A Directional Multilevel Complex-Space Fast Multipole Algorithm for Electrically Large Problems of Various Dimensions

Guang-Yu Zhu, Wei-Dong Li, *Member, IEEE*, Wei E. I. Sha, *Senior Member, IEEE*, Hou-Xing Zhou, and Wei Hong, *Fellow, IEEE*

***Abstract*— A directional multilevel complex-space fast multipole algorithm (DMLCSFMA) is proposed for solving electrically large problems of various dimensions. This algorithm implements a high-frequency generalization of the well-known mid-frequency multilevel fast multipole algorithm (MLFMA). It is established by exploring the fundamental connection between the conventional MLFMA and the recently developed directional fast multipole algorithms (D-FMAs), as well as the plane wave expansion induced from the complex source beam [Gaussian beam (GB)]. Different from the conventional MLFMA which exhibits the complexity of $O(N^2)$ for certain situations such as the quasi-1-D elongated object, the proposed high-frequency generalized version is capable of achieving a stable complexity of $O(N \log N)$, irrespective of the dimensional features of the objects. Besides, the proposed algorithm also manifests itself as a spectral counterpart of the traditional D-FMAs. However, unlike the traditional D-FMAs which leverage the equivalent source-based sampling expansions, the proposed algorithm is established using the plane wave-based exponential expansions. Thus, the feasibility of building a D-FMA with analytically diagonalized translators is also demonstrated in this work. Several numerical examples are provided to illustrate the complexity and accuracy of the proposed algorithm.**

***Index Terms*— Electrically large problems, fast multipole algorithm (FMA), Fraunhofer far-field condition, Gaussian beam (GB), ray physics, wave physics.**

Manuscript received 16 November 2022; revised 22 October 2023; accepted 18 December 2023. Date of publication 24 January 2024; date of current version 9 February 2024. This work was supported in part by the Fundamental Research Funds for the Central Universities under Grant 2242022k30003; and in part by the National Science Foundation of China under Grant 62293492, Grant 62131008, Grant 61975177, and Grant U20A20164. *(Corresponding author: Guang-Yu Zhu.)*

Guang-Yu Zhu is with the State Key Laboratory of Millimeter Waves, Southeast University, Nanjing 210096, China, and also with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China (e-mail: gyzhu_china@163.com).

Wei-Dong Li and Wei Hong are with the State Key Laboratory of Millimeter Waves and the Frontiers Science Center for Mobile Information Communication and Security, Southeast University, Nanjing 210096, China (e-mail: wdli@seu.edu.cn; weihong@seu.edu.cn).

Wei E. I. Sha is with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China (e-mail: weisha@zju.edu.cn).

Hou-Xing Zhou is with the State Key Laboratory of Millimeter Waves, Southeast University, Nanjing 210096, China (e-mail: zhouhx1301@163.com).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TAP.2024.3354816.

Digital Object Identifier 10.1109/TAP.2024.3354816

I. Introduction

THE fast multipole algorithm (FMA) is among the "Top Ten Algorithms of the Century" due to its great influence on modern science and engineering [1]. It was originally designed for the rapid simulation of the N-body problems [2], [3] with the smooth kernels [4], [5], [6], and used for the calculation of the gravitational and electrostatic interactions [5], [6]. Over the past decades, many new versions of the FMA have developed, and the effective range of applicability of this algorithm has been extended significantly [14], [15], [16], [17].

To accelerate the solution of the electrodynamic problems with the oscillatory kernel [4], an FMA with diagonalized translators is developed [7], [8], and it reduces the computational complexity of an integral formulation iterative solver [18] from $O(N^2)$ to $O(N^{3/2})$, where N denotes the number of unknowns. With the interpolation and anterpolation technique [9], a multilevel (ML) extension of the diagonal FMA [7], [8], widely known as the MLFMA, is later developed [10], [11], and it can further reduce the complexity to $O(N \log N)$ for many electrodynamic problems. These FMAs are also known as the mid-frequency FMAs [21], and have been widely used to solve large-scale radiation and scattering problems characterized by the wave physics [15].

To alleviate the subwavelength breakdown of the MLFMA, low-frequency FMAs are subsequently developed [10], [20]. Based on the same box-tree framework, the low-frequency FMAs have been hybridized with the mid-frequency FMA [10] in a unified manner, yielding broadband FMAs [20], [21]. These broadband schemes can effectively capture both the wave physics and the circuit physics [4], [16]. They further enrich the basic theory of the FMA and greatly extend the scope of applications of the FMA. In particular, they provide efficient and stable solutions for many real-world engineering electromagnetic (EM) problems that show a multiscale nature [16].

Although the FMAs capturing low-frequency and mid-frequency physics have been firmly established and widely adopted, it is not the case for the FMAs exploiting high-frequency physics [4]. As we know, ray-inspired FMAs [35], [36], [37], [38], [39], such as the ray-propagation FMA (RPFMA) [35] and fast far-field approximation (FAFFA) [37], reduce the complexity of the nonnested diagonal FMA [7], [8] from $O(N^{3/2})$ to $O(N^{4/3})$. Therefore, a natural idea

is to integrate the essentials of these ray-inspired algorithms [35], [36], [37], [38], [39] into the ML framework of the MLFMA [10] so as to gain further improvement on the complexity of the MLFMA [10]. Some attempts along this path include FAFFA-MLFMA [24], RPFMA-MLFMA [25], and recent developments based on the complex source beams [Gaussian beams (GBs)] [42], [43], [44], [45], [46]. In contrast to the literature on the nonnested implementations [35], [36], [37], [38], [39], the papers [24], [25], [46] on employing directional windowed translators [35], [37], [45] within the ML framework of the MLFMA [10] only reported the achievable optimization on the constant factors in front of the asymptotic complexity factor, while the improvement in the order of the asymptotic complexity was however not reported. Besides, although achieving some efficiency optimizations compared to the mid-frequency MLFMA, in practice the ray-enhanced counterparts [24], [25] of the MLFMA are rarely adopted in large-scale computing. This is unlike the conventional mid-frequency MLFMA [10] whose parallelization strategies have been actively studied [22], [23].

Recently, more variants of the FMA are developed [27], [28], [29], [30], [31], [32], [33], [34]. Specifically, for the static cases with smooth kernels, several kernel-independent FMAs (KI-FMA) [27], [28], [29] are developed. Some of these FMAs leverage the spatial sampling expansions [27], [28], while some are based on the spectral Fourier expansions [29]. Furthermore, for the dynamic cases with the oscillatory kernels, several directional FMAs (D-FMA) [30], [31], [32], [33], [34] are developed. These D-FMAs are established based on a kind of directional admissible far-field condition [30]. With such far-field conditions, several expansions [26], [27], [28] that were originally utilized to build fast algorithms for the static cases are shown to be equally useful when constructing fast algorithms for the dynamic cases. Nevertheless, among the various kernel expansions, only the spatial domain expansions [30], [31], [32], [33], [34] such as the Chebyshev interpolations [33] are exploited to establish the D-FMAs. To the best of our knowledge, there has been no report on building a D-FMA with the spectral domain expansions such as the plane wave expansions [10], [43], [45]. Moreover, although the D-FMAs [30], [31], [32], [33], [34] and the plane wave-based mid-frequency FMAs [7], [10] are both designed for the dynamic cases, the connection between them has never been explored yet.

Motivated by the above investigations, a directional ML complex-space FMA (DMLCSFMA) is proposed in this work. To be more clear, the contributions and significations of this study are elaborated from the two complementary perspectives as follows.

1) On the one hand, the proposed algorithm is a high-frequency generalized version of the conventional mid-frequency MLFMA [10]. It demonstrates a systematic way to further incorporate ray physics into the framework of the MLFMA [10]. The conventional MLFMA [10] has the limitation that its computational complexity becomes $O(N^2)$ for certain situations such as the elongated quasi-1-D object [33]. In contrast, the proposed high-frequency generalized version of the MLFMA is capable of achieving stable $O(N \log N)$ complexity, irrespective of the dimensional features of the objects. Thus, the proposed algorithm achieves further improvement on the order of the asymptotic complexity, compared to the MLFMA [10], and correspondingly extends the effective range of applicability of the MLFMA [10].
2) On the other hand, the proposed algorithm also manifests itself as a spectral counterpart of the traditional spatial D-FMAs [30], [31], [32], [33], [34]. Previously, algorithms of similar type were constructed based on the random sampling [30], Chebyshev interpolation [33], and adaptive cross approximation [34], respectively. They all employ spatial sampling expansions based on the concept of equivalent sources. Instead, plane waves-based exponential expansions [43], [45] are adopted herein to implement the overall algorithm. Thus, the feasibility of establishing a spectral D-FMA is demonstrated. Furthermore, different from the traditional spatial D-FMAs [30], [31], [32], [33], [34] which depend heavily on the matrix compression techniques and are thus relatively algebra-biased, the proposed spectral counterpart is established based on the analytic expansion of the Green's functions [43], [45]. Hence, clearer and deeper physical insights can also be gained.

Preliminary results of this study were reported in our conference paper [49]. Herein, we significantly extend our research in both theoretical and experimental aspects. Throughout the article, we use λ to denote the wavelength in free space. Unless otherwise stated, we use electrical size to characterize the size of the object.

II. Basic Formulations

In this section, the complex source beam-induced plane wave expansion is first introduced. Then, several unique properties of this expansion are discussed.

A. Plane Wave Expansion Induced From Complex Source Beam

Based on the two well-known identities, i.e., the addition theorem for the dynamic Green's function and the integral expression of the spherical Bessel function [8], [53], the following expansion can be conventionally obtained:

$$\frac{e^{ik|\boldsymbol{r}+\boldsymbol{d}|}}{|\boldsymbol{r}+\boldsymbol{d}|} = \frac{ik}{4\pi}\sum_{v=0}^{\infty} i^v (2v+1) \mathrm{h}_v^{(1)}(kr) \int_{\Omega} d^2\hat{\boldsymbol{k}} e^{ik\hat{\boldsymbol{k}}\cdot\boldsymbol{d}} \mathrm{P}_v\left(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}}\right) \tag{1}$$

where $\boldsymbol{k} = k\hat{\boldsymbol{k}}$, the integral in the above is defined on the unit sphere Ω, $\mathrm{h}_v^{(1)}(z)$ is the first kind of Hankel function of vth order, $\mathrm{P}_v(z)$ is the Legendre polynomial of vth order, $\mathbf{r}$ and $\boldsymbol{d}$ are real vectors with their lengths described by d and r, respectively. The region of validity for the above expansion (1) is described by $d < r$.

Suppose that we have two well-separated groups with radius R. The centers of the two groups are denoted by $\boldsymbol{r}_m$ and $\boldsymbol{r}_n$,

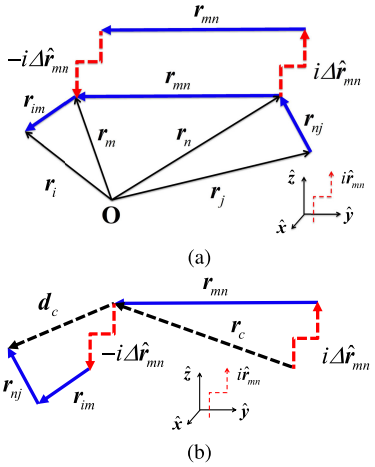


Fig. 1. Geometric settings for the complex source beam (GB) induced plane wave expansion. In the above, the complex-space extension has been introduced. Here, we treat $i\hat{\boldsymbol{r}}_{mn}$ as an extra dimension. Note that, when the beam parameter Δ is set to zero, the above reduces to the conventional settings for the mid-frequency MLFMA. (a) Configuration of the translation vectors. (b) Equivalent representation.

respectively. Then, considering a field point $\boldsymbol{r}_i$ within group m and a source point $\boldsymbol{r}_j$ within group n, and referring to Fig. 1(a), we have the following geometric relations [45]:

$$\begin{aligned}\boldsymbol{r}_{ij} &= \boldsymbol{r}_i - \boldsymbol{r}_j = \boldsymbol{r}_i - \boldsymbol{r}_m + \boldsymbol{r}_m - \boldsymbol{r}_n + \boldsymbol{r}_n - \boldsymbol{r}_j \\ &= \left[\boldsymbol{r}_{im} + \left(i\Delta\hat{\boldsymbol{r}}_{mn} + \boldsymbol{r}_m\right)\right] - \left[\left(i\Delta\hat{\boldsymbol{r}}_{mn} + \boldsymbol{r}_n\right) - \boldsymbol{r}_{nj}\right] \\ &= \boldsymbol{r}_{im} + \boldsymbol{r}_{nj} - i\Delta\hat{\boldsymbol{r}}_{mn} + \boldsymbol{r}_{mn} + i\Delta\hat{\boldsymbol{r}}_{mn} \\ &= \left(\boldsymbol{r}_{im} + \boldsymbol{r}_{nj} - i\Delta\hat{\boldsymbol{r}}_{mn}\right) + (r_{mn} + i\Delta)\hat{\boldsymbol{r}}_{mn} \end{aligned} \tag{2}$$

where by "adding one term $i\Delta\hat{\boldsymbol{r}}_{mn}$, and subtracting the same term $i\Delta\hat{\boldsymbol{r}}_{mn}$," complex coordinate extension is introduced. Here, Δ is referred to as the beam parameter. Accordingly, denoting

$$\boldsymbol{d}_c = \boldsymbol{r}_{im} + \boldsymbol{r}_{nj} - i\Delta\hat{\boldsymbol{r}}_{mn} \tag{3}$$

$$\boldsymbol{r}_c = r_c\hat{\boldsymbol{r}}_{mn} = (r_{mn} + i\Delta)\hat{\boldsymbol{r}}_{mn} \tag{4}$$

as depicted in Fig. 1(b), and substituting into (1), we have

$$\begin{aligned}\frac{e^{ik|\boldsymbol{r}_{ij}|}}{|\boldsymbol{r}_{ij}|} = \frac{ik}{4\pi}\sum_{v=0}^{\infty} i^v(2v+1)\mathrm{h}_v^{(1)}(k\{r_{mn}+i\Delta\}) \\ \int_{\Omega} d^2\hat{\boldsymbol{k}} e^{i\boldsymbol{k}\cdot\left(\boldsymbol{r}_{im}+\boldsymbol{r}_{nj}-i\Delta\hat{\boldsymbol{r}}_{mn}\right)}\mathrm{P}_v\left(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}}_{mn}\right)\end{aligned} \tag{5}$$

which holds for $\Delta \geq 0$ and $\boldsymbol{r}_{im} + \boldsymbol{r}_{nj} \in \Phi$, where Φ denotes the region of validity for expansion (5). Here, Φ is described by the interior of a body of revolution and depends on both $\boldsymbol{r}_{mn}$ and Δ [43], [45].

For convenience, the left-hand side of (5) is denoted as $\mathrm{G}(r_{ij})$ where $r_{ij} = |\boldsymbol{r}_{ij}|$. Truncating the infinite series in the right-hand side of (5), followed by exchanging the order of summation and integration, a plane wave expansion is then obtained as follows:

$$\mathrm{G}\left(r_{ij}\right) \approx \int_{\Omega} d^2\hat{\boldsymbol{k}}\beta\left(\hat{\boldsymbol{k}}, \boldsymbol{r}_{im}, \boldsymbol{r}_{nj}\right)\alpha_{mn}\left(\hat{\boldsymbol{k}}, \boldsymbol{r}_{mn}, \Delta\right) \tag{6}$$

where

$$\beta\left(\hat{\boldsymbol{k}}, \boldsymbol{r}_{im}, \boldsymbol{r}_{nj}\right) = e^{ik\hat{\boldsymbol{k}}\cdot\left(\boldsymbol{r}_{im}+\boldsymbol{r}_{nj}\right)} \tag{7}$$

is the radiation pattern, and

$$\begin{aligned}\alpha_{mn}\left(\hat{\boldsymbol{k}}, \boldsymbol{r}_{mn}, \Delta\right) = \frac{ik}{4\pi}e^{k\Delta\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}}_{mn}} \\ \sum_{v=0}^{V} i^v(2v+1)\mathrm{h}_v^{(1)}(k\{r_{mn}+i\Delta\})\mathrm{P}_v\left(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}}_{mn}\right)\end{aligned} \tag{8}$$

is referred to as the GB translator, where V denotes the truncation length. Such a translator reduces to the conventional translator adopted in the MLFMA [10] when $\Delta = 0$. In practice, to maintain the numerical stability, (8) is further recast into

$$\begin{aligned}\alpha_{mn}\left(\hat{\boldsymbol{k}}, \boldsymbol{r}_{mn}, \Delta\right) = \frac{ik}{4\pi}e^{k\Delta\left(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}}_{mn}-1\right)} \\ \sum_{v=0}^{V} i^v(2v+1)\tilde{\mathrm{h}}_v^{(1)}(k\{r_{mn}+i\Delta\})\mathrm{P}_v\left(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}}_{mn}\right)\end{aligned} \tag{9}$$

where $\tilde{\mathrm{h}}_v^{(1)}(Z) = \mathrm{h}_v^{(1)}(Z)e^{\mathrm{Im}(Z)}$ is the normalized spherical Hankel function [43].

B. Far-Field Setting and Beam Parameter

To facilitate subsequent discussions, here we consider a specific geometric far-field setting between two well-separated groups, together with a particular rule for choosing the beam parameter Δ. Motivation for these assumptions will be discussed later in Section II-E.

Concretely, consider the specific far-field setting where the central distance r_{mn} between the field group m and the source group n is set by

$$r_{mn} = \mathrm{f}_{\mathrm{QFF}}(R) := C_{\mathrm{QFF}} \cdot R^2 = O\left(R^2\right) \tag{10}$$

where C_{QFF} is a prespecified constant, and the r_{mn} considered here is thus a quadratic function of the group radius R. Then, for such a specific geometric setting, consider the particular rule for choosing Δ as follows:

$$\Delta = \sigma \cdot \mathrm{f}_{\mathrm{QFF}}(R) = \sigma \cdot C_{\mathrm{QFF}} \cdot R^2 = O\left(R^2\right) \tag{11}$$

where σ is a prespecified positive constant, and here Δ is configured to be linearly proportional to $\mathrm{f}_{\mathrm{QFF}}(R)$, thus a quadratic function of R as well.

In the following, two unique properties of the complex source beam induced plane wave expansion (6) will be further discussed, for the specified settings (10) and (11) mentioned above.

C. Truncation Length

When $\Delta = 0$, the conventional closed-form excess bandwidth formula [15] provides a good estimate of the required truncation length V for a prescribed tolerance ε. However, with the complex-space extension (i.e., $\Delta > 0$), such a closed-form formula is no longer applicable [42]. Hence, when $\Delta > 0$, the required truncation length V is in general determined by directly finding the smallest integer V that satisfies

$$\left|\mathrm{G}\left(r_{ij}\right) - \sum_{v=0}^{V}\mathrm{U}_v\left(\Delta, r_{mn}, \hat{\boldsymbol{r}}_{mn}, \boldsymbol{r}_{im}, \boldsymbol{r}_{nj}\right)\right| \leq \left|\mathrm{G}\left(r_{ij}\right)\right| \cdot \varepsilon \tag{12}$$

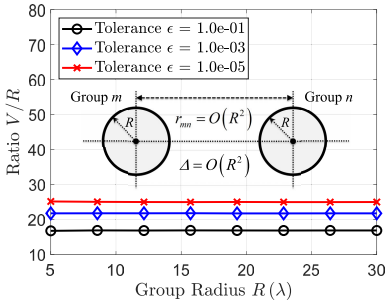


Fig. 2. Changing trends of the ratio V/R for increasing R. Here, the far-field setting r_{mn} is as described in (10), with Δ chosen by the particular rule (11). The truncation length V is directly calculated from (12) for a given truncation tolerances ε. Clearly, as R increases, $V/R = O(1)$, implying that $V = O(R)$. Besides, the trends for different truncation tolerances ε are illustrated.

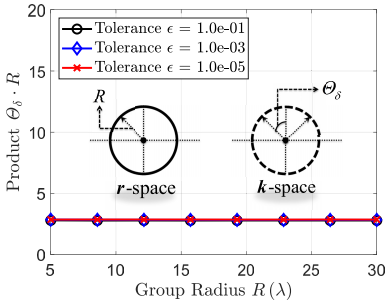


Fig. 3. Changing trends of the product $\Theta_\delta \cdot R$ for increasing R. Here, Δ is chosen by the particular rule (11). The cut-off range Θ_δ is numerically calculated from (15), with a cut-off threshold $\delta = 10^{-15}$. Clearly, as R increases, $\Theta_\delta \cdot R = O(1)$, thus implying $\Theta_\delta = O(1/R)$. Besides, the trends for different truncation tolerances ε are illustrated.

for all possible locations of $\boldsymbol{r}_{im}$ and $\boldsymbol{r}_{nj}$, where

$$\mathrm{U}_v\big(\Delta, r_{mn}, \hat{\boldsymbol{r}}_{mn}, \boldsymbol{r}_{im}, \boldsymbol{r}_{nj}\big) = ik(-1)^v(2v+1)\,\mathrm{h}_v^{(1)}(kr_c)\mathrm{j}_v(kd_c)\mathrm{P}_v\big(\hat{\boldsymbol{d}}_c \cdot \hat{\boldsymbol{r}}_{mn}\big) \quad (13)$$

with the length $d_c = (\boldsymbol{d}_c \cdot \boldsymbol{d}_c)^{1/2}$ being a complex length [43]. The square root is defined to have a nonnegative real part and the branch cut along the negative real axis.

Nevertheless, for two interacting groups characterized by the specific far-field setting (10), with Δ chosen by the particular rule (11), it can be shown that the required truncation length V still follows the same leading order behavior as the conventional excess bandwidth formula [15], namely

$$V = O(R). \quad (14)$$

That is, the required truncation length V is still linearly proportional to the group radius R, even if the in-group vector $\boldsymbol{r}_{im} + \boldsymbol{r}_{nj}$ has been augmented with $i\Delta\hat{\boldsymbol{r}}_{mn}$, where $\Delta = O(R^2)$. Some theoretical explanations on the above property (14) are given in Appendix A. Here, some numerical results are calculated, for verifying (14). For clarity, the ratios between V and R are illustrated in Fig. 2. As expected, such ratios are shown to be almost constant order for increasing R, thus implying (14) above.

D. Range of Window

As a result of the complex-space extension, the GB translator $\alpha_{mn}(\Theta)$ exhibits an obvious windowed feature in the angular spectrum domain, where $\Theta \in [0, \pi]$ is the angle between $\hat{\boldsymbol{k}}$ and $\hat{\boldsymbol{r}}_{mn}$. For two interacting groups, only those plane waves $\hat{\boldsymbol{k}}$ near the direction of $\hat{\boldsymbol{r}}_{mn}$ have prominent contributions and needs to be retained for achieving accurate far-field translations. In general, the corresponding prominent range Θ_δ can be determined by finding the smallest Θ subject to [43]

$$|\alpha_{mn}(\Theta)| \le \max_{\Theta' \in [0,\pi]} |\alpha_{mn}(\Theta')| \cdot \delta \quad (15)$$

where δ is a prescribed threshold for cutting off the window.

In particular, for two interacting groups characterized by (10), it can be shown that, with the Δ chosen by (11), the prominent range Θ_δ asymptotically follows:

$$\Theta_\delta = O(1/R). \quad (16)$$

Some analysis to reveal the above property is given in Appendix B. To verify the above property, the changing trends of the product $\Theta_\delta \cdot R$ for increasing R are here numerically investigated, as shown in Fig. 3. As expected, $\Theta_\delta \cdot R = O(1)$, thus implying (16) above.

E. Further Discussion

The conditions (10) and (11) are so far introduced as arbitrary assumptions. To gain deeper insight into the nature of the problem, a motivation for the assumptions (10) and (11) are further discussed as follows. The main feature of the "Gaussian translator" (8) is the exponential factor

$$\exp\big(k\Delta\hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{r}}_{mn}\big) \approx \exp(k\Delta)\exp\big(-k\Delta\Theta^2/2\big) \quad (17)$$

i.e., an approximate Gaussian dependence on the angle Θ between $\hat{\boldsymbol{k}}$ and $\hat{\boldsymbol{r}}_{mn}$. The resulting GB distribution in the plane-wave directions $\hat{\boldsymbol{k}}$ has the width

$$\Theta_{\mathrm{GB}} \sim \frac{1}{\sqrt{k\Delta}} = \sqrt{\frac{\lambda}{2\pi\Delta}}. \quad (18)$$

If one wants to interpret this beam as emitted by an aperture, the aperture diameter, say $\bar{R}$, has to be $\bar{R} \sim (2\pi\lambda\Delta)^{1/2}$. Conversely, a source (cluster) of size $\bar{R}$ can emit a beam of "diffraction-limited" width

$$\Theta_d \sim \frac{\lambda}{\pi\bar{R}} \quad (19)$$

characterized by the far-field (diffractive or Rayleigh) range

$$\bar{R}_d \sim \frac{\bar{R}}{\Theta_d} \sim \frac{\pi\bar{R}^2}{\lambda}. \quad (20)$$

One can then identify the angles and far-field ranges, $\Theta_{\mathrm{GB}} = \Theta_d$ and $r_{mn} = \bar{R}_d$, provided the conditions (10) and (11) hold, with

$$C_{\mathrm{QFF}} \sim \frac{1}{\lambda} \quad \text{and} \quad \sigma \sim \frac{1}{2\pi^2}. \quad (21)$$

The assumptions (10) and (11) are thus equivalent to the "directional parabolic separation condition" of [30] or the "cluster separation condition (9)" in [33].

III. ML Algorithm

In this section, the essentials of the conventional mid-frequency MLFMA are briefly reviewed first. Then, the basic principle and technical key points of the proposed high-frequency generalized version are introduced.

A. Conventional Mid-Frequency MLFMA

The MLFMA [10] is a box-tree-based ML algorithm, and diagonal translations are performed between well-separated (nonneighboring) boxes at each involved level. Usually, the one-box buffer scheme [13] is adopted, and the nearest distance between well-separated boxes is then one box-width. In this case, the central distance r_{mn} between the well-separated boxes satisfies the following far-field condition:

$$r_{mn} \geq \mathrm{f}_{\mathrm{LFF}}(R) := C_{\mathrm{LFF}} \cdot R \tag{22}$$

where $\mathrm{f}_{\mathrm{LFF}}(R)$ is linear as a function of R and C_{LFF} is a constant. With the above linear condition (22), the far-field boxes exist at the levels $l = 2, \ldots, L$, where $L \sim \log W$ and W denotes the edge length of the root box at level $l = 0$.

The plane wave expansion adopted by the MLFMA is a special case of the expansion (6), i.e., expansion (6) with $\Delta \equiv 0$. Hence, with the conventional quadrature rules [15] (i.e., Gaussian quadrature rule along θ and trapezoidal quadrature rule along ϕ), the discretized expansion for the MLFMA can be formulated as

$$\mathrm{G}(r_{ij}) \approx \sum_{\hat{\boldsymbol{k}}_q \mapsto \Omega} w_q \, \beta(\hat{\boldsymbol{k}}_q) \alpha_{mn}(\hat{\boldsymbol{k}}_q, \boldsymbol{r}_{mn}, \Delta \equiv 0) \tag{23}$$

where $\{\Omega : \theta \in [0, \pi], \phi \in [0, 2\pi]\}$ denotes the whole sphere. In the above, $\hat{\boldsymbol{k}}_q \mapsto \Omega$ denotes that the required quadrature nodes $\hat{\boldsymbol{k}}_q$ cover the whole sphere Ω. In principle, the number of the required nodes is determined by the truncation length V and is typical of $O(V^2)$ on the sphere [8]. Since $\hat{\boldsymbol{k}}_q$ of all directions are involved, the far-field translation procedures based on (23) can be referred to as the full translations. By performing such full translations for all the levels involving well-separated boxes (i.e., levels $l = 2, \ldots, L$), all the far-field interactions are thus taken into account [15].

Another important element of the MLFMA is the interpolation of the outgoing waves and anterpolation of the incoming waves, which can be formulated as

$$\mathcal{B}\left(\hat{\boldsymbol{k}}_q^{(l-1)}\right) = \sum_{\hat{\boldsymbol{k}}_q^{(l)}} \mathrm{h}\left(\hat{\boldsymbol{k}}_q^{(l-1)}, \hat{\boldsymbol{k}}_q^{(l)}\right) \mathcal{B}\left(\hat{\boldsymbol{k}}_q^{(l)}\right) \tag{24}$$

where $\hat{\boldsymbol{k}}_q^{(l)}$ and $\hat{\boldsymbol{k}}_q^{(l-1)}$ denote the quadrature nodes at level l and parent level $l-1$, $\mathrm{h}(\cdot, \cdot)$ denotes the interpolation kernel, and $\mathcal{B}(\cdot)$ denotes the outgoing waves in the aggregation process or the incoming waves in the disaggregation process.

Moreover, the computational cost for performing the far-field translations at level l can be described concretely by

$$T^{(l)} = N_{\mathrm{LBox}}^{(l)} \times M_{\mathrm{TList}}^{(l)} \times K_{\mathrm{TPair}}^{(l)} \tag{25}$$

where $N_{\mathrm{LBox}}^{(l)}$ denotes the total number of boxes at level l, $M_{\mathrm{TList}}^{(l)}$ denotes the maximum length of the interaction list for boxes at level l, $K_{\mathrm{TPair}}^{(l)}$ denotes the translation content between well-separated boxes at level l. It is well known that, if a 3-D electrodynamic problem modeled by the surface integral equation (SIE) is solved using the MLFMA, then $T^{(l)}$ is characterized by the following relation [15]:

$$M_{\mathrm{TList}}^{(l)} \sim O(1), \quad N_{\mathrm{LBox}}^{(l)} \times K_{\mathrm{TPair}}^{(l)} \sim O(N) \tag{26}$$

where $l = 2, \ldots, L$. Clearly, $T^{(l)} \sim O(N)$ for varying l. In this case, an $O(N \log N)$ complexity is obtained when all $l = 2, \ldots, L$ levels are taken into account, where $L = O(\log N)$. However, one fundamental limitation of the MLFMA is that such $O(N \log N)$ complexity is not always achievable for objects of different dimensional features, and the MLFMA exhibits $O(N^2)$ complexity for certain situations such as the extremely elongated object [33].

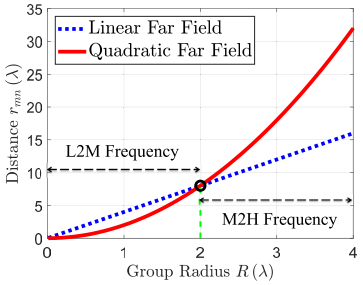


Fig. 4. Function curves for the linear far-field (LFF) $\mathrm{f}_{\mathrm{LFF}}(R) := C_{\mathrm{LFF}} \cdot R$ and the quadratic far-field (QFF) $\mathrm{f}_{\mathrm{QFF}}(R) := C_{\mathrm{QFF}} \cdot R^2$. A crossover point can be seen from the figure. Physically, the range below the crossover point can be viewed as the low-to-mid (L2M) frequency regime where the circuit and wave physics dominate, while the range above the crossover point can be viewed as the mid-to-high (M2H) frequency regime where the wave and ray physics dominate.

B. High-Frequency Generalization of the MLFMA

In the following, we present a systematic way to generalize the above mid-frequency diagonal MLFMA, by full exploitation of the high-frequency ray physics. To achieve this goal, referring to (10), the following QFF condition is further considered:

$$r_{mn} \geq \mathrm{f}_{\mathrm{QFF}}(R) := C_{\mathrm{QFF}} \cdot R^2. \tag{27}$$

With the above quadratic condition (27), the far-field boxes exist only at the levels $l = L, \ldots, \Gamma$, where $\Gamma \sim \log(W)^{1/2} \approx L/2$. This means the number of the levels involving far-field boxes is asymptotically half of that under the LFF condition (22).

To facilitate subsequent discussions, the function curves for both the linear function $\mathrm{f}_{\mathrm{LFF}}(R)$ and the quadratic function $\mathrm{f}_{\mathrm{QFF}}(R)$ are plotted in Fig. 4. For these two functions, a crossover point exists at

$$R_{\mathrm{M2H}} = C_{\mathrm{LFF}}/C_{\mathrm{QFF}}. \tag{28}$$

Suppose that the levels whose group radii satisfy $R^{(l)} < R_{\mathrm{M2H}}$ correspond to $l = L, \ldots, L_\Delta + 1$. Accordingly, the rest of the levels whose group radii satisfy $R^{(l)} \geq R_{\mathrm{M2H}}$ correspond to $l = L_\Delta, \ldots, \Gamma, \ldots, 2, \ldots, 0$. Then, in this work, the full translations used by the conventional mid-frequency MLFMA are applied only for the first few lower levels $l = L, \ldots, L_\Delta + 1$ where the LFF conditions (22) are adopted,

and the windowed translations presented in Section II are applied for the rest of the higher levels $l = L_\Delta, \ldots, \Gamma$ where the QFF conditions (27) are adopted. To be more clear, the proposed arrangements can be summarized as follows:

$$\overbrace{\underbrace{L, \ldots, L_\Delta + 1}_{\text{Full Translations}}}^{\text{Linear Far Field}}, \overbrace{\underbrace{L_\Delta, \cdots\cdots\cdots, \Gamma}_{\text{Windowed Translations}}}^{\text{Quadractic Far Field}}, \cdots, 2, \cdots, 0. \quad (29)$$

Thus, for two interacting far-field groups at level $l = L_\Delta, \ldots, \Gamma$, only the quadrature nodes $\hat{\boldsymbol{k}}_q$ within a spherical cap are required for achieving the far-field translations. Consequently, instead of (23), the following expansion is adopted at level $l = L_\Delta, \ldots, \Gamma$, namely

$$\mathrm{G}(r_{ij}) \approx \sum_{\hat{\boldsymbol{k}}_q \mapsto \Psi} w_q \, \beta(\hat{\boldsymbol{k}}_q) \alpha_{mn}\big(\Theta\langle \hat{\boldsymbol{k}}_q, \hat{\boldsymbol{r}}_{mn}\rangle, r_{mn}, \Delta\big) \quad (30)$$

where $\{\Psi : \Theta \in [0, \Theta_\delta]\}$ denotes the spherical cap and the beam parameter Δ is configured with the particular choice (11) introduced in Section II-B. In the above, $\hat{\boldsymbol{k}}_q \mapsto \Psi$ denotes that the required quadrature nodes $\hat{\boldsymbol{k}}_q$ are confined within the spherical cap Ψ. This spherical cap Ψ is revolved around the direction vector $\hat{\boldsymbol{r}}_{mn}$, with the spanning angle determined by the cut-off range Θ_δ.

Furthermore, invoking (14) and (16), the number of the quadrature nodes $\hat{\boldsymbol{k}}_q$ required to accomplish accurate far-field translations can be estimated as follows:

$$K_{mn} = O\left(V^2 \cdot \frac{(\Theta_\delta)^2}{4\pi}\right) = O\left(R^2 \cdot \left(\frac{1}{R}\right)^2\right) = O(1) \quad (31)$$

where V^2 indicates the order of the total number of nodes on the whole sphere, $(\Theta_\delta)^2/4\pi$ indicates the ratio of the spherical cap and the whole sphere. The above says that, to achieve accurate translations under the QFF condition (27), the required quadrature nodes (translation contents) do not grow for increasingly larger groups, and maintain constant order as function of the group radius R. Physically, the above relation (31) is a manifestation of the limited degree of freedom (DoF) [50], [51] of the scattered fields, where the fields are radiated from a source antenna with an aperture of radius $O(R)$ and observed within an area of radius $O(R)$, and the distance between the source and field is subject to the Fraunhofer far-field condition [54] which is $O(R^2)$.

Correspondingly, instead of (26), the computational cost $T^{(l)}$ for varying l are here characterized by [30], [31], [32], [33], [34]

$$K^{(l)}_{\mathrm{TPair}} \sim O(1), \quad N^{(l)}_{\mathrm{LBox}} \times M^{(l)}_{\mathrm{TList}} \sim O(N) \quad (32)$$

where $l = L_\Delta, \ldots, \Gamma$. In accordance with (31), the translation content $K^{(l)}_{\mathrm{TPair}}$ remains constant order for varying l. Note that, although a different kind of workload distributions is adopted at the higher levels $l = L_\Delta, \ldots, \Gamma$, the windowed translations (30) can be unified seamlessly with the full translations (23) performed at the lower levels $l = L, \ldots, L_\Delta + 1$. In fact, the windowed translations approach the full translations at the transition level $l = L_\Delta$, where the distance of the QFF (27) reduces roughly to that of the LFF (22), and meanwhile the area of the spherical cap Ψ enlarges nearly to that of the whole sphere Ω.

C. Quadrature Nodes and Interpolation Samples

With the above high-frequency evolvement, relevant procedures involving the numerical integration and the spherical interpolation need to be reconsidered and modified, compared to the conventional mid-frequency MLFMA.

Referring to (30) and (31), for given a source group n and a far-field group m, only a subset $\hat{\boldsymbol{k}}_q \mapsto \Psi$ of all the spherical samples (discrete plane wave directions) on Ω are necessary to achieve the windowed translations. Suppose that this subset is denoted as $\mathcal{K}_{mn}$. Since a source group generally interacts with more than one far-field group, e.g., M far-field groups (i.e., far-field group m with $m = 1, \ldots, M$), the spherical samples required for the given source group n are thus determined by the union $\bar{\mathcal{K}}_n$ of M subsets, namely, $\bar{\mathcal{K}}_n = \bigcup \mathcal{K}_{mn}$ with $m = 1, \ldots, M$. In many cases (e.g., the 1D surface problems discussed in Section IV), this union subset $\bar{\mathcal{K}}_n$ can be much smaller than the universal set containing all the spherical samples on Ω. Therefore, to avoid unnecessary computations so as to truly achieve improved asymptotic complexity, a compact and sparsified paradigm should be considered, in which only the necessary union subset (e.g., $\bar{\mathcal{K}}_n$) is involved throughout the whole computational process and the complementary subset containing unnecessary spherical samples on Ω is never involved. To achieve this goal, the rotation and interpolation techniques [40], [41] are further incorporated. Correspondingly, the modified procedures are as follows.

1) Coordinate System Rotation: Except for the original coordinates system $\mathrm{O}_\perp$, a rotated coordinates system $\mathrm{O}'_\perp$ [40], [41] is defined for each given translation direction $\boldsymbol{r}_{mn}$. The z'-axis of $\mathrm{O}'_\perp$ is defined to be parallel to $\boldsymbol{r}_{mn}$, i.e., $\hat{z}' \parallel \boldsymbol{r}_{mn}$. For clarity, in the following, the prime superscript is used to denote that the variables are evaluated in the rotated coordinate system $\mathrm{O}'_\perp$. Correspondingly, the Green's function expansion (6) can be recast in $\mathrm{O}'_\perp$ as follows:

$$\mathrm{G}(r_{ij}) \approx \int_{\Omega'} \beta'(\hat{\boldsymbol{k}}') \alpha'_{mn}(\hat{\boldsymbol{k}}' \cdot \hat{z}', r_{mn}, \Delta) d^2\hat{\boldsymbol{k}}' \quad (33)$$

where Ω' denotes the whole unit spherical surface.

2) Quadrature on Spherical Cap: Due to the windowing effect of the GB translator, the required integral domain in (33) can be reduced from Ω' to Ψ', where Ψ' denotes the spherical cap characterized by the cut-off window. Then, numerical integration is evaluated only on such spherical cap Ψ' (i.e., reduced integral domain), rather than on the whole sphere Ω', yielding

$$\mathrm{G}(r_{ij}) \approx \sum_{\hat{\boldsymbol{k}}'_{qc} \mapsto \Psi'} w'_{qc} \, \beta'(\hat{\boldsymbol{k}}'_{qc}) \alpha'_{mn}(\hat{\boldsymbol{k}}'_{qc}) \quad (34)$$

where $\hat{\boldsymbol{k}}'_{qc} \mapsto \Psi'$ denotes that the quadrature nodes $\hat{\boldsymbol{k}}'_{qc}$ lie within the spherical cap Ψ'.

In the above, the range of the cap Ψ' is described precisely by $\theta' \in [0, \Theta_\delta]$ and $\phi' \in [0, 2\pi]$. In this case, the conventional Gaussian and trapezoidal quadrature rules [15] for the MLFMA are still appropriate choices and can also be applied here with respect to θ' and ϕ', respectively. Illustrations of the spherical cap Ψ' and corresponding quadrature nodes are shown in Fig. 5. Furthermore, to integrate the spherical cap

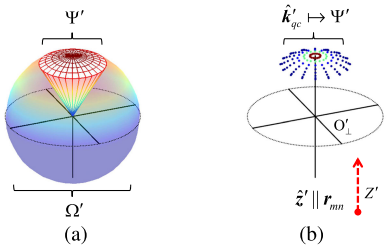


Fig. 5. Illustration of the spherical cap Ψ' on the sphere Ω', and the corresponding quadrature nodes $\hat{k}'_{qc}$ on Ψ'. Here, instead of performing the numerical integration on the whole sphere Ω in the original coordinate $O_\perp$, the proposed algorithm only needs quadrature on the spherical cap Ψ' in the rotated coordinate $O'_\perp$. (a) Spherical cap. (b) Quadrature nodes.

region Ψ' accurately, the required quadrature orders maintain constant order for increasing R, as implied from (31).

3) Interpolation of Radiation Pattern: The quadrature nodes $\hat{k}'_{qc}$ in (34) are defined for a given far-field translation direction (i.e., a given r_{mn}), thus direction-dependent. Therefore, for a different translation direction, a different set of quadrature nodes $\hat{k}'_{qc}$ needs to be defined. However, to establish a FMA, the far-field translations along different translation directions should share a common set of spherical samples. To this end, the interpolation technique [40], [41] is further incorporated, and a discrete plane wave expansion allowing for cycling information is then derived as follows.

Denoting $\hat{k}_{qc}$ in $O_\perp$ as the corresponding spherical sample whose direction is parallel to that of $\hat{k}'_{qc}$ in $O'_\perp$, and invoking the definition of the radiation pattern function β as in (7), it is easy to show that [41]

$$\beta'(\hat{k}'_{qc}) = \beta(\hat{k}_{qc}). \tag{35}$$

Moreover, denoting $\hat{k}_s$ as a set of spherical samples on Ω, and using appropriate interpolation scheme, the radiation pattern function β sampled at $\hat{k}_{qc}$ can be expanded as

$$\beta(\hat{k}_{qc}) = \sum_{\hat{k}_s} \mathrm{h}(\hat{k}_{qc}, \hat{k}_s)\beta(\hat{k}_s) \tag{36}$$

where $\mathrm{h}(\cdot,\cdot)$ denotes the interpolation kernel. Then, with (35) and (36), the expansion (34) can be transformed into a different form by the following manipulations:

$$\begin{aligned}
\mathrm{G}(r_{ij}) &\approx \sum_{\hat{k}'_{qc} \mapsto \Psi'} w'_{qc}\alpha'_{mn}(\hat{k}'_{qc})\beta'(\hat{k}'_{qc}) \\
&= \sum_{\hat{k}'_{qc} \mapsto \Psi'} w'_{qc}\alpha'_{mn}(\hat{k}'_{qc})\beta(\hat{k}_{qc}) \\
&= \sum_{\hat{k}'_{qc} \mapsto \Psi'} w'_{qc}\alpha'_{mn}(\hat{k}'_{qc}) \sum_{\hat{k}_s} \mathrm{h}(\hat{k}_{qc}, \hat{k}_s)\beta(\hat{k}_s) \\
&= \sum_{\hat{k}_s} \beta(\hat{k}_s) \sum_{\hat{k}'_{qc} \mapsto \Psi'} w'_{qc}\alpha'_{mn}(\hat{k}'_{qc})\mathrm{h}(\hat{k}_{qc}, \hat{k}_s) \\
&= \sum_{\hat{k}_s} \beta(\hat{k}_s)\alpha_{mn}(\hat{k}_s)
\end{aligned} \tag{37}$$

where exchanging the order of summations has been performed and

$$\alpha_{mn}(\hat{k}_s) := \sum_{\hat{k}'_{qc} \mapsto \Psi'} w'_{qc}\,\alpha'_{mn}(\hat{k}'_{qc})\mathrm{h}(\hat{k}_{qc}, \hat{k}_s). \tag{38}$$

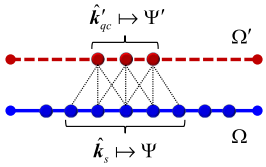


Fig. 6. Mapping relations between the quadrature nodes $\hat{k}'_{qc}$ and the interpolation samples $\hat{k}_s$. With the local interpolation, each $\hat{k}'_{qc}$ is associated with a constant number of $\hat{k}_s$, as depicted by the dot lines. The spherical area covered by the spherical cap Ψ is similar to that covered by Ψ'. Here, Ω' and Ω denote the whole spheres in the rotated and original coordinate systems, respectively.

Different from (34), the expansion (37) now expands the Green's function in terms of the interpolation samples $\hat{k}_s$ defined in $O_\perp$. With such expansion, the far-field translations along different translation directions (i.e., different r_{mn}) can thus share a common set of spherical samples (i.e., $\hat{k}_s$ in $O_\perp$).

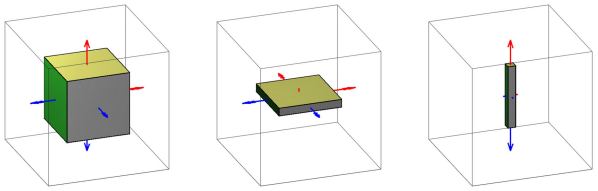
In the above, two types of spherical samples are involved, including the quadrature nodes $\hat{k}'_{qc}$ and the interpolation samples $\hat{k}_s$. As we know, to interpolate the bandlimited radiation pattern function $\beta(\cdot)$ accurately at some $\hat{k}_{qc}$ (corresponding to each $\hat{k}'_{qc}$) in the context of the MLFMA [15], the required local interpolation stencil samples $\hat{k}_s$ around $\hat{k}_{qc}$ do not need to grow for increasing R [12], meaning that each $\hat{k}'_{qc}$ is only associated with constant number of $\hat{k}_s$. Meanwhile, the required $\hat{k}'_{qc} \mapsto \Psi'$ is of constant order for increasing R. Hence, combining these two aspects, it can be known that the interpolation samples $\hat{k}_s$ required for interpolating all the quadrature nodes $\hat{k}'_{qc}$ within Ψ' are still concentrated within a spherical cap region (denoted here by $\hat{k}_s \mapsto \Psi$) and also maintain constant order for increasing R. For clarity, an illustration of the mapping relation between $\hat{k}'_{qc}$ and $\hat{k}_s$ is depicted in Fig. 6. Notably, to achieve such a relation, the distributions of $\hat{k}_s$ on the sphere are not critical, and we are free to choose the spherical interpolation schemes and corresponding interpolation samples. Hence, in addition to the commonly-used Lagrange interpolation [15], the well-established optimal spherical interpolation scheme [52] can also be adopted.

Besides, with the above modifications, the outgoing waves during the aggregation, as well as the incoming waves during the disaggregation, are thus sampled at the spherical samples $\hat{k}_s$. Consequently, instead of (24), the interpolations and anterpolations are correspondingly reformulated as follows:

$$\mathcal{B}\left(\hat{k}_s^{(l-1)}\right) = \sum_{\hat{k}_s^{(l)}} \mathrm{h}\left(\hat{k}_s^{(l-1)}, \hat{k}_s^{(l)}\right)\mathcal{B}\left(\hat{k}_s^{(l)}\right). \tag{39}$$

Here, $\hat{k}_s$ are the interpolation samples employed in (36) and (37). They are a different set of spherical samples relative to the quadrature nodes (i.e., $\hat{k}'_{qc}$). This is unlike (24), in which the outgoing and incoming waves $\mathcal{B}(\cdot)$ are sampled at the quadrature nodes [i.e., $\hat{k}_q$ in (24)], and these quadrature nodes also play the role of the interpolation samples.

4) Remark: Although the modified procedures introduced above are primarily meant to realize the nonredundant workflow of the higher levels $l = L_\Delta, \ldots, \Gamma$ where the windowed translations are employed, such modified procedures



(a) (b) (c)

Fig. 7. Illustration of the volume problems with three different types of dimensional features. These problems are classified here as volume problems because the considered object of certain dimensional features fills the space of corresponding characteristic dimensions in a densely packed manner. This is in contrast to the surface problems (discussed later) where only the boundary areas of the above solid objects are utilized to model the problems and thus the occupied spaces are more sparsely populated. (a) Expanded in 3-D. (b) Expanded in 2-D. (c) Expanded in 1-D.

themselves are however universal and also applicable to the lower levels $l = L, \ldots, L_\Delta + 1$ where the full translations are employed. With this in mind, the modified procedures can thus be adopted for all the $l = L, \ldots, \Gamma$ levels involved. Consequently, unified procedures can be obtained for the overall ML algorithm.

IV. Complexity Analysis

The proposed algorithm inherits the diagonal translation feature of the MLFMA and the workload distributions of the D-FMAs. The asymptotic complexity of the proposed algorithm is therefore similar to that of the D-FMAs [30], [33]. Relevant complexities are thus only briefly analyzed herein, for a better understanding and for the completeness of the article.

The object is discretized conventionally with the quasi-uniform mesh, yielding N unknowns. The width of the root box of the box-tree is here denoted by W. Clearly, in the usual box-tree setting, $R^{(l)} = O(W/2^l)$ where $l = L, \ldots, 0$, and $R^{(L)} = O(W^0) = O(1)$ where $L = O(\log W)$. For notational simplicity, in the following, we omit the superscript and use R to express $R^{(l)}$ without compromising understanding. Then, the total number of nonempty boxes at level l is characterized by $N^{(l)}_{\text{LBox}} = O((W/R)^\tau) = O((2^l)^\tau)$.

A. Complexity Analysis for the Volume Problems

Depending on the dimensional features of the objects, three types of problems are first considered, as shown in Fig. 7. For simplicity, these problems are categorized here as the volume problems, since the object of certain dimensional features fills the space of corresponding characteristic dimensions in a densely packed manner. For such volume problems, we have

$$N = O(W^\tau), \qquad \tau = 1, 2, 3 \tag{40}$$

where τ characterizes the dimensional features of the object.

The complexity of the conventional mid-frequency MLFMA is first outlined for the above volume problems. Referring to (25), the translation cost at level $l = L, \ldots, 2$ is described by

$$T^{(l)} = O\big((W/R)^\tau \cdot 1 \cdot R^2\big) = O\big(W^\tau \cdot R^{2-\tau}\big). \tag{41}$$

Invoking the relation between W and R, the above becomes

$$T^{(l)} = O\Big(W^\tau \cdot \big(W/2^l\big)^{2-\tau}\Big) = O\Big(W^2 \cdot \big(2^{\tau-2}\big)^l\Big). \tag{42}$$

Then, summing up the costs of all the involved levels yields

$$\sum_{l=L}^{2} T^{(l)} = \begin{cases} O\big(W^2 \cdot W\big) = O\big(W^3\big) = O(N), & \tau = 3 \\ O\big(W^2 \cdot \log W\big) = O(N \log N), & \tau = 2 \\ O\big(W^2 \cdot 1\big) = O\big(W^2\big) = O\big(N^2\big), & \tau = 1. \end{cases} \tag{43}$$

Moreover, when the usual local interpolations are adopted, the complexities for the aggregations and disaggregations are the same as those for the translations. The complexities for both the 3-D and 2-D volume problems are therefore $O(N)$ and $O(N \log N)$, respectively. However, the complexity for the 1-D volume problems (i.e., elongated objects) is $O(N^2)$, which is unfavorable in practice.

The complexity of the proposed algorithm is then outlined. For the proposed algorithm, the total cost is dominated by the translation cost at the higher levels $l = L_\Delta, \ldots, \Gamma$. Referring to (25), the translation cost for each of these higher levels is described by

$$T^{(l)} = O((W/R)^\tau \cdot R^\tau \cdot 1) = O(W^\tau) = O(N). \tag{44}$$

Notably, in (44), cancelation happens in terms of R, or more precisely, in terms of R^τ, thus making such computational cost dimensionless. Consequently, for all the three types of volume problems considered (i.e., $\tau = 1, 2, 3$), the total cost summing up all the involved levels $l = L, \ldots, \Gamma$ always follows $O(N \log N)$.

B. Complexity Analysis for the Surface Problems

In practice, we often encounter another category of problems, where only the boundary areas of the above solid objects are utilized to model the problems and thus the occupied spaces are more sparsely populated. This category of problems is thus referred to as the surface problems and is also considered herein. For such surface problems, we have

$$N = O\big(W^{\tau-1}\big), \qquad \tau = 1, 2, 3. \tag{45}$$

Following similar argumentations for the volume problems, it can be easily shown that the MLFMA has $O(N \log N)$ complexity for the 3-D surface problem and $O(N^2)$ complexity for the 2-D surface problem, while the proposed algorithm has $O(N \log N)$ complexity for both cases.

Notably, for the 1-D surface problem, N no longer scales with respect to W, i.e., $N = O(W^0) = O(1)$. Due to such degeneracy, this case is analyzed separately and paid special attention. A typical example of this degenerate case is the interaction between two objects separated by some distance $\tilde{W}$. Note that, with the usual box-tree grouping, $\tilde{W} \propto W$. When $\tilde{W} \to \infty$, these two objects can be simply treated as two points. In the following, the complexity of such two points interaction problems is analyzed without loss of generality. Besides, only the complexity with respect to W needs to be studied, since N is a constant in this case.

When the MLFMA is used, the far-field translations are performed only between two well-separated groups at level $l = 2$, yielding

$$\sum_{l=L}^{2} T^{(l)} = T^{(2)} = O\big(R^2\big) = O\big(W^2\big). \tag{46}$$

Moreover, for this particular case, the cost for the aggregations and disaggregations is characterized by

$$\sum_{l=L}^{2} T_{\mathrm{K}}^{(l)} = O\left(\sum_{l=L}^{2} \left(R^{(l)}\right)^2\right) = O\left(\sum_{l=L}^{2} \left(\frac{W}{2^l}\right)^2\right) = O\left(W^2\right) \tag{47}$$

where $T_{\mathrm{K}}^{(l)}$ denotes the interpolation and anterpolation cost for level l and is in principle characterized by the number of the samples involved at level l. Thus, considering both aspects above, the total cost is finally $O(W^2)$, which is clearly unfavorable in practice.

When the proposed algorithm is used, the far-field translations are performed only between two well-separated groups at level $l = \Gamma$, yielding

$$\sum_{l=L}^{\Gamma} T^{(l)} = T^{(\Gamma)} = O\left(R^0\right) = O(1). \tag{48}$$

Then, the cost for the aggregations and disaggregations is considered. Since the samples $\hat{\boldsymbol{k}}_s$ involved at level Γ are here determined by the samples necessary for the two-group translations at level Γ, the interpolation and anterpolation costs $T_{\mathrm{K}}^{(l)}$ for level $l = \Gamma$ are thus described by

$$T_{\mathrm{K}}^{(\Gamma)} = T^{(\Gamma)} = O\left(W^0\right) = O(1) \tag{49}$$

and meanwhile the involved samples $\hat{\boldsymbol{k}}_s$ at level Γ are thus confined within a cone region. Moreover, since the sampling densities are reduced by almost a half along each of the two dimensions (i.e., θ and ϕ) of the spherical surface when descending the levels, the interpolation and anterpolation costs between adjacent levels are therefore related by

$$T_{\mathrm{K}}^{(l+1)} \approx T_{\mathrm{K}}^{(l)}/4, \qquad l = L-1, \cdots, \Gamma. \tag{50}$$

With (50) in mind, the cost for the aggregations and disaggregations of all the involved levels is then described by

$$\sum_{l=L}^{\Gamma} T_{\mathrm{K}}^{(l)} \approx T_{\mathrm{K}}^{(\Gamma)} \cdot \left[1 + \frac{1}{4} + \left(\frac{1}{4}\right)^2 + \cdots \left(\frac{1}{4}\right)^{L-\Gamma}\right] \tag{51}$$

which implies

$$T_{\mathrm{K}}^{(\Gamma)} < \sum_{l=L}^{\Gamma} T_{\mathrm{K}}^{(l)} < 2 \cdot T_{\mathrm{K}}^{(\Gamma)}. \tag{52}$$

Combining (49) and (52), the cost for the aggregations and disaggregations thus follows:

$$\sum_{l=L}^{\Gamma} T_{\mathrm{K}}^{(l)} = O\left(T_{\mathrm{K}}^{(\Gamma)}\right) = O\left(W^0\right) = O(1). \tag{53}$$

Then, taking both (48) and (53) into account, the total cost is finally $O(1)$. Besides, since N is a constant in this case, we thus have $O(N \log N) = O(1)$. Consequently, for the 1-D surface problems, the proposed algorithm can also be regarded as following the complexity of $O(N \log N)$. In other words, for the 1-D/2-D/3-D volume and surface problems considered, a complexity of $O(N \log N)$ can always be achieved.

V. Numerical Examples

For the proposed algorithm, several parameters are specified as follows. The width w_L of the bottom level box is set to 0.2λ. The number of LFF levels used is denoted by P_{LFF}, and P_{LFF} is here set to 4. The truncation tolerance ϵ is set to 0.1. The constant σ is set to 0.05, and the threshold δ is set to 1.0E-3. More discussions on the choice of the parameters (w_L, P_{LFF}, and σ) and relevant constants (C_{LFF} and C_{QFF}) are given in Appendix C.

The combined field integral equation (CFIE) is here employed to characterize the scattering problem [15]. The objects considered are discretized with the triangular mesh [55]. The edge length of the mesh is by convention set to about 0.1λ. The method of moments (MoMs) [18] with the Rao–Wilton–Glisson (RWG) functions [19] discretization is then adopted routinely. The conjugate gradient stabilized (CGS) method [56] is used to solve the MoM matrix equation. The iteration residual is set to 1.0E-3. The block diagonal preconditioner is employed to improve the convergence [10].

A. Complexity and Accuracy

In the following, the proposed fast matrix-vector product (MVP) is tested for problems of various dimensional features. The MVP results obtained using the brute-force method (i.e., dense MVP) are used as the reference to calculate the relative errors.

First, the 1-D surface problem is considered. Concretely, the two points interaction problem is studied. For this problem, we always have $N = 2$, and the computational performance with respect to the distance $\tilde{W}$ between these two points is investigated. The computational time and relative errors with respect to $\tilde{W}$ are illustrated in Fig. 8(a) and (b), respectively. It can be seen that the MVP time follows the trend of $O(1)$ complexity, as expected. Moreover, it is seen that, as $\tilde{W}$ increases, the relative errors maintain almost constant order.

Then, we consider the surface problems where N increases with respect to W, i.e., the 2-D and 3-D surface problems. Clearly, these problems are typical N-body problems with highly oscillatory kernels. Without loss of generality, here we use the point distributions to characterize the object considered. The point distributions and corresponding computational results are illustrated in Fig. 8. In these figures, we use XYZ to characterize the geometric range spanned by the object. The spacing between the points along each dimension is 2.0 λ. Here, denoting the number of points discretized along each cartesian coordinate direction as $\bar{p}$, then $N = \bar{p}^{\tau} - (\bar{p} - 2)^{\tau}$ points are employed to characterize the object, where $\tau = 2, 3$. The MVP results for both the 2-D and 3-D surface problems considered are shown in Fig. 8. We can see that the computational time follows the trend of $O(N \log N)$ complexity. Furthermore, we can also see that the relative errors maintain almost constant order for increasingly larger N.

The computational performances for the volume problems of various dimensional features are further tested. Here, $N = \bar{p}^{\tau}$ points are employed to characterize the objects, where $\tau = 1, 2, 3$. From Fig. 9, we can see that, similar to the results for the surface problems, the computational time for the volume

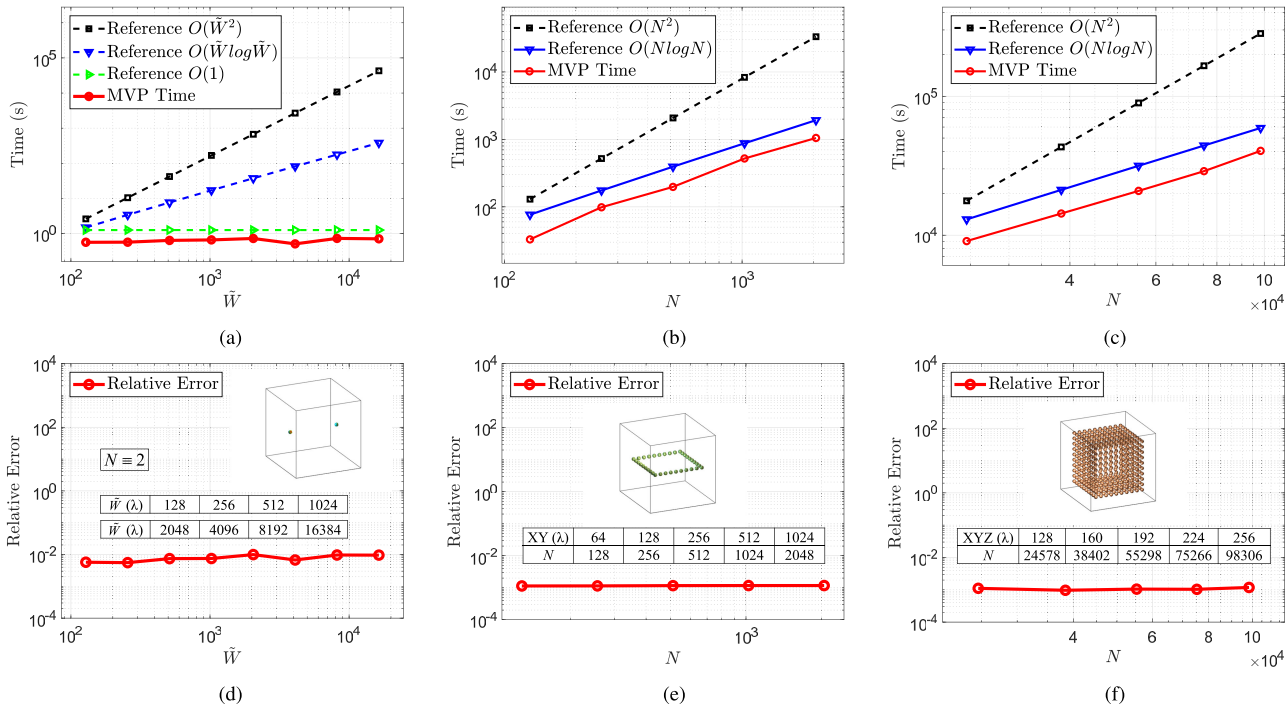


Fig. 8. Time and accuracy for the surface problems of various dimensions. (a) Computational time 1-D. (b) Computational time 2-D. (c) Computational time 3-D. (d) Relative error 1-D. (e) Relative error 2-D. (f) Relative error 3-D.

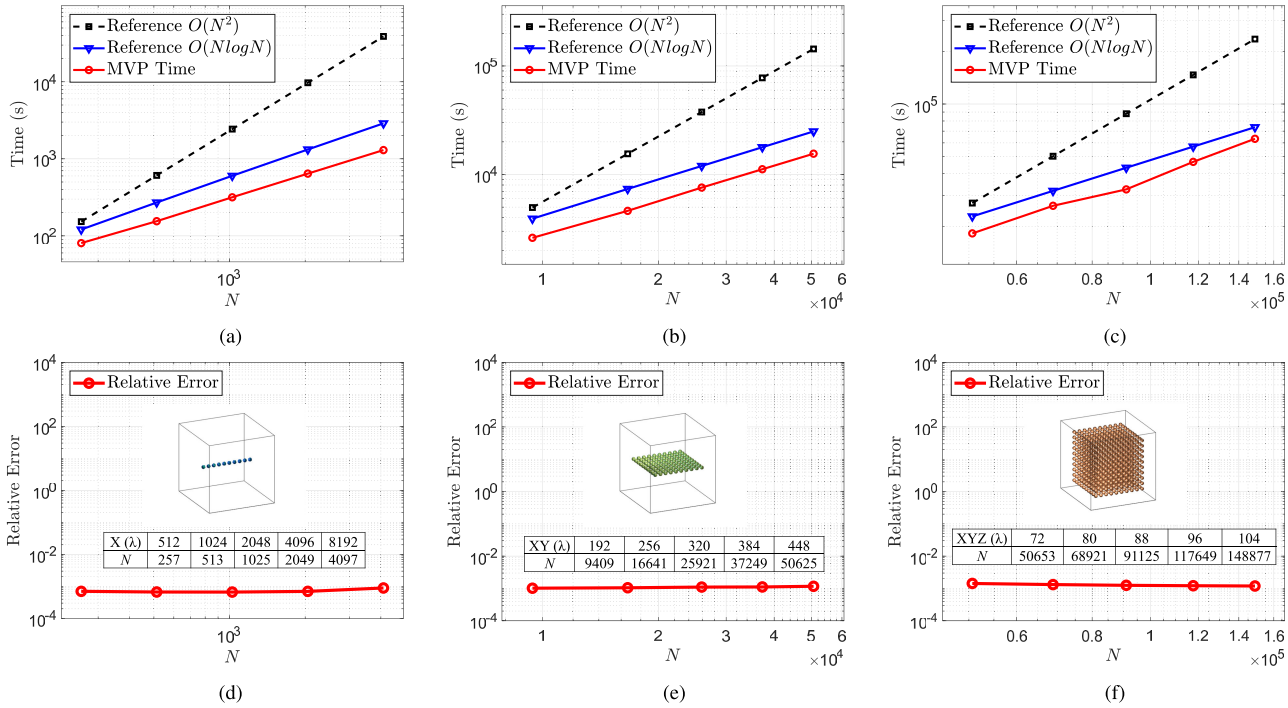


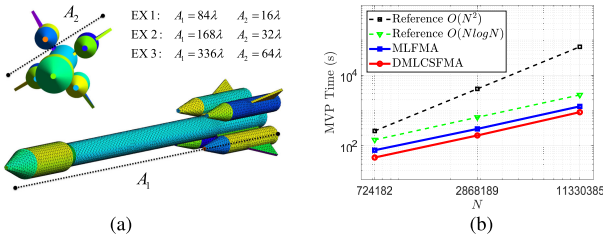
Fig. 9. Time and accuracy for the volume problems of various dimensions. (a) Computational time 1-D. (b) Computational time 2-D. (c) Computational time 3-D. (d) Relative error 1-D. (e) Relative error 2-D. (f) Relative error 3-D.

problems follows the trend of $O(N \log N)$ complexity and the corresponding relative errors maintain almost constant order for increasingly larger N.

B. Application to Scattering Problems

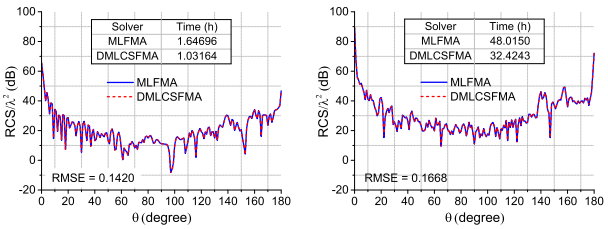
To demonstrate the performance for realistic problems, three typical scenarios are investigated in the following.

1) A rocket model is first considered. The height and the maximum diameter of the rocket are denoted by the A_1 and A_2, respectively. Here, three cases corresponding to different electrical sizes are studied, as shown in Fig. 10(a). Clearly, such a series of examples correspond to a commonly encountered scenario, i.e., a given object illuminated by incident waves of different incident frequencies f. Suppose that the height A_1 and the maximum diameter A_2 of the rocket model are



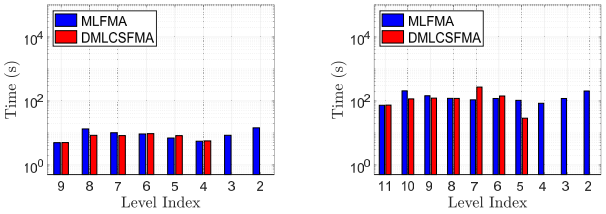
(a) (b)

Fig. 10. Geometric settings of the rocket model and corresponding MVP times. (a) Geometric settings. (b) MVP time.



(a) (b)

Fig. 11. RCS results and solution times for the rocket model. (a) RCS EX1. (b) RCS EX3.

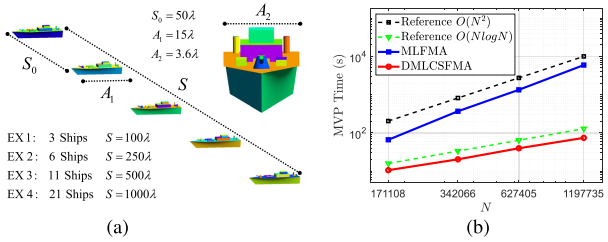


(a) (b)

Fig. 12. Distribution of the computational work for the rocket model. (a) Workload distributions EX1. (b) Workload distributions EX3.

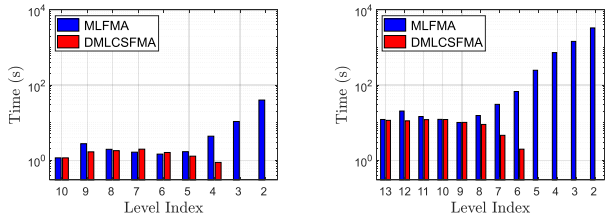
84 and 16 m, respectively. Then, case 1 corresponds to the case with the incident frequency $f = 300$ MHz. Meanwhile, cases 2 and 3 correspond to the cases with the incident frequencies $f = 600$ MHz and $f = 1.2$ GHz, respectively.

The rocket is obviously a 3-D object. Moreover, the scattering problem is here characterized by the SIE (i.e., the CFIE adopted herein). Thus, for increasing f, the number of unknowns N should be proportional to the surface area of this 3-D object. This means that, for increasing f, N should be proportional to the square of the span of the object [i.e., $N = O({A_1}^2) = O({A_2}^2)$], whatever the shape of the object (i.e., even if the ratio A_1/A_2 is very large and the object shape is highly elongated). Consequently, for the three abovementioned cases with increasing N (induced by the increase in f), the resultant MVP times of both the MLFMA and the DMLCSFMA should follow the trend of the complexity of the 3-D surface problem discussed in Section IV-B. That is, both algorithms should exhibit $O(N \log N)$ complexity, for the rocket examples considered. In Fig. 10(b), the resultant MVP times of both algorithms are illustrated, and the numerical results are in good agreement with the theoretical predictions. Besides, in Fig. 11, the RCS results and solution times for the rocket model are illustrated. Note that, in this scenario, both the MLFMA and the DMLCSFMA show good performance in practice. Moreover, the workload distributions for some of the considered settings are shown in Fig. 12.



(a) (b)

Fig. 13. Geometric settings of the ship model (a fleet of ships) and corresponding MVP times. (a) Geometric settings. (b) MVP time.



(a) (b)

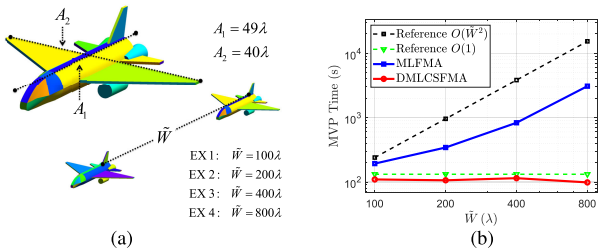
Fig. 14. Distribution of the computational work for the ship model (a fleet of ships). (a) Workload distributions EX1. (b) Workload distributions EX4.

2) A scenario consisting of a fleet of ships is considered. The length and width of a single ship are denoted by A_1 and A_2, respectively. Here, four cases corresponding to different numbers of ships are studied, as depicted in Fig. 13(a). In particular, all the ships are arranged along a line (i.e., a single dimension), with a separation distance of S_0 between any two adjacent ships. If the span of the fleet of ships along the direction of arrangement is denoted by S and such fleet of ships is treated as a whole object, then such an object has the length S and the width A_1. When S is sufficiently larger than A_1, such an object can be viewed as an object of elongated shape. Since A_1 is fixed, the object shape then becomes more and more elongated, as S increases (i.e., when more ships are involved).

Suppose that the number of ships considered is denoted by M_U. For the fleet of ships, the total number of unknowns N is obviously proportional to M_U. Moreover, the ships are arranged along a line. Thus, as M_U increases, N increases linearly with respect to the span S. That is, the increase in N is only caused by the increase in the object span along a single dimension. Then, one can identify that the scenario considered here matches the 1-D volume problem discussed in Section IV-A. That is, for the four cases considered here, the MVP times of the MLFMA should follow the $O(N^2)$ complexity, while that of the DMLCSFMA should follow the $O(N \log N)$ complexity. In Fig. 13(b), the resultant MVP times of both algorithms are illustrated, and the numerical results are in good agreement with the theoretical predictions. Besides, the workload distributions for some of the considered settings are shown in Fig. 14.

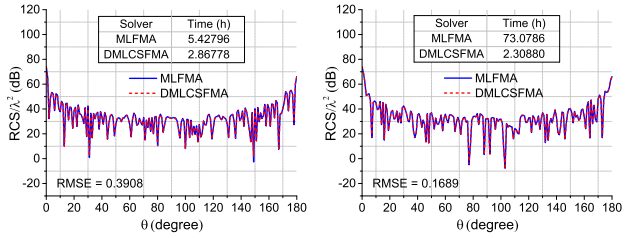
3) A scenario consisting of two aircraft is finally considered. The length and width of a single aircraft are denoted by A_1 and A_2, respectively. Here, four cases corresponding to different central distance $\tilde{W}$ between the two aircraft are studied, as shown in Fig. 15(a).

Obviously, as $\tilde{W}$ changes, the two aircraft remain unchanged and the total number of unknowns N thus remains unchanged. With this in mind, one can then identify that the scenario considered here matches the 1-D surface problem discussed



(a) (b)

Fig. 15. Geometric settings of the aircraft model (two aircraft) and corresponding MVP times. (a) Geometric settings. (b) MVP time.



(a) (b)

Fig. 16. RCS results and solution times for the aircraft model (two aircraft). (a) RCS EX1. (b) RCS EX4.

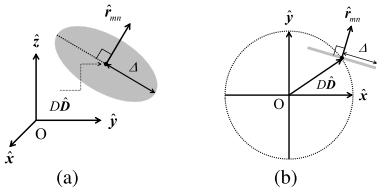
in Section IV-B. This implies that, as $\tilde{W}$ increases, the MVP times of the MLFMA should ultimately follow the trend of $O(\tilde{W}^2)$ complexity, while that of the DMLCSFMA should follow the trend of $O(1)$ complexity. In Fig. 15(b), the resultant MVP times of both algorithms are illustrated. It can be clearly seen that, as $\tilde{W}$ increases, the numerical results approach the theoretical predictions, as expected. Besides, in Fig. 16, the RCS results and solution times for the aircraft model are illustrated. In particular, for the two cases considered, the solution times of the DMLCSFMA are reduced by 47.16% and 96.84%, respectively, compared with that of the MLFMA. Hence, for this scenario, the improvement can be very prominent.

C. Further Discussion

To gain further insight, the distribution of the computational work across the involved levels in the MLFMA and the DMLCSFMA are illustrated, in Fig. 12 for the rocket examples and in Fig. 14 for the ship examples. It is seen that the reduced costs of the DMLCSFMA (relative to the MLFMA) are mainly attributable to the reduced costs at the relatively higher levels (e.g., levels 2 and 3) where the boxes are relatively larger. In contrast, at the relatively lower levels (e.g., levels 9 and 10) where the boxes are relatively smaller, the costs of both algorithms are very similar, if not identical. These facts imply that the $(size)^2$ complexity of FMA is due to the expensive and inefficient operations on the largest boxes, where a large fraction of radiation directions are "wasted."

VI. Conclusion

In this work, a DMLCSFMA is developed for accelerating the calculation of the electrically large problems of various dimensions. This algorithm implements a high-frequency generalization of the conventional mid-frequency MLFMA. In particular, a stable complexity of $O(N \log N)$ can always be achieved, for objects of different dimensional features.



(a) (b)

Fig. 17. Illustration of the complex source of the GB pattern. Here, $D\hat{\boldsymbol{D}}$ defines the location of a real point source. The sources of the corresponding GB lie on a circular disk (grey region) of radius Δ, centered at $D\hat{\boldsymbol{D}}$, with normal $\hat{\boldsymbol{r}}_{mn}$. (a) Three-dimensional view. (b) Two-dimensional view.

Furthermore, being a spectral counterpart of the traditional equivalent source-based D-FMAs, the proposed algorithm also demonstrates the feasibility of building a plane wave-based D-FMA. Both basic principles and implementation details are thoroughly discussed.

Besides, this work potentially opens up new opportunities to drastically upgrade the parallel efficiency of the state-of-the-art parallel MLFMA [22], [23]. Effective parallelization strategies of the MLFMA (such as the simple, the hybrid, the hierarchical, and the ternary approach) [23] are closely related to the ML structure of the MLFMA, and depend heavily on the proper partitioning of the box-tree and the corresponding group far-field patterns. Since the box-tree levels involved in the DMLCSFMA are almost half of that of the MLFMA and the far-field patterns involved are also sparsified, there should be greater space for further optimization (especially on the partitioning strategy), and improved parallel efficiency can thus be expected.

Appendix A

Here, we give more detailed explanations of the asymptotic behavior of the truncation length V mentioned in (14). To this end, consider the following radiation pattern (i.e., GB pattern) [42], [47]:

$$\beta\left(\hat{\boldsymbol{k}}, \boldsymbol{D}_c\right) = e^{-ik\hat{\boldsymbol{k}}\cdot\boldsymbol{D}_c} \tag{54}$$

where $\boldsymbol{D}_c = D\hat{\boldsymbol{D}} + i\Delta\hat{\boldsymbol{r}}_{mn}$ is a complex-valued spatial vector, with $D = 2R$ and Δ as suggested in (11). Here, the real part $D\hat{\boldsymbol{D}}$ defines the location of a real point source, which can be enclosed by a sphere of radius D. Meanwhile, the imaginary part $i\Delta\hat{\boldsymbol{r}}_{mn}$ induces the GB [43]. The sources of the corresponding GB lie on a circular disk of radius Δ, centered at $D\hat{\boldsymbol{D}}$, with normal $\hat{\boldsymbol{r}}_{mn}$, as shown in Fig. 17. It is known that the above radiation pattern (54) can be expanded into a spherical harmonics series [12], [47]. Then, a truncation length N_{D_c} can be correspondingly determined for a given tolerance ε. Note that this truncation length is also an indication of the bandwidth of the radiation pattern.

Specifically, for a purely real point source $\boldsymbol{D} = D\hat{\boldsymbol{D}}$ (i.e., without considering $i\Delta\hat{\boldsymbol{r}}_{mn}$), the corresponding truncation length N_D for a given ε can be determined by the smallest integer that satisfies [12], [47]

$$|e^{-ikD\,\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{D}}} - \sum_{v=0}^{N_D}(2v+1)i^{-v}\mathrm{j}_v(kD)\mathrm{P}_v\left(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{D}}\right)| \le \varepsilon \tag{55}$$

for arbitrary $\hat{\boldsymbol{D}}$. In this well-known case, a good closed-form formula is available and is given by [15]

$$N_D \approx kD + C(kD)^{1/3} \tag{56}$$

where $C = (-3\ln\varepsilon)^{2/3}/2$. Clearly, in this case, N_D is roughly linearly proportional to D.

However, for a purely imaginary source $\boldsymbol{D}_{i\Delta} = i\Delta\hat{\boldsymbol{r}}_{mn}$, the corresponding truncation length $N_{i\Delta}$ for a given ε is instead determined by the smallest integer that satisfies [47]

$$|e^{k\Delta\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}}_{mn}} - \sum_{v=0}^{N_{i\Delta}} (2v+1) i^{-v} \mathrm{j}_v(ik\Delta)\mathrm{P}_v\big(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}}_{mn}\big)| \le \varepsilon \cdot e^{k\Delta} \tag{57}$$

for arbitrary $\hat{\boldsymbol{r}}_{mn}$. It has been shown that $N_{i\Delta}$ here follows a square-root relation with respective to Δ, namely [42], [47], [48]

$$N_{i\Delta} \approx \gamma\sqrt{k\Delta} \tag{58}$$

where $\gamma = (-2\ln\varepsilon)^{1/2}$. Furthermore, using the idea of physical equivalence, the above can also be rewritten as

$$N_{i\Delta} \approx k \cdot \tilde{D}_{i\Delta}, \quad \tilde{D}_{i\Delta} = \gamma\sqrt{\frac{\Delta}{k}}. \tag{59}$$

Comparing (56) with (59), $\tilde{D}_{i\Delta}$ can thus be viewed as the real effective radius corresponding to the purely imaginary source $\boldsymbol{D}_{i\Delta} = i\Delta\hat{\boldsymbol{r}}_{mn}$.

Based on the above, by considering the composite radius resulted from both D and $\tilde{D}_{i\Delta}$, the effective radius of the sphere enclosing the general complex source $\boldsymbol{D}_c = D\hat{\boldsymbol{D}} + i\Delta\hat{\boldsymbol{r}}_{mn}$ can be further obtained as follows [42]:

$$\tilde{D}_c = \sqrt{D^2 + \gamma^2\frac{\Delta}{k} + 2\gamma\sqrt{\frac{\Delta}{k}}D|\sin\vartheta|} \tag{60}$$

where ϑ is the angle between $\hat{\boldsymbol{D}}$ and $\hat{\boldsymbol{r}}_{mn}$. It is noticed that the above achieves the maximum when $\vartheta = \pm\pi/2$, i.e., $\hat{\boldsymbol{D}}\perp\hat{\boldsymbol{r}}_{mn}$. In this case, we have an upper bound of the effective radius $\tilde{D}_c$, namely

$$\tilde{D}_c^{\sup} = D + \gamma\sqrt{\frac{\Delta}{k}}. \tag{61}$$

Consequently, an upper bound for N_{D_c} is thus given by [42]

$$N_{D_c}^{\sup} \approx k\tilde{D}_c^{\sup} + C\big(k\tilde{D}_c^{\sup}\big)^{1/3} \tag{62}$$

which is obtained by replacing D in (56) with $\tilde{D}_c^{\sup}$ in (61).

Besides, from the above, we can easily identify

$$N_{D_c}^{\sup} \approx N_D + N_{i\Delta}. \tag{63}$$

It is worth mentioning that such a conclusion can also be reached straightforwardly from the perspective of the Fourier relation, by noticing that

$$\beta\big(\hat{\boldsymbol{k}}, \boldsymbol{D}_c\big) = \beta\big(\hat{\boldsymbol{k}}, \boldsymbol{D}\big)\cdot\beta\big(\hat{\boldsymbol{k}}, \boldsymbol{D}_{i\Delta}\big) \tag{64}$$

which implies that the bandwidth of the radiation pattern β with the complex source $\boldsymbol{D}_c$ is the sum of the bandwidths with a purely real source $\boldsymbol{D}$ and with a purely imaginary source $\boldsymbol{D}_{i\Delta}$.

Interestingly, it is found that, using the Δ suggested in (11), formula (61) can be further simplified, yielding

$$\tilde{D}_c^{\sup} = D + \gamma\sqrt{\frac{\sigma\cdot C_{\mathrm{QFF}}\cdot(D/2)^2}{k}} = \rho\cdot D \tag{65}$$

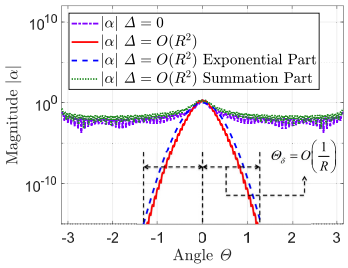


Fig. 18. Magnitudes of the translator function α_{mn} and relevant subparts. The magnitudes for both $\Delta = 0$ and $\Delta = O(R^2)$ are plotted. Moreover, the relevant exponential part and the summation part are also shown above. Here, the central distance between the source group n and the field group m is set to $r_{mn} = \mathrm{f}_{\mathrm{QFF}}(R) := C_{\mathrm{QFF}}\cdot R^2$. Particularly, with the suggested $\Delta = O(R^2)$, the cut-off range Θ_δ of the spectral window is inversely proportional to the group radius R, i.e., $\Theta_\delta = O(1/R)$.

where $\rho = 1 + \gamma/2\cdot(\sigma\cdot C_{\mathrm{QFF}}/k)^{1/2}$ is a constant. The above says, with the suggested Δ, the corresponding upper bound $\tilde{D}_c^{\sup}$ is an explicit linear function of the real radius D, namely, $\tilde{D}_c^{\sup} \propto D$. Notice that here the suggested quadratic setting of Δ with respect to D happens to balance out the intrinsic square-root relation related to Δ, thus producing an explicit linear function of D. Then, substituting (65) into (62), we can easily obtain

$$N_{D_c}^{\sup} = O\big(\tilde{D}_c^{\sup}\big) = O(D). \tag{66}$$

Moreover, it can be noticed that N_D in (56) is actually a lower bound of N_{D_c}. Hence, we also have

$$N_{D_c}^{\inf} = N_D = O(D). \tag{67}$$

Consequently, considering the fact that $N_{D_c}^{\inf} \le N_{D_c} \le N_{D_c}^{\sup}$ and taking both (66) and (67) into account, we thus have

$$N_{D_c} = O(D) = O(R) \tag{68}$$

where $D = 2R$, meaning that, similar to N_D for the real point source case, the truncation length N_{D_c} for the complex source case is still of linear order with respect to D. Then, noticing that the truncation length V is in principle characterized by N_{D_c} [12], [42], [43], the asymptotic behavior of V is thus obtained, which is $V = O(R)$.

APPENDIX B

In the following, we study the expression of the windowed translator function α_{mn} in (9), in order to reveal the asymptotic behavior of the spectral window.

To this end, referring to (10), consider the representative far-field setting $r_{mn} = \mathrm{f}_{\mathrm{QFF}}(R)$. In this case, according to (11), the suggested choice of the beam parameter can also be expressed as $\Delta = \sigma\cdot r_{mn}$. Correspondingly, the expression of the GB translator α_{mn} in (9) can be rewritten as

$$\alpha_{mn}\big(\hat{\boldsymbol{k}}, \boldsymbol{r}_{mn}, \Delta\big) = \frac{ik}{4\pi}e^{k\Delta(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}}_{mn}-1)} \sum_{v=0}^{W} i^v(2v+1)\tilde{\mathrm{h}}_v^{(1)}\big(k^\dagger r_{mn}\big)\mathrm{P}_v\big(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}}_{mn}\big) \tag{69}$$

where $k^\dagger = k(1+i\sigma)$. Interestingly, it can be noticed that the summation part in (69) takes the form of the conventional

translator (α_{mn} with $\Delta = 0$) with a lossy wavenumber $k^{\dagger}$, and exhibits undecaying magnitude behavior similar to the conventional translator, as shown in Fig. 18. Furthermore, the magnitude of the windowed translator function α_{mn} is dominated by the exponential part $e^{k\Delta(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}}_{mn}-1)}$ in (69). Thus, to yield an appropriate estimate of the asymptotic behavior of the window for increasing R, we can proceed solely based on this simple exponential factor.

Specifically, since $\cos\Theta \approx 1 - \Theta^2/2$ for $\Theta \to 0$, the exponential factor in (69) can be approximated as follows:

$$e^{k\Delta(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}}_{mn}-1)} = e^{k\Delta(\cos\Theta-1)} \approx e^{-k\Delta\Theta^2/2}. \tag{70}$$

Hence, for a given cut-off threshold δ, the range of the cut-off window can be estimated by solving

$$e^{-k\Delta\Theta^2/2} \geq \delta. \tag{71}$$

Then, substituting the suggested Δ of (11) into (71), we thus arrive at

$$\Theta \leq \Theta_\delta = \frac{C_\delta}{R} = O\left(\frac{1}{R}\right) \tag{72}$$

where $C_\delta = (-2\ln\delta/(k\sigma C_{\mathrm{QFF}}))^{1/2}$ is a constant, which means, with the suggested Δ, the range of the cut-off window characterized by Θ_δ is inversely proportional to R.

Appendix C

First, the value of C_{LFF} is discussed. Here, the commonly-used setting where the well-separated boxes are in the nearest case one-box apart (i.e., 1-box buffer scheme [13]) is considered. For such a setting, the nearest central distance between well-separated boxes is $r_{mn} = 2w$, where w denotes the width of the box. Accordingly, referring to (22), C_{LFF} can be obtained by solving $2 \cdot w = C_{\mathrm{LFF}} \cdot R$, yielding

$$C_{\mathrm{LFF}} = 4/\sqrt{3} \approx 2.3094010767585 \tag{73}$$

where the fact that $R = (3)^{1/2}w/2$ has been used.

Then, we discuss a rule for choosing C_{QFF}. Traditionally, for the MLFMA, the width w_L of the bottom level box is given, which is commonly set to be around 0.2–0.5 λ. Around such a scale, low-frequency physics begins to diminish and mid-frequency wave physics begins to prevail. Here, for the DMLCSFMA, we need to further specify the number of LFF levels employed, i.e., P_{LFF}. Correspondingly, the group radius at the transition level L_Δ can be expressed as

$$R_{L_\Delta} = R_L \cdot 2^{L-L_\Delta} = \frac{\sqrt{3}}{2} w_L \cdot 2^{L-L_\Delta} = \frac{\sqrt{3}}{2} w_L \cdot 2^{P_{\mathrm{LFF}}}. \tag{74}$$

Recall that the crossover point (28) between the LFF and the QFF is $R_{\mathrm{M2H}} = C_{\mathrm{LFF}}/C_{\mathrm{QFF}}$. To guarantee a smooth transition from the highest LFF level (i.e., level $L_\Delta + 1$) to the lowest QFF level (i.e., level L_Δ), it is suggested to make sure

$$R_{L_\Delta} = R_{\mathrm{M2H}}. \tag{75}$$

With the above constraint, C_{QFF} can be correspondingly determined, yielding

$$C_{\mathrm{QFF}} = \frac{2C_{\mathrm{LFF}}}{\sqrt{3} w_L \cdot 2^{P_{\mathrm{LFF}}}}. \tag{76}$$

Generally, ray physics becomes prominent at the scale of several wavelengths. Hence, if we choose $w_L = 0.2\lambda$, then setting $P_{\mathrm{LFF}} = 4$ can yield an appropriate transition level L_Δ, at which the group radius is $R_{L_\Delta} = 3.2\lambda$. Correspondingly, in this case, we have $C_{\mathrm{QFF}} \approx 0.83$. Note that this value is very close to the estimate given in (21) derived from a different perspective.

Next, we discuss the choice of σ. Roughly speaking, the applicable range of σ can be derived from the convergence condition of the complex-space extended addition theorem [43]. Specifically, according to [43] and [45], the upper bound of the applicable Δ is given by

$$\Delta_{\max} = \frac{r_{mn}^2 - D^2}{2D} \tag{77}$$

where $D = 2R$. Moreover, from (11), we have $\sigma = \Delta/r_{mn}$ with $r_{mn} = C_{\mathrm{QFF}}R^2$. Then, combining the above aspects, it can be shown that

$$\sigma_{\max} = \frac{C_{\mathrm{QFF}}R}{4} - \frac{1}{C_{\mathrm{QFF}}R}. \tag{78}$$

Notice that $\sigma_{\max}$ in (78) is an increasing function of R. Thus, to make sure a specified σ is valid for all possible R (i.e., $R^{(l)}$ where $l = L_\Delta, \ldots, \Gamma$), the minimum of $\sigma_{\max}$ within the considered range of R should be determined. Clearly, such minimum is reached when $R = R_{L_\Delta}$. Considering that $R_{L_\Delta} = R_{\mathrm{M2H}} = C_{\mathrm{LFF}}/C_{\mathrm{QFF}}$ as suggested in (75), the minimum of $\sigma_{\max}$ is thus given by

$$\sigma_{\max} = \sigma_{\max}(R_{L_\Delta}) = \frac{C_{\mathrm{LFF}}}{4} - \frac{1}{C_{\mathrm{LFF}}}. \tag{79}$$

Finally, substituting C_{LFF} in (73) into (79) above yields $\sigma_{\max} = (3)^{1/2}/12 \approx 0.1443$. In practice, to maintain the numerical stability [43], the σ adopted should not be too close to $\sigma_{\max}$. Numerical experiments show that the estimate $\sigma = 1/(2\pi^2) \approx 0.05$ given in (21) can be an appropriate choice and is adopted in this article.

Acknowledgment

The authors would like to thank the anonymous reviewers for their helpful suggestions.

References

[1] B. A. Cipra, "The best of the 20th century: Editors name top 10 algorithms," *SIAM News*, vol. 33, no. 4, pp. 1–22, 2000.
[2] J. Barnes and P. Hut, "A hierarchical O(NlogN) force calculation algorithm," *Nature*, vol. 324, pp. 446–449, Dec. 1986.
[3] L. Hernquist, "Hierarchical N-body methods," *Comput. Phys. Commun.*, vol. 48, no. 1, pp. 107–115, Jan. 1988.
[4] W. C. Chew, "Computational electromagnetics: The physics of smooth versus oscillatory fields," *Phil. Trans. Roy. Soc. London. Ser. A, Math., Phys. Eng. Sci.*, vol. 362, no. 1816, pp. 579–602, Mar. 2004.
[5] V. Rokhlin, "Rapid solution of integral equations of classical potential theory," *J. Comput. Phys.*, vol. 60, no. 2, pp. 187–207, Sep. 1985.
[6] L. Greengard and V. Rokhlin, "A fast algorithm for particle simulations," *J. Comput. Phys.*, vol. 73, no. 2, pp. 325–348, Dec. 1987.
[7] V. Rokhlin, "Diagonal forms of translation operators for the Helmholtz equation in three dimensions," *Appl. Comput. Harmon. Anal.*, vol. 1, no. 1, pp. 82–93, Dec. 1993.
[8] R. Coifman, V. Rokhlin, and S. Wandzura, "The fast multipole method for the wave equation: A pedestrian prescription," *IEEE Antennas Propag. Mag.*, vol. 35, no. 3, pp. 7–12, Jun. 1993.

[9] A. Brandt, "Multilevel computations of integral transforms and particle interactions with oscillatory kernels," *Comput. Phys. Commun.*, vol. 65, nos. 1–3, pp. 24–38, Apr. 1991.

[10] J. Song, C.-C. Lu, and W. C. Chew, "Multilevel fast multipole algorithm for electromagnetic scattering by large complex objects," *IEEE Trans. Antennas Propag.*, vol. 45, no. 10, pp. 1488–1493, Oct. 1997.

[11] M. Hidayetoglu and W. C. Chew, "On computational complexity of the multilevel fast multipole algorithm in various dimensions," in *Proc. IEEE Int. Symp. Antennas Propag. (APSURSI)*, Jun. 2016, pp. 1559–1560, doi: 10.1109/APS.2016.7696486.

[12] S. Koc, J. Song, and W. C. Chew, "Error analysis for the numerical evaluation of the diagonal forms of the scalar spherical addition theorem," *SIAM J. Numer. Anal.*, vol. 36, no. 3, pp. 906–921, Jan. 1999.

[13] S. Ohnuki and W. C. Chew, "Error minimization of multipole expansion," *SIAM J. Sci. Comput.*, vol. 26, no. 6, pp. 2047–2065, Jan. 2005.

[14] W. Cho Chew, J.-M. Jin, C.-C. Lu, E. Michielssen, and J. M. Song, "Fast solution methods in electromagnetics," *IEEE Trans. Antennas Propag.*, vol. 45, no. 3, pp. 533–543, Mar. 1997.

[15] W. C. Chew, J.-M. Jin, E. Michielssen, and J. Song, *Fast and Efficient Algorithms in Computational Electromagnetics*. Boston, MA, USA: Artech House, 2001.

[16] W. C. Chew and L. J. Jiang, "Overview of large-scale computing: The past, the present, and the future," *Proc. IEEE*, vol. 101, no. 2, pp. 227–241, Feb. 2013.

[17] W. C. Chew et al., "New trends in computational electromagnetics," in *New Trends in Computational Electromagnetics*. London, U.K.: Institution of Engineering and Technology, 2020, pp. 13–91.

[18] R. F. Harrington, *Field Computation by Moment Methods*. New York, NY, USA: MacMillian, 1968.

[19] S. Rao, D. Wilton, and A. Glisson, "Electromagnetic scattering by surfaces of arbitrary shape," *IEEE Trans. Antennas Propag.*, vol. AP-30, no. 3, pp. 409–418, May 1982.

[20] L. J. Jiang and W. C. Chew, "A mixed-form fast multipole algorithm," *IEEE Trans. Antennas Propag.*, vol. 53, no. 12, pp. 4145–4156, Dec. 2005.

[21] T. Xia, L. L. Meng, Q. S. Liu, H. H. Gan, and W. C. Chew, "A low-frequency stable broadband multilevel fast multipole algorithm using plane wave multipole hybridization," *IEEE Trans. Antennas Propag.*, vol. 66, no. 11, pp. 6137–6145, Nov. 2018.

[22] S. Velamparambil, W. C. Chew, and J. M. Song, "10 Million unknowns: Is it that big?" *IEEE Antennas Propag. Mag.*, vol. 45, no. 2, pp. 43–58, Sep. 2003.

[23] M.-L. Yang, B.-Y. Wu, H.-W. Gao, and X.-Q. Sheng, "A ternary parallelization approach of MLFMA for solving electromagnetic scattering problems with over 10 billion unknowns," *IEEE Trans. Antennas Propag.*, vol. 67, no. 11, pp. 6965–6978, Nov. 2019.

[24] W. C. Chew, T. J. Cui, and J. M. Song, "A FAFFA-MLFMA algorithm for electromagnetic scattering," *IEEE Trans. Antennas Propag.*, vol. 50, no. 11, pp. 1641–1649, Nov. 2002.

[25] T. J. Cui, W. C. Chew, G. Chen, and J. Song, "Efficient MLFMA, RPFMA, and FAFFA algorithms for EM scattering by very large structures," *IEEE Trans. Antennas Propag.*, vol. 52, no. 3, pp. 759–770, Mar. 2004.

[26] M. Bebendorf, "Approximation of boundary element matrices," *Numerische Math.*, vol. 86, no. 4, pp. 565–589, Oct. 2000.

[27] L. Ying, G. Biros, and D. Zorin, "A kernel-independent adaptive fast multipole algorithm in two and three dimensions," *J. Comput. Phys.*, vol. 196, no. 2, pp. 591–626, May 2004.

[28] W. Fong and E. Darve, "The black-box fast multipole method," *J. Comput. Phys.*, vol. 228, no. 23, pp. 8712–8725, Dec. 2009.

[29] B. Zhang, J. Huang, N. P. Pitsianis, and X. Sun, "A Fourier-series-based kernel-independent fast multipole method," *J. Comput. Phys.*, vol. 230, no. 15, pp. 5807–5821, Jul. 2011.

[30] B. Engquist and L. Ying, "Fast directional multilevel algorithms for oscillatory kernels," *SIAM J. Sci. Comput.*, vol. 29, no. 4, pp. 1710–1737, Jan. 2007.

[31] B. Engquist and L. Ying, "A fast directional algorithm for high frequency acoustic scattering in two dimensions," *Commun. Math. Sci.*, vol. 7, no. 2, pp. 327–345, 2009.

[32] P. Tsuji and L. Ying, "A fast directional algorithm for high-frequency electromagnetic scattering," *J. Comput. Phys.*, vol. 230, no. 14, pp. 5471–5487, Jun. 2011.

[33] M. Messner, M. Schanz, and E. Darve, "Fast directional multilevel summation for oscillatory kernels based on Chebyshev interpolation," *J. Comput. Phys.*, vol. 231, no. 4, pp. 1175–1196, Feb. 2012.

[34] M. Bebendorf, C. Kuske, and R. Venn, "Wideband nested cross approximation for Helmholtz problems," *Numerische Math.*, vol. 130, no. 1, pp. 1–34, May 2015.

[35] R. L. Wagner and W. C. Chew, "A ray-propagation fast multipole algorithm," *Microw. Opt. Technol. Lett.*, vol. 7, no. 10, pp. 435–438, Jul. 1994.

[36] R. Coifman, V. Rokhlin, and S. Wandzura, "Faster single-stage Multipole Method for the wave equation," *10th Annu. Rev. Prog. Appl. Computat. Electromagn.*, Monterey, CA, Mar. 1994, pp. 19–24.

[37] C.-C. Lu and W. C. Chew, "Fast far field approximation for calculating the RCS of large objects," *Microw. Opt. Technol. Lett.*, vol. 8, no. 5, pp. 238–241, Apr. 1995.

[38] R. J. Burkholder and D.-H. Kwon, "High-frequency asymptotic acceleration of the fast multipole method," *Radio Sci.*, vol. 31, no. 5, pp. 1199–1206, Sep. 1996.

[39] V. Rokhlin, "Sparse diagonal forms for translation operators for the Helmholtz equation in two dimensions," *Appl. Comput. Harmon. Anal.*, vol. 5, no. 1, pp. 36–67, Jan. 1998.

[40] B. Hu, W. C. Chew, E. Michielssen, and J. Zhao, "Fast inhomogeneous plane wave algorithm for the fast analysis of two-dimensional scattering problems," *Radio Sci.*, vol. 34, no. 4, pp. 759–772, Jul. 1999.

[41] B. Hu, W. C. Chew, and S. Velamparambil, "Fast inhomogeneous plane wave algorithm for the analysis of electromagnetic scattering," *Radio Sci.*, vol. 36, no. 6, pp. 1327–1340, Nov. 2001.

[42] T. B. Hansen, "Translation operators based on Gaussian beams for the fast multipole method in two dimensions," *Wave Motion*, vol. 50, no. 4, pp. 793–808, Jun. 2013.

[43] T. B. Hansen, "Translation operator based on Gaussian beams for the fast multipole method in three dimensions," *Wave Motion*, vol. 50, no. 5, pp. 940–954, Jul. 2013.

[44] T. B. Hansen, "Numerical properties of a Gaussian translation operator for the 2D FMM," *IEEE Trans. Antennas Propag.*, vol. 62, no. 6, pp. 3119–3129, Jun. 2014.

[45] T. B. Hansen, "Exact plane-wave expansion with directional spectrum: Application to transmitting and receiving antennas," *IEEE Trans. Antennas Propag.*, vol. 62, no. 8, pp. 4187–4198, Aug. 2014.

[46] T. B. Hansen and O. Borries, "Gaussian translation operator in a multilevel scheme," *Radio Sci.*, vol. 50, no. 8, pp. 754–763, Aug. 2015.

[47] T. B. Hansen, "Array realization of complex-source beam," *Wave Motion*, vol. 57, pp. 207–218, Sep. 2015.

[48] T. B. Hansen, "Conformal array realizations of the near and far fields of a complex point source," *Wave Motion*, vol. 72, pp. 1–12, Jul. 2017.

[49] G.-Y. Zhu, H.-X. Zhou, and W. Hong, "A method to construct directional multilevel algorithm in angular spectral domain," in *Proc. Int. Conf. Microw. Millim. Wave Technol. (ICMMT)*, May 2018, pp. 1–2, doi: 10.1109/ICMMT.2018.8563769.

[50] O. Bucci and G. Franceschetti, "On the spatial bandwidth of scattered fields," *IEEE Trans. Antennas Propag.*, vol. AP-35, no. 12, pp. 1445–1455, Dec. 1987.

[51] O. M. Bucci and G. Franceschetti, "On the degrees of freedom of scattered fields," *IEEE Trans. Antennas Propag.*, vol. 37, no. 7, pp. 918–926, Jul. 1989.

[52] O. M. Bucci, C. Gennarelli, and C. Savarese, "Optimal interpolation of radiated fields over a sphere," *IEEE Trans. Antennas Propag.*, vol. 39, no. 11, pp. 1633–1643, Nov. 1991.

[53] F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark, *NIST Handbook of Mathematical Functions*. New York, NY, USA: Cambridge Univ. Press, 2010.

[54] J. D. Kraus and R. J. Marhefka, *Antennas: For All Applications*, 3rd ed. New York, NY, USA: McGraw-Hill, 2001.

[55] C. Geuzaine and J. Remacle, "Gmsh: A 3-D finite element mesh generator with built-in pre- and post-processing facilities," *Int. J. Numer. Methods Eng.*, vol. 79, no. 11, pp. 1309–1331, Sep. 2009.

[56] R. Barrett et al., *Templates for Solution Linear Systems: Building Blocks for Iterative Methods*. Philadelphia, PA, USA: SIAM, 1994.