

Error Control in MLFMA with Multiple-Precision Arithmetic

Mert Kalfa^{1,2}, Özgür Ergül³, Vakur B. Ertürk¹,

¹Department of Electrical and Electronics Engineering, Bilkent University, Ankara, TR-06800, Turkey

²ASELSAN Research Center, ASELSAN Inc., Ankara, TR-06370, Turkey

³Department of Electrical and Electronics Engineering, Middle East Technical University, Ankara, TR-06800, Turkey
E-mail: kalfa@ee.bilkent.edu.tr

Abstract—We present a new error control method that provides the truncation numbers as well as the required digits of machine precision for the translation operator of the multilevel fast multipole algorithm (MLFMA). The proposed method is valid for all frequencies, whereas the previous studies on error control are valid only for high-frequency problems (i.e., electrically large translation distances). When combined with a multiple-precision implementation of MLFMA, the proposed method can be used to solve low-frequency problems that are problematic with a fixed-precision implementation. Numerical results in the form of optimal truncation numbers and machine precisions for a variety of box sizes and desired relative error thresholds are presented and compared with the methods or numerical surveys available in the literature.

Index Terms—multilevel fast multipole algorithm, diagonalization, low-frequency breakdown, multiple-precision arithmetic, error analysis.

I. INTRODUCTION

The multilevel fast multipole algorithm (MLFMA) [1] is able to achieve $O(N \log N)$ complexity for N unknowns, enabling the solution of extremely large electromagnetic problems. MLFMA's computational efficiency comes from the diagonalization of the Green's function that allows the computation of the interactions between basis and testing functions in a group by group manner. However, diagonalization requires an expansion of the Green's function in terms of spherical harmonics (corresponding to spherical Hankel functions) that leads to the well-known low-frequency breakdown of MLFMA [2]. Moreover, even at high-frequencies, the spherical Hankel functions exhibit numerical stability problems for large orders. Therefore, for a desired error threshold, the summation over the spherical harmonics must be truncated at an optimum point, which makes selecting the truncation number a critical part of MLFMA implementations.

Most of the existing work on error control of MLFMA focuses on selecting the optimal truncation number. In practice, the excess bandwidth formula [3]–[5] is widely used to select the optimum truncation number. However, the excess bandwidth formula is derived using the large argument approximation of the Hankel functions [6], which makes it valid only for large translation distances, i.e., high-frequency problems. For low-frequencies or electrically small translation distances, there are many different methods available in the literature

to treat the low-frequency breakdown [7]–[14]. Most methods either completely change the formulation or require the solver to be re-coded, while introducing increased complexity. Recently, a simple and elegant alternative to the treatment of the low-frequency problem is proposed and demonstrated in [15], where a multiple-precision-arithmetic (MPA) based implementation is shown to handle overflowing summations. Authors in [15] determined the required truncation numbers and machine precisions through extensive simulations since the excess bandwidth formula fails to give meaningful results for electrically small translation distances.

This study presents an error control method for a MPA-based MLFMA that is valid for all frequencies. The proposed method can be used to determine the required truncation numbers as well as the required machine precisions given the translation distance and the desired relative error threshold. Throughout this paper, $e^{-i\omega t}$ time convention, where $\omega = 2\pi f$ and f is the operating frequency, is assumed and suppressed.

II. FORMULATION

The Gegenbauer's addition theorem for the free-space Green's function is given as

$$\frac{\exp(ik|\vec{w} + \vec{v}|)}{4\pi|\vec{w} + \vec{v}|} = \frac{ik}{4\pi} \sum_{t=0}^{\infty} (-1)^t (2t+1) j_t(kv) h_t^{(1)}(kw) P_t(\hat{w} \cdot \hat{v}), \quad (1)$$

where j_t and $h_t^{(1)}$ are the spherical Bessel and Hankel functions of the first kind, respectively. P_t is the Legendre polynomial, while $w = |\vec{w}|$ and $v = |\vec{v}|$ are the translation and shift vectors, respectively. When a truncation number of τ is used in (1), the relative error with respect to the free-space Green's function can be found as

$$\hat{\epsilon} = kR \left| \sum_{t=\tau+1}^{\infty} (-1)^t (2t+1) j_t(kv) h_t^{(1)}(kw) P_t(\hat{w} \cdot \hat{v}) \right|, \quad (2)$$

where $R = |\vec{w} + \vec{v}|$. Assuming that the leading term of (2) dominates, the relative error can be approximated as

$$\hat{\epsilon} \approx kR(2\tau+3) \left| j_{\tau+1}(kv) h_{\tau+1}^{(1)}(kw) P_{\tau}(\hat{w} \cdot \hat{v}) \right|. \quad (3)$$

The large argument approximation of the Bessel and Hankel functions in (3) is used in the derivation of the excess bandwidth formula, which makes it only valid for electrically large translation distances. To have a valid approximation throughout all frequencies, we use the large order approximation given as [6]:

$$J_t(t \operatorname{sech} \gamma_j) \approx \frac{\exp[t(\tanh \gamma_j - \gamma_j)]}{\sqrt{2\pi t \tanh \gamma_j}}, \quad (4)$$

$$Y_t(t \operatorname{sech} \gamma_y) \approx -i \frac{\exp[t(\gamma_y - \tanh \gamma_y)]}{\sqrt{\frac{1}{2}\pi t \tanh \gamma_y}}. \quad (5)$$

We convert (4) and (5) into the spherical form and substitute into (3), which is then solved for the optimal truncation number for a given translation vector and desired relative error threshold.

After finding the optimum number of harmonics, the far-zone interaction between boxes are computed using the diagonal form of the Green's function as

$$\frac{\exp(ik|\vec{w} + \vec{v}_1 + \vec{v}_2|)}{4\pi|\vec{w} + \vec{v}_1 + \vec{v}_2|} = \frac{ik}{(4\pi)^2} \int d^2\hat{k} \beta(\vec{k}, \vec{v}_1) \alpha_\tau(\vec{k}, \vec{w}) \beta(\vec{k}, \vec{v}_2), \quad (6)$$

where the shift operator (β) is defined as

$$\beta(\vec{k}, \vec{v}) = \exp(i\vec{k} \cdot \vec{v}) \quad (7)$$

and the translation operator (α_τ) is defined as

$$\alpha_\tau(\vec{k}, \vec{w}) = \sum_{t=0}^{\tau} (i)^t (2t+1) h_t^{(1)}(kw) P_t(\hat{k} \cdot \hat{w}). \quad (8)$$

When computing (6) the machine precision of the computing platform must be able to handle all of the elementary functions and their summations, integrations, products, etc. To determine the required machine precision for a given truncation number, we numerically calculate each term at the highest possible order (i.e., the truncation number) and convert them to decimal digits via base-10 logarithm. Then we calculate the required digits for each combination of the terms of (6) in a similar fashion. Note that, to estimate the worst case in terms of the required machine precision, we assume all the terms within a summation or integration add up coherently, i.e., we substitute the summation in (8) with a multiplication by $(\tau+1)$ and the integration in (6) with a multiplication by $2(\tau+1)^2$ that is the number of integration points in the Gauss-Legendre sampling given in [16].

III. NUMERICAL RESULTS

To demonstrate and validate the proposed error control formulation, the required truncation numbers and machine precisions were calculated for several cases and are listed in Table I. Note that the translation vector is taken as one-box-buffer along the y-direction, i.e., $\vec{w} = [0 \ 2a \ 0]^T$, where a is the box edge length and superscript T is the vector transpose operator. Source and observation boxes in a one-box-buffer scheme are illustrated in Fig. 1. The diagonal form of the

TABLE I
TRUNCATION NUMBERS (τ) AND DIGITS OF MACHINE PRECISION (MP)
FOR SEVERAL DESIRED RELATIVE ERROR THRESHOLDS AND BOX
SIZES (a) WHEN $\vec{w} = [0 \ 2a \ 0]^T$

Box Size (a)	Error Threshold	τ	MP
8λ	10^{-5}	115	13
2λ	10^{-5}	69	32
$\lambda/32$	10^{-3}	34	73
$\lambda/2048$	10^{-2}	20	81

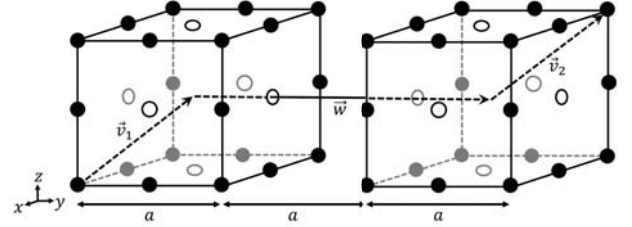


Fig. 1. Source and observation boxes in a one-box-buffer scheme.

Green's function in (6) were computed for each row of Table I for a range of truncation numbers and machine precisions around the optimum values. The results are provided in Figs. 2–5. Note that MPA is implemented in MATLAB using a commercially available toolbox [17].

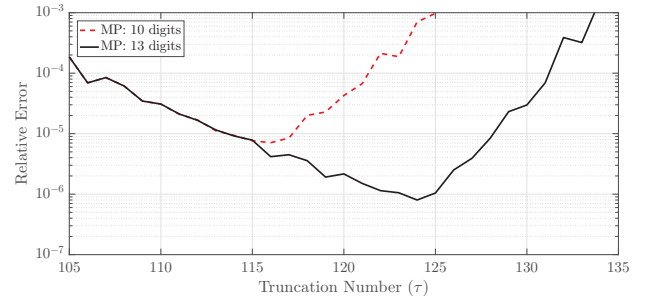


Fig. 2. Relative error of the diagonalized Green's function at different machine precisions for $a = 8\lambda$ and $\vec{w} = [0 \ 2a \ 0]^T$.

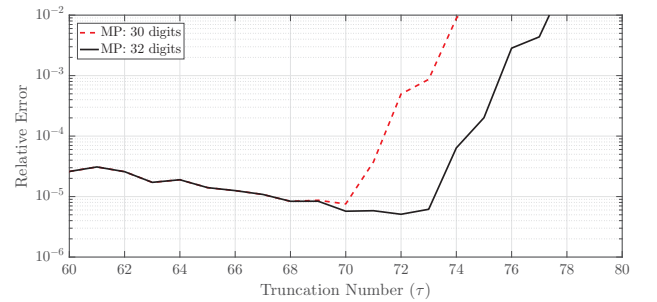


Fig. 3. Relative error of the diagonalized Green's function at different machine precisions for $a = 2\lambda$ and $\vec{w} = [0 \ 2a \ 0]^T$.

Note that in Figs. 2–5 the relative errors for smaller machine precisions are also given to illustrate the critical points. As clearly seen in Figs. 2–5, the truncation numbers provided by

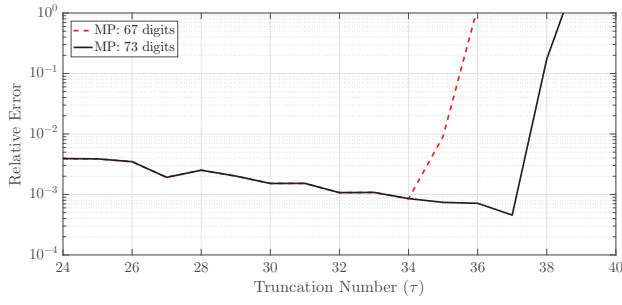


Fig. 4. Relative error of the diagonalized Green's function at different machine precisions for $a = \lambda/32$ and $\vec{w} = [0 \ 2a \ 0]^T$.

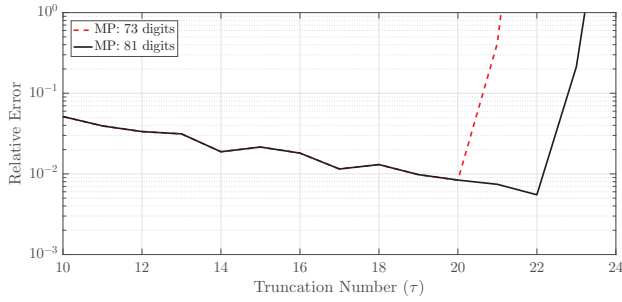


Fig. 5. Relative error of the diagonalized Green's function at different machine precisions for $a = \lambda/2048$ and $\vec{w} = [0 \ 2a \ 0]^T$.

the proposed method achieves the desired error thresholds, while the given machine precisions in Table I are large enough so that numerical stability problems do not occur at or close to the listed truncation points. Also note that the method is tested for electrically large and small boxes, which demonstrates its ability to work at all frequencies or for box sizes.

IV. CONCLUSION

A novel error control formulation for MPA implementation of MLFMA is presented. Unlike the previous works on the error control that use the large argument approximation of the Bessel and Hankel functions, the proposed formulation uses the large order approximation that makes it valid for all frequencies. The implementation of MPA also elegantly solves the well-known low-frequency breakdown of MLMFA.

Currently, MPA operations can easily be implemented with open-source or commercial libraries without changing any of the formulation. However, a software implementation of the higher precision arithmetic introduces a significant overhead in terms of processing-time, which must be addressed for a practical implementation.

REFERENCES

- [1] 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.
- [2] L. Greengard, J. Huang, V. Rokhlin, and S. Wadzura, "Accelerating fast multipole methods for the Helmholtz equation at low frequencies," *IEEE Comput. Sci. Eng.*, vol. 5, no. 3, pp. 32–38, Jul.–Sep. 1998.
- [3] S. Koc, J. M. Song, and W. C. Chew, "Error analysis for the numerical evaluation of the diagonal forms of the scalar spherical addition theorem," *SIAM Journal Numer. Anal.*, vol. 36, no. 3, pp. 906–921, 1999.
- [4] J. M. Song and W. C. Chew, "Error analysis for the truncation of multipole expansion of vector Green's functions," *IEEE Microwave and Wireless Components Lett.*, vol. 11, no. 7, pp. 311–313, 2001.
- [5] W. C. Chew, J.-M. Jin, E. Michielssen, and J. M. Song, "Fast and efficient algorithms in computational electromagnetics," Artech House, Boston, 2001.
- [6] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. New York: Dover, 1964.
- [7] J.-S. Zhao and W. C. Chew, "Three dimensional multilevel fast multipole algorithm from static to electrodynamic," *Microw. Opt. Technol. Lett.*, vol. 26, no. 1, pp. 43–48, Jul. 2000.
- [8] J.-S. Zhao and W. C. Chew, "Applying matrix rotation to the three-dimensional low-frequency multilevel fast multipole algorithm," *Microw. Opt. Technol. Lett.*, vol. 26, no. 2, pp. 105–110, Jul. 2000.
- [9] J.-S. Zhao and W. C. Chew, "Applying LF-MLFMA to solve complex PEC structures," *Microw. Opt. Technol. Lett.*, vol. 28, no. 3, pp. 155–160, Feb. 2001.
- [10] Y.-H. Chu and W. C. Chew, "A multilevel fast multipole algorithm electrically small composite structures," *Microw. Opt. Technol. Lett.*, vol. 43, no. 3, pp. 202–207, Nov. 2004.
- [11] Ö. Ergül and L. Gürel, "Efficient solutions of metamaterial problems using a low-frequency multilevel fast multipole algorithm," *Prog. Electromagn. Res.*, vol. 108, pp. 81–99, 2010.
- [12] L. J. Jiang and W. C. Chew, "Low-frequency fast inhomogeneous planewave algorithm (LF-FIPWA)," *Microw. Opt. Technol. Lett.*, vol. 40, no. 2, pp. 117–122, Jan. 2004.
- [13] I. Bogaert, J. Peeters, and F. Olyslager, "A nondirective plane wave MLFMA stable at low frequencies," *IEEE Trans. Antennas Propag.*, vol. 56, no. 12, pp. 3752–3767, Dec. 2008.
- [14] I. Bogaert and F. Olyslager, "A low frequency stable plane wave addition theorem," *J. Comput. Phys.*, vol. 228, no. 4, pp. 1000–1016, Mar. 2009.
- [15] Ö. Ergül and B. Karaosmanoğlu, "Low-frequency fast multipole method based on multiple-precision arithmetic," *IEEE Antennas Wireless Propag. Lett.*, vol. 13, pp. 975–978, 2014.
- [16] Ö. Ergül and L. Gürel, *The Multilevel Fast Multipole Algorithm for Solving Large-Scale Computational Electromagnetics Problems*. Wiley-IEEE Press, 2014.
- [17] "Multiprecision computing toolbox for MATLAB," Accessed Aug. 11, 2017 [Online]. Available: <http://www.advanpix.com>