The HKU Scholars Hub The University of Hong Kong 杳港大学学術庫


| Title | Truncation error analysis of multipole expansion |
| :---: | :--- |
| Author（s） | Ohnuki，S；Chew，WC |
| Citation | SIAM Journal On Scientific Computing，2003，v． 25 n．4，p．1293－ <br> 1306 |
| Issued Date | 2003 |
| URL | http：／／hdl．handle．net／10722／182739 |
| Rights | Creative Commons：Attribution 3.0 Hong Kong License |

# TRUNCATION ERROR ANALYSIS OF MULTIPOLE EXPANSION* 

SHINICHIRO OHNUKI ${ }^{\dagger}$ AND WENG CHO CHEW ${ }^{\dagger}$


#### Abstract

The multilevel fast multipole algorithm is based on the multipole expansion, which has numerical error sources such as truncation of the addition theorem, numerical integration, and interpolation/anterpolation. Of these, we focus on the truncation error and discuss its control precisely. The conventional selection rule fails when the buffer size is small compared to the desired numerical accuracy. We propose a new approach and show that the truncation error can be controlled and predicted regardless of the number of buffer sizes.


Key words. fast multipole method, multilevel fast multipole algorithm, error analysis, truncation error

AMS subject classifications. 65G50, 78A40

DOI. 10.1137/S1064827502412668

1. Introduction. Electromagnetic scattering problems with a large number of unknowns can be treated with the recently developed fast algorithms, such as the fast multipole method (FMM) [1, 2, 3] and the multilevel fast multipole algorithm (MLFMA) $[4,5,6]$. These methods are based on the multipole expansion, which has the following three error sources in its numerical implementation: the truncation of the addition theorem, numerical integration, and interpolation/anterpolation. Therefore, error control is one of the most important issues and has been discussed by many authors $[2,6,7,8,9,10,11]$. Of the three error sources, the error due to numerical integration and interpolation/anterpolation is the most easily handled, since all the signals treated in these algorithms are band-limited. Hence, there are ways to fully control these errors [10, 11]. The most difficult issue is that of controlling the truncation error and this still remains a complex problem [12]. Although these fast algorithms were shown to be error controllable in previous discussions, those authors assumed the condition that the buffer size is always large enough compared to the desired numerical accuracy. To satisfy this condition, we need to enlarge the buffer size for increasing the desired accuracy. This implementation becomes expensive, especially for high accuracy requirements, since the fields due to elements in the buffer boxes are computed by the traditional method of moments.

In this paper, we assume the condition that the number of buffer sizes is fixed regardless of the desired numerical accuracy. Under this condition, the conventional selection rule fails and the minimum error is no longer bounded by machine precision. To overcome this difficulty, we propose a new approach to predict and control the truncation error precisely. This approach also has a powerful potential to predict the error convergence process for a given interaction and the minimum error for a given box size and buffer size. This information reveals the error controllable region of the multipole expansion for small buffer cases.

[^0]2. Error sources. The 0th order of the Hankel function can be expressed in the Fourier space by using the integral representation of the Bessel function [11, 13]:
\[

$$
\begin{equation*}
H_{0}^{(1)}\left(k \rho_{j i}\right) \doteq \frac{1}{2 \pi} \int_{0}^{2 \pi} d \alpha \tilde{\beta}_{j l^{\prime}}(\alpha) \tilde{\alpha}_{l^{\prime} l}(\alpha) \tilde{\beta}_{l i}(\alpha) \tag{1}
\end{equation*}
$$

\]

where

$$
\begin{equation*}
\tilde{\alpha}_{l^{\prime} l}(\alpha)=\sum_{p=-P}^{P} H_{p}^{(1)}\left(k \rho_{l^{\prime} l}\right) e^{-i p\left(\phi_{l^{\prime} l}-\alpha+\pi / 2\right)} \tag{2}
\end{equation*}
$$

and

$$
\begin{equation*}
\tilde{\beta}_{j l^{\prime}}(\alpha)=e^{-i k \rho_{j l^{\prime}} \cos \left(\alpha-\phi_{j l^{\prime}}\right)}, \quad \tilde{\beta}_{l i}(\alpha)=e^{-i k \rho_{l i} \cos \left(\alpha-\phi_{l i}\right)} \tag{3}
\end{equation*}
$$

It can be expressed by its discretized form [11]

$$
\begin{equation*}
H_{0}^{(1)}\left(k \rho_{j i}\right) \doteq \frac{1}{Q} \sum_{q=1}^{Q} \tilde{\beta}_{j l^{\prime}}\left(\alpha_{q}\right) \tilde{\alpha}_{l^{\prime} l}\left(\alpha_{q}\right) \tilde{\beta}_{l i}\left(\alpha_{q}\right) \tag{4}
\end{equation*}
$$

Equation (4) can be written as products of a matrix and vectors as

$$
\begin{equation*}
H_{0}^{(1)}\left(k \rho_{j i}\right) \doteq \hat{\boldsymbol{\beta}}_{j l^{\prime}}^{t} \cdot \hat{\overline{\boldsymbol{\alpha}}}_{l^{\prime} l} \cdot \hat{\boldsymbol{\beta}}_{l i} . \tag{5}
\end{equation*}
$$

For the multilevel case, the sampling rate should be changed in the process of aggregation and disaggregation. Considering this process, (5) can be replaced with a new factorization using the interpolation matrices $\bar{I}_{m}$ and the anterpolation matrices $\bar{I}_{m}^{t}$ $(m=1,2, \ldots, n)[11]$,

$$
\begin{align*}
H_{0}^{(1)}\left(k \rho_{j i}\right) \doteq \hat{\boldsymbol{\beta}}_{j J_{1}}^{t} & \cdot \overline{\boldsymbol{I}}_{1}^{t} \cdot \hat{\overline{\boldsymbol{\beta}}}_{J_{1} J_{2}} \cdot \overline{\boldsymbol{I}}_{2}^{t}  \tag{6}\\
& \cdots \hat{\overline{\boldsymbol{\beta}}}_{J_{n-1} J_{n}} \cdot \overline{\boldsymbol{I}}_{n}^{t} \cdot \hat{\overline{\boldsymbol{\beta}}}_{J_{n} l^{\prime}} \cdot \hat{\overline{\boldsymbol{\alpha}}}_{l^{\prime} l} \cdot \hat{\overline{\boldsymbol{\beta}}}_{l I_{n}} \cdot \overline{\boldsymbol{I}}_{n} \cdot \hat{\overline{\boldsymbol{\beta}}}_{I_{n} I_{n-1}} \\
& \cdots \overline{\boldsymbol{I}}_{2} \cdot \hat{\boldsymbol{\beta}}_{I_{2} I_{1}} \cdot \overline{\boldsymbol{I}}_{1} \cdot \hat{\boldsymbol{\beta}}_{I_{1} i} .
\end{align*}
$$

In this numerical implementation, we need to truncate the translation operator $\tilde{\alpha}_{l^{\prime} l}$ and perform numerical integration and interpolation/anterpolation. Of the existing error sources, the numerical integration error and the interpolation/anterpolation error can be fully controlled $[10,11,12]$. However, the choice of the truncation number $P$ in (2) still remains a complex issue. We will focus our discussion on the error which comes from the truncation process of the discretized form in (4).
3. Selection rules. The truncation number $P$ has been selected by either the semi-empirical formula $[2,6,7,8,9]$ or the excess bandwidth formula $[11,12,14,15$, $16,17]$. Of these two selection rules, the excess bandwidth formula can be applied for a wider range of group sizes [11]. We will explain the excess bandwidth formula briefly and propose a new approach to predict and control the error.
3.1. Excess bandwidth formula. One conventional way to predict and control truncation error for the multipole expansion is to use the excess bandwidth formula $[11,12,14,15,16,17]$. Chew et al. [11] and Song and Chew [17] recently proposed the following refined formula:

$$
\begin{equation*}
P \simeq k D+1.8 d_{0}^{2 / 3}(k D)^{1 / 3} \tag{7}
\end{equation*}
$$

where $k D$ is the maximum diameter of a group size, and $d_{0}$ is the desired number of digits of accuracy. If we assume a square box as a group, the above can be rewritten, using the box size $k a$, as

$$
\begin{align*}
P & \simeq \sqrt{2} k a+1.8 d_{0}^{2 / 3}(\sqrt{2} k a)^{1 / 3}  \tag{8}\\
& \simeq 1.4 k a+2.0 d_{0}^{2 / 3}(k a)^{1 / 3} . \tag{9}
\end{align*}
$$

This equation is transformed into

$$
\begin{equation*}
d_{0} \simeq\left[\frac{P-1.4 k a}{2.0(k a)^{1 / 3}}\right]^{3 / 2} . \tag{10}
\end{equation*}
$$

Equation (10) implies that the number of digits of accuracy is determined by the truncation number $P$. This expression represents the convergence rate of the integrand in (1) with $P$. We perform a numerical experiment with finite double precision-the value $d_{0}$ is bounded by this accuracy.
3.2. New approach. If the truncation number $P$ becomes larger and larger, the integral representation of Hankel function (4) tends to diverge. This is caused by the Hankel functions involved in the translation operator (2). Therefore, another evaluation is needed to predict and control the error in this case.

When $p \rightarrow \infty$ and $k \rho_{l^{\prime} l} \sim O(p)$, the Hankel function has an approximate form [18]

$$
\begin{equation*}
H_{p}^{(1)}\left(k \rho_{l^{\prime} l}\right) \sim-i \frac{e^{p \ln \left\{\left[p+\sqrt{p^{2}-\left(k \rho_{l^{\prime}}\right)^{2}}\right] / k \rho_{l^{\prime}}\right\}-\sqrt{p^{2}-\left(k \rho_{l^{\prime}}\right)^{2}}}}{\sqrt{\frac{1}{2} \pi \sqrt{p^{2}-\left(k \rho_{l^{\prime} l}\right)^{2}}}} . \tag{11}
\end{equation*}
$$

The above diverges rapidly for $p>k \rho_{l^{\prime} l}$, and the following formula can be derived:

$$
\begin{align*}
d_{1} & \simeq\left[\frac{P-k \rho_{l^{\prime} l}}{1.8\left(k \rho_{l^{\prime} l}\right)^{1 / 3}}\right]^{3 / 2},  \tag{12}\\
& =\left\{\frac{P-(n+1) k a}{1.8[(n+1) k a]^{1 / 3}}\right\}^{3 / 2}, \tag{13}
\end{align*}
$$

where $n$ is the number of buffer boxes. Here, $d_{1}$ is related to the value of the Hankel function $\left[O\left(10^{d_{1}}\right)\right]$ for $p=P$. When the integral representation of the Hankel function in (4) is evaluated, the number of digits is lost by approximately $d_{1}$ in (13) because of cancellation and roundoff. Therefore, the value $d_{0}-d_{1}$ can be considered as the true number of digits of accuracy. This is what we call the new approach. This approach considers both the convergence rate of the Bessel functions and the divergence rate of the Hankel functions with the truncation number $P$.
4. Error control and prediction. To make a comparison of the truncation error for a 10 -box buffer case and a 2 -box buffer case, error control and prediction for the multipole expansion will be discussed. In this section, we clarify the topic of applying the excess bandwidth formula and the new approach to the selection of the truncation number $P$.
4.1. 10-box buffer case. The absolute relative error between the original 0th order of the Hankel function and its discretized form (4) is investigated for the elements shown in Figure 1. The box size is assumed to be $k a=5$ and the distance between


Fig. 1. Distribution of elements to study the relative error.


Fig. 2. Relative error due to the truncation. A 10-box buffer is assumed between two boxes.
two box centers $(n+1) k a=11 k a$. The dots in Figure 2 show the computational result and the solid line shows the prediction. We can confirm that both results are in good agreement. This error for changing the truncation number $P$ is classified into the following four regions, and the error in each region is controlled and predicted in the following ways:
(a-1) Region I $[P<1.4 k a]$.
Control: The truncation number should be increased.
Prediction: The value $10^{0}$ is plotted.
(a-2) Region II $\left[1.4 k a<P<1.4 k a+12.3(k a)^{1 / 3}\right]$.
Control: The error can be controlled by the excess bandwidth formula in (7). The number of digits of accuracy $d_{0}$ varies from $d_{0}=0$ to 15 .

Prediction: Using the transformed excess bandwidth formula in (10), the value $10^{-d_{0}}$ is plotted for changing the truncation number $P$.


FIG. 3. Relative error due to the truncation. A 2-box buffer is assumed between two boxes.
(a-3) Region III $\left[1.4 k a+12.3(k a)^{1 / 3}<P<(n+1) k a\right]$.
Control: Considering roundoff error, any number in this region can be chosen for the completely converged value. The absolute relative error is bounded by approximately $10^{-15}$ because of finite double precision accuracy. Prediction: The error is predicted as $10^{-15}$.
(a-4) Region IV $[(n+1) k a<P]$.
Control: $\quad$ The error becomes larger and tends to diverge for large $P$. The truncation number in this region cannot be selected.
Prediction: Since the Hankel function tends to diverge in this region, the value $d_{0}-d_{1}$ can be considered as the true number of digits of accuracy using the new approach. The value $10^{-\left(15-d_{1}\right)}$ is plotted.
The above classification shows that the error can be controlled by using the excess bandwidth formula when the truncation number $P$ is selected as $1.4 k a<P<1.4 k a+$ $12.3(k a)^{1 / 3}$. If we select $d_{1}=0$ in this region, the new approach gives the same result as the excess bandwidth formula. When $P>(n+1) k a$, the new approach can be applicable to error prediction.
4.2. 2-box buffer case. The absolute relative error for the elements in Figure 1 is investigated. The number of buffer boxes is reduced from $n=10$ to $n=2$. The box size is $k a=5$ and the distance between two box centers is $(n+1) k a=3 k a$. Figure 3 is a plot of the absolute relative error for these parameters. The dots show the computational results and the solid line shows the prediction. Both results are in good agreement. This result for the 2-box buffer case also is classified into the following four regions. However, the way to control and predict error is different from the 10 -box buffer case:
(b-1) Region I $[P<1.4 k a]$.
Control: The truncation number should be increased.
Prediction: The value $10^{0}$ is plotted.
(b-2) Region $\mathrm{II}^{\prime}$ [1.4ka $\left.<P<(n+1) k a\right]$.
Control: The error can be controlled by the excess bandwidth formula in (7). The selection of $d_{0}$ varies from 0 to $\left\{[(n-0.4) k a] /\left[2.0(k a)^{1 / 3}\right]\right\}^{3 / 2}$, which is less than 15.
Prediction: The value $10^{-d_{0}}$ is plotted for changing the truncation number $P$.
(b-3) Region III' $\left[(n+1) k a<P<1.4 k a+12.3(k a)^{1 / 3}\right]$.
Control: The truncation number is selected to enable the condition $d_{0}-$ $d_{1} \simeq \log (1 / e r r)$ to be satisfied. But the general solution does not exist, so the truncation number $P$ for the desired error level err is determined by the linear interpolation using $\left(P_{1}, \operatorname{err}_{1}\right)$ and $\left(P_{2}, \operatorname{err}_{2}\right)$,

$$
\begin{equation*}
P \simeq \frac{P_{2}-P_{1}}{\left[\log \left(1 / e r r_{2}\right)-\log \left(1 / e r r_{1}\right)\right]}\left[\log (1 / e r r)-\log \left(1 / e r r_{1}\right)\right]+P_{1} \tag{14}
\end{equation*}
$$

where $\operatorname{err}_{1}$ is the absolute relative error for the truncation number $P_{1} \simeq$ $(n+1) k a$ and can be found by the transformed excess bandwidth formula, and $e r r_{2}$ for $P_{2} \simeq 1.4 k a+12.3 k a$ can be found by the new approach.
Prediction: Using the new approach, the value $d_{0}-d_{1}$ can be considered as the true number of digits of accuracy in this region. Therefore, the error is predicted by $10^{-\left(d_{0}-d_{1}\right)}$, where $d_{0}$ and $d_{1}$ are given by (10) and (13), respectively.
(b-4) Region IV $\left[1.4 k a+12.3(k a)^{1 / 3}<P\right]$.
Control: $\quad$ The error becomes larger and tends to diverge for large $P$. The truncation number in this region cannot be selected.
Prediction: The value $10^{-\left(15-d_{1}\right)}$ is plotted, since $d_{0}$ is bounded by 15 for $P>1.4 k a+12.3(k a)^{1 / 3}$.
The above classification shows that the error can be controlled when the truncation number $P$ is chosen as $1.4 k a<P<1.4 k a+12.3(k a)^{1 / 3}$. To control the error, the truncation number is selected by the excess bandwidth formula for $1.4 k a<P<$ $(n+1) k a$ and by the new approach for $(n+1) k a<P<1.4 k a+12.3(k a)^{1 / 3}$.
4.3. Large buffer and small buffer. To classify the error property into four regions, we need to consider the following three parameters: the maximum box size $1.4 k a$, the value $1.4 k a+12.3(k a)^{1 / 3}$ for $d_{0}=15$ in the excess bandwidth formula, and the distance between two box centers $(n+1) k a$, where $n$ is the number of buffer boxes. The condition $1.4 k a<1.4 k a+12.3(k a)^{1 / 3}$ is always satisfied for $k a>0$. Also $1.4 k a<(n+1) k a$ is always true, since buffer regions are assumed to be at least one box $(n \geq 1)$. The uncertain condition is the difference between $1.4 k a+12.3(k a)^{1 / 3}$ and $(n+1) k a$. Considering this fact, we propose the following definitions, which depend on machine precision, large and small buffer sizes:
(c-1) Large buffer: $1.4 k a+12.3(k a)^{1 / 3}<(n+1) k a$.
The distance between two box centers $(n+1) k a$ is larger than the value $1.4 k a+$ $12.3(k a)^{1 / 3}$ for $d_{0}=15$ in the excess bandwidth formula. If this condition is satisfied, the number of digits of accuracy $d_{0}$ in (9) can be selected from 0 to 15 and the error is completely controlled by using the excess bandwidth formula.
(c-2) Small buffer: $(n+1) k a<1.4 k a+12.3(k a)^{1 / 3}$.


Fig. 4. Four types of distributions to study the relative error. The elements are placed at $i$ and $j$. The box centers are indicated by $l$ and $l^{\prime}$.

The value $1.4 k a+12.3(k a)^{1 / 3}$ is larger than the distance between two box centers $(n+1) k a$. The conventional excess bandwidth does not hold when the truncation number $P$ is larger than $(n+1) k a$. Therefore, the error should be predicted and controlled by using the new approach when $P>(n+1) k a$.
5. Distribution of elements. In this section, we discuss the numerical error of the multipole expansion for the four types of interactions shown in Figure 4. The elements are placed at $i$ and $j$, and the box centers are fixed at $l$ and $l^{\prime}$. The absolute relative error between the original 0th order of the Hankel function and its discretized form (4) is investigated for changing the truncation number $P$ when the box size is fixed as $k a=5$.

Figure 5 shows the absolute relative error of cases 1-4 in Figure 4 for changing the truncation number $P$. As an example of a large buffer case, the number of buffer sizes is selected to be $n=10$. The dots indicate the numerical results and the solid line shows the prediction by the excess bandwidth formula for $P<(n+1) k a$ and by the new approach for $P>(n+1) k a$. We can see that the error of case 1 is always


Fig. 5. Relative error due to the distribution of elements. Four types of distribution from Figure 4 are studied. A 10-box buffer is assumed between two boxes.
the largest and this value is almost the same as the predicted one. Therefore, the solid line can be considered as the error bound. If we select $d_{1}=0$ in (13), the new approach gives the same prediction.

Figure 6 shows the same numerical experiment for a small buffer case $(n=1)$. For this parameter, the condition $1.4 k a<(n+1) k a<1.4 k a+12.3(k a)^{1 / 3}$ is satisfied. The solid line shows the prediction using the new approach. It can be considered as the proper error bound for any truncation numbers. On the other hand, the dashed-dotted line from the excess bandwidth formula does not indicate the proper error bound for all the truncation numbers. It gives the correct error bound only for $1.4 k a<P<(n+1) k a$. When $(n+1) k a<P<P_{1}$, there is a possibility that the computational error becomes larger than the predicted one. For example, if the truncation number is selected as $P=20$ and we have case 1 or 2 as the real interaction, the error becomes larger than the prediction. For $P>P_{1}$, the computational error is always larger than the predicted one.
6. Error convergence process. Using the new approach, the convergence process of the absolute relative error for changing the truncation number $P$ can be predicted. The interactions shown in Figure 4 are considered. The box size is fixed as $k a=5$ and the buffer size is assumed to be one box.

The number of digits of accuracy given by the new approach is $d_{0}-d_{1}$. This expression can be treated by $d_{0}$ in (10) and $d_{1}$ in (13) separately. The $d_{0}$ part stands for the convergence rate of the Bessel functions in the multipole expansion. Therefore, this value is determined by the placement of elements, namely $i$ and $j$. For conventional use, the maximum box size is chosen to predict the error bound.


Fig. 6. Relative error due to the distribution of elements. Four types of distribution from Figure 4 are studied. A 1-box buffer is assumed between two boxes.

To predict the convergence rate for each interaction, we need to substitute $\delta k a$ for the box size $k a$ in (10), where $\delta$ is the interaction parameter with values $0 \leq \delta \leq 1$ to estimate all the possible pairs between two boxes. In Figure 7, dots show our numerical experiment and solid lines indicate the prediction by the new approach. Both results are in good agreement.

Next, we will discuss the same prediction for changing the distance between two box centers $(n+1) k a$. The placement of elements is fixed as the case $1(\delta=1)$ in Figure 4. For this selection, $k a$ in the $d_{0}$ part should be fixed, since the placement of elements is the same. We need only consider the $d_{1}$ part for changing buffer sizes. The $d_{1}$ part represents the divergence rate of the Hankel function of the translation operator in (2). This is determined by the distance between the two box centers $(n+1) k a$. If we predict the convergence process for changing the buffer sizes $n$, this value in $d_{1}$ should be changed. Figure 8 is a plot of the numerical experiment and the prediction. Both results are in good agreement.

Considering these results, the new approach is applicable for any truncation number $P$ when we use the modified $d_{0}$ and $d_{1}$ as follows:

$$
\begin{align*}
& d_{0} \simeq\left\{\begin{array}{cl}
0, & P<1.4 \delta k a, \\
{\left[\frac{P-1.4 \delta k a}{2.0(\delta k a)^{1 / 3}}\right]^{3 / 2},} & 1.4 \delta k a<P<1.4 \delta k a+12.3(\delta k a)^{1 / 3}, \\
15 & \text { otherwise } .
\end{array}\right.  \tag{15}\\
& d_{1} \simeq\left\{\begin{array}{cl}
0, & P<(n+1) k a, \\
{\left[\frac{P-(n+1) k a}{1.8[(n+1) k a]^{1 / 3}}\right]^{3 / 2}} & \text { otherwise } .
\end{array}\right.
\end{align*}
$$



Fig. 7. Prediction of the error convergence process. Four types of distribution from Figure 4 are studied. A 1-box buffer is assumed between two boxes.


Fig. 8. Prediction of the error convergence process. The number of buffer sizes is changed from 1-box to 4-box. The placement of elements is case 1 from Figure 4.
7. Minimum error. We need to consider the meaning of the minimum error carefully. If a buffer size is large enough, the error can be predicted by using the excess bandwidth formula. In this case, the minimum error is defined by machine precision. Since we assume finite double precision in our computation, the minimum error is bounded by this accuracy with some roundoff errors. This is the conventional definition of the minimum error for the multipole expansion.

However, the true minimum error is different from the above definition when the buffer size is small. It is determined by the relationship between the convergence rate $d_{0}$ and the divergence rate $d_{1}$. Since it is worse than machine precision, prediction of the minimum error is very important for knowing the limitation of the numerical accuracy. Using this information, it is possible to estimate the minimum error for a given box size and buffer size, or needed box sizes and buffer sizes, to control error under the desired accuracy. This is another advantage of using the new approach.

In Figure 8, we have seen that the minimum error for case $1(\delta=1)$ from Figure 4 is found to be

$$
\begin{equation*}
P_{\min } \simeq 1.4 k a+12.3(k a)^{1 / 3} \tag{17}
\end{equation*}
$$

The new approach gives the number of digits of accuracy for this choice as

$$
\begin{equation*}
d_{\min } \simeq 15-\left[\frac{P_{\min }-(n+1) k a}{1.8[(n+1) k a]^{1 / 3}}\right]^{3 / 2} \tag{18}
\end{equation*}
$$

To find the minimum errors for all the possible interactions between two boxes, $k a$ in (17) should be replaced with $\delta k a$, that is,

$$
\begin{equation*}
P_{\min } \simeq 1.4 \delta k a+12.3(\delta k a)^{1 / 3} \tag{19}
\end{equation*}
$$

where $\delta$ is the interaction parameter with values $0 \leq \delta \leq 1$.
Figure 9 is a plot of the computational result for the truncation number $P$ changing with $\delta$. The solid line is the prediction of the minimum error and the dots indicate the convergence process for cases $1-4$ of Figure 4 . We can see that the minimum value for each case is very close to the predicted line. When $P$ is equal to $1.4 k a+12.3(k a)^{1 / 3}$, the value gives the minimum error for the worst interaction $(\delta=1)$.

Figure 10 shows the prediction of the minimum error for changing buffer sizes from $n=1$ to $n=4$. When the buffer size is increased for the fixed box size, the minimum error becomes smaller. Finally, it converges to machine precision, which is the same as the prediction given by the excess bandwidth formula.
8. Error range. The new approach can predict the error convergence process and the minimum error for all the possible interactions between two boxes. Using this information, we can define the error range inside which the error of the multipole expansion is always found. When the box size $k a=5$ and the buffer size is 1-box, the error with the truncation number $P$ is found inside the hatched area in Figure 11. The error is bounded by the slowest convergence line for the worst interaction $(\delta=1)$ and the minimum error line. When the truncation number is selected as $P=1.4 k a+12.3(k a)^{1 / 3}$, these two lines meet at the intersection indicated by the dot. If $P$ is further increased, both lines merge and tend to diverge.


Fig. 9. Prediction of the minimum error for all possible pairs between two boxes. A 1-box buffer is assumed between two boxes.


Fig. 10. Prediction of the minimum error for changing buffer sizes.


Fig. 11. Error range of the multipole expansion. The interaction is assumed to be case 1 from Figure 4. A 1-box buffer is assumed between two boxes.
9. Conclusions. We have studied the truncation error of the multipole expansion. The conventional selection rule using the excess bandwidth formula becomes worse when the buffer size is small compared to the desired numerical accuracy. To overcome this difficulty, we have proposed a new approach for error prediction and control regardless of the buffer size. Using this approach, it is possible to predict the error convergence process for a given interaction and the minimum error for a given box size and buffer size.

## REFERENCES

[1] V. Rokhlin, Rapid solution of integral equation of scattering theory in two dimensions, J. Comput. Phys., 86 (1990), pp. 414-439.
[2] R. Coifman, V. Rokhlin, and S. Wandzura, The fast multipole method for the wave equation: A pedestrian prescription, IEEE Antennas Propagation Mag., 35 (1993), pp. 7-12.
[3] C. C. Lu and W. C. Chew, Fast algorithm for solving hybrid integral equation, Proc. IEE-H, 140 (1993), pp. 455-460.
[4] C. C. Lu and W. C. Chew, A multilevel algorithm for solving a boundary integral equation of wave scattering, Micro. Opt. Tech. Lett., 7 (1994), pp. 466-470.
[5] J. M. Song and W. C. Chew, Multilevel fast-multipole algorithm for solving combined field integral equations of electromagnetic scattering, Micro. Opt. Tech. Lett., 10 (1995), pp. 14-19.
[6] J. M. Song, C.-C. Lu, and W. C. Chew, Multilevel fast multipole algorithm for electromagnetic scattering by large complex objects, IEEE Trans. Antennas and Propagation, 45 (1997), pp. 1488-1493.
[7] J. M. Song and W. C. Chew, Fast multipole method solution using parametric geometry, Micro. Opt. Tech. Lett., 7 (1994), pp. 760-765.
[8] M. F. Gyure and M. A. Stalzer, A prescription for the multilevel Helmholtz FMM, IEEE Comput. Sci. Engrg., 5 (1998), pp. 39-47.
[9] B. Dembart and E. Yip, The accuracy of fast multipole methods for Maxwell's equations, IEEE Comput. Sci. Engrg., 5 (1998), pp. 48-56.
[10] 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., 36 (1999), pp. 906921.
[11] W. C. Chew, J.-M. Jin, E. Michielssen, and J. M. Song, eds., Fast and Efficient Algorithms in Computational Electromagnetics, Artech House, Boston, London, 2001.
[12] S. Ohnuki and W. C. Chew, Numerical Accuracy of Multipole Expansion for 2-D MLFMA, Research report CCEM-5-02, Center for Computational Electromagnetics, University of Illinois, Urbana, IL, 2002.
[13] W. C. Chew, Waves and Fields in Inhomogeneous Media, IEEE/OUP Series on Electromagnetic Wave Theory, Oxford University Press, Oxford, UK, IEEE Press, New York, 1994.
[14] W. J. Wiscombe, Improved Mie scattering algorithms, Applied Optics, 19 (1980), pp. 15051509.
[15] O. M. Bucci and G. Franceschetti, On the spatial bandwidth of scattered fields, IEEE Trans. Antennas and Propagation, 35 (1987), pp. 1445-1455.
[16] V. Rokhlin, Sparse Diagonal Forms for Translation Operations for the Helmholtz Equation in Two Dimensions, Research report YALEU/DCS/RR-1095, Department of Computer Science, Yale University, New Haven, CT, 1995.
[17] 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., 11 (2001), pp. 311-313.
[18] M. Abramowitz and I. A. Stegun, Handbook of Mathematical Functions, Dover Publications, New York, 1972.


[^0]:    *Received by the editors August 9, 2002; accepted for publication (in revised form) June 18, 2003; published electronically December 19, 2003. This work was supported by AFOSR under an MURI grant F49620-96-1-0025, a contract from WPAFB via SAIC, and the Kajima Foundation's assistance for research abroad. Part of this work was presented at the IEEE Antennas and Propagation Society International Symposium, San Antonio, TX, June 16-21, 2002.
    http://www.siam.org/journals/sisc/25-4/41266.html
    ${ }^{\dagger}$ Center for Computational Electromagnetics and Electromagnetics Laboratory, Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 618012991 (ohnuki@sunchew.ece.uiuc. edu, w-chew@uiuc.edu).

