Convergence of the discrete dipole approximation.
I. Theoretical analysis

Maxim A. Yurkin

Faculty of Science, Section Computational Science, University of Amsterdam, Kruislaan 403, 1098 SJ, Amsterdam, The Netherlands, and Institute of Chemical Kinetics and Combustion, Siberian Branch of the Russian Academy of Sciences, Institutskaya 3, 630090 Novosibirsk, Russia

Valeri P. Maltsev

Institute of Chemical Kinetics and Combustion, Siberian Branch of the Russian Academy of Sciences, Institutskaya 3, 630090 Novosibirsk, Russia, and Novosibirsk State University, Pirogova Street 2, 630090 Novosibirsk, Russia

Alfons G. Hoekstra

Faculty of Science, Section Computational Science, University of Amsterdam, Kruislaan 403, 1098 SJ, Amsterdam, The Netherlands

We perform a rigorous theoretical convergence analysis of the discrete dipole approximation (DDA). We prove that errors in any measured quantity are bounded by a sum of a linear term and a quadratic term in the size of a dipole d when the latter is in the range of DDA applicability. Moreover, the linear term is significantly smaller for cubically than for noncubically shaped scatterers. Therefore, for small d, errors for cubically shaped particles are much smaller than for noncubically shaped ones. The relative importance of the linear term decreases with increasing size; hence convergence of DDA for large enough scatterers is quadratic in the common range of d. Extensive numerical simulations are carried out for a wide range of d. Finally, we discuss a number of new developments in DDA and their consequences for convergence. © 2006 Optical Society of America

OCIS codes: 290.5850, 260.2110, 000.4430.

1. INTRODUCTION

The discrete dipole approximation (DDA) is a well-known method to solve the light-scattering problem for arbitrary shaped particles. Since its introduction by Purcell and Pennypacker[1] (PP), it has been improved constantly. The formulation of DDA summarized by Draine and Flatau[2] more than ten years ago is still the one most widely used for many applications,[3] partly owing to the publicly available high-quality and user-friendly code DDSCAT.[4] Although modern improvements of DDA (as discussed in detail in Subsection 2.F) exist, they are still in the research stage because they are not widely used in real applications.

DDA directly discretizes the volume of the scatterer and hence is applicable to arbitrary shaped particles. However, the drawback of this discretization is the extreme computational complexity of DDA, although it is significantly decreased by advanced numerical techniques.[2,5] That is why the usual application strategy for DDA is single computation, where a discretization is chosen on the basis of available computational resources and some empirical estimates of the expected errors.[3,4] These error estimates are based on a limited number of benchmark calculations[3] and hence are external to the light-scattering problem under investigation. Such error estimates have evident drawbacks; however, no better alternative is available. Some results of analytical analysis of errors in computational electromagnetics are known, e.g., Refs. 6 and 7; however, they typically consider the surface-integral equations. To the best of our knowledge, such analysis has not been done for volume-integral equations (such as DDA).

Usually errors in DDA are studied as a function of the size parameter of the scatterer x (at a constant or few different total numbers of dipoles N), e.g., Refs. 2 and 8. Only a small number of papers directly present errors versus discretization parameter (e.g., d—the size of a single dipole).[9–17] The range of d typically studied in those papers is limited to a five-times difference between minimum and maximum values, with the exception of two papers[11,12] where it is 15 times. Those plots of errors versus discretization parameter are always used to illustrate the performance of a new DDA formulation and compare it with others. No conclusions about the convergence properties of DDA, as a function of d, have been made on the basis of these plots. To our knowledge, no theoretical analysis of DDA convergence has been performed; only a few limited empirical studies have appeared in the literature.

In this paper we perform a theoretical analysis of DDA convergence when refining the discretization (Section 2). We derive rigorous theoretical bounds on the error in any measured quantity for any scatterer. In Section 3 we present extensive numerical results of DDA computations

for five different scatterers using many different discretizations. These results are discussed in Section 4 to support conclusions of the theoretical analysis. We formulate the conclusions of the paper in Section 5. In a follow-up paper[18] (which from now on we refer to as Paper 2), the theoretical convergence results are used for an extrapolation technique to increase the accuracy of DDA computations.

2. THEORETICAL ANALYSIS

In this section we analyze theoretically the errors of DDA computations. We formulate the volume-integral equation for the internal electric field and its operator counterpart in Subsection 2.A and its discretization in Subsection 2.B. Subsection 2.C contains integral and discretized formulas for measured quantities that are the final goal of any light-scattering simulation. We derive the main results in Subsection 2.D, where we consider errors of the traditional DDA formulation[2] without shape errors, which are considered separately in Subsection 2.E. Finally, in Subsection 2.F we discuss some recent DDA improvements from the viewpoint of our convergence theory.

A. Integral Equation
Throughout this paper we assume the $\exp(-i\omega t)$ time dependence of all fields. The scatterer is assumed dielectric but not magnetic (magnetic permittivity $\mu=1$), and the electric permittivity is assumed isotropic [nonisotropic permittivity will significantly complicate the derivations but will not principally change the main conclusion of Section 2—expressions (70) and (87)].

The general form of the integral equation governing the electric field inside the dielectric scatterer is the following[19,20]:

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}^{inc}(\mathbf{r}) + \int_{V \setminus V_0} \mathrm{d}^3 r' \bar{\mathbf{G}}(\mathbf{r},\mathbf{r}')\chi(\mathbf{r}')\mathbf{E}(\mathbf{r}') + \mathbf{M}(V_0,\mathbf{r})$$
$$- \bar{\mathbf{L}}(\partial V_0,\mathbf{r})\chi(\mathbf{r})\mathbf{E}(\mathbf{r}), \qquad (1)$$

where $\mathbf{E}^{inc}(\mathbf{r}), \mathbf{E}(\mathbf{r})$ are the incident and total electric fields at location $\mathbf{r}$ and $\chi(\mathbf{r})=(\epsilon(\mathbf{r})-1)/4\pi$ is the susceptibility of the medium at point $\mathbf{r}$ [$\epsilon(\mathbf{r})$ is relative permittivity]. V is the volume of the particle (more generally, the volume that contains all points where the susceptibility is not zero), and V_0 is a smaller volume such that $V_0 \subset V$, $\mathbf{r} \in V_0 \setminus \partial V_0$. $\bar{\mathbf{G}}(\mathbf{r},\mathbf{r}')$ is the free-space dyadic Green's function, defined as

$$\bar{\mathbf{G}}(\mathbf{r},\mathbf{r}') = (k^2\bar{\mathbf{I}} + \hat{\nabla}\hat{\nabla})g(R)$$
$$= g(R)\left[k^2\left(\bar{\mathbf{I}} - \frac{\hat{R}\hat{R}}{R^2}\right) - \frac{1 - ikR}{R^2}\left(\bar{\mathbf{I}} - 3\frac{\hat{R}\hat{R}}{R^2}\right) \right], \qquad (2)$$

where $\bar{\mathbf{I}}$ is the identity dyadic, $k=\omega/c$ is the free-space wave vector, $\mathbf{R}=\mathbf{r}-\mathbf{r}'$, $R=|\mathbf{R}|$, $\hat{R}\hat{R}$ is a dyadic defined as $\hat{R}\hat{R}_{\mu\nu}=R_\mu R_\nu$ (μ, ν are Cartesian components of the vector or tensor), and $g(R)$ is the scalar Green's function

$$g(R) = \frac{\exp(ikR)}{R}. \qquad (3)$$

$\mathbf{M}$ is the following integral associated with the finiteness of the exclusion volume V_0:

$$\mathbf{M}(V_0,\mathbf{r}) = \int_{V_0} \mathrm{d}^3 r'(\bar{\mathbf{G}}(\mathbf{r},\mathbf{r}')\chi(\mathbf{r}')\mathbf{E}(\mathbf{r}') - \bar{\mathbf{G}}^s(\mathbf{r},\mathbf{r}')\chi(\mathbf{r})\mathbf{E}(\mathbf{r})),$$
$$(4)$$

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

$$\bar{\mathbf{G}}^s(\mathbf{r},\mathbf{r}') = \hat{\nabla}\hat{\nabla}\frac{1}{R} = -\frac{1}{R^3}\left(\bar{\mathbf{I}} - 3\frac{\hat{R}\hat{R}}{R^2}\right). \qquad (5)$$

$\bar{\mathbf{L}}$ is the so-called self-term dyadic:

$$\bar{\mathbf{L}}(\partial V_0,\mathbf{r}) = -\oint_{\partial V_0} \mathrm{d}^2 r' \frac{\hat{n}'\hat{R}}{R^3}, \qquad (6)$$

where $\hat{n}'$ is an external (as viewed from $\mathbf{r}$) normal to the surface ∂V_0 at point $\mathbf{r}'$.

Equation (1) can be rewritten in operator form as follows:

$$\widetilde{\mathbf{A}} \cdot \widetilde{\mathbf{E}} = \widetilde{\mathbf{E}}^{inc}, \qquad (7)$$

where $\widetilde{\mathbf{E}} \in H_1=L^1(V \to \mathbf{C}^3)$ represents functions from V to $\mathbf{C}^3$ that have finite L^1 norm and $\widetilde{\mathbf{E}}^{inc} \in H_2$ is a subspace of H_1 containing all functions that satisfy Maxwell equations in free space. $\widetilde{\mathbf{A}}$ is a linear operator: $H_1 \to H_2$. Although the Sobolev norm is physically sounder (based on the finiteness of energy of the electric field),[6,21] we use the L^1 norm. A detailed discussion of all assumptions made for the electric field is performed in Subsection 2.D.

B. Discretization
To solve Eq. (1) numerically, a discretization is done in the following way.[20] Let $V=\cup_{i=1}^N V_i$, $V_i \cap V_j=\emptyset$ for $i \neq j$. N denotes the number of subvolumes (dipoles). Assuming $\mathbf{r} \in V_i$ and choosing $V_0=V_i$, we can rewrite Eq. (1) as

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}^{inc}(\mathbf{r}) + \sum_{j \neq i} \int_{V_j} \mathrm{d}^3 r' \bar{\mathbf{G}}(\mathbf{r},\mathbf{r}')\chi(\mathbf{r}')\mathbf{E}(\mathbf{r}')$$
$$+ \mathbf{M}(V_i,\mathbf{r}) - \bar{\mathbf{L}}(\partial V_i,\mathbf{r})\chi(\mathbf{r})\mathbf{E}(\mathbf{r}). \qquad (8)$$

The set of Eq. (8) (for all i) is exact. Further, one fixed point $\mathbf{r}_i$ inside each V_i (its center) is chosen, and $\mathbf{r}=\mathbf{r}_i$ is set.

The usual approximation[20] considers $\mathbf{E}$ and χ constant inside each subvolume:

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}(\mathbf{r}_i) = \mathbf{E}_i, \quad \chi(\mathbf{r}) = \chi(\mathbf{r}_i) = \chi_i \quad \text{for } \mathbf{r} \in V_i. \qquad (9)$$

Equation (8) can then be rewritten as

$$\mathbf{E}_i = \mathbf{E}_i^{inc} + \sum_{j \neq i} \bar{\mathbf{G}}_{ij} V_j \chi_j \mathbf{E}_j + (\bar{\mathbf{M}}_i - \bar{\mathbf{L}}_i)\chi_i \mathbf{E}_i, \qquad (10)$$

where $\mathbf{E}_i^{inc}=\mathbf{E}^{inc}(\mathbf{r}_i)$, $\bar{\mathbf{L}}_i=\bar{\mathbf{L}}(\partial V_i,\mathbf{r}_i)$,

$$\bar{\mathbf{M}}_i = \int_{V_i} \mathrm{d}^3 r' (\bar{\mathbf{G}}(\mathbf{r}_i, \mathbf{r}') - \bar{\mathbf{G}}^s(\mathbf{r}_i, \mathbf{r}')), \qquad (11)$$

$$\bar{\mathbf{G}}_{ij} = \frac{1}{V_j} \int_{V_j} \mathrm{d}^3 r' \bar{\mathbf{G}}(\mathbf{r}_i, \mathbf{r}'). \qquad (12)$$

A further approximation, which is used in almost all formulations of DDA, is

$$\bar{\mathbf{G}}_{ij}^{(0)} = \bar{\mathbf{G}}(\mathbf{r}_i, \mathbf{r}_j). \qquad (13)$$

This assumption is made implicitly by all formulations that start by replacing the scatterer with a set of point dipoles, as was done originally by PP.[1] For a cubical (as well as spherical) cell V_i with $\mathbf{r}_i$ located at the center of the cell, $\bar{\mathbf{L}}_i$ can be calculated analytically, yielding[22]

$$\bar{\mathbf{L}}_i = \frac{4\pi}{3}\bar{\mathbf{I}}. \qquad (14)$$

Equation (10), together with Eqs. (13) and (14) and completely neglecting $\bar{\mathbf{M}}_i$, is equivalent to the original DDA by PP.[1] The diagonal terms in Eq. (10) are then equivalent to the well-known Clausius–Mossotti polarizability for point dipoles. Modifications introduced by other DDA prescriptions are discussed in Subsection 2.F.

In matrix notation, Eq. (10) reads as

$$\bar{\mathbf{A}}^d \mathbf{E}^d = \mathbf{E}^{inc,d}, \qquad (15)$$

where $\mathbf{E}^d$, $\mathbf{E}^{inc,d}$ are elements of $(\mathbf{C}^3)^N$ [vectors of size N where each element is a three-dimensional (3D) complex vector] and $\bar{\mathbf{A}}^d$ is a $N \times N$ matrix where each element is a 3×3 tensor. d is the size of one dipole. In operator notation Eq. (8) (for $\mathbf{r} = \mathbf{r}_i$) is as follows:

$$(\widetilde{\mathbf{A}}\widetilde{\mathbf{E}})(\mathbf{r}_i) = \widetilde{\mathbf{E}}^{inc}(\mathbf{r}_i) = \mathbf{E}_i^{inc,d}. \qquad (16)$$

We define the discretization error function as

$$\mathbf{h}_i^d = (\widetilde{\mathbf{A}}\widetilde{\mathbf{E}})(\mathbf{r}_i) - (\bar{\mathbf{A}}^d \mathbf{E}^{0,d})_i, \qquad (17)$$

where $\mathbf{E}^{0,d}$ is the exact field at the centers of the dipoles, $\mathbf{E}_i^{0,d} = \widetilde{\mathbf{E}}(\mathbf{r}_i)$, in contrast to $\mathbf{E}^d$ that is only an approximation obtained from the solution of Eq. (15) [here we neglect the numerical error that appears from the solution of Eq. (15) itself, which is acceptable if this error is controlled to be much less than other errors]. Using Eqs. (15)–(17), one can immediately obtain the error in internal fields due to discretization $\delta \mathbf{E}^d$:

$$\delta \mathbf{E}^d = \mathbf{E}^d - \mathbf{E}^{0,d} = (\bar{\mathbf{A}}^d)^{-1} \mathbf{h}^d. \qquad (18)$$

C. Measured Quantities
After having determined the internal electric fields, we can calculate scattered fields and cross sections. Scattered fields are obtained by taking the limit $r \to \infty$ of the integral in Eq. (1) (see, e.g., Ref. 23):

$$\mathbf{E}^{sca}(\mathbf{r}) = \frac{\exp(\mathrm{i}kr)}{-\mathrm{i}kr}\mathbf{F}(\mathbf{n}), \qquad (19)$$

where $\mathbf{n} = \mathbf{r}/r$ is the unit vector in the scattering direction and $\mathbf{F}$ is the scattering amplitude:

$$\mathbf{F}(\mathbf{n}) = -\mathrm{i}k^3(\bar{\mathbf{I}} - \hat{n}\hat{n})\sum_i \int_{V_i} \mathrm{d}^3 r' \exp(-\mathrm{i}k\mathbf{r}' \cdot \mathbf{n})\chi(\mathbf{r}')\mathbf{E}(\mathbf{r}'). \qquad (20)$$

All other differential scattering properties, such as the amplitude and Mueller scattering matrices, and asymmetry parameter $\langle \cos \theta \rangle$ can be easily derived from $\mathbf{F}(\mathbf{n})$, calculated for two incident polarizations.[24] We consider an incident polarized plane wave:

$$\mathbf{E}^{inc}(\mathbf{r}) = \mathbf{e}^0 \exp(\mathrm{i}\mathbf{k} \cdot \mathbf{r}), \qquad (21)$$

where $\mathbf{k} = k\mathbf{a}$, $\mathbf{a}$ is direction of incidence, and $|\mathbf{e}^0| = 1$ is assumed. The scattering and extinction cross sections (C_{sca}, C_{ext}) are derived from the scattering amplitude[23]:

$$C_{sca} = \frac{1}{k^2} \int \mathrm{d}\Omega |\mathbf{F}(\mathbf{n})|^2, \qquad (22)$$

$$C_{ext} = \frac{4\pi}{k^2} \mathrm{Re}(\mathbf{F}(\mathbf{a}) \cdot \mathbf{e}^{0*}), \qquad (23)$$

where * denotes complex conjugation. The expression for the absorption cross section (C_{abs}) directly uses the internal fields[23]:

$$C_{abs} = 4\pi k \sum_i \int_{V_i} \mathrm{d}^3 r' \, \mathrm{Im}(\chi(\mathbf{r}'))|\mathbf{E}(\mathbf{r}')|^2. \qquad (24)$$

Since only values of the internal field in the centers of dipoles are known, Eqs. (20) and (24) are approximated by (PP)

$$\mathbf{F}(\mathbf{n}) = -\mathrm{i}k^3(\bar{\mathbf{I}} - \hat{n}\hat{n})\sum_i V_i \chi_i \mathbf{E}_i^d \exp(-\mathrm{i}k\mathbf{r}_i \cdot \mathbf{n}), \qquad (25)$$

$$C_{abs} = 4\pi k \sum_i V_i \, \mathrm{Im}(\chi_i)|\mathbf{E}_i^d|^2. \qquad (26)$$

Corrections to Eq. (26) are discussed in Subsection 2.F.

Both Eqs. (20) (for each component) and (24) can be generalized as $\widetilde{\phi}(\widetilde{\mathbf{E}})$ (a functional that is not necessarily linear), which is approximated as

$$\widetilde{\phi}(\widetilde{\mathbf{E}}) = \phi^d(\mathbf{E}^d) + \delta\phi^d, \qquad (27)$$

where $\phi^d(\mathbf{E}^d)$ corresponds to Eq. (25) or (26), and the error $\delta\phi^d$ consists of two parts:

$$\delta\phi^d = [\widetilde{\phi}(\widetilde{\mathbf{E}}) - \phi^d(\mathbf{E}^{0,d})] + [\phi^d(\mathbf{E}^{0,d}) - \phi^d(\mathbf{E}^d)]. \qquad (28)$$

The first one comes from discretization [similar to Eq. (17)], and the second comes from errors in the internal fields.

D. Error Analysis
In this subsection we perform error analysis for the PP formulation of DDA. Improvements of DDA are further discussed in Subsection 2.F.

We assume cubical subvolumes with size d. We also assume that the shape of the particle is exactly described by these cubical subvolumes (we call this cubically shaped scatterer). Moreover, χ is a smooth function inside V (exact assumptions on χ are formulated below). An extension of the theory to shapes that do not satisfy these conditions is presented in Subsection 2.E. If there are several regions with different values of χ (smooth inside each region), the analysis is still valid, but interfaces inside V should be considered the same way as the outer boundary of V. We further fix the geometry of the scattering problem and incident field. Therefore we will be interested only in variation of discretization (which is characterized by the single parameter d); for reasons that will become clear in the sequel, we assume that $kd < 2$ (this bound is not limiting, since otherwise DDA is generally inapplicable[2]).

We switch to dimensionless parameters by assuming $k=1$, which is equivalent to measuring all the distances in units of $1/k$. The unit of the electric field can be chosen arbitrary but constant. In all further derivations we will use two sets of constants: γ_i and c_i. $\gamma_1 - \gamma_{13}$ are basic constants that do not depend on the discretization d but do depend directly on all other problem parameters—size parameter $x = kR_{eq}$ (R_{eq} is the volume-equivalent radius), m, shape, and incident field—or some of them. On the contrary, $c_1 - c_{94}$ are auxiliary values that either are numerical constants or can be derived in terms of constants γ_i. Although the dependencies of c_i on γ_i are not explicitly derived in this paper, one can easily obtain them following the derivations of this section. That is the main motivation for using such a vast amount of constants instead of an order-of-magnitude formalism. However, such explicit derivation has limited application because, as we will see further, constants in the final result depend on almost all basic constants. Qualitative analysis of these dependencies will be performed at the end of this subsection. It should be noted that the main theoretical results concerning DDA convergence [boundedness of errors by a quadratic function, cf. expression (70)] can be formulated and applied without consideration of any constants (which is simpler). However, our full derivation enables us to make additional conclusions related to the behavior of specific error terms.

The total number of dipoles used to discretize the scatterer is

$$N = \gamma_1 d^{-3}. \qquad (29)$$

We assume that the internal field $\widetilde{\mathbf{E}}$ is at least four times differentiable and all these derivatives are bounded inside V:

$$|\mathbf{E}(\mathbf{r})| \leq \gamma_2, \quad |\partial_\mu \mathbf{E}(\mathbf{r})| \leq \gamma_3, \quad |\partial_\mu \partial_\nu \mathbf{E}(\mathbf{r})| \leq \gamma_4,$$

$$|\partial_\mu \partial_\nu \partial_\rho \mathbf{E}(\mathbf{r})| \leq \gamma_5, \quad |\partial_\mu \partial_\nu \partial_\rho \partial_\tau \mathbf{E}(\mathbf{r})| \leq \gamma_6$$

$$\text{for } \mathbf{r} \in V \text{ and } \forall \, \mu, \nu, \rho, \tau. \quad (30)$$

This assumption is acceptable, since there are no inter-

faces inside V; therefore $\widetilde{\mathbf{E}}$ should be a smooth function. $|\cdot|$ denotes the Euclidian (L^2) norm, which is used for all 3D objects: vectors and tensors. We use the L^1 norm, $\|\cdot\|_1$, for N-dimensional vectors and matrices as well as for functions and operators. Expressions (30) immediately imply that $\widetilde{\mathbf{E}} \in L^1(V)$. We require that χ satisfies expressions (30) with constants $\gamma_7 - \gamma_{11}$. Further, we will state an estimate for the norm of $\bar{\mathbf{G}}(\mathbf{R})$ and its derivatives. One can easily obtain from Eq. (2) that for $R > 1$ $\bar{\mathbf{G}}(\mathbf{R})$ satisfies expressions (30) (with constants $c_1 - c_5$), while for $R \leq 2$

$$|\bar{\mathbf{G}}(\mathbf{R})| \leq c_6 R^{-3}, \quad |\partial_\mu \bar{\mathbf{G}}(\mathbf{R})| \leq c_7 R^{-4},$$

$$|\partial_\mu \partial_\nu \bar{\mathbf{G}}(\mathbf{R})| \leq c_8 R^{-5}, \quad |\partial_\mu \partial_\nu \partial_\rho \bar{\mathbf{G}}(\mathbf{R})| \leq c_9 R^{-6},$$

$$|\partial_\mu \partial_\nu \partial_\rho \partial_\tau \bar{\mathbf{G}}(\mathbf{R})| \leq c_{10} R^{-7} \quad \text{for } \forall \, \mu, \nu, \rho, \tau. \quad (31)$$

Next, we state two auxiliary facts that will be used later. Let V_c be a cube with size d and with its center at the origin and $f(\mathbf{r})$ be a four-times differentiable function inside V_c. Then

$$\left| \frac{1}{d^3} \int_{V_c} \mathrm{d}^3 r f(\mathbf{r}) - f(\mathbf{0}) \right| \leq c_{11} d^2 \max_{\mu\nu, \mathbf{r} \in V_c} |\partial_\mu \partial_\nu f(\mathbf{r})|, \quad (32)$$

$$\left| \frac{1}{d^3} \int_{V_c} \mathrm{d}^3 r f(\mathbf{r}) - f(\mathbf{0}) \right| \leq \frac{d^2}{24} |(\nabla^2 f(\mathbf{r}))_{\mathbf{r}=\mathbf{0}}|$$

$$+ c_{12} d^4 \max_{\mu\nu\rho\tau, \mathbf{r} \in V_c} |\partial_\mu \partial_\nu \partial_\rho \partial_\tau f(\mathbf{r})|. \quad (33)$$

Expressions (32) and (33) are the corollary of expanding f into a Taylor series. Odd orders of the Taylor expansion vanish because of cubical symmetry.

Our first goal is to estimate $\|\mathbf{h}^d\|_1$. Starting from Eq. (17), we write $\mathbf{h}_i^d$ as

$$\mathbf{h}_i^d = \sum_{j \neq i} \left(\int_{V_j} \mathrm{d}^3 r' \bar{\mathbf{G}}(\mathbf{r}_i, \mathbf{r}') \mathbf{P}(\mathbf{r}') - d^3 \bar{\mathbf{G}}_{ij}^{(0)} \mathbf{P}_j \right) + \mathbf{M}(V_i, \mathbf{r}_i),$$

$$(34)$$

where we have introduced the polarization vector for conciseness:

$$\mathbf{P}(\mathbf{r}) = \chi(\mathbf{r}) \mathbf{E}(\mathbf{r}), \quad \mathbf{P}_i = \mathbf{P}(\mathbf{r}_i). \quad (35)$$

It is evident that $\mathbf{P}(\mathbf{r})$ also satisfies expressions (30) (with constants $c_{13} - c_{17}$). We start by estimating $|\mathbf{M}(V_i, \mathbf{r}_i)|$. Substituting a Taylor expansion of $\mathbf{P}(\mathbf{r})$,

$$\mathbf{P}(\mathbf{R}) = \mathbf{P}(\mathbf{0}) + \sum_\rho R_\rho (\partial_\rho \mathbf{P})(\mathbf{0}) + \frac{1}{2} \sum_{\rho\tau} R_\rho R_\tau (\partial_\rho \partial_\tau \mathbf{P})(\widetilde{\mathbf{r}}(\rho, \tau, \mathbf{R})),$$

$$(36)$$

where $0 \leq \widetilde{r}_\mu \leq R_\mu$, into Eq. (4) gives

$$\mathbf{M}(V_i,\mathbf{r}_i) = \int_{V_i} d^3R(\bar{\mathbf{G}}(\mathbf{R}) - \bar{\mathbf{G}}^s(\mathbf{R}))\mathbf{P}_i$$
$$+ \frac{1}{2}\int_{V_i} d^3R\bar{\mathbf{G}}(\mathbf{R})\sum_{\rho\tau} R_\rho R_\tau(\partial_\rho\partial_\tau\mathbf{P})(\tilde{\mathbf{r}}(\rho,\tau,\mathbf{R})).$$

(37)

The norms of these two terms can be estimated as

$$\left| \int_{V_i} d^3R(\bar{\mathbf{G}}(\mathbf{R}) - \bar{\mathbf{G}}^s(\mathbf{R}))\mathbf{P}_i \right| = \left| \frac{2}{3}\bar{\mathbf{I}}\mathbf{P}_i \int_{V_i} d^3R g(R) \right| \le c_{18}d^2,$$

(38)

$$\left| \int_{V_i} d^3R\bar{\mathbf{G}}(\mathbf{R})\sum_{\rho\tau} R_\rho R_\tau(\partial_\rho\partial_\tau\mathbf{P})(\tilde{\mathbf{r}}(\rho,\tau,\mathbf{R})) \right|$$
$$\le 3c_{15}\int_{V_i} d^3R|\bar{\mathbf{G}}(\mathbf{R})|R^2 \le c_{19}d^2.$$

(39)

Expression (38) follows directly from the definitions in Eqs. (2) and (5). To derive expression (39), we used expressions (31) and the fact that $\sum_{\rho\tau}|R_\rho R_\tau| \le 3R^2$. Finally, expressions (37)–(39) lead to

$$|\mathbf{M}(V_i,\mathbf{r}_i)| \le c_{20}d^2.$$

(40)

To estimate the sum in Eq. (34), we consider separately three cases: (1) dipole j lies in a complete shell of dipole i (we define the shell below); (2) j lies in a distant shell of dipole i—$R_{ij}=|\mathbf{r}_j-\mathbf{r}_i|>1$; and (3) all j that fall between the first two cases (see Fig. 1). We define the first shell $[S_1(i)]$ of a cubical dipole as a set of dipoles that touch it (including touching in one point only). The second shell $[S_2(i)]$ is a set of dipoles that touch the outer surface of the first shell, and so on. The lth shell $[S_l(i)]$ is then a set of all dipoles that lie on the boundary of the cube with size $(2l+1)d$ and center coinciding with the center of the original dipole. We call a shell complete if all its elements lie inside the volume of the scatterer V. A shell is called a distant shell if all its elements satisfy $R_{ij}>1$; i.e., if its order $l>K_{max}=[1/d]$. Let $K(i)$ be the order of the first incomplete shell, which is an indicator of how close dipole i is to the surface. We demand $K(i)\le K_{max}$ to separate cases (1)
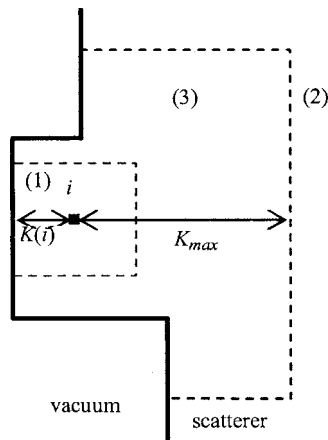

Fig. 1. Partition of the scatterer's volume into three regions relative to dipole i.

and (2) described above. All j that fall in the third case satisfy $R_{ij}<2$ (the exact value of this constant—slightly larger than $\sqrt{3}$—depends on d). The number of dipoles in a shell S_l (which can be incomplete)—$n_s(l)$—can be estimated as

$$n_s(l) \le (2l+1)^3 - (2l-1)^3 \le c_{21}l^2.$$

(41)

The sum of the error over all dipoles that lie in complete shells is then

$$\sum_{l=1}^{K(i)-1} \sum_{j\in S_l(i)} \left(\int_{V_j} d^3r'\bar{\mathbf{G}}(\mathbf{r}_i,\mathbf{r}')\mathbf{P}(\mathbf{r}') - d^3\bar{\mathbf{G}}_{ij}^{(0)}\mathbf{P}_j \right).$$

(42)

Since each shell in expression (42) is complete, it can be divided into pairs of dipoles that are symmetric over the center of the shell (j and $-j$). For convenience we set $\mathbf{r}=\mathbf{0}$. The inner sum in expression (42) can then be rewritten as

$$\frac{1}{2}\sum_{j\in S_l(i)} \left(\int_{V_j} d^3r'\bar{\mathbf{G}}(\mathbf{r}')(\mathbf{P}(\mathbf{r}') + \mathbf{P}(-\mathbf{r}')) - d^3\bar{\mathbf{G}}_{ij}^{(0)}(\mathbf{P}_j + \mathbf{P}_{-j}) \right).$$

(43)

Further, we introduce the auxiliary function

$$\mathbf{u}(\mathbf{r}') = \frac{1}{2}(\mathbf{P}(\mathbf{r}') + \mathbf{P}(-\mathbf{r}')) - \mathbf{P}(\mathbf{0}),$$

(44)

which satisfies the following inequalities [follows from expressions (30) for $\mathbf{P}(\mathbf{r})$ and Taylor series]:

$$|\mathbf{u}(\mathbf{r})| \le c_{22}r^2, \quad |\partial_\mu\mathbf{u}(\mathbf{r})| \le c_{23}r,$$
$$|\partial_\mu\partial_\nu\mathbf{u}(\mathbf{r})| \le c_{24} \quad \text{for } \forall \mu,\nu.$$

(45)

Then expression (43) is equivalent to

$$\sum_{j\in S_l(i)} \left(\int_{V_j} d^3r'\bar{\mathbf{G}}(\mathbf{r}')\mathbf{u}(\mathbf{r}') - d^3\bar{\mathbf{G}}_{ij}^{(0)}\mathbf{u}_j \right)$$
$$+ \sum_{j\in S_l(i)} \left(\int_{V_j} d^3r'\bar{\mathbf{G}}(\mathbf{r}') - d^3\bar{\mathbf{G}}_{ij}^{(0)} \right)\mathbf{P}_i, \quad (46)$$

where $\mathbf{u}_j=\mathbf{u}(\mathbf{r}_j)$. To estimate the first term, we apply expression (32) to the whole function under the integral. Using expressions (31) and (45), one may obtain

$$\max_{\mu\nu,\mathbf{r}'\in V_j} |\partial_\mu\partial_\nu(\bar{\mathbf{G}}(\mathbf{r}')\mathbf{u}(\mathbf{r}'))| \le c_{25}R_{ij}^{-3}$$

(47)

and hence

$$\sum_{j\in S_l(i)} \left(\int_{V_j} d^3r'\bar{\mathbf{G}}(\mathbf{r}')\mathbf{u}(\mathbf{r}') - d^3\bar{\mathbf{G}}_{ij}^{(0)}\mathbf{u}_j \right)$$
$$\le \sum_{j\in S_l(i)} c_{26}d^5R_{ij}^{-3} \le c_{27}d^2l^{-1}, \quad (48)$$

where we have used expression (41) and $R_{ij}\ge ld$ for $j\in S_l(i)$.

It is straightforward to show that

$$\sum_{j\in S_l(i)} \int_{V_j} d^3r'\bar{\mathbf{G}}(\mathbf{r}') = \frac{2}{3}\bar{\mathbf{I}}\sum_{j\in S_l(i)} \int_{V_j} d^3r'g(r'),$$

(49)

$$\sum_{j\in S_l(i)} \overline{\mathbf{G}}_{ij}^{(0)} = \frac{2}{3}\overline{\mathbf{I}} \sum_{j\in S_l(i)} g(R_{ij}). \tag{50}$$

The derivation is based on Eq. (2) and the equivalence $\hat{R}\hat{R}/R^2 \Leftrightarrow \frac{1}{3}\overline{\mathbf{I}}$ in all sums and integrals that satisfy cubical symmetry. Then the second part of expression (46) is transformed to

$$\left| \sum_{j\in S_l(i)} \left(\int_{V_j} \mathrm{d}^3r'\,\overline{\mathbf{G}}(\mathbf{r}') - d^3\overline{\mathbf{G}}_{ij}^{(0)} \right)\mathbf{P}_i \right|$$

$$\leq c_{28} \sum_{j\in S_l(i)} \left| \int_{V_j} \mathrm{d}^3r'\,g(\mathbf{r}') - d^3 g(R_{ij}) \right| \leq c_{29}d^4 l + c_{30}d^2 l^{-3}, \tag{51}$$

where we apply expression (33) to derive the second inequality and use the identity $\nabla^2 g(r) = -g(r)$ and the following inequalities:

$$|g(R)| \leq c_{31}R^{-1}, \quad |\partial_\mu \partial_\nu \partial_\rho \partial_\tau g(R)| \leq c_{32}R^{-5} \quad \text{for } \forall\, \mu,\nu,\rho,\tau. \tag{52}$$

Substituting expressions (48) and (51) into expression (42), one can obtain

$$\sum_{l=1}^{K(i)-1} \sum_{j\in S_l(i)} \left(\int_{V_j} \mathrm{d}^3r'\,\overline{\mathbf{G}}(\mathbf{r}_i,\mathbf{r}')\mathbf{P}(\mathbf{r}') - d^3\overline{\mathbf{G}}_{ij}^{(0)}\mathbf{P}_j \right)$$

$$\leq (c_{33} + c_{34}\ln K(i))d^2, \tag{53}$$

using the fact that $K(i)d \leq 1$.

We now consider the second part of the sum in Eq. (34) (where $R_{ij} > 1$). We first apply expression (32), then use expressions (30) for $\mathbf{P}(\mathbf{r})$ and $\overline{\mathbf{G}}(\mathbf{r})$, and finally invoke Eq. (29):

$$\sum_{j,R_{ij}>1} \left(\int_{V_j} \mathrm{d}^3r'\,\overline{\mathbf{G}}(\mathbf{r}_i,\mathbf{r}')\mathbf{P}(\mathbf{r}') - d^3\overline{\mathbf{G}}_{ij}^{(0)}\mathbf{P}_j \right)$$

$$\leq \sum_{j,R_{ij}>1} c_{35}d^5 \leq Nc_{35}d^5 \leq c_{36}d^2. \tag{54}$$

To analyze the third part of the sum in Eq. (34), we again sum over shells; however, since they are incomplete, we cannot use symmetry considerations. We apply expression (33) to the whole function under the integral and proceed analogously to the derivation of expression (51). Using the identity

$$\nabla^2\overline{\mathbf{G}}(\mathbf{r}) = -\overline{\mathbf{G}}(\mathbf{r}) \tag{55}$$

(since we have assumed $k=1$), we obtain

$$|\nabla^2(\overline{\mathbf{G}}(\mathbf{r})\mathbf{P}(\mathbf{r}))_{\mathbf{r}=\mathbf{R}_{ij}}| \leq c_{37}R_{ij}^{-4}, \tag{56}$$

$$\max_{\mu\nu\rho\tau,\mathbf{r}'\in V_j} |\partial_\mu\partial_\nu\partial_\rho\partial_\tau(\overline{\mathbf{G}}(\mathbf{r}')\mathbf{P}(\mathbf{r}'))| \leq c_{38}R_{ij}^{-7}, \tag{57}$$

which leads to

$$\sum_{j\in S_l(i)} \left(\int_{V_j} \mathrm{d}^3r'\,\overline{\mathbf{G}}(\mathbf{r}_i,\mathbf{r}')\mathbf{P}(\mathbf{r}') - d^3\overline{\mathbf{G}}_{ij}^{(0)}\mathbf{P}_j \right) \leq c_{39}dl^{-2} + c_{40}l^{-5} \tag{58}$$

and then analogously to expression (53),

$$\sum_{l=K(i)}^{K_{\max}} \sum_{j\in S_l(i)} \left(\int_{V_j} \mathrm{d}^3r'\,\overline{\mathbf{G}}(\mathbf{r}_i,\mathbf{r}')\mathbf{P}(\mathbf{r}') - d^3\overline{\mathbf{G}}_{ij}^{(0)}\mathbf{P}_j \right)$$

$$\leq c_{41}dK^{-1}(i) + c_{42}K^{-4}(i). \tag{59}$$

Collecting expressions (40), (53), (54), and (59), we finally obtain

$$|\mathbf{h}_i^d| \leq c_{41}dK^{-1}(i) + c_{42}K^{-4}(i) + (c_{43} + c_{44}\ln K(i))d^2. \tag{60}$$

Then

$$\|\mathbf{h}^d\|_1 = \sum_{i=1}^{N} |\mathbf{h}_i^d| \leq (c_{43} + c_{44}\ln K_{\max})Nd^2$$

$$+ \sum_{K=1}^{K_{\max}} n(K)(c_{41}dK^{-1} + c_{42}K^{-4}), \tag{61}$$

where $n(K)$ is the number of dipoles whose order of the first incomplete shell is equal to K. It is clear that

$$n(K) \leq n(1) \leq \gamma_{12}Nd, \tag{62}$$

where γ_{12} is the surface-to-volume ratio of the scatterer. Finally, we obtain

$$\|\mathbf{h}^d\|_1 \leq N[(c_{43} - c_{45}\ln d)d^2 + c_{46}d]. \tag{63}$$

The last term in expression (63) is mostly determined by dipoles that lie on the surface (or few dipoles deep) because it comes from the K^{-4} term in expression (61) (which rapidly decreases when moving from the surface). We define surface errors as those associated with the linear term in expression (63). Our numerical simulation (see Subsection 3.B) shows that this term is small compared with other terms for typical values of d; however, it is always significant for small enough values of d.

From Eq. (18) we directly obtain

$$\|\delta\mathbf{E}^d\|_1 \leq \|(\overline{\mathbf{A}}^d)^{-1}\|_1\|\mathbf{h}^d\|_1. \tag{64}$$

We assume that a bounded solution of Eq. (7) uniquely exists for any $\widetilde{\mathbf{E}}^{inc} \in H_2$; moreover, we assume that if $\|\widetilde{\mathbf{E}}^{inc}\|_1 = 1$, then $\|\widetilde{\mathbf{E}}\|_1 \leq \gamma_{13}$. These assumptions are equivalent to the fact that $\|\widetilde{\mathbf{A}}^{-1}\|_1$ exists and is finite (the operator $\widetilde{\mathbf{A}}^{-1}$ is bounded). Because $\overline{\mathbf{A}}^d$ is a discretization of $\widetilde{\mathbf{A}}$, one would expect that

$$\lim_{d\to 0}\|(\overline{\mathbf{A}}^d)^{-1}\|_1 = \|\widetilde{\mathbf{A}}^{-1}\|_1 = \gamma_{13}. \tag{65}$$

Although Eq. (65) seems intuitively correct, its rigorous proof, even if feasible, lies outside the scope of this paper. For an intuitive understanding, one may consult the paper by Rahola,[25] where he studied the spectrum of the discretized operator (for scattering by a sphere) and showed

that it does converge to the spectrum of the integral operator with decreasing d. It should, however, be noted that convergence of the spectrum implies only the convergence of the spectral (L^2) norm of the operator and not necessarily the convergence of the L^1 norm. Therefore Eq. (65) should be considered an assumption. It implies that there exists a d_0 such that

$$\text{for } d < d_0 \quad \|(\bar{\mathbf{A}}^d)^{-1}\|_1 \leq c_{47}, \qquad (66)$$

where c_{47} is an arbitrary constant larger than γ_{13} (although d_0 depends on its choice). For example, $c_{47} = 2\gamma_{13}$ should lead to a rather large d_0 (a rigorous estimate of d_0 does not seem feasible). Therefore $\|\delta\mathbf{E}^d\|_1$ satisfies the same constraint as $\|\mathbf{h}^d\|_1$ [expression (63)] but with constants $c_{48}-c_{50}$.

Next, we estimate the errors in the measured quantities and start with the discretization error [first part in Eq. (28)]. Examining Eqs. (20) and (24), one can see that expression (32) may be directly applied, leading to

$$|\tilde{\phi}(\tilde{\mathbf{E}}) - \phi^d(\mathbf{E}^{0,d})| \leq \sum_i c_{51} d^5 \leq c_{52} d^2. \qquad (67)$$

The second part in Eq. (28) is estimated as

$$|\phi^d(\mathbf{E}^{0,d}) - \phi^d(\mathbf{E}^d)| \leq \sum_i c_{53} d^3 |\delta\mathbf{E}_i^d| \leq c_{53} d^3 \|\delta\mathbf{E}^d\|_1$$

$$\leq (c_{54} - c_{55} \ln d) d^2 + c_{56} d, \qquad (68)$$

where we used Eq. (29). The estimation of the error for C_{abs} additionally uses the fact $|\delta\mathbf{E}_i^d|^2 \leq c_{57} |\delta\mathbf{E}_i^d|$ ($c_{57} = \max\limits_{d<2,i} |\delta\mathbf{E}_i^d|$).

By combining expressions (67) and (68), we obtain the final result of this subsection:

$$|\delta\phi^d| \leq (c_{58} - c_{55} \ln d) d^2 + c_{56} d. \qquad (69)$$

It is important to remember that the derivation was performed for constant x, m, shape, and incident field. There are 13 basic constants ($\gamma_1 - \gamma_{13}$). γ_1 [Eq. (29)] characterizes the total volume of the scatterer; hence it depends only on x. $\gamma - \gamma_{11}$ [expressions (30) for $\chi(\mathbf{r})$] can be easily obtained given the function $\chi(\mathbf{r})$; moreover, it is completely trivial in the common case of homogenous scatterers. γ_{12} [surface-to-volume ratio, expression (62)] depends on the shape of the scatterer and is inversely proportional to x. It is not feasible (except for certain simple shapes) to obtain the values of constants $\gamma_2 - \gamma_6$ [expressions (30)], since an exact solution for the internal fields is required. These constants definitely depend on all the parameters of the scattering problem. Moreover, these dependencies can be rapidly varying, especially near the resonance regions. The same is true for γ_{13} [L^1 norm of the inverse of the integral operator, Eq. (65)]. Finally, there is the important constant d_0 that also depends on all the parameters; however, one may expect it to be large enough (e.g., $d_0 \geq 2$) for most of the problems—then its variation can be neglected.

Before proceeding, we introduce the discretization parameter $y = |m|kd$. We employ the commonly used formula as proposed by Draine[8]; however, the exact dependence on m is not important because all the conclusions are still valid for constant m. Replacing d by y does not significantly change the dependence of the constants in expres-

sion (69), since they all already depend on m through the basic constants $\gamma_2 - \gamma_{11}$, γ_{13}. This leads to

$$|\delta\phi^y| \leq (c_{59} - c_{60} \ln y) y^2 + c_{61} y. \qquad (70)$$

It is not feasible to make any rigorous conclusions about the variation of the constants in expression (70) with varying parameters because all these constants depend on $\gamma_2 - \gamma_6$, γ_{13}, which, in turn, depend in a complex way on the parameters of the scattering problem. However, we can make one conclusion about the general trend of this dependency.

Following the derivation of expression (70), one can observe that c_{61} is proportional to γ_{12}, whereas c_{59} and c_{60} do not directly depend on it (at least part of the contributions to them is independent of γ_{12}). Therefore the general trend will be a decrease of the ratio c_{61}/c_{59} with increasing x (when all other parameters are fixed). This is a mathematical justification of the intuitively evident fact that surface errors are less significant for larger particles.

In the analysis of the results of the numerical simulations (Subsection 3.B), we will neglect the variation of the logarithm. Expression (70) then states that error is bounded by a quadratic function of y (for $d \leq d_0$). However, keep in mind that our derivation does not lead to an optimal error estimation; i.e., it overestimates the error and can be improved. For example, the constants $\gamma_2 - \gamma_6$ are usually largest inside a small volume fraction of the scatterer (near the surface or some internal resonance regions), whereas in the rest of the scatterer the internal electric field and its derivatives are bounded by significantly smaller constants. However, the order of the error is estimated correctly, as we will see in the numerical simulations.

It is important to note that expression (70) does not imply that $\delta\phi^y$ (which is a signed value) actually depends on y as a quadratic function, but we will see later that it is the case for small enough y (Subsection 3.B, see detailed discussion in Paper 2). Moreover, the coefficients of linear and quadratic terms for $\delta\phi^y$ may have different signs, which may lead to zero error for nonzero y (however, this y, if it exists, is unfortunately different for each measured quantity).

E. Shape Errors
In this subsection we extend the error analysis as presented in Subsection 2.D to shapes that cannot be described exactly by a set of cubical subvolumes.

We perform the discretization the same way as in Subsection 2.B, but some of the V_i are not cubical (for $i \in \partial V$, which denotes that dipole i lies on the boundary of the volume V). We set $\mathbf{r}_i$ to be still in the center of the cube (circumscribing V_i), not to break the regularity of the lattice. The standard PP prescription uses equal volumes ($V_i = d^3$) in Eqs. (10), (14), (25), and (26); i.e., the discretization changes the shape of the particle a little bit. We will estimate the errors introduced by these boundary dipoles. These errors should then be added to those obtained in Subsection 2.D. We start by estimating $\|\mathbf{h}^d\|_1$. First, we consider $\mathbf{h}_i^d$ for $i \notin \partial V$:

$$\mathbf{h}_i^d = \sum_{j \in \partial V} \left(\int_{V_j} \mathrm{d}^3 r' \overline{\mathbf{G}}(\mathbf{r}_i, \mathbf{r}') \mathbf{P}(\mathbf{r}') - d^3 \overline{\mathbf{G}}_{ij}^{(0)} \mathbf{P}_j \right), \quad (71)$$

which is just a reduction of Eq. (34). For $i \in \partial V$, $\mathbf{h}_i^d$ is the same plus the error in the self-term:

$$\mathbf{h}_i^d = \sum_{\substack{j \in \partial V \\ j \neq i}} \left(\int_{V_j} \mathrm{d}^3 r' \overline{\mathbf{G}}(\mathbf{r}_i, \mathbf{r}') \mathbf{P}(\mathbf{r}') - d^3 \overline{\mathbf{G}}_{ij}^{(0)} \mathbf{P}_j \right) + \mathbf{M}(V_i, \mathbf{r}_i)$$

$$- \left(\overline{\mathbf{L}}(\partial V_i, \mathbf{r}_i) - \frac{4\pi}{3} \overline{\mathbf{I}} \right) \chi_i \mathbf{E}_i. \quad (72)$$

Let us define

$$\mathbf{h}_{ij}^{sh} = \int_{V_j} \mathrm{d}^3 r' \overline{\mathbf{G}}(\mathbf{r}_i, \mathbf{r}') \mathbf{P}(\mathbf{r}') - d^3 \overline{\mathbf{G}}_{ij}^{(0)} \mathbf{P}_j, \quad (73)$$

$$\mathbf{h}_{ii}^{sh} = \mathbf{M}(V_i, \mathbf{r}_i) - \left(\overline{\mathbf{L}}(\partial V_i, \mathbf{r}_i) - \frac{4\pi}{3} \overline{\mathbf{I}} \right) \chi_i \mathbf{E}_i. \quad (74)$$

We estimate each of the terms in Eq. (73) separately (since there is actually no significant cancellation and the error is of the same order of magnitude as the values themselves) using expressions (30) for $\mathbf{P}(\mathbf{r})$ and $\overline{\mathbf{G}}(\mathbf{r})$ and expressions (31). This leads to

$$|\mathbf{h}_{ij}^{sh}| \leq \begin{cases} c_{62} d^3 R_{ij}^{-3}, & R_{ij} < 2 \\ c_{63} d^3, & R_{ij} > 1 \end{cases}. \quad (75)$$

To estimate $\mathbf{h}_{ii}^{sh}$, we assume that the surface of the scatterer is a plane on the scale of the size of the dipole. A finite radius of curvature only changes the constants in the following expressions. We will prove that

$$|\mathbf{h}_{ii}^{sh}| \leq c_{64}; \quad (76)$$

therefore we do not need to consider the third term in Eq. (74) (coming from the unity tensor) at all, since it is bounded by a constant:

$$\mathbf{M}(V_i, \mathbf{r}_i) = \int_{V_i} \mathrm{d}^3 r' (\overline{\mathbf{G}}(\mathbf{r}_i, \mathbf{r}') - \overline{\mathbf{G}}^s(\mathbf{r}_i, \mathbf{r}')) \mathbf{P}(\mathbf{r}')$$

$$+ \int_{V_i} \mathrm{d}^3 r' \overline{\mathbf{G}}^s(\mathbf{r}_i, \mathbf{r}') (\mathbf{P}(\mathbf{r}') - \mathbf{P}(\mathbf{r}_i)) \quad (77)$$

The function in the first integral is always bounded by $c_{65} |\mathbf{r}' - \mathbf{r}_i|^{-2}$. If $\mathbf{r}_i \in V_i$, the same is true for the second integral, and hence

$$|\mathbf{M}(V_i, \mathbf{r}_i)| \leq c_{66} d. \quad (78)$$

If $\mathbf{r}_i \notin V_i$, we introduce an auxiliary point $\mathbf{r}''$ that is symmetric to $\mathbf{r}_i$, over the particle surface and apply the identity

$$\mathbf{P}(\mathbf{r}') - \mathbf{P}(\mathbf{r}_i) = (\mathbf{P}(\mathbf{r}') - \mathbf{P}(\mathbf{r}'')) + (\mathbf{P}(\mathbf{r}'') - \mathbf{P}(\mathbf{r}_i)) \quad (79)$$

to the second integral in Eq. (77). Using a Taylor expansion of $\mathbf{P}$ near $\mathbf{r}''$ and the fact that $|\mathbf{r}' - \mathbf{r}''| \leq |\mathbf{r}' - \mathbf{r}_i|$ for $\mathbf{r}' \in V_i$, one can show that

$$|\mathbf{M}(V_i, \mathbf{r}_i)| \leq c_{67} d + c_{68} \left| \int_{V_i} \mathrm{d}^3 r' \overline{G}^s(\mathbf{r}_i, \mathbf{r}') \right|, \quad (80)$$

where the remaining integral can be proven to be equal to $-\overline{\mathbf{L}}(\partial V_i, \mathbf{r}_i)$. The last proof left [see expressions (74) and (80)] is to demonstrate that $\overline{\mathbf{L}}(\partial V_i, \mathbf{r}_i)$ is bounded by a constant. The only potential problem may come from the subsurface of ∂V_i that is part of the particle surface (because it may be close to $\mathbf{r}_i$). This subsurface is assumed planar. We will calculate the integral in Eq. (6) over the infinite plane $\mathbf{r}' - \mathbf{r}_i = \boldsymbol{\rho} + \mathbf{r}$ such that $\boldsymbol{\rho} \cdot \mathbf{r} = 0$. Then $\mathbf{n}' = \pm \boldsymbol{\rho}/\rho$ and

$$\overline{\mathbf{L}}(\text{infinite plane}, \mathbf{r}_i) = \mp \int_{\mathbf{R}^2} \mathrm{d}^2 r \frac{\hat{\rho}\hat{\rho}}{\rho(\rho^2 + r^2)^{3/2}} = \mp 2\pi \frac{\hat{\rho}\hat{\rho}}{\rho^2}, \quad (81)$$

which is bounded. The rest of the integral (over the part of the cube surface) is bounded by a constant, which is a manifestation of a more general fact that (by its definition) $\overline{\mathbf{L}}(\partial V_i, \mathbf{r}_i)$ does not depend on the size but only on the shape of the volume. Finally, we have

$$|\overline{\mathbf{L}}(\partial V_i, \mathbf{r}_i)| \leq c_{69}, \quad (82)$$

which, together with expressions (74), (78), and (80), proves expression (76).

Using expressions (75) and (76), we obtain

$$\|\mathbf{h}^d\|_1 \leq \sum_i \sum_{j \in \partial V} |\mathbf{h}_{ij}^{sh}| + \sum_{j \in \partial V} |\mathbf{h}_{ii}^{sh}| \leq \sum_{j \in \partial V} \left(\sum_{l=1}^{K_{\max}} c_{62} n_s(l) l^{-3} \right.$$

$$\left. + c_{70} \right) \leq N d (c_{71} - c_{72} \ln d), \quad (83)$$

where we have changed the order of the summation in the double sum and split the summation over cubical shells for $l \leq K_{\max}$ and $l > K_{\max}$. Then we have grouped everything into one sum over boundary dipoles. Expressions (41) and (62) were used in the last inequality. Combining expressions (63) and (83), one can obtain the total estimate of $\|\mathbf{h}^d\|_1$ for any scatterer:

$$\|\mathbf{h}^d\|_1 \leq N[(c_{43} - c_{45} \ln d) d^2 + (c_{73} - c_{72} \ln d) d]. \quad (84)$$

Using expression (66), we immediately obtain the same estimate for $\|\delta \mathbf{E}^d\|_1$.

The derivation of the errors in the measured quantities is slightly modified, compared with Subsection 2.D, by the presence of the shape errors. Expressions (67) and (68) are changed to

$$|\tilde{\phi}(\tilde{\mathbf{E}}) - \phi^d(\mathbf{E}^{0,d})| \leq \sum_i c_{51} d^5 + \sum_{i \in \partial V} c_{74} d^3 \leq c_{52} d^2 + c_{75} d,$$

$$(85)$$

$$|\phi^d(\mathbf{E}^{0,d}) - \phi^d(\mathbf{E}^d)| \leq c_{53} d^3 \|\delta \mathbf{E}^d\|_1 \leq (c_{54} - c_{55} \ln d) d^2$$

$$+ (c_{76} - c_{77} \ln d) d. \quad (86)$$

The second term in expression (85) comes from surface dipoles for which errors are the same order as the values themselves. Finally, the generalization of expression (70) is

$$|\delta\phi^y| \leq (c_{59} - c_{60} \ln y)y^2 + (c_{78} - c_{79} \ln y)y. \quad (87)$$

The shape errors reinforce the surface errors (the linear term of discretization error), and, although both of them generally decrease with increasing size parameter x, one may expect the linear term in expression (87) to be significant up to higher values of y than in expression (70).

All the derivations in this subsection can, in principle, be extended to interfaces inside the particle, i.e., when a surface, which cannot be described exactly as a surface of a set of cubes, separates two regions where $\chi(\mathbf{r})$ varies smoothly. Two parts of the cubical dipole on the interface should be considered separately the same way as was done above. This will, however, not change the main conclusion of this subsection—expression (87)—but only the constants.

F. Different Discrete Dipole Approximation Formulations

In this subsection we discuss how different DDA formulations modify the error estimates derived in Subsections 2.D and 2.E.

Most of the improvements of PP proposed in the literature are concerned with the self-term $\mathbf{M}(V_i, \mathbf{r}_i)$. They are the radiative reaction correction[8] (RR), the digitized Green's function,[23] the formulation by Lakhtakia,[26,27] the a_1-term method,[28,29] the lattice dispersion relation[30] (LDR), the formulation by Peltoniemi[31] (PEL), and the corrected LDR.[32] All of them provide an expression for $\mathbf{M}(V_i, \mathbf{r}_i)$ that is of order d^2 (except for RR, which is of order d^3). For instance, LDR is equivalent to

$$\mathbf{M}(V_i, \mathbf{r}_i) = [(b_1 + b_2 m^2 + b_3 m^2 S)d^2 + (2/3)\mathrm{i}d^3]\mathbf{P}_i \quad (88)$$

(remember that we assumed $k=1$), where b_1, b_2, b_3 are numerical constants and S is a constant that depends only on the propagation and polarization vectors of the incident field. However, none of these formulations can exactly evaluate the integral in expression (39) because the variation of the electric field is not known beforehand (PEL solves this problem but only for a spherical dipole). Therefore they (it is hoped) decrease the constant in expression (40), thus decreasing the overall error in the measured quantities. However, these formulations are not expected to change the order of the error from d^2 to some higher order.

We do not analyze the improvements by Rahmani *et al.*[33,34] and the surface-corrected LDR,[17] as they are limited to certain particle shapes.

There exist two improvements of the interaction term in PP: filtered coupled dipoles[12] (FCD) and integration of Green's tensor[35] (IT). A rigorous analysis of FCD errors is beyond the scope of this paper, but it seems that FCD is not designed to reduce the linear term in expression (63) that comes from the incomplete (nonsymmetric) shells. This is because FCD employs sampling theory to improve the accuracy of the overall discretization for regular cubical grids. FCD does not improve the accuracy of a single $\bar{\mathbf{G}}_{ij}$ calculation (approximation of an integral over one subvolume).

IT, which numerically evaluates the integral in Eq. (12), has a more pronounced effect on the error estimate. Consider dipole j from the lth shell (incomplete) of dipole i, then

$$\int_{V_j} \mathrm{d}^3 r' \bar{\mathbf{G}}(\mathbf{r}_i, \mathbf{r}') \mathbf{P}(\mathbf{r}') - d^3 \bar{\mathbf{G}}_{ij} \mathbf{P}_j$$

$$= \int_{V_j} \mathrm{d}^3 r' \bar{\mathbf{G}}(\mathbf{r}_i, \mathbf{r}')(\mathbf{P}(\mathbf{r}') - \mathbf{P}_j)$$

$$\leq c_{80} \int_{V_j} \mathrm{d}^3 r' r'^2 \max_{\mu, r' \in V_j} |\partial \mu \bar{\mathbf{G}}(\mathbf{r}_i, \mathbf{r}')|$$

$$+ c_{81} \int_{V_j} \mathrm{d}^3 r' |\bar{\mathbf{G}}(\mathbf{r}_i, \mathbf{r}')| r'^2 \leq c_{82} d l^{-4}. \quad (89)$$

Here we have used Eq. (36) and a Taylor expansion of Green's tensor up to the first order. Expression (89) states that the second term in expression (58) is completely eliminated and so is the linear term in expressions (69) and (70) (surface errors). Therefore convergence of DDA with IT for cubically shaped scatterers is expected to be purely quadratic (neglecting the logarithm). However, for noncubically shaped scatterers the linear term reappears, owing to the shape errors. Both IT and FCD also modify the self-term; however, the effect is basically the same as for the other formulations.

Several papers aimed to reduce shape errors.[10,11,36] The first one—the generalized semi-analytical method[10]—modifies the whole DDA scheme, and the other two propose averaging of the susceptibility over the boundary dipoles. We will analyze here weighted discretization (WD) by Piller,[11] which is probably the most advanced method to reduce shape errors available today.

WD modifies the susceptibility and self-term of the boundary subvolume. We slightly modify the definition of the boundary subvolume used in Subsections 2.B and 2.E to automatically take into account interfaces inside the scatterer. We define V_i to be always cubical but with a possible interface inside. The particle surface, crossing the subvolume V_i, is assumed planar and divides the subvolume into two parts: the principal volume V_i^p (containing the center) and the secondary volume V_i^s with susceptibilities $\chi_i^p \equiv \chi_i$, χ_i^s and electric fields $\mathbf{E}_i^p \equiv \mathbf{E}_i$, $\mathbf{E}_i^s$, respectively. The electric fields are considered constant inside each part and related to each other via the boundary-condition tensor $\bar{\mathbf{T}}_i$:

$$\mathbf{E}_i^s = \bar{\mathbf{T}}_i \mathbf{E}_i. \quad (90)$$

In WD the susceptibility of the boundary subvolume is replaced by an effective one, defined as

$$\bar{\chi}_i^e = (V_i^p \chi_i^p \bar{\mathbf{I}} + V_i^s \chi_i^s \bar{\mathbf{T}}_i)/d^3, \quad (91)$$

which gives the correct total polarization of the cubical dipole. The effective self-term is directly evaluated starting from Eq. (4), considering χ and $\mathbf{E}$ constant inside each part,

$$\bar{\mathbf{M}}(V_i,\mathbf{r}) = \left(\int_{V_i^p} d^3r'(\bar{\mathbf{G}}(\mathbf{r}_i,\mathbf{r}') - \bar{\mathbf{G}}^s(\mathbf{r}_i,\mathbf{r}'))\chi_i^p \right. $$
$$\left. + \int_{V_i^s} d^3r'(\bar{\mathbf{G}}(\mathbf{r},\mathbf{r}') - \bar{\mathbf{G}}^s(\mathbf{r}_i,\mathbf{r}'))\chi_i^s\bar{\mathbf{T}}_i \right)\mathbf{E}_i. \tag{92}$$

Piller evaluated the integrals in Eq. (92) numerically.[11]

To take a smooth variation of the electric field and susceptibility into account, we define $\chi_i^s = \chi(\mathbf{r}'')$ ($\mathbf{r}''$ is defined in Subsection 2.E), and $\bar{\mathbf{T}}_i$ is calculated at the surface between $\mathbf{r}_i$ and $\mathbf{r}''$. $\mathbf{P}_i^p \equiv \mathbf{P}_i$, and $\mathbf{P}_i^s = \chi_i^s\mathbf{E}_i^s = \chi_i^s\bar{\mathbf{T}}_i\mathbf{E}_i$. Then

$$|\mathbf{P}(\mathbf{r}'') - \mathbf{P}_i^s| \leq c_{83}\min_{\mathbf{r}\in V_i^s}|\mathbf{r} - \mathbf{r}_i|, \tag{93}$$

where we have assumed that expressions (30) for $\chi(\mathbf{r})$ and $\mathbf{E}(\mathbf{r})$ are also valid in V_i^s.

We start estimating errors of WD with $\mathbf{h}_{ij}^{sh}$ [cf. Eq. (73)]:

$$\mathbf{h}_{ij}^{sh} = \int_{V_j^p} d^3r'(\bar{\mathbf{G}}(\mathbf{r}_i,\mathbf{r}')\mathbf{P}(\mathbf{r}') - \bar{\mathbf{G}}_{ij}^{(0)}\mathbf{P}_j^p)$$
$$+ \int_{V_j^s} d^3r'(\bar{\mathbf{G}}(\mathbf{r}_i,\mathbf{r}')\mathbf{P}(\mathbf{r}') - \bar{\mathbf{G}}_{ij}^{(0)}\mathbf{P}_j^s). \tag{94}$$

Using Taylor expansions of $\mathbf{P}(\mathbf{r}')$ near $\mathbf{r}_j$ and $\mathbf{r}''$ in V_j^p and V_j^s, correspondingly, and expression (93), one may find that the main contribution comes from the derivative of Green's tensor, leading to [cf. expression (75)]

$$|\mathbf{h}_{ij}^{sh}| \leq \begin{cases} c_{84}d^4R_{ij}^{-4}, & R_{ij} < 2 \\ c_{85}d^4, & R_{ij} > 1 \end{cases}. \tag{95}$$

$\mathbf{h}_{ii}^{sh}$ is the following [cf. Eq. (74)]:

$$\mathbf{h}_{ii}^{sh} = (\mathbf{M}(V_i,\mathbf{r}_i) - \bar{\mathbf{L}}(\partial V_i,\mathbf{r}_i)\mathbf{P}_i^p) - \left(\int_{V_i^p} d^3r'(\bar{\mathbf{G}}(\mathbf{r}_i,\mathbf{r}') \right.$$
$$- \bar{\mathbf{G}}^s(\mathbf{r}_i,\mathbf{r}'))\mathbf{P}_i^p + \int_{V_i^s} d^3r'(\bar{\mathbf{G}}(\mathbf{r},\mathbf{r}') - \bar{\mathbf{G}}^s(\mathbf{r}_i,\mathbf{r}'))\mathbf{P}_i^s$$
$$\left. - \bar{\mathbf{L}}(\partial V_i,\mathbf{r}_i)\bar{\chi}_i^e\mathbf{E}_i \right)$$
$$= \int_{V_i^p} d^3r'\,\mathbf{G}(\mathbf{r}_i,\mathbf{r}')(\mathbf{P}(\mathbf{r}') - \mathbf{P}_i^p)$$
$$+ \int_{V_i^s} d^3r'\,\bar{\mathbf{G}}(\mathbf{r}_i,\mathbf{r}')(\mathbf{P}(\mathbf{r}') - \mathbf{P}_i^s)$$
$$+ \int_{V_i^s} d^3r'\,\bar{\mathbf{G}}^s(\mathbf{r}_i,\mathbf{r}')(\mathbf{P}_i^s - \mathbf{P}_i^p)$$
$$+ \bar{\mathbf{L}}(\partial V_i,\mathbf{r}_i)\bar{\chi}_i^e\mathbf{E}_i - \bar{\mathbf{L}}(\partial V_i,\mathbf{r}_i)\mathbf{P}_i^p. \tag{96}$$

The first two integrals can be easily shown to be $\leq c_{86}d$ [cf. Eq. (77)], and the third one is transformed to $\bar{\mathbf{L}}$ the same way as in expression (80), thus

$$|\mathbf{h}_{ii}^{sh}| \leq c_{86}d + |\bar{\mathbf{L}}(\partial V_i^p,\mathbf{r}_i)\mathbf{P}_i^p + \bar{\mathbf{L}}(\partial V_i^s,\mathbf{r}_i)\mathbf{P}_i^s - \bar{\mathbf{L}}(\partial V_i,\mathbf{r}_i)\bar{\chi}_i^e\mathbf{E}_i|, \tag{97}$$

where the second term comes from the fact that averaged $\bar{\mathbf{L}}\mathbf{P}$ is not the same as $\bar{\mathbf{L}}$ times averaged $\mathbf{P}$. This error depends on the geometry of the interface inside V_i and generally is of order unity. For example, if the plane interface is described as $z = z_i + \epsilon$, taking the limit $\epsilon \to 0$ gives the error $|2\pi(\mathbf{P}_i^p - \mathbf{P}_i^s)_z|$ [using Eq. (81)]. Therefore WD does not principally improve the error estimate of $\mathbf{h}_{ii}^{sh}$, given by expression (76), although it may significantly decrease the constant. On the other hand, since $\bar{\mathbf{L}}(\partial V_i^p,\mathbf{r}_i)$ and $\bar{\mathbf{L}}(\partial V_i^s,\mathbf{r}_i)$ can be (analytically) evaluated for a cube intersected by a plane, WD can be further improved to reduce the error in $\mathbf{h}_{ii}^{sh}$ to linear in d, which is a subject of future research.

Proceeding analogously to the derivation of expression (83), one can obtain

$$\|\mathbf{h}^d\|_1 \leq \sum_{j\in\partial V}\left(\sum_{l=1}^{K_{\max}} c_{84}n_s(l)l^{-4} + c_{87} + c_{88}d \right) \leq c_{89}Nd. \tag{98}$$

It can be shown that for the scattering amplitude [Eq. (25)] the error estimate given by expression (85) can be improved, since WD correctly evaluates the zeroth order of value for the boundary dipoles, leading to

$$|\tilde{\phi}(\tilde{\mathbf{E}}) - \phi^d(\mathbf{E}^{0,d})| \leq \sum_i c_{51}d^5 + \sum_{i\in\partial V} c_{90}d^4 \leq c_{91}d^2. \tag{99}$$

In his original paper[11] Piller did not specify the expression that should be used for C_{abs}. Direct application of the susceptibility provided by WD into Eq. (26) does not reduce the order of error when compared with the exact Eq. (24) (except when $\chi_i^s = 0$), since they are not linear functions of the electric field. However, if we consider separately V_i^p and V_i^s [which is equivalent to replacing $V_i\,\mathrm{Im}((\bar{\chi}_i^e\mathbf{E}_i)\cdot\mathbf{E}_i^*)$ by $V_i^p\,\mathrm{Im}(\chi_i^p)|\mathbf{E}_i|^2 + V_i^s\,\mathrm{Im}(\chi_i^s)|\bar{\mathbf{T}}_i\mathbf{E}_i|^2$], the same estimate as in expression (99) can be derived for C_{abs}.

Using expressions (98) and (99), and the first part of expression (86), one can derive the final error estimate for WD:

$$|\delta\phi^y| \leq (c_{92} - c_{93}\ln y)y^2 + c_{94}y, \tag{100}$$

where the constant before the linear term, as compared with expression (87), does not contain a logarithm and is expected to be significantly smaller because several factors contributing to it are eliminated in WD. Although WD has a potential for improving, it does not seem feasible to completely eliminate the linear term in the shape error. The accuracy of evaluation of the interaction term over the boundary dipole [cf. Eq. (94)] can be improved by IT over V_i^p and V_i^s separately but that would ruin the block-Toeplitz structure of the interaction matrix and hinder the fast Fourier transform–based algorithm for the solution of linear equations.[5] Since there is no comparable alternative to fast Fourier transform nowadays, this method seems inapplicable.

Minor modifications of the expression for C_{abs} are possible. Draine[8] proposed a modification of Eq. (26) that was widely used afterward and that was further modified by Chaumet *et al.*[35] However, for many cases these expressions are equivalent, and, even when they are not, the difference is of order d^3, which is neglected in our error analysis.

3. NUMERICAL SIMULATIONS

A. Discrete Dipole Approximation

The basics of the DDA method were summarized by Draine and Flatau.[2] In this paper we use the LDR prescription for dipole polarizability,[30] which is most widely used nowadays, e.g., in the publicly available code DDSCAT 6.1.[4] We also employ dipole size correction[8] for noncubically shaped scatterers to ensure that the cubical approximation of the scatterer has the correct volume; this is believed to diminish shape errors, especially for small scatterers.[2] We use a standard discretization scheme as described in Subsection 2.E, without any improvements for boundary dipoles. It is important to note that all the conclusions are valid for any DDA implementation but with a few changes for specific improvements, as discussed in Subsection 2.F.

Our code—AMSTERDAM DDA[37]—is capable of running on a cluster of computers (parallelizing a single DDA computation), which allows us to use a practically unlimited number of dipoles, since we are not limited by the memory of a single computer.[38,39] We used a relative error of residual $<10^{-8}$ as a stopping criterion for the iterative solution of the DDA linear system. Tests suggest that the relative error of the measured quantities due to the iterative solver is then $<10^{-7}$ (data not shown) and hence can be neglected (total relative errors in our simulations are $>10^{-6} \div 10^{-5}$—see Subsection 3.B). More details about our code can be found in Paper 2. All DDA simulations were carried out on the Dutch national computer cluster LISA.[40]

B. Results

We study five test cases: one cube with $kD=8$; three spheres with $kD=3, 10, 30$; and a particle obtained by a cubical discretization of the $kD=10$ sphere using 16 dipoles per D (total 2176 dipoles, x equal to that of a sphere; see detailed description in Paper 2). By D we denote the diameter of a sphere or the edge size of a cube. All scatterers are homogenous with $m=1.5$. Although DDA errors significantly depend on m (see, e.g., Ref. 14), we limit ourselves to one single value and study effects of size and shape of the scatterer.

The maximum number of dipoles per D (n_D) was 256. The values of n_D that we used are of the form $\{4,5,6,7\}$ (2^p) (p is an integer), except for the discretized sphere, where all n_D are multiples of 16 (this is required to exactly describe the shape of the particle composed from a number of cubes). The minimum values for n_D were 8 for the $kD=3$ sphere; 16 for the cube, the $kD=10$ sphere, and the discretized sphere; and 40 for the $kD=30$ sphere.

All the computations use a direction of incidence parallel to one of the principal axes of the cubical dipoles. The scattering plane is parallel to one of the faces of the cubi-

cal dipoles. In this paper we show results only for the extinction efficiency Q_{ext} (for incident light polarized parallel to one of the principal axes of the cubical dipoles) and phase function $S_{11}(\theta)$ as the most commonly used in applications. However, the theory applies to any measured quantity. For instance, we have also confirmed it for other Mueller matrix, elements (data not shown).

Exact results of $S_{11}(\theta)$ for all five test cases are shown in Fig. 2. For spheres this is the result of Mie theory (the relative accuracy of the code we used[24] is at least $<10^{-6}$) and for the cube and discretized sphere an extrapolation over the five finest discretizations (the extrapolation technique is presented in Paper 2, together with all details of obtaining these results, including their estimated errors). We use such "exact" results because analytical theory is unavailable for these shapes and because errors of the best discretization are larger than that of the extrapolation. Their use as references for computing real errors (difference between the computed and the exact values) of single DDA calculations is justified because all these real errors are significantly larger than the errors of the references themselves (see Paper 2; in general, real errors obtained this way have an uncertainty of reference error). Exact values of Q_{ext} for all test cases are presented in Table 1.

In the following we show the results of DDA convergence. Figures 3 and 4 present relative errors (absolute values) of S_{11} at different angles θ and maximum error over all θ versus y in a log–log scale. In many cases the maximum errors are reached at an exact backscattering direction, then these two sets of points overlap. Deep minima that happen at intermediate values of y for some
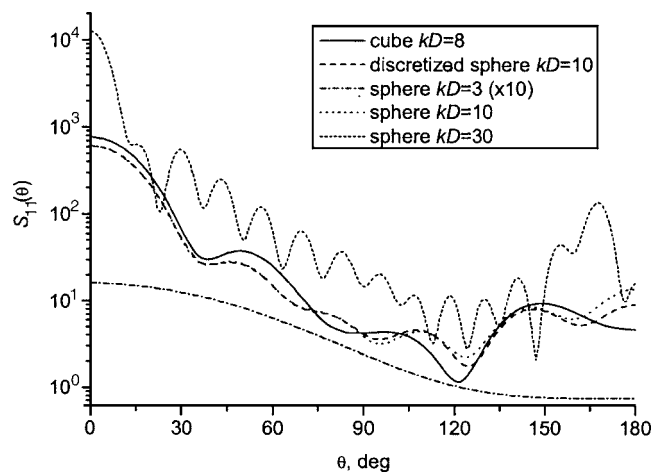

Fig. 2. $S_{11}(\theta)$ for all five test cases in logarithmic scale. The result for the $kD=3$ sphere is multiplied by 10 for convenience.

Table 1. Exact Values of Q_{ext} for the Five Test Cases

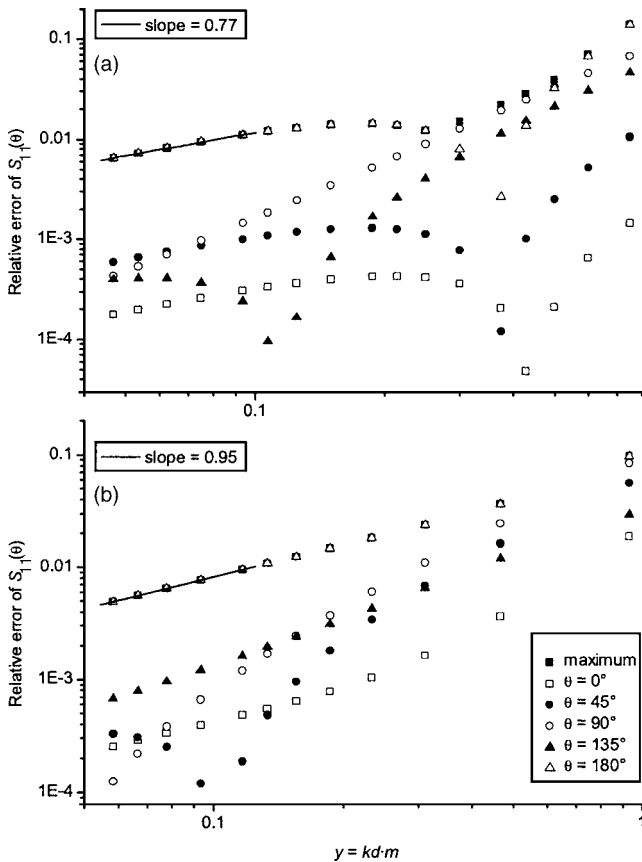| Particle	Q_{ext}
$kD=8$ cube	4.490
Discretized $kD=10$ sphere	3.916
$kD=3$ sphere	0.753
$kD=10$ sphere	3.928
$kD=30$ sphere	1.985

Fig. 3.   Relative errors of S_{11} at different angles θ and maximum over all θ versus y for (a) the $kD=8$ cube, (b) the cubical discretization of the $kD=10$ sphere. A log–log scale is used. A linear fit of maximum over θ errors is shown ($m=1.5$).

values of θ (and also sometimes for Q_{ext}—Fig. 5) are due to the fact that the differences between simulated and reference values change sign near these values of y (see Paper 2 for a detailed description of this behavior). The solid lines are linear fits to all or some points of maximum error. The slopes of these lines are depicted in the figures. Figure 5 shows relative errors of Q_{ext} for all five studied cases in a log–log scale. A linear fit through the five finest discretizations of the $kD=3$ sphere is shown. More results of these numerical simulations are presented in Paper 2.

4. DISCUSSION

Convergence of DDA for cubically shaped particles (Fig. 3) shows the following trends. All curves have linear and quadratic parts (the nonmonotonic behavior of errors for some θ is also a manifestation of the fact that the signed difference can be approximated by a sum of linear and quadratic terms that have different signs). The transition between these two regimes occurs at different y (which indicates the relative importance of linear and quadratic coefficients). Although for maximum errors that are close to those of the backscattering direction the linear term is significant for larger y, it is much smaller and not significant in the whole range of y studied for side scattering ($\theta=90°$). Results of DDA convergence for spheres (Fig. 4) show a different behavior for different sizes. Errors for the small ($kD=3$) sphere converge purely linear [except

for a small deviation of errors of $S_{11}(90°)$ for large values of y]. Similar results are obtained for the $kD=10$ sphere but with significant oscillations superimposed on the general trend. Convergence for the large ($kD=30$) sphere is quadratic or even faster in the range of y studied, also with significant oscillations.

Comparing Figs. 3 and 4 [especially Figs. 3(b) and 4(b) showing results for almost the same particles], one can deduce the following differences in DDA convergence for cubically and noncubically shaped scatterers. The linear term for cubically shaped scatterers is significantly smaller, resulting in smaller total errors, especially for small y. All these conclusions, together with the size dependence of the significance of the linear term in the total
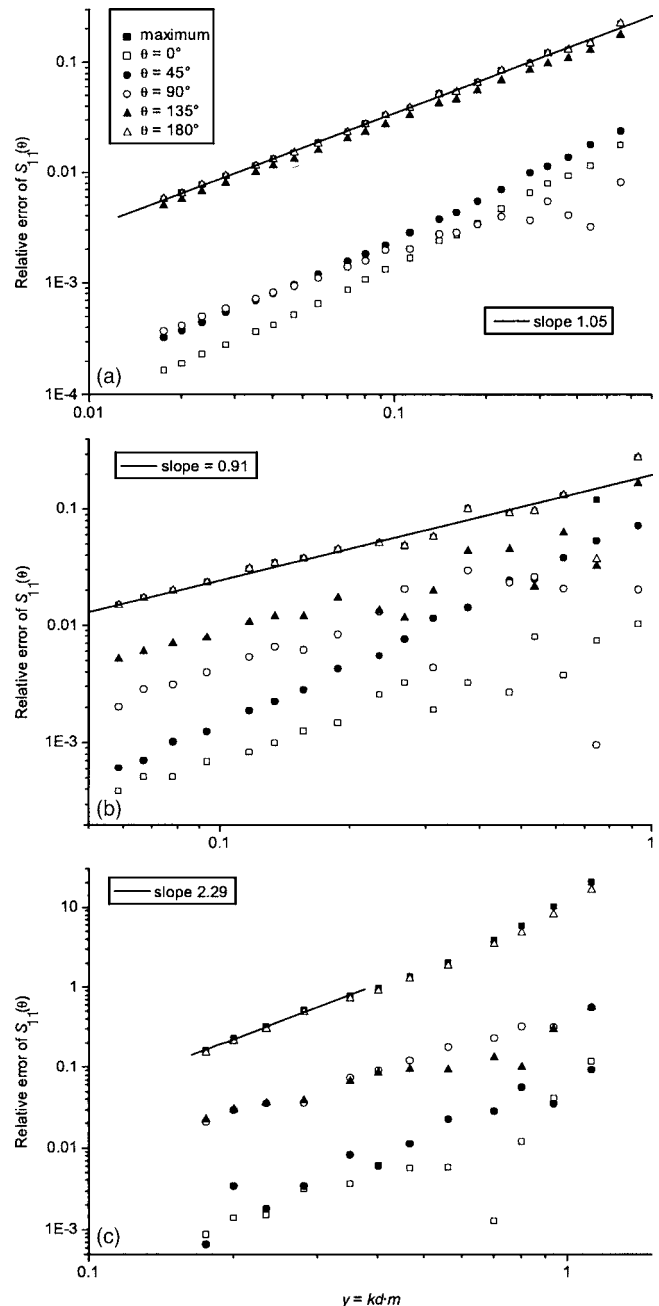

Fig. 4.   Same as Fig. 3 but for (a) $kD=3$, (b) $kD=10$, and (c) $kD=30$ spheres.
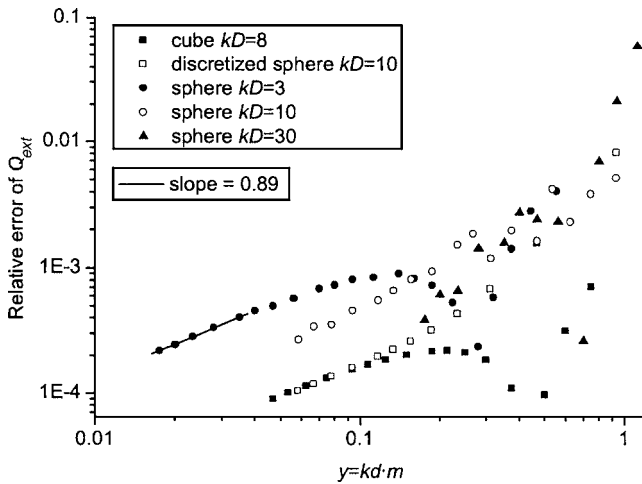
Fig. 5.   Relative errors of Q_{ext} versus y for all five test cases. A log–log scale is used. A linear fit through five finest discretizations of the $kD=3$ sphere is shown.

errors, are in perfect agreement with the theoretical predictions made in Subsections 2.D and 2.E. Errors for noncubically shaped particles exhibit quasi-random oscillations that are not present for cubically shaped particles. This can be explained by the sharp variations of shape errors with changing y (discussed in detail in Paper 2). Oscillations for the $kD=3$ sphere [Fig. 4(a)] are very small (but still clearly present), which is due to the small size of the particle and hence featurelessness of its light-scattering pattern—the surface structure is not that important, and one may expect rather small shape errors. Results for Q_{ext} (Fig. 5) fully support the conclusions. Errors of Q_{ext} for the large sphere at small values of y are unexpectedly smaller than for smaller spheres. This feature requires further study before any firm conclusions are made; however, there is definitely no similar tendency for $S_{11}(\theta)$ (cf. Fig. 4).

We have also studied a $kD=8$ porous cube that was obtained by dividing a cube into 27 smaller cubes and then removing randomly nine of them. All the conclusions are the same as those reported for the cube but with slightly higher overall errors (data not shown).

In this paper we have used a traditional DDA formulation[2] for numerical simulations. However, as we showed in Subsection 2.E, several modern improvements of DDA (namely, IT and WB) should significantly change its convergence behavior. IT should completely eliminate the linear term for cubically shaped scatterers, which should improve the accuracy especially for small y. WD should significantly decrease shape and hence total errors for noncubically shaped particles; moreover, it should significantly decrease the amplitude of quasi-random error oscillations because it takes into account the location of the interface inside the boundary dipoles. Numerical testing of DDA convergence using IT and WD is a subject of a future study.

5. CONCLUSION

To the best of our knowledge, we conducted for the first time a rigorous theoretical convergence analysis of DDA. In the range of DDA applicability ($kd<2$), errors are

bounded by a sum of a linear term and a quadratic term in the discretization parameter y; the linear term is significantly smaller for cubically than for noncubically shaped scatterers. Therefore for small y, errors for cubically shaped particles are much smaller than for noncubically shaped ones. The relative importance of the linear term decreases with increasing size; hence convergence of DDA for large enough scatterers is quadratic in the common range of y. All these conclusions were verified by extensive numerical simulations.

Moreover, these simulations showed that errors are not only bounded by a quadratic function (as predicted in Section 2) but actually can be (with good accuracy) described by a quadratic function of y. This fact provides a basis for the extrapolation technique presented in Paper 2.

Our theory predicts that modern DDA improvements (namely, IT and WD) should significantly change the convergence of DDA; however, numerical testing of these predictions is left for future research.

REFERENCES

1. E. M. Purcell and C. R. Pennypacker, "Scattering and adsorption of light by nonspherical dielectric grains," Astrophys. J. **186**, 705–714 (1973).
2. B. T. Draine and P. J. Flatau, "Discrete-dipole approximation for scattering calculations," J. Opt. Soc. Am. A **11**, 1491–1499 (1994).
3. B. T. Draine, "The discrete dipole approximation for light scattering by irregular targets," in *Light Scattering by Nonspherical Particles, Theory, Measurements, and Applications*, M. I. Mishchenko, J. W. Hovenier, and L. D. Travis, eds. (Academic, 2000), pp. 131–145.
4. B. T. Draine and P. J. Flatau, "User guide for the discrete dipole approximation code DDSCAT 6.1," http://xxx.arxiv.org/abs/astro-ph/0409262 (2004).
5. J. J. Goodman, B. T. Draine, and P. J. Flatau, "Application of fast-Fourier-transform techniques to the discrete-dipole approximation," Opt. Lett. **16**, 1198–1200 (1991).
6. G. C. Hsiao and R. E. Kleinman, "Mathematical foundations for error estimation in numerical solutions of integral equations in electromagnetics," IEEE Trans. Antennas Propag. **45**, 316–328 (1997).
7. K. F. Warnick and W. C. Chew, "Error analysis of the moment method," IEEE Antennas Propag. Mag. **46**, 38–53 (2004).
8. B. T. Draine, "The discrete-dipole approximation and its application to interstellar graphite grains," Astrophys. J. **333**, 848–872 (1988).
9. J. I. Hage, J. M. Greenberg, and R. T. Wang, "Scattering from arbitrarily shaped particles: theory and experiment," Appl. Opt. **30**, 1141–1152 (1991).
10. F. Rouleau and P. G. Martin, "A new method to calculate the extinction properties of irregularly shaped particles," Astrophys. J. **414**, 803–814 (1993).

11. N. B. Piller, "Influence of the edge meshes on the accuracy of the coupled-dipole approximation," Opt. Lett. **22**, 1674–1676 (1997).

12. N. B. Piller and O. J. F. Martin, "Increasing the performance of the coupled-dipole approximation: a spectral approach," IEEE Trans. Antennas Propag. **46**, 1126–1137 (1998).

13. N. B. Piller, "Coupled-dipole approximation for high permittivity materials," Opt. Commun. **160**, 10–14 (1999).

14. A. G. Hoekstra, J. Rahola, and P. M. A. Sloot, "Accuracy of internal fields in volume integral equation simulations of light scattering," Appl. Opt. **37**, 8482–8497 (1998).

15. S. D. Druger and B. V. Bronk, "Internal and scattered electric fields in the discrete dipole approximation," J. Opt. Soc. Am. B **16**, 2239–2246 (1999).

16. Y. L. Xu and B. A. S. Gustafson, "Comparison between multisphere light-scattering calculations: rigorous solution and discrete-dipole approximation," Astrophys. J. **513**, 894–909 (1999).

17. M. J. Collinge and B. T. Draine, "Discrete-dipole approximation with polarizabilities that account for both finite wavelength and target geometry," J. Opt. Soc. Am. A **21**, 2023–2028 (2004).

18. M. A. Yurkin, V. P. Maltsev, and A. G. Hoekstra, "Convergence of the discrete dipole approximation. II. An extrapolation technique to increase the accuracy," J. Opt. Soc. Am. A **23**, 2592–2601 (2006).

19. A. Lakhtakia, "Strong and weak forms of the method of moments and the coupled dipole method for scattering of time-harmonic electromagnetic-fields," Int. J. Mod. Phys. C **3**, 583–603 (1992).

20. F. M. Kahnert, "Numerical methods in electromagnetic scattering theory," J. Quant. Spectrosc. Radiat. Transf. **79**, 775–824 (2003).

21. C. P. Davis and K. F. Warnick, "On the physical interpretation of the Sobolev norm in error estimation," Appl. Comput. Electromagn. Soc. J. **20**, 144–150 (2005).

22. A. D. Yanghjian, "Electric dyadic Green's function in the source region," Proc. IEEE **68**, 248–263 (1980).

23. G. H. Goedecke and S. G. O'Brien, "Scattering by irregular inhomogeneous particles via the digitized Green's function algorithm," Appl. Opt. **27**, 2431–2438 (1988).

24. C. F. Bohren and D. R. Huffman, *Absorption and Scattering of Light by Small Particles* (Wiley, 1983).

25. J. Rahola, "On the eigenvalues of the volume integral operator of electromagnetic scattering," SIAM (Soc. Ind. Appl. Math) J. Sci. Comput. (USA) **21**, 1740–1754 (2000).

26. A. Lakhtakia and G. W. Mulholland, "On two numerical techniques for light scattering by dielectric agglomerated structures," J. Res. Natl. Inst. Stand. Technol. **98**, 699–716 (1993).

27. J. I. Hage and J. M. Greenberg, "A model for the optical properties of porous grains," Astrophys. J. **361**, 251–259 (1990).

28. C. E. Dungey and C. F. Bohren, "Light scattering by nonspherical particles: a refinement to the coupled-dipole method," J. Opt. Soc. Am. A **8**, 81–87 (1991).

29. H. Okamoto, "Light scattering by clusters: the A1-term method," Opt. Rev. **2**, 407–412 (1995).

30. B. T. Draine and J. J. Goodman, "Beyond Clausius–Miossotti—wave propagation on a polarizable point lattice and the discrete dipole approximation," Astrophys. J. **405**, 685–697 (1993).

31. J. I. Peltoniemi, "Variational volume integral equation method for electromagnetic scattering by irregular grains," J. Quant. Spectrosc. Radiat. Transf. **55**, 637–647 (1996).

32. D. Gutkowicz-Krusin and B. T. Draine, "Propagation of electromagnetic waves on a rectangular lattice of polarizable points," http://xxx.arxiv.org/abs/astro-ph/0403082 (2004).

33. A. Rahmani, P. C. Chaumet, and G. W. Bryant, "Coupled dipole method with an exact long-wavelength limit and improved accuracy at finite frequencies," Opt. Lett. **27**, 2118–2120 (2002).

34. A. Rahmani, P. C. Chaumet, and G. W. Bryant, "On the importance of local-field corrections for polarizable particles on a finite lattice: application to the discrete dipole approximation," Astrophys. J. **607**, 873–878 (2004).

35. P. C. Chaumet, A. Sentenac, and A. Rahmani, "Coupled dipole method for scatterers with large permittivity," Phys. Rev. E **70**, 036606 (2004).

36. K. F. Evans and G. L. Stephens, "Microwave radiative transfer through clouds composed of realistically shaped ice crystals. Part 1. Single scattering properties," J. Atmos. Sci. **52**, 2041–2057 (1995).

37. "Amsterdam DDA," http://www.science.uva.nl/research/scs/Software/adda.

38. A. G. Hoekstra, M. D. Grimminck, and P. M. A. Sloot, "Large scale simulations of elastic light scattering by a fast discrete dipole approximation," Int. J. Mod. Phys. C **9**, 87–102 (1998).

39. M. A. Yurkin, K. A. Semyanov, P. A. Tarasov, A. V. Chemyshev, A. G. Hoekstra, and V. P. Maltsev, "Experimental and theoretical study of light scattering by individual mature red blood cells by use of scanning flow cytometry and discrete dipole approximation," Appl. Opt. **44**, 5249–5256 (2005).

40. "Description of the national computer cluster Lisa," http://www.sara.nl/userinfo/lisa/description/mdex.html(2005).