optimized for particular aspect ratios of the scatterer. This approach is compatible with the FFT acceleration and has been discussed in a number of papers [5,8,14,17,18]. Unfortunately, this discussion has mostly been theoretical without even test simulations.

In particular, Gutkowicz-Krusin and Draine [5] considered an infinite rectangular lattice of point dipoles and derived their polarizability (up to second order of kd) from the requirement that it should give the correct result for a plane wave in a homogeneous medium. We further denote this approach as corrected lattice dispersion relation (corrected LDR or CLDR), since in the specific case of cubic dipoles it is a minor correction to the previously derived LDR [4]. Only this specific case is implemented in the latest release of the DDSCAT code [19] and has been used for published simulations. We note, however, that traces of support for general rectangular lattice can be found in the code, in particular, in the input shape file. Similar rectangular-lattice sums, but for magneto-electric dipoles, were considered by Landy and Smith [18]. However, the simulation results were also presented only for cubic lattice. Massa et al. [20] proposed to use polarizability of the small rectangular parallelepiped, computed under assumption that electric field inside it is constant [21], which is reasonable for a volume element inside a larger scatterer. The same formulae were earlier derived by Tsang et al. [17]. Again, the simulation results were presented [20] only for a cubic lattice, when it is equivalent to the independently derived “integration of Green’s tensor” (IGT) formulation of the dipole polarizability [22] (computed up to the second order of kd). The problem is that for non-cubic dipoles the IGT expression for polarizability is markedly different from that based on abovementioned lattice sums even in the limit kd → 0.

A different approach was proposed by Chaumet et al. [8], who introduced the (full) IGT formulation of the DDA both for the interaction between dipoles and for their polarizability. Those formulae can be directly employed for both cubic and general rectangular dipoles, but the simulation results were presented only for the former. The IGT polarizability formula was given as an integral [8], but it should approximately equal the abovementioned closed-form expression of (second-order) IGT. Therefore, in the following we do not distinguish between those two. Tsang et al. [17] also mentioned that using IGT interaction (at least for nearby dipoles) can improve the accuracy, but implied that point–dipole interaction can be used as well (as in [20]). Recently, Agha et al. [14] used IGT-DDA (albeit naming it a dyadic Green’s technique) with certain numerical optimizations to compute integrals. To the best of our knowledge, that is the only paper that actually presented simulation results using the rectangular (non-cubic) dipoles. Use of the latter resulted in up to 3-fold faster simulation for 20 × 20 × 100 nm³ silver nanorod (with comparable accuracy).

In the current paper we, first, theoretically analyze the different approaches to use rectangular dipoles in the DDA, including spectral properties of the interaction matrix, and solve the controversy between the IGT and lattice-sums polarizabilities. Second, we implement both possible formulations in the open-source code ADDA [23] and test their relative performance. In particular, we show that the slow convergence of the iterative solver becomes a serious issue for lattice-sums approach applied to large refractive indices. Third, we apply the rectangular-dipole DDA to several sample oblate or prolate particle and show that it gives large acceleration (up to 100 times) compared to the standard DDA.

Finally, we note that the considered rectangular dipoles imply rectangular lattice of such dipoles and are, thus, different from cubic lattices of anisotropic (including rectangular) dipoles, which can be used to model either homogeneous scatterers with anisotropic refractive index [16,23,24] or inhomogeneous metamaterials [25]. Both latter cases were successfully treated with the standard DDA codes.

2. Rigorous derivation of the DDA

For completeness we recall the standard derivation of the DDA from the integral equation for the electric field E [2,22]. Assuming exp(−iωt) time dependence and non-magnetic scatterer, the integral equation is given as

\[
E(\mathbf{r}) = E^{\text{inc}}(\mathbf{r}) + \int_{V, V_0} d^3r \mathbf{G}(\mathbf{r}, \mathbf{r'}) \mathbf{E}(\mathbf{r'})
\]

\[+ M(V_0, \mathbf{r}) - L(\partial V_0, \mathbf{r}), \]

(1)

where \( E^{\text{inc}}(\mathbf{r}) \) and \( \mathbf{E}(\mathbf{r}) \) are the incident and total electric field at location \( \mathbf{r} \). \( \mathbf{G}(\mathbf{r}, \mathbf{r'}) \) is the free space Green’s tensor, defined as

\[
\mathbf{G}(\mathbf{r}, \mathbf{r'}) = \mathbf{G}(\mathbf{R}) = \frac{\exp(ik\mathbf{R})}{R} \begin{bmatrix} 1 & -i k R & 1 - i k R \end{bmatrix} - \frac{1}{R^2} \begin{bmatrix} 1 & -i k R & 1 - i k R \end{bmatrix} - \frac{i k R}{R^2} \mathbf{I} - \frac{3}{R^2} \mathbf{I} - \frac{3}{R^2} \mathbf{I}.
\]

(2)

where \( \mathbf{R} = \mathbf{r} - \mathbf{r'} \), \( R = |\mathbf{R}| \), \( \hat{\mathbf{R}} \) is a dyadic defined as \( \hat{R}_{\mu\nu} = R_\mu R_\nu \) (\( \mu \) and \( \nu \) are the Cartesian components of the vector or tensor), \( \mathbf{I} \) is the unity tensor, \( \mathbf{M} \) is the following integral associated with the finiteness of the
exclusion volume \( V_0 \)

\[
\mathbf{M}(V_0, \mathbf{r}) = \int_{V_0} d^3 r \left( \mathbf{G}^{st}(\mathbf{r}, \mathbf{r}') \mathbf{X}(\mathbf{r}') \mathbf{E}(\mathbf{r}') - \mathbf{G}^{st}(\mathbf{r}, \mathbf{r}') \mathbf{X}(\mathbf{r}) \mathbf{E}(\mathbf{r}) \right),
\]

where \( \mathbf{G}^{st}(\mathbf{r}, \mathbf{r}') \) is the static limit \((k \to 0)\) of \( \mathbf{G}(\mathbf{r}, \mathbf{r}') \):

\[
\mathbf{G}^{st}(\mathbf{r}, \mathbf{r}') = \frac{1}{R^3} \left( I - \frac{3 \mathbf{R} \mathbf{R}}{R^2} \right),
\]

\( \mathbf{I} \) is the so-called self-term tensor, related to the excluded singularity of Green's tensor:

\[
\mathbf{L}(\partial V_0, \mathbf{r}) = -\int_{\partial V_0} d^2 r \frac{\hat{n} \cdot \mathbf{R}}{R^3},
\]

where \( \hat{n} \) is an external normal to the surface \( \partial V_0 \) at point \( \mathbf{r}' \).

To solve Eq. (1) we divide \( V \) into subvolumes (dipoles) \( V_i, i = 1, \ldots, N, \) centered around \( \mathbf{r}_i \), and assume both \( \mathbf{X} \) and \( \mathbf{E} \) are constant within each dipole, leading to

\[
\mathbf{E}_i = \mathbf{E}_i^{inc} + \sum_{j \neq i} \mathbf{G}_{ij} \mathbf{X}_j \mathbf{E}_j + (\mathbf{M}_i - \mathbf{L}_i) \mathbf{X}_i \mathbf{E}_i,
\]

where \( \mathbf{E}_i = \mathbf{E}(\mathbf{r}_i), \mathbf{X}_i = \mathbf{X}(\mathbf{r}_i), \mathbf{E}_i^{inc} = \mathbf{E}^{inc}(\mathbf{r}_i), \mathbf{L}_i = \mathbf{L}(\partial V_i, \mathbf{r}_i), \mathbf{M}_i = \int d^3 r' \left( \mathbf{G}(\mathbf{r}_i, \mathbf{r}') - \mathbf{G}^{st}(\mathbf{r}_i, \mathbf{r}') \right).

\[
\mathbf{G}^{st}_{ij} \equiv \frac{1}{V_i} \int d^3 r \mathbf{G}(\mathbf{r}, \mathbf{r}') - \mathbf{G}^{st}(\mathbf{r}, \mathbf{r}')
\]

Such formulation is equivalent to the method of moments [26] applied to Eq. (1), using unit pulse and delta-function (point-matching) as basis and testing functions, respectively. And that is exactly the IGT formulation of the DDA [8]. Thus, we confirm that IGT is a direct consequence of Maxwell equations for any set of discretization elements \( V_i \). The only approximation used is that of slow variation of \( \mathbf{X} \) and \( \mathbf{E} \), which is reasonable if small enough dipoles are used. In particular, the solution of Eq. (6) should converge to the true one with refining discretization [10].

Let us further define the self-term Green's tensor as

\[
\mathbf{G}^{st}_{ii} \equiv (\mathbf{M}_i - \mathbf{L}_i)/V_i,
\]

dipole polarization as \( \mathbf{P}_i = V_i \mathbf{X} \mathbf{E}_i \), and its polarizability as

\[
\alpha_i = \left( V_i \mathbf{X}_i \right)^{-1} \mathbf{G}^{st}_{ii} = V_i \mathbf{X}_i \left( \mathbf{I} - \mathbf{G}^{st}_{ii} V_i \mathbf{X}_i \right)^{-1}.
\]

Using these definitions, Eq. (6) can be rewritten in the standard DDA form:

\[
\mathbf{G}^{st}_{ii} \mathbf{P}_i - \sum_{j \neq i} \mathbf{G}^{st}_{ij} \mathbf{P}_j = \mathbf{E}_i^{inc}.
\]

Further on we consider \( V_i \) to be equal rectangular parallelepipeds \((d_1 \times d_2 \times d_3)\) placed on a rectangular cuboid lattice. To describe the dipole size with a single variable we use \( d = \max(d_1, d_2, d_3) \). Corresponding expressions for \( \mathbf{L}_i \) and \( \mathbf{M}_i \) are given in Appendix A.

3. Set of point dipoles

To relate the IGT formulation to that corresponding to point dipoles we need to take a closer look at the commonly used approximation of replacing \( \mathbf{G}^{st}_{ij} \) (cf. Eq. (8)) by

\[
\mathbf{G}^{st}_{ij} = \mathbf{G}(\mathbf{r}_i, \mathbf{r}_j).
\]

This approximation is reasonable when the dipoles \( i \) and \( j \) are far apart but can be very wrong (relative error of order 1) for a specific pair of nearby, especially touching, dipoles. The reason why this approach works fine for cubic dipoles is discussed in [10]. Briefly, the single value of \( \mathbf{G}^{st}_{ij} \) is not important, rather the sum in Eq. (11) is. Generally (for most dipoles) the dipole \( i \) is at least several dipoles far from the particle surface or any internal interface. Then we can divide the sum into that over cubic shells around the dipole \( i \) and consider only a few of the closest shells, where the difference between \( \mathbf{G}^{st}_{ij} \) and \( \mathbf{G}^{st}_{ii} \) is expected to be the largest. Those shells are inside a homogeneous medium and corresponding distances are much smaller than the wavelength \( \lambda \), then as first approximation \( \mathbf{G} \approx \mathbf{G}^{st} \) and \( \mathbf{P} \approx \mathbf{const} \). In this case, the sum in Eq. (11) over nearby shells boils down to the sum over \( \mathbf{G}^{st}_{ij} \) alone, which is either an integral over cubically symmetric volume or a sum over cubically symmetric set of points of \( \mathbf{G}^{st}(\mathbf{r}_i, \mathbf{r}_j) \). Both of those equal zero due to symmetry and are thus equal to each other. To conclude, the difference between \( \mathbf{G}^{st}_{ij} \) and \( \mathbf{G}^{st}_{ii} \) causes only \( O(kd) \) differences in the solution of the DDA equations for cubic dipoles [10].

The above immediately explains what goes wrong for rectangular dipoles – the shells are no longer cubically symmetric and, therefore, the \( O(1) \) differences between \( \mathbf{G}^{st}_{ij} \) and \( \mathbf{G}^{st}_{ii} \) do not cancel out. Hence, combining point–dipole Green’s tensor with the polarizability, defined by Eqs. (9) and (10), (which was implied by Massa et al. [20]) leads to completely wrong results. We did confirm it by simulations – the error increases with aspect ratio of the dipoles, but does not decrease to zero with refining discretization (data not shown).

Still, a dense enough set of point dipoles placed on any regular grid seems an appropriate approximation of a homogeneous medium. The critical question then is that of dipole polarizability, which must be different from that in the IGT framework. Following [5], let us consider a test problem of infinite homogeneous media, where a plane incident wave is propagating. Since our goal here is to illustrate the phenomenon, we limit ourselves to the static limit \((k \to 0)\), while the general case up to the third order of \( kd \) is discussed in Appendix B. Physically, this case corresponds to a small sphere \( V \) in a constant incident field \( \mathbf{E}^{inc} \): the total field \( \mathbf{E} \) inside sphere is also constant. First, consider Eq. (1) with spherical exclusion volume and \( r \) in the center of \( V \). Then, the integral is zero due to symmetry, \( \mathbf{M} = 0 \), and \( \mathbf{L} = (4\pi/3)\mathbf{I} \), leading to the well-known solution:

\[
\mathbf{E} = \left( 1 + \frac{4\pi}{3} \mathbf{X} \right)^{-1} \mathbf{E}^{inc} = 3(\mathbf{I} + 2\mathbf{I})^{-1} \mathbf{E}^{inc}.
\]

The same solution should be obtained for a set of identical point dipoles [Eq. (11)], which implies

\[
V_d \sum_{j} \mathbf{G}^{st}_{ij} = -\frac{4\pi}{3} \mathbf{I},
\]

where \( i \) corresponds to the dipole in the center and \( V_d \) is the dipole volume. Since the particle is assumed much larger than \( d \), we can effectively extend the sum in Eq. (14)
to an infinite dipole grid. Note, however, that the obtained sum is conditionally convergent and the result may depend on the order of summation, see e.g. [27]. Thus, we consider it as a sum inside a sphere (or a cube), which size is then increased to infinity (conforming to the spherical shape of the original particle). Eq. (14) is the main ingredient in explaining the existing controversy – expression for the polarizability (or \( \mathbf{c}_{ij} \)) must conform to the expression for the interaction part (\( \mathbf{G}_{ij} \)).

For cubic dipoles, Eq. (14) is trivial. In both IGT and point–dipole formulations, the sum over any cubic shell of dipoles is zero due to symmetry (see discussion above), leaving only \( V_d \mathbf{C}_{ij} = -4\pi/3 \mathbf{I} \), which in turn implies the standard Clausius–Mosotti (CM) polarizability [2]. Note, however, that considering the actual particle (finite grid) in Eq. (14) results in polarizability corrections that are significant for dipoles near the surface [28].

For rectangular dipoles, the situation is markedly different. In the IGT formulation we can still use that the integral in Eq. (1) is zero (for spherical ferrofluid), such that \( \min \) corrections of \( O((V^2)_{ijkl}) \) for all dipoles, \( \mathbf{c}_{ij} \), matrix \( \mathbf{c}_{ij} \) accumulates for all dipoles inside \( V_c \). Such compensation is only possible if \( R_s \) is small enough, so that variation of dipole polarization can be neglected. And it comes at a cost of effectively enlarging the dipole size, since the point–dipole DDA equations are correct at a scale \( R_s \) instead of \( d \). Hence, one can expect that the accuracy of the point–dipole formulation with rectangular dipoles would be worse than that of the IGT for the same discretization. Another problem associated with the point–dipole formulation is discussed in Section 4.

4. Spectral properties

Following [29], let us consider the (eigen-)spectrum of the interaction matrix and operator. Note that these spectra are completely different from the wavelength spectrum. To simplify discussion, we consider only homogeneous scatterers with isotropic \( \varepsilon \) in this section, as well as in all simulations further on. However, the case of general tensor \( \varepsilon \) is also tractable [30]. First, we rewrite Eq. (6) in the matrix form as

\[
\mathbf{B} \mathbf{E} = (\mathbf{I} + (\varepsilon - 1)\mathbf{Z}) \mathbf{E} = \mathbf{E}^{\text{inc}},
\]

where \( \mathbf{E} \) and \( \mathbf{E}^{\text{inc}} \) are arrays compounded from \( \mathbf{E}_j \) and \( \mathbf{E}^{\text{inc}}_j \) for all dipoles, \( \mathbf{Z} = -V_d \mathbf{G}/(4\pi) \), matrix \( \mathbf{G} \) includes the values of Green’s tensor, including the diagonal ones, \( \mathbf{I} \) is the identity matrix. Matrix \( \mathbf{Z} \) depends only on the shape and size of the scatterer and is independent of \( \varepsilon \). Complex-symmetric matrices \( \mathbf{B} \) and \( \mathbf{Z} \) are finite discretizations of corresponding infinite-dimensional singular operators, cf. Eq. (1), which we further denote as \( \mathbf{B} \) and \( \mathbf{Z} \), respectively.

Budko and Samokhin [30] proved that in the static limit, all eigenvalues of \( \mathbf{Z} \) are real values between 0 and 1 (so-called essential spectrum). The latter is supplemented by morphological-dependent resonances (e.g., Mie resonances [31]) for larger scatterers [30,32]. The analysis of such resonances, if possible at all, will be inherently limited to a particular particle shape. Hence, for the remainder of this section we consider only the static case as it provides general insights on the difference between the DDA formulations. In this case, the matrix \( \mathbf{B} \) is Hermitian and \( \mathbf{Z} \) is real symmetric, both independent of the particle size. The latter has real eigenvalues \( \{z_n\} \), which

where in the middle we used

\[
\sum_{j \in V_i} \mathbf{G}^{\text{st},0}_{ij} = \sum_{j \in V_i} \mathbf{G}^{\text{st},0}_{ij} - \sum_{j \in V_i} \mathbf{G}_{ij}^{\text{st}} \approx \frac{4\pi}{3V_d} \mathbf{I}
\]

\[
- \sum_{j \in V_i} \mathbf{G}_{ij}^{\text{st}} \approx \sum_{j \in V_i} \mathbf{G}^{\text{st},0}_{ij},
\]

(18)

which follows from Eq. (14) applied both to \( \mathbf{G}^{\text{st}}_{ij} \) and \( \mathbf{G}^{\text{st},0}_{ij} \). If we set, e.g., \( R_s = \sqrt{d} \) \( \min(\lambda, D) \) and take the limit of infinite discretization \( (d \rightarrow 0) \), then all equalities in Eqs. (17) and (18) become exact ones, i.e., Eq. (11) and hence its solution for point–dipole formulation converges to that of IGT. More detailed analysis, including a small fraction of dipoles near the interfaces and tracking the errors as orders of \( kd \), can be performed similar to that in [10]. However, we leave it for future studies.

In other words, the polarizability, defined by Eq. (16), compensates the difference between \( \mathbf{G}^{\text{st}}_{ij} \) and \( \mathbf{G}_{ij}^{\text{st}} \). If we set, e.g., \( R_s = \sqrt{d} \) \( \min(\lambda, D) \) and take the limit of infinite discretization \( (d \rightarrow 0) \), then all equalities in Eqs. (17) and (18) become exact ones, i.e., Eq. (11) and hence its solution for point–dipole formulation converges to that of IGT. More detailed analysis, including a small fraction of dipoles near the interfaces and tracking the errors as orders of \( kd \), can be performed similar to that in [10]. However, we leave it for future studies.

In other words, the polarizability, defined by Eq. (16), compensates the difference between \( \mathbf{G}^{\text{st}}_{ij} \) and \( \mathbf{G}_{ij}^{\text{st}} \). If we set, e.g., \( R_s = \sqrt{d} \) \( \min(\lambda, D) \) and take the limit of infinite discretization \( (d \rightarrow 0) \), then all equalities in Eqs. (17) and (18) become exact ones, i.e., Eq. (11) and hence its solution for point–dipole formulation converges to that of IGT. More detailed analysis, including a small fraction of dipoles near the interfaces and tracking the errors as orders of \( kd \), can be performed similar to that in [10]. However, we leave it for future studies.

In other words, the polarizability, defined by Eq. (16), compensates the difference between \( \mathbf{G}^{\text{st}}_{ij} \) and \( \mathbf{G}_{ij}^{\text{st}} \). If we set, e.g., \( R_s = \sqrt{d} \) \( \min(\lambda, D) \) and take the limit of infinite discretization \( (d \rightarrow 0) \), then all equalities in Eqs. (17) and (18) become exact ones, i.e., Eq. (11) and hence its solution for point–dipole formulation converges to that of IGT. More detailed analysis, including a small fraction of dipoles near the interfaces and tracking the errors as orders of \( kd \), can be performed similar to that in [10]. However, we leave it for future studies.

In other words, the polarizability, defined by Eq. (16), compensates the difference between \( \mathbf{G}^{\text{st}}_{ij} \) and \( \mathbf{G}_{ij}^{\text{st}} \). If we set, e.g., \( R_s = \sqrt{d} \) \( \min(\lambda, D) \) and take the limit of infinite discretization \( (d \rightarrow 0) \), then all equalities in Eqs. (17) and (18) become exact ones, i.e., Eq. (11) and hence its solution for point–dipole formulation converges to that of IGT. More detailed analysis, including a small fraction of dipoles near the interfaces and tracking the errors as orders of \( kd \), can be performed similar to that in [10]. However, we leave it for future studies.

In other words, the polarizability, defined by Eq. (16), compensates the difference between \( \mathbf{G}^{\text{st}}_{ij} \) and \( \mathbf{G}_{ij}^{\text{st}} \). If we set, e.g., \( R_s = \sqrt{d} \) \( \min(\lambda, D) \) and take the limit of infinite discretization \( (d \rightarrow 0) \), then all equalities in Eqs. (17) and (18) become exact ones, i.e., Eq. (11) and hence its solution for point–dipole formulation converges to that of IGT. More detailed analysis, including a small fraction of dipoles near the interfaces and tracking the errors as orders of \( kd \), can be performed similar to that in [10]. However, we leave it for future studies.

In other words, the polarizability, defined by Eq. (16), compensates the difference between \( \mathbf{G}^{\text{st}}_{ij} \) and \( \mathbf{G}_{ij}^{\text{st}} \). If we set, e.g., \( R_s = \sqrt{d} \) \( \min(\lambda, D) \) and take the limit of infinite discretization \( (d \rightarrow 0) \), then all equalities in Eqs. (17) and (18) become exact ones, i.e., Eq. (11) and hence its solution for point–dipole formulation converges to that of IGT. More detailed analysis, including a small fraction of dipoles near the interfaces and tracking the errors as orders of \( kd \), can be performed similar to that in [10]. However, we leave it for future studies.

In other words, the polarizability, defined by Eq. (16), compensates the difference between \( \mathbf{G}^{\text{st}}_{ij} \) and \( \mathbf{G}_{ij}^{\text{st}} \). If we set, e.g., \( R_s = \sqrt{d} \) \( \min(\lambda, D) \) and take the limit of infinite discretization \( (d \rightarrow 0) \), then all equalities in Eqs. (17) and (18) become exact ones, i.e., Eq. (11) and hence its solution for point–dipole formulation converges to that of IGT. More detailed analysis, including a small fraction of dipoles near the interfaces and tracking the errors as orders of \( kd \), can be performed similar to that in [10]. However, we leave it for future studies.
are also called generalized depolarization factors (form-factors), in analogy to the electrostatic problem for ellipsoid [33]. This spectrum is critical for the solution of Eq. (19) – if $\varepsilon$ is close or equal to an eigenvalue of $B$ the problem is ill-posed or solution does not exist at all. This condition corresponds to $\varepsilon = 1 - 1/z$. The interval $z \in [0,1]$ is safe, as it corresponds to physically unrealistic cases of negative $\varepsilon$, including extreme cases of 0 and $\infty$ that require separate consideration. However, any other real value of $z$ corresponds to a finite real positive $\varepsilon$.

As noted above, spectrum of $Z$ belongs to [0,1], implying that Eq. (1) is solvable for any physical $\varepsilon$. But any finite-dimensional approximation $\{z_n\}$ is only expected to converge to the operator spectrum with refining discretization. Numerical tests [29,30,32,33] proved that values of $z_n$ approximately fall into [0,1]; however, strict compliance does not necessarily take place. The point–dipole formulation of the DDA with cubic dipoles, equivalent to the CM one in the static limit, leads to “spill-out” of the spectrum by less than 0.1 [33], which should not cause problems except for high-contrast dielectric problems ($\varepsilon > 10$ or $0 < \varepsilon < 0.1$).

To analyze the spectra in a meaningful way one should account for largely different contribution of different eigenvalues. The weights $f_n$ depend on the particular scattering quantity of interest; in the following we consider those corresponding to the orientation-averaged absorption cross section ($C_{\text{abs}}$) normalized by particle volume [29,33]:

$$\frac{\langle C_{\text{abs}} \rangle}{kV} = \frac{1}{3N} \sum_{n=1}^{3N} \frac{f_n}{1/(\varepsilon - 1) + z_n} = \frac{1}{\varepsilon - 1} \sum_{n=1}^{3N} f_n \frac{dz}{1/(\varepsilon - 1) + z_n}, \quad (20)$$

where $f_n = \frac{1}{3N} \sum_{\mu=1}^{3} (\mathbf{u}_n, \mathbf{e}_\mu)^2$, $f(z) = \sum_{n=1}^{3N} f_n \delta(z - z_n)$, (21)

where $\mathbf{u}_n$ is an eigenvector corresponding to eigenvalue $z_n$, $\mathbf{e}_\mu$ is the vector combined of $N$ copies of unit vector along the $\mu$-axis. The function $f(z)$ fully describes the shape of the particle and is expected to converge to the one corresponding to $Z$ with refining discretization. In particular, this limit is $\delta(z - 1/3)$ for a spherical particle.

We have computed $f(z)$ for a sphere and a cube, using two DDA formulations and two types of dipoles – cubic and rectangular ones (elongated two-fold along the $z$-axis). For that we created a program in Mathematica, independent of the code described in Section 5 but based on the same formulae in the static limit. We used 16 dipoles along the $x$-axis. Unfortunately, significant increase of this number, e.g., to study the convergence of $f(z)$, is computationally prohibitive, since the complexity of spectral analysis is $O(N^3)$ in contrast to $O(N \log N)$ of a single DDA solution. We have further smoothed $f(z)$ by simple averaging with a bin width of 0.1; the results are presented in Fig. 1. First, note the general dependence of $f(z)$ on shape. For a sphere it is concentrated in the vicinity of $z = 1/3$, while for a cube it has large values over a large fraction of interval [0,1], falling off exponentially only close to 0 and 1. Apart from that the behavior of $f(z)$ for different DDA variants is similar for both shapes; thus, further conclusions are expected to be shape-independent.

All four DDA variants agree for the largest values of $f(z)$ and hence should have comparable accuracy in DDA simulations for $\varepsilon$ far from resonance regions (see below). The tails of $f(z)$ near $z = 0$ and 1 are systematically larger for rectangular dipoles, which can be partly explained by the twice smaller number of dipoles. The CM formulation for cubic dipoles leads to a small spill-out of the spectrum in line with above-mentioned published results. By contrast, the spectrum of $Z$ for the IGT is always within the interval [0,1], both for cubic and rectangular dipoles. This shows that IGT does a better job at approximating the spectrum of $Z$, which is not totally unexpected. However, a rigorous proof of such spectrum-preserving property is not readily available. In this respect, the IGT is similar to the “filtered coupled dipoles” (FCD) formulation of the DDA for cubic dipoles [29].

Finally, the most important and striking feature in Fig. 1 is the wide tails of $f(z)$ for CM formulation with rectangular dipoles. Although the absolute values of $f(z)$ are small on the tails (decay exponentially), they present a large problem for corresponding values of $\varepsilon$ in terms of both accuracy (cf. Eq. (20)) and convergence of the iterative solver. For instance, such common refractive index as $m=1.5$ (corresponding to $z = -0.8$) is already very hard for such DDA formulation. Moreover, the spectrum of $B$ is now not a line from 1 to $\varepsilon$, but significantly extends on both ends. According to the analysis of Rahola [32], this leads to significantly slower convergence of the iterative solver, even for values of $\varepsilon$ far from the abovementioned resonances (see also Fig. 4(b) below).
The rigorous derivation of the spectral spill-out width as a function of dipole aspect ratios is a challenging task outside the scope of this paper. However, this width is comparable to the correction to the diagonal of the matrix $Z$ by lattice sums (Eq. (B7)). In particular, we computed $f(z)$ for the same sphere using elongated dipoles with different aspect ratios from 1:1:1 (cubic) to 1:1:3. The result in Fig. 2 shows that the left boundary is well approximated by $z = R_0(3)$ (Table B1), depicted as vertical lines. This correlation can be partly explained by the fact that lattice-sum correction is designed to compensate the inherent flaws of the point–dipole formulation (Section 3). Since this compensation is not perfect due to finite discretization, there should be a residual spectral spill-out with the width comparable to the spectrum of the correction. The largest of the latter in absolute value is $R_0(3)$. Note also that the right tails of $f(z)$ in Fig. 2 are approximately symmetric to the left ones with respect to $z = 1/3$.

The main practical consequence is that the spectral spill-out (and all associated problems) increases with increasing dipole aspect ratio. As shown in Appendix B, absolute value of components of $R_0$ increases without limit in this case, that results in the spectrum of $Z$ occupying the whole real line. By contrast, the diagonal elements of $Z$ in the IGT formulation are always in the range [0,1], even allowing the limit of infinitely thin plate- or needle-like dipoles (Appendix A). This indicates that off-diagonal terms do not have principal flaws that would require drastic compensation.

To finalize this section we note that there exist at least two ideas to alleviate the spectral spill-out of point–dipole DDA formulation. First, the iterative solver may produce reasonable result even for singular linear system, if the iterations are stopped early enough [34]. Thus, the applicability domain in terms of refractive index can be enlarged if relatively large threshold for the convergence of the iterative solver is used, e.g., $10^{-2}$ instead of usually used $10^{-5}$, and the corresponding errors are acceptable. Second, filtering of Green’s function has the potential to remove the spectral spill-out altogether, but it requires an extension of the FCD formulation, including derivations of all formulae [35], to a general rectangular grid of dipoles. Both these ideas are interesting topics for future research.

5. Software implementation

We have implemented the DDA formulations with rectangular dipoles, described in Sections 2 and 3, in the open-source DDA code ADDA. The details of the code, including specific formulae used to calculate far-field scattering quantities, are given in [23]. The main addition from the user perspective is the new command line option "-rect_dip …" specifying the aspect ratio of the dipoles, e.g., “1 1 2” in ADDA will then automatically use such dipoles in shape generation and further calculation. The rest of the user control is accomplished by existing ADDA functionality with clarification that command line options specifying dipole size, like "-dpl …", now apply to the size along the x-axis (the argument equals $\lambda/d_1$). The interaction term is calculated using either the point–dipole formula (Eq. (12)) or the IGT (Eq. (8)). The latter is implemented with minor changes in the existing code that was previously used for cubic dipoles, and was originally developed in [8]. The only drawback of this IGT implementation is that it requires a lot of computational time (comparable
to 500 iterations of the iterative solver). However, this problem can be alleviated by calculating the integral accurately only for nearby dipoles (within the cutoff distance $R_{\text{IGT}}$) and using $G_{ij}^0$ for larger distances. This is achieved by $\text{“-int igt < R}_{\text{IGT}}/d>\text{”}$ and allows one to obtain the result very close to that of the complete IGT, but only requiring several-iterations-worth of extra computational time. This approach can be partly justified by analysis similar to that including a sphere $V_s$ in Section 3. The particular optimal $R_{\text{IGT}}$ depends on the problem at hand; however, the satisfactory first guess could be $R_{\text{IGT}}=3d$ that is further denoted as IGT3 (see simulation results in Sections 6 and 7).

Currently, ADDA supports three polarizability formulations for rectangular dipoles: CM, CLDR, and IGT approximation, correct in the second order of $kd$ – IGT(SO). The CM is based on Eqs. (10), (16), and (B7), while CLDR additionally includes corrections of up to the second order of $kd$ – Eq. (B17). Both are based on the lattice sums and are currently limited to a set of aspect ratios tabulated in [5] (see Appendix B). By contrast, IGT(SO) polarizability is based on explicit Eqs. (9), (A1) and (A3) and supports any aspect ratios.

The philosophy of ADDA is to give full control of DDA formulation to the user. This implies that one can use definitely wrong combination of polarizability and interaction terms (e.g., CM and IGT, respectively). ADDA will produce a warning but continue simulations in such cases. When using rectangular dipoles, we recommend using one of two combinations: CLDR polarizability with point–dipole interaction or IGT(SO) polarizability with IGT interaction. We further denote them as the CLDR and IGT formulations of the DDA, specified by command line options $\text{“-pol cldr -int poi” and “-pol igt_so -int igt …”}$ (with optional cutoff range), respectively. We study their performance in the following sections. Moreover, in the static limit, considered in Section 4, the CLDR formulation is equivalent to the CM.

**Fig. 4.** Contour plot of $N_{\text{iter}}$ (in log scale) versus $m$ for $kd_0=1$ sphere using the CLDR (a, b) and IGT (c, d) DDA formulations with cubic (a, c) and rectangular 1:1:2 (b, d) dipoles. Black regions mostly include the cases where the iterative solver was stopped after maximum allowed number of iterations.
The described code is freely available from the development branch of ADDA.\textsuperscript{1} It has been tested, in particular by simulations presented in the following sections, and is suitable for general scattering problems. However, the new functionality has not yet been fully integrated with all other ADDA parts. For instance, the rectangular dipoles are incompatible with calculation of radiation forces and with scatterers near plane substrate [36]. Moreover, the format of the shape file does not explicitly specify the dipole aspect ratios. In the future we plan to remove these limitations and make the rectangular-dipoles feature available in the main ADDA package.\textsuperscript{2} A detailed description of its usage will then be provided in the ADDA manual.

Finally, we note that ADDA can run on a cluster of processors with distributed memory, which allows one to handle very large problems. This feature is fully integrated with rectangular dipoles. In particular, the most time- and memory-consuming simulations of the following sections were run on the compute cluster of the Novosibirsk State University.

6. Test simulations

First, we perform standard convergence tests with increasing number of dipoles [10,29]. For that we consider a sphere and a cube, both with $kD_0=8$, where $D_0$ is the diameter or edge size, respectively, and discretized using from 8 to 128 dipoles per particle dimension along the $x$-axis ($n_x$). For a cube both propagation and polarization direction of the incident wave is along the cube edges. The rectangular dipoles have aspect ratios 1:1:2, but we also consider cubic dipoles for comparison. For both dipole shapes several formulations are considered (CM, CLDR, IGT, and IGT3). The value of refractive index is taken as $i_e=1.5$, which is close to the region, where point-dipole formulations are expected to have singularities due to spectral spill-out (Section 4). We show relative errors of extinction efficiency $Q_{\text{ext}}$ in Fig. 3, other scattering quantities behave similarly (data not shown). For the sphere the reference value is obtained using the Mie theory [31], for the cube – using the extrapolation technique based on five DDA simulations (CLDR with cubic dipoles) with $n_x$ between 128 and 256. The latter resulted in $Q_{\text{ext}}=4.2927$ with estimated uncertainty of $2 \times 10^{-4}$. For spheres we use the volume correction, which ensures that volume of the dipole set is equal to that of sphere (a default option in ADDA) [23].

Overall, Fig. 3 demonstrates the convergence of all DDA formulations with increasing number of dipoles down to very good accuracies. That is sufficient to verify the correctness of both theory and implementation; however, there are a few features worth discussing. First, the dips in some curves for spheres are artefacts of chosen representation, which discards the sign of the difference, i.e., they are related to the function $Q_{\text{ext}}(n_x)$ crossing the true value at some finite $n_x$ (see [10,37] for details).

Second, rectangular dipoles result in worse accuracy than cubic ones, which is totally expected for such symmetric shapes. The goal of this test is to show that rectangular dipoles are viable for any particle, while the cases, where rectangular dipoles do provide advantage, are discussed in Section 7.

The comparative performance of IGT3 and IGT is a bit trickier to explain. For rectangular dipoles IGT3 converges to a slightly wrong value, which is similar to the problem of combining point–dipole Green’s tensor with the IGT polarizability (see Section 3), but is much less pronounced. By further refining discretization up to $n_x=512$ (data not shown), we estimated the residual relative error to be within 0.05% for both test particles. In principle, this can be fixed by adjusting the polarizability expression to match this particular interaction formulation, but we leave it for a future study. By contrast, for cubic dipoles the difference between IGT3 and IGT does decrease to zero with dipole size. Moreover, if this difference is opposite in sign to the error of the IGT itself, IGT3 may seem more accurate than IGT, as is the case for the cube.

As a second test, we study the performance of the iterative solver in terms of number of iterations $N_{\text{iter}}$ as a function of $m$. We consider a small sphere ($kD_0=1$) to relate the results with conclusions of Section 4. For wavelength-sized and larger particles $N_{\text{iter}}$ is in most cases larger than the following values for the same $m$, especially for large $i_e$. The sphere was discretized with $n_x=32$ with either cubic or rectangular (1:1:2) dipoles. The default iterative solver (quasi-minimal residual) and convergence threshold ($10^{-2}$) were used. However, the iterations were stopped after 3N iterations even if not converged. Thus, $N_{\text{iter}}$ was effectively truncated at this level, which is 9.8 $\times$ 10\textsuperscript{4} and 4.9 $\times$ 10\textsuperscript{4} for cubic and rectangular dipoles, respectively. Both real and imaginary parts of $m$ were varied from 0 to 10 in steps of 0.5 with exception of the imaginary axis, where Re($m$)=0.05 (instead of 0) was used, and $m=1+0.001i$ instead of exactly 1. The results for CLDR and IGT formulations are shown in Fig. 4. They are almost identical to the results of CM and IGT3, respectively; thus, the data for the latter two are not shown.

The convergence of the iterative solver for the CLDR is significantly slower than that for the IGT even for cubic dipoles, especially for large real $m$. However, the most prominent feature of Fig. 4 is the failure of the CLDR for rectangular dipoles for most of the considered values of $m$, in agreement with analysis of Section 4. The particular applicability domain of CLDR depends on dipole aspect ratios and computational constrains, but it is definitely inferior for many applications. The overall pattern for the IGT with rectangular dipoles is very similar to that with cubic dipoles, $N_{\text{iter}}$ is from 1 to 3 times larger for all values of $m$ except that close to the imaginary axis. We can conclude that IGT formulation of rectangular dipoles is reliable in the wide range of refractive indices, while CLDR (or any other point–dipole one) is generally inferior to the IGT, both in terms of accuracy and computational speed. Finally, $N_{\text{iter}}$ (and Fig. 4) does depend on the specific iterative solver. However, our experience with cubic dipoles (data not shown) suggests that general trends in dependence of $N_{\text{iter}}$ on $m$ are the same for all iterative solvers, at least for those implemented in ADDA [36].

\textsuperscript{1}http://code.google.com/p/a-dda/source/browse/branches/rectangular_dipole.

\textsuperscript{2}http://code.google.com/p/a-dda/.
incident wave propagating along the z-axis, and their relative errors are shown in Fig. 5. To describe the accuracy by a single number we use root-mean-square relative error (RMSRE) over the angular range $[0, 90\degree]$, which we further denote as $\eta$.

First, note the rather unexpected result – even for cubic dipoles CLDR is much less accurate ($\eta=9.1\%$) than IGT3 (0.7%) because all dipoles are boundary ones (see discussion in Section 3). Thus, it is always recommended to use IGT for very thin particles. Second, using 2:2:1 dipoles leads to almost the same accuracy ($\eta=0.6\%$) but does it 7 times faster (4 times smaller $N$ and 1.4 times smaller $N_{\text{iter}}$). Here and further we consider only the time of the iterative solver, since it is a major part of the total DDA simulation time (except for the smallest problems) but is mostly independent of the secondary simulation parameters (e.g., number of scattering angles). Third, using 5:5:1 dipoles leads to a dramatic speed increase (42 times), but at the expense of lesser accuracy ($\eta=2.7\%$). In general, comparison of simulation speed of two methods is meaningful only when their accuracies are the same, e.g., [39]. However, the standard cubic-dipoles DDA cannot be accelerated for graphene plate even at expense of accuracy (dipole size is limited by the plate thickness). Thus, if moderate errors can be tolerated the acceleration due to rectangular dipoles is 42 and 200 times. Note also that decrease of $N$ does not only accelerate computations, but also decrease memory requirements proportionally. In particular, this makes larger particles or better accuracy achievable on a given hardware.

Finally, we once more emphasize that using CLDR with rectangular dipoles is completely impractical for such problems, since such formulation combines slow convergence of the iterative solver with very bad accuracy ($\eta$ more than 100%). The latter can only be improved by using at least 20 dipoles per plate thickness (data not shown).

The second example is a red blood cell (RBC) – also an oblate particle, but with all dimensions larger than $\lambda$. We used the shape mode, built into ADDA [36], with diameter $7.65 \mu m$, minimum and maximum thicknesses $0.69$ and $2.44 \mu m$, respectively, and diameter at the maximum thickness $5.46 \mu m$, as has been considered in [39]. The wavelength in the vacuum is $0.633 \mu m$, which corresponds to $\lambda=0.473 \mu m$ in the medium (buffered saline). The relative refractive index of the RBC is $1.045 + 8 \times 10^{-3}i$, and its symmetry axis lies in the $yz$-plane constituting a $30^\circ$ angle with the $z$-axis (incident propagation direction). Same as above, we compute $S_{11}$ in the $yz$-plane for scattering angles up to $90^\circ$ and use $\eta$ as an integral accuracy measure. The latter is calculated against a

![Fig. 5. Light scattering pattern for $9 \times 9 \times 0.02 \mu m^3$ graphene plate: (a) reference results obtained by the IGT3 with 2 nm cubic dipoles, (b) relative errors of four DDA formulations with cubic and rectangular dipoles (in log scale).](image-url)
reference obtained with cubic CLDR using very fine discretization (1400 dipoles per RBC diameter).

Similar to graphene, we tried cubic and rectangular (2:2:1) dipoles. However, here the variation of discretization level is possible in wide ranges for both dipole shapes. Thus, we need to compare the computational times to reach the same accuracy. This time depends on many factors, including the details of Green’s tensor integration in the IGT, but the most general factor is the product \( N N_{\text{iter}} \), which becomes the dominant one (simulation time is proportional to it) if \( N_{\text{iter}} \) is large enough. Therefore, we varied the number of dipole per RBC diameter from 150 to 300 for cubic and from 100 to 250 for rectangular dipoles. The resulting dependence of \( \eta \) on \( N N_{\text{iter}} \) is shown in Fig. 6. First, the value of \( \eta \) is less than 15% (generally acceptable) even for the coarsest used discretizations. Second, the whole curves for rectangular dipoles are almost the same for CLDR and IGT3 formulations, except that \( N_{\text{iter}} \) for IGT3 is systematically smaller. Note that satisfactory performance of CLDR with rectangular dipoles is due to \( m \) being close to 1. Third, the results for rectangular dipoles are generally comparable to that for cubic dipoles. However, there is a systematic superiority of rectangular dipoles for finer discretizations – up to twice better accuracy for the same computational requirements.

To summarize the above results we note that the magnitude of variation (derivative) of the electric field depends on the smaller of the two characteristic scales: the wavelength (isotropic) and particle dimensions (can be anisotropic). Thus, when one particle dimension is much smaller than both \( \lambda \) and another particle dimension, the derivative of the electric field is strongly anisotropic, which warrants the use of rectangular dipoles in the DDA. We have illustrated that by oblate particles, but we expect similar conclusions to hold for prolate needle-like particles (e.g., nanorods), for which thickness is much smaller than \( \lambda \).

8. Conclusion

This paper is devoted to the use of rectangular lattice of dipoles in the DDA. We theoretically analyzed two different approaches, namely the IGT and point–dipole one (based on lattice-sum polarizability), deriving them from the volume-integral equation for the electric field. Although both approaches are viable, we showed that the expressions for the polarizability and interaction terms must conform to each other. This resolves the existing controversy in the literature.

We also studied the spectrum of the interaction matrix in the static limit and found significant differences between the two DDA formulations. In the IGT formulation the spectrum of the interaction matrix is always limited to the physical bounds of the spectrum of the integral scattering operator. By contrast, the point–dipole formulations (CM or CLDR) cause the spectrum to extend beyond these physical bounds. Although this spectral spill-out can often be neglected for cubic dipoles, its width increases unboundedly with dipole aspect ratio. Practically, this implies unphysical resonances for real positive \( m \) and overall deterioration of the simulation accuracy and speed with increasing absolute value of \( m \). The slowdown is caused by slower convergence of the iterative solver, which has been confirmed by simulations. Thus, the applicability of point–dipole DDA formulations with rectangular dipoles is severely limited to the region of \( m \) close to 1, with specific bounds depending on dipole aspect ratios and particle size.

We implemented IGT, CM, and CLDR formulations with rectangular dipoles in the open-source code ADDA to make it available to other researchers. We tested the correctness of implementation on a number of simple test cases – all formulations converge with refining discretization. We further illustrated the virtues of the IGT formulation with rectangular dipoles by simulations of light scattering by a graphene sheet, with thickness much smaller than \( \lambda \). First, we found that use of IGT is critical even for cubic dipoles, since point–dipole (CLDR) formulation applied to two-dipoles-per-thickness discretization resulted in more than 10 times larger errors. Second, use of IGT with rectangular dipoles leads to more than 100 times faster simulations keeping the satisfactory accuracy. The required memory can be reduced by the same factor allowing one to handle larger scattering problems with a desktop computer.

We also applied the DDA with rectangular dipoles to a red blood cell – also an oblate particle, but with thickness larger than \( \lambda \). Unfortunately, we noticed only marginal improvement in comparison with the cubic dipoles. Therefore, we conclude that use of rectangular dipoles is expected to be especially beneficial for strongly oblate and prolate particles, for which the smallest dimension is much smaller than \( \lambda \). This class includes many potential applications related to plate- or needle-like particles, e.g., in nanotechnology.

Finally, we stress that the DDA with rectangular dipoles owes its success to the IGT formulation, which has been largely underestimated by the DDA community. This formulation is currently implemented only in a single available DDA code (ADDA) and has been used only in a few papers. This makes further development and optimization of IGT a relevant task for future research. It may include optimization of numerical integration of Green’s tensor and developing approximate expressions based on expansions in orders of
kd and/or lookup (interpolation) tables. Also, the success of IGT with rectangular lattice of dipoles justifies the term “rectangular dipole” (volume element of the IGT), which we adopted in this paper.

Acknowledgements

This work was partly supported by the Russian Science Foundation (Grant no. 14-15-00155). M.A. Yu. also acknowledges the support by the Stipend of the President of Russian Federation for young scientists.

Appendix A: Self-term expressions for rectangular dipoles

In this appendix we provide the formulae for $\mathbf{L}$ and $\mathbf{M}$, in the IGT formulation of the DDA with rectangular dipoles, which were independently derived (in different forms) by Tsang et al. [17] and Massa et al. [20]. Since we assume all dipoles to be the same ($d_1 \times d_2 \times d_3$), we omit the index $i$ in this appendix for simplicity. Tensor $\mathbf{L}$ can be obtained analytically

$$L_{\mu\nu} = 2\delta_{\mu\nu} \Omega_{\mu},$$

(A1)

where $\delta_{\mu\nu}$ is the Kronecker symbol and $\Omega_{\mu}$ is the solid angle of rectangular dipole face perpendicular to the $\mu$-axis as observed from the dipole center ($\mu, \nu$ are Cartesian indices):

$$\Omega_{\mu} = 4 \arcsin \left( \frac{d_\mu}{\sqrt{d_\mu^2 + d_\nu^2 + d_\tau^2}} \right) = 4 \arctan \left( \frac{V_\mu}{d_\mu^2 D} \right),$$

(A2)

where $V_\mu = d_1 d_2 d_3$ is dipole volume, and $D = \sqrt{d_\mu^2 + d_\nu^2 + d_\tau^2}$ - its diagonal.

The integral in the expression for $\mathbf{M}$ (Eq. (7)) can be evaluated in closed form only if the integrand is expanded in powers of $kd$ (or, equivalently, $kD$) up to third order, which leads to

$$M_{\mu\nu} = \delta_{\mu\nu} \left( \frac{L^2}{2} \beta_{\mu} + \frac{2}{3} i k^3 V_\mu \right),$$

(A3)

$$\beta_{\mu} = \int_{V_\mu} d^3 r \frac{1}{r^2} \left( 1 + \frac{r^2}{k^2} \right) \left( \sum_{\nu=\mu} \frac{V_\nu}{d_\nu} \frac{\log (D + d_\nu)}{D - d_\nu} - \frac{1}{2} \frac{d_\nu^2}{D^2} \Omega_{\nu} \right) + \frac{2V_\mu}{d_\mu} \log \left( \frac{D + d_\mu}{D - d_\mu} \right),$$

(A4)

The latter is a simplified form of the expression given in the supplementary data of [21]. Since the main non-trivial part of $\mathbf{M}$ comes from terms of order $(kd)^2$ and due to historical reasons, we name the corresponding formulation of the DDA self-term (polarizability) “second-order IGT” IGT(SO) to distinguish it from precise evaluation of Eq. (7), which was done in [8]. However the difference between these two polarizability formulations is $O(k^4 d V_\mu)$, which is minor (especially for strongly elongated dipoles, when $V_\mu \ll d^3$) since $kd$ must be $\ll 1$ for the assumption of small variation of electric field inside the dipole to be valid.

It is instructive to analyze the behavior of $\mathbf{M}$ and $\mathbf{L}$ when the dipole is strongly elongated. In particular, let us analyze the case of plate-like and needle-like dipoles.

Direct implications of Eqs. (A1) and (A2) are

$$d_1 \ll d_2, d_3 \Rightarrow \Omega_1 = 4\pi, \quad \Omega_2 = \Omega_3 = 0,$$

(A5)

$$d_1, d_2 \ll d_3 \Rightarrow \Omega_1 = 8 \arctan (d_2/d_1), \quad \Omega_2 = 8 \arctan (d_1/d_2), \quad \Omega_3 = 0.$$ 

(A6)

In particular, $\Omega_{\mu\nu}$ is always in the range $[0, 4\pi]$. By contrast, Eqs. (A3) and (A4) imply

$$M = O \left( k^2 V_\mu \log \left( \frac{D}{D - d} \right) \right),$$

(A7)

which is $O((kd)^2)$ in the limit of one dipole dimension much smaller than the other. As discussed above, $kd$ must be small in all DDA simulations, so the $\mathbf{M}$ term can be neglected altogether for any extreme-elongation cases.

Appendix B: Lattice sums

To evaluate $C_{I}^{\alpha} = \log (Q)$ in Eq. (16) we perform derivations similar to that in [5]. First, let us define the reciprocal lattice $Q = q V_d/(2\pi)$, and its normalized variant $\mathbf{Q} = q V_d/(2\pi)$, where $\mathbf{n} = (n_1, n_2, n_3)$ is an integer vector. The regularized static Green’s tensor [35] (with extracted singularity) is defined as

$$C^{\alpha}(\mathbf{r}) = \left\{ \begin{array}{ll} 0, & \mathbf{r} \in V', \\
C^{\alpha}(\mathbf{r}), & \text{otherwise}, \end{array} \right.$$ 

(B1)

where $V'$ is the infinitesimal spherical exclusion domain (implying the limit $\mathbf{r} \to 0$). Then its Fourier transform is given by

$$F \left[ C^{\alpha}(\mathbf{r}) \right] (\mathbf{k}) = \lim_{\Omega \to 0} \int_{V'} d^3 r C^{\alpha}(\mathbf{r}) \exp (i \mathbf{k} \cdot \mathbf{r})$$

$$= \lim_{r \to 0} \int_{r}^{\infty} \frac{dR}{R} \int_{1}^{R} d \tau \exp (i \mathbf{k} \cdot \mathbf{r}) \int_{0}^{2\pi} d \varphi \left( \frac{2R^3}{R^2 - 1} \right)$$

$$= \pi \left( \mathbf{k} - \frac{3\mathbf{k}^2}{R^2} \right) \lim_{r \to 0} \int_{1}^{R} d \tau \exp (i \mathbf{k} \cdot \mathbf{r}) (1 - 3\tau^2)$$

$$= 4\pi \left( \frac{3}{2} \mathbf{k} - \frac{3\mathbf{k}^2}{R^2} \right),$$

(B2)

where $t = \cos (\theta)$, $\theta$ is the angle between $\mathbf{k}$ and $\mathbf{R}$, $\varphi$ is the azimuthal angle in plane perpendicular to $\mathbf{k}$, $u = \mathbf{R}$. The integral over $\varphi$ is calculated by expanding $\mathbf{R}$ into components parallel and perpendicular to $\mathbf{k}$: $\mathbf{R} = u \mathbf{k} + \tau \sqrt{1 - t^2}$, where $\mathbf{\tau} = u \mathbf{k}$ is the unit vector perpendicular to $\mathbf{k}$, and using that the averaged tensor $\hat{\tau} \hat{\tau}$ is proportional to the projector on the perpendicular plane:

$$\int_{0}^{2\pi} d \varphi \: \hat{\tau} \hat{\tau} = \pi \left( \mathbf{k} - \frac{3\mathbf{k}^2}{R^2} \right).$$

(B3)

The same result (Eq. (B2)) can also be obtained starting with the Fourier transform of $1/R$ [35]:

$$F[1/R](\mathbf{k}) = 4\pi/k^2,$$

(B4)

and using equivalence of $\hat{\mathbf{V}} \hat{\mathbf{V}} -> -\mathbf{k} \mathbf{k}$ under the Fourier transform. However, this requires rigorous consideration of integration–differentiation interchange, which can be done only using the generalized functions. In particular, the latter would lead to appearance of the term $\mathbf{I}$ in the final part of Eq. (B2).
Next, we apply Poisson summation formula, e.g., [40], together with Eq. (B1) and obtain:

\[
\sum_{j \neq i} C^0_{ij} = \frac{1}{V_d} \sum_n F(C^0_n|q) = \frac{4\pi}{V_d} \delta_{\mu \nu} R_0(\mu),
\]

where \( R_0(\mu) = \sum_n \left( \frac{1}{n^2} \frac{q^2}{Q^2} \right) = \sum_n \frac{n^2}{n^2} - \sum_n \frac{Q_j^2}{Q_j^2}, \)

(B6)

where the right-hand side of Eq. (B5) defines the tensor element-wise, and the value of the summand at the origin of the reciprocal lattice is taken direction-averaged, i.e., \( q^2/Q^2 = 1/3 \). The latter corresponds to the spherical outer boundary in evaluation of the original conditionally convergent sum (Section 3). Alternatively, the origin can be excluded from the summation, since it does not change the final result (both in Eq. (B6) and in the following sums). The sum \( R_0(\mu) \) is exactly the same as the one derived in [5] by summing vector potentials instead of electric fields. We do not actually evaluate those lattice sums although efficient algorithms are available [27]. Instead we refer to [5], where the values of \( R_0(\mu) \) are calculated for a number of dipole aspect ratios – one is unity, and the other two are chosen from a set \( 1, 1.5, 2, 3 \). We provide a subset of those values in Table B1, only for aspect ratios that we used in simulations.

To finalize the static case we combine Eqs. (16) and (B5) into

\[
V_d C^0_{\mu \nu} = -4\pi \delta_{\mu \nu} (1/3 + R_0(\mu)),
\]

(B7)

and consider the limit of strongly elongated dipoles. For the latter we notice that

\[
\sum_{j \neq i} C^0_{ij} = O \left( \min(d_1, d_2, d_3) \right)^3,
\]

(B8)

i.e., the magnitude of the sum is determined by the lattice points closest to the origin (see also the discussion in Section 3). Therefore, \( R_0(\mu) \) increases without bounds with increasing dipole elongation. More specifically, if one of the dipole dimensions (\( d_i \)) is much larger than the other, then the corresponding summation index (\( n_\mu \)) can be set to 0 (neglecting non-zero values of \( n_\mu \)). Thus, the sum in Eq. (B8) reduces to 2D or 1D one. For instance,

\[
d_1 < d_2, d_3 \Rightarrow R_0(1) = \frac{V_d}{4\pi} \sum_{n_\mu} \frac{2}{(n_\mu d_1)^3} = \frac{\zeta(3)}{\pi} \frac{d_2 d_3}{d_1^3},
\]

(R0)1

\[
R_0(2) = R_0(3) = -\frac{R_0(1)}{2},
\]

(B9)

Table B1

<table>
<thead>
<tr>
<th>( d_1 )</th>
<th>( d_2 )</th>
<th>( d_3 )</th>
<th>( R_0(1) )</th>
<th>( R_0(2) )</th>
<th>( R_0(3) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1.5</td>
<td>0.20426</td>
<td>0.20426</td>
<td>-0.40851</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0.38545</td>
<td>0.38545</td>
<td>-0.77090</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>3</td>
<td>0.74498</td>
<td>0.74498</td>
<td>-1.48995</td>
</tr>
</tbody>
</table>

\[
d_1 = d_2 \ll d_3 \Rightarrow R_0(3) = -\frac{V_d}{4\pi} \sum_{n_\mu} \frac{1}{(n_\mu^2 + n_\nu^2)} = -\frac{\zeta(3/2)\beta(3/2)}{2\pi} \frac{d_3}{d_1^3},
\]

(B10)

where \( \zeta \) and \( \beta \) are the Riemann zeta and Dirichlet beta functions, respectively [40]. The 2D sums in Eq. (B10) for general case of \( d_1 \neq d_2 \) are also discussed in [40].

Gutkowicz-Krusin and Draine [5] also derived the dynamic corrections to the above formula, which constitutes the CLDR formulation for general rectangular lattice. Here, we present their main formulæ for completeness. First, define additional lattice sums

\[
R_1 = \sum_n \frac{1}{n^2} - \sum_n \frac{1}{Q^2} = \sum_{\mu, \nu} R_2(\mu, \nu),
\]

(B11)

\[
R_2(\mu) = \sum_n \frac{n^2}{n^2} - \sum_n \frac{Q^2}{Q^2} = \sum_{\nu} R_3(\mu, \nu),
\]

(B12)

\[
R_3(\mu, \nu) = \sum_n \frac{n^2}{n^2} - \sum_n \frac{Q^2}{Q^2},
\]

(B13)

which are all expressed in terms of six independent components of \( R_3 \). These sums are tabulated in [5] for the same dipole aspect ratios as \( R_0 \) (see above). In particular, they all vanish for cubic lattice. The general expression for the self-term for a plane wave propagating in the direction, given by unit vector \( a \), is

\[
V_d C^0_{\mu \nu} = V_d C^0_{\mu \nu} - \frac{1}{k^2} V_d^{2/3} (\delta_{\mu \nu} N_\mu + m^2 a_\mu a_\nu K_{\mu \nu}) + \frac{2}{3} k^2 V_d \delta_{\mu \nu} + O ((kd)^4),
\]

(B14)

where

\[
N_\mu = c_3 + m^2 c_2 \left( 1 - 3a^2_\mu \right) - m^2 c_3 a^2_\mu - R_1 - (m^2 - 1) R_2(\mu) - 8m^2 a^2_\mu R_3(\mu, \nu) + 4m^2 \sum_{\nu} a^2_\mu R_3(\mu, \nu),
\]

(B15)

\[
K_{\mu \nu} = c_3 + R_1 - 2R_2(\mu) - 2R_2(\nu) + 8R_3(\mu, \nu) + c_4,
\]

(B16)

where \( c_1 = -5.9424219, c_2 = 0.5178819, c_3 = 4.0069747 \) are the LDR constants [4] and \( c_4 \) is an arbitrary constant. One can, in principle, associate \( V_d C^0_{\mu \nu} \) with \( -L_i \) and the remaining terms in Eq. (B14) with \( M_i \) (cf. Eq. (9)). However, this would only add ambiguity since \( L_i \) and \( M_i \) have specific (IGT-related) definitions that are generally incompatible with point–dipole formulation.

If one sets \( c_4 = -c_3 \) then \( C^0_{\mu \nu} \) is diagonal for a cubic lattice (for any \( a \) [5] – this is an accepted definition of cubic CLDR in current DDA implementations [19,23]. We extend this setting to rectangular lattice, which justifies the use of the same name. Unfortunately, \( C^0_{\mu \nu} \) is generally non-diagonal for rectangular lattice, which is incompatible with the current data structure inside ADDA that supports only diagonal polarizabilities. Therefore, we additionally discard non-diagonal values and obtain

\[
[C^0_{\mu \nu}|_{\mu \nu} = \delta_{\mu \nu} C^0_{\mu \nu}|_{\mu \nu} \Big|_{c_4 = -c_3},
\]

(B17)
Note that the latter approximation is exact for propagation vectors along any axis (then matrix \(a_{ij}\) is diagonal), which includes all simulations in this paper.

As a side note we propose a different way to diagonalize \(G_{ii}\), that is by averaging over \(a\). Then \(a_{ij}a_{ij} = a_{ii}\) and no limitations remain on the choice of \(c_i\). Thus, we choose the most obvious value \(c_i = 0\) and obtain:

\[
V_d \text{tr} G_{ii}^{0,av} = V_d \left( \text{tr} G_{ii} \mid c_i = 0 \right)_a = \delta_{ii} \left\{ -4 \pi \left( \frac{1}{3} + \text{R} (\mu) \right) - \frac{1}{\pi} k^2 V_d^{3/2} \left[ c_1 + \left( \frac{m^2}{3} - 1 \right) R_1 \right] - \left( m^2 - 1 \right) R_2 (\mu) + \frac{2}{3} k^2 V_d \right\}. \tag{B18}
\]

The physical sense of such averaging stems from the fact that the field inside (even large) finite scatterer is not expected to be that of a single plane wave, but rather can be approximated by a collection of plane waves with different directions (cf. a geometric-optics viewpoint). For cubic lattice \(G_{ii}\) is similar to that of ICT (cf. Eq. (A3)); in particular, it does not depend on \(m\). Note also that it differs from direction-averaged LDR [4] due to inaccuracies in the original LDR derivation. For rectangular lattice the dependence of \(G_{ii}\) on \(m\) is also relatively weak – the trace of the tensor is independent of \(m\).

References


