SIAM J. Sci. Comput.
Vol. 26, No. 6, pp. 2047–2065

ERROR MINIMIZATION OF MULTIPOLE EXPANSION*

SHINICHIRO OHNUKI† AND WENG CHO CHEW‡

Abstract. In this paper, we focus on the truncation error of the multipole expansion for the fast multipole method and the multilevel fast multipole algorithm. When the buffer size is large enough, the error can be controlled and minimized by using the conventional selection rules. On the other hand, if the buffer size is small, the conventional selection rules no longer hold, and the new approach which we have recently proposed is needed. However, this method is still not sufficient to minimize the error for small buffer cases. We clarify this fact and show that the information about the placement of true worst-case interaction is needed. A novel algorithm to minimize the truncation error is proposed.

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

AMS subject classifications. 78A40, 65G50

DOI. 10.1137/S1064827502417970

1. Introduction.

Electromagnetic scattering problems with a large number of unknowns can be treated by the recent development of 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 some error sources in numerical implementation. Although these errors are shown in the previous discussions to be fully controlled, we have pointed out that the conventional selection rules to control the truncation error do not hold under some conditions [7, 8].

In this paper, we will consider this issue in detail and investigate the condition to minimize the truncation error of multipole expansion. When the buffer size is large enough compared to the machine precision, the error can be minimized by using the conventional selection rules. Namely, the total error from all the interaction pairs is minimized if the error from the assumed worst-case interaction is minimized. When the buffer size is small compared to the machine precision, it becomes a totally different issue. To estimate the error bound, the truncation number is usually selected for the assumed worst-case interaction in the nearest box pair. Although it is possible to minimize the error for this interaction by using the new approach [7, 8], the total error usually cannot be minimized. We will discuss this topic carefully and show that the information about the placement of the true worst-case interaction is also needed. Considering this fact, a novel algorithm to minimize the error will be proposed.

*Received by the editors November 15, 2002; accepted for publication (in revised form) November 10, 2004; published electronically July 13, 2005. This work was supported by AFOSR under MURI grant F49620-96-1-0025, by a contract from WPAFB via SAIC, and by the Kajima Foundation's assistance for research abroad.
http://www.siam.org/journals/sisc/26-6/41797.html

†Department of Electrical Engineering, College of Science and Technology, Nihon University, Tokyo 101-8308, Japan (ohnuki@ele.cst.nihon-u.ac.jp).

‡Department of Electrical and Computer Engineering, Center for Computational Electromagnetics and Electromagnetics Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801-2991 (w-chew@uiuc.edu).

2. Error control method. The 0th order Hankel function can be expressed in the Fourier space by using the integral representation of the Bessel function [9, 10]:

$$H_0^{(1)}(k\rho_{ji}) \doteq \frac{1}{2\pi}\int_0^{2\pi} d\alpha \tilde{\beta}_{jl'}(\alpha)\tilde{\alpha}_{l'l}(\alpha)\tilde{\beta}_{li}(\alpha), \tag{1}$$

where

$$\tilde{\alpha}_{l'l}(\alpha) = \sum_{p=-P}^{P} H_p^{(1)}(k\rho_{l'l})e^{-ip(\phi_{l'l}-\alpha+\frac{\pi}{2})} \tag{2}$$

and

$$\tilde{\beta}_{jl'}(\alpha) = e^{-ik\rho_{jl'}\cos(\alpha-\phi_{jl'})}, \qquad \tilde{\beta}_{li}(\alpha) = e^{-ik\rho_{li}\cos(\alpha-\phi_{li})}. \tag{3}$$

This two-level expression can be extended to the multilevel one [10]. Among numerical error sources due to this expression, we focus on the error which comes from the truncation process in (2).

One conventional way to select the truncation number is to use the excess bandwidth formula [10, 11, 12, 13, 14]. Song and Chew recently proposed the following refined formula [10, 14]:

$$P \simeq kD + 1.8 d_0^{2/3}(kD)^{1/3}, \tag{4}$$

where kD is the maximum diameter of a group size to assume the worst-case interaction, 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 ka as

$$P \simeq \sqrt{2}ka + 1.8 d_0^{2/3}(\sqrt{2}ka)^{1/3}. \tag{5}$$

Considering the property of the Bessel functions and finite double precision accuracy, this equation can be transformed into

$$d_0 \simeq \begin{cases} 0, & P < \sqrt{2}ka, \\ \left[\dfrac{P-\sqrt{2}ka}{2.0(ka)^{1/3}}\right]^{3/2}, & \sqrt{2}ka < P < \sqrt{2}ka + 12.3(ka)^{1/3}, \\ 15, & \sqrt{2}ka + 12.3(ka)^{1/3} < P. \end{cases} \tag{6}$$

When the buffer size is large enough compared to the machine precision, the refined excess bandwidth formula gives the proper truncation number.

If the buffer size is small compared to the machine precision, we also need to evaluate the divergence rate of the Hankel functions involved in the translation operator (2). Considering the property of the Hankel functions, it can be estimated by [7, 8]

$$d_1 \simeq \begin{cases} 0, & P < (n+1)ka, \\ \left[\dfrac{P-(n+1)ka}{1.8[(n+1)ka]^{1/3}}\right]^{3/2}, & (n+1)ka < P, \end{cases} \tag{7}$$

where n is the number of buffer boxes. Here, d_1 is related to the value of the Hankel function $[O(10^{+d_1})]$ for $p = P$. Therefore, the true number of digits of accuracy is

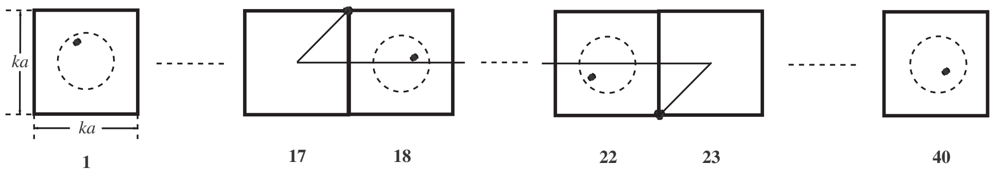


FIG. 1. *Forty contiguous boxes to study the absolute relative error. The enforced worst-case interaction is placed between box* 17 *and box* 23*. The other elements are randomly distributed inside dotted circles.*

given by $d_0 - d_1$ for small buffer cases. This is what we call the new approach. If $d_0 - d_1$ is plotted for changing the truncation number P, the error convergence process can be predicted.

Conventionally, we assume that the worst-case interaction exists in the nearest box pair to estimate the error bound. Therefore, the maximum diameter $\sqrt{2}ka$ is selected in (5) and (6). This selection relates to predicting the slowest convergence process. Later, we will show that it is also important to estimate the convergence process for some specific interactions which are different from the worst case. To consider any interactions in a box pair, it is convenient to replace $\sqrt{2}ka$ for the maximum diameter by $\sqrt{2}\delta ka$, where δ is the interaction parameter which depends on distribution of elements ($0 \leq \delta \leq 1$).

Considering this parameter, the minimum error for any interactions is found when the truncation number is selected for $d_0 = 15$ in (5),

$$P_{\min} \simeq \sqrt{2}\delta ka + 12.3(\delta ka)^{1/3}. \tag{8}$$

For this choice, the new approach predicts the number of digits of accuracy,

$$d_{\min} \simeq 15 - \left[\frac{P_{\min} - (n+1)ka}{1.8[(n+1)ka]^{1/3}}\right]^{3/2}. \tag{9}$$

To find the minimum error for all the possible interactions, δ in (8) varies from 0 to 1. This numerical result is called the minimum error line.

3. Error minimized state. In this section, we will clarify the general property of the error convergence process for changing the truncation number P and investigate the condition to minimize the truncation error. The absolute relative error between the original 0th order of Hankel function and its integral representation form in (1) is studied when there is a row of 40 contiguous boxes. For simple discussion, each box has only one element which is randomly distributed inside a circle whose radius $kr = 0.4ka$ by way of example. Here, ka is the box size and is fixed as $ka = 5$.

3.1. Large buffer case. The buffer size to apply the multipole expansion is assumed to be 5 boxes large. For this parameter, the condition $\sqrt{2}ka < \sqrt{2}ka + 12.3(ka)^{1/3} < (n+1)ka$ is satisfied. Here, $\sqrt{2}ka$ is the maximum box size, $\sqrt{2}ka + 12.3(ka)^{1/3}$ is for $d_0 = 15$ in the excess bandwidth formula, and $(n+1)ka$ is the distance between two box centers determined by the number of buffer sizes n. The elements are randomly distributed in all boxes except for boxes 17 and 23. We enforce the worst-case interaction ($\delta = 1$) for this pair, as shown in Figure 1. The slowest convergence line for this pair and the minimum error line for a 5-box buffer case,

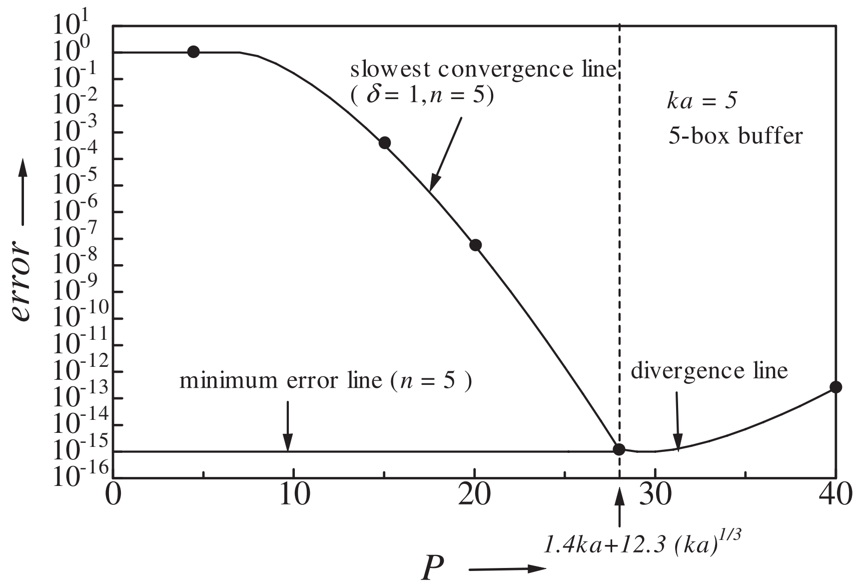


FIG. 2. *The slowest convergence line and the minimum error line for a* 5*-box buffer.*

as shown in Figure 2, are predicted by using the new approach. Considering this information, we will investigate the error distribution for all the possible pairs.

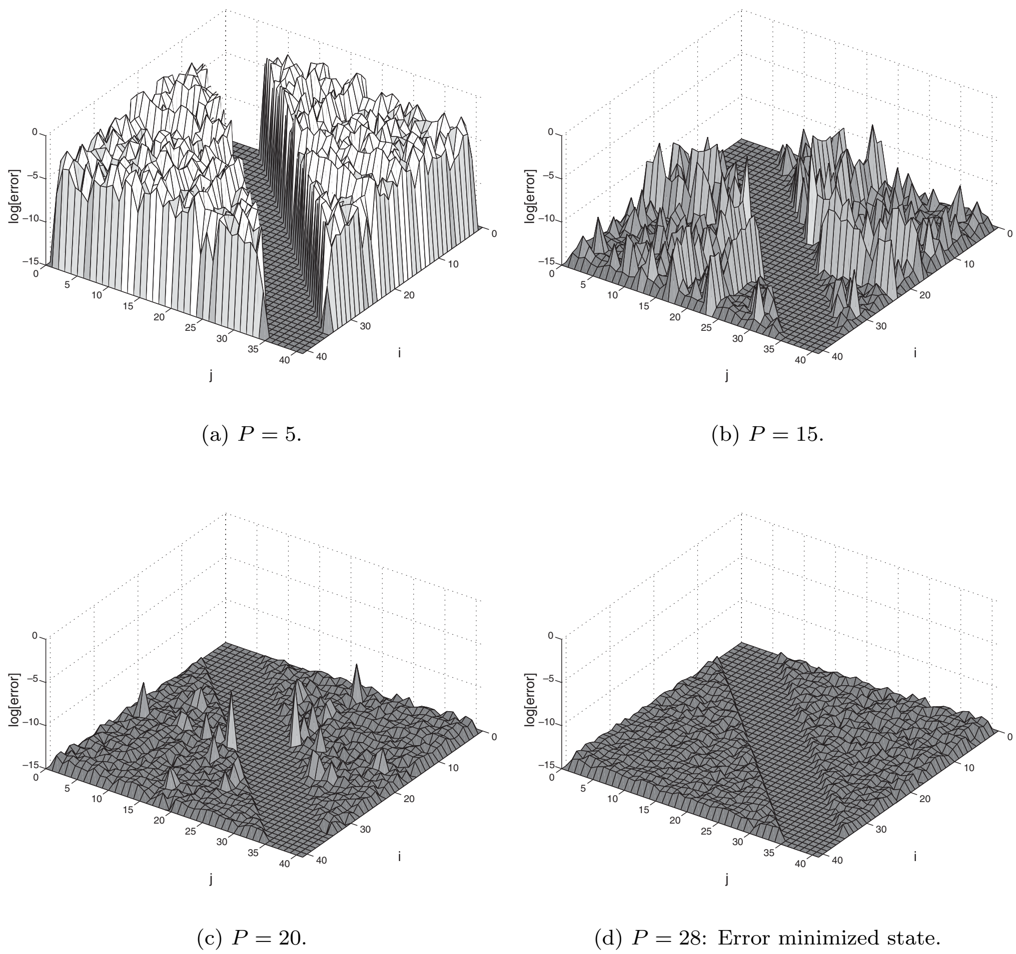
Figure 3 visualizes the error distribution for the four selections of P. The x-axis shows the box number which varies from $i = 1$ to 40, and the y-axis shows the other box number from $j = 1$ to 40 which makes a pair of the box i. The z-axis shows the absolute relative error of the multipole expansion. We can see the following properties for changing P:

(a-1) $P = 5$.
All the errors except those in the diagonal region are almost the same level and large. The multipole expansion is not applied to interactions in the diagonal region which are treated by the traditional method of moments. Therefore, the errors are always zero.

(a-2) $P = 15$.
To increase the truncation number from $P = 5$ to 15, all the errors become smaller. The dominant error comes from the worst-case interaction between boxes 17 and 23.

(a-3) $P = 20$.
The errors for almost all the interactions converge. These values are a little bit worse than the machine precision because of the roundoff error. For the worst-case interaction and some related pairs, the errors are still recognized.

(a-4) $P = 28$: Error minimized state.
Considering the roundoff error, the errors for all the interactions completely converge. We define this condition as the error minimized state when the buffer size is large. The truncation number for this state is selected by $P \simeq \sqrt{2}ka + 12.3(ka)^{1/3}$ using the excess bandwidth formula when $d_0 = 15$.



(a) $P = 5$.
(b) $P = 15$.
(c) $P = 20$.
(d) $P = 28$: Error minimized state.

FIG. 3. *Error-convergence process for a large buffer case.*

3.2. Small buffer case. We will perform the same numerical experiment for small buffer cases. Compared with the large buffer case, the difference of the error minimized state will be clarified. The buffer size to apply the multipole expansion is reduced from five boxes to one box. For this parameter, the condition $\sqrt{2}ka < (n+1)ka < \sqrt{2}ka + 12.3(ka)^{1/3}$ is satisfied. Here, we consider two types of the worst-case interaction. One is in a nearest box pair, and the other is not in a nearest box pair. The error minimized state is different under these two conditions.

3.2.1. Worst-case interaction in a nearest box pair. The elements are randomly distributed in all the boxes except for boxes 19 and 21. We enforce the worst-case interaction ($\delta = 1$) for this box pair, as shown in Figure 4. The slowest convergence line for the worst-case interaction and the minimum error line for a 1-box buffer case are predicted by the new approach in Figure 5. Considering this information, we will investigate the distribution of errors in the contiguous 40-box case. The five truncation numbers indicated by dots are selected, and the distribu-

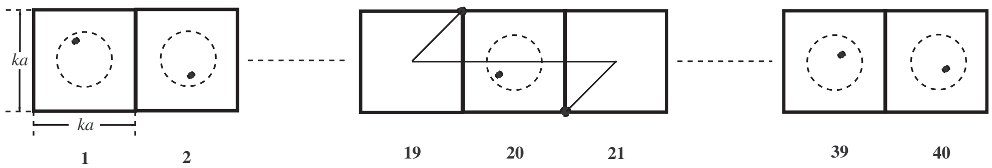


FIG. 4. *Forty contiguous boxes to study the absolute relative error. The enforced worst-case interaction is placed between box* 19 *and box* 21*. The other elements are randomly distributed inside dotted circles.*

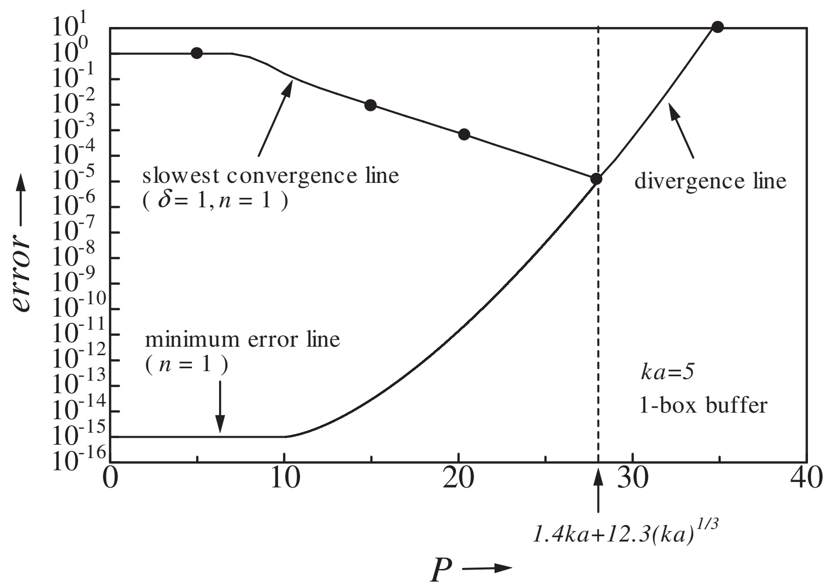


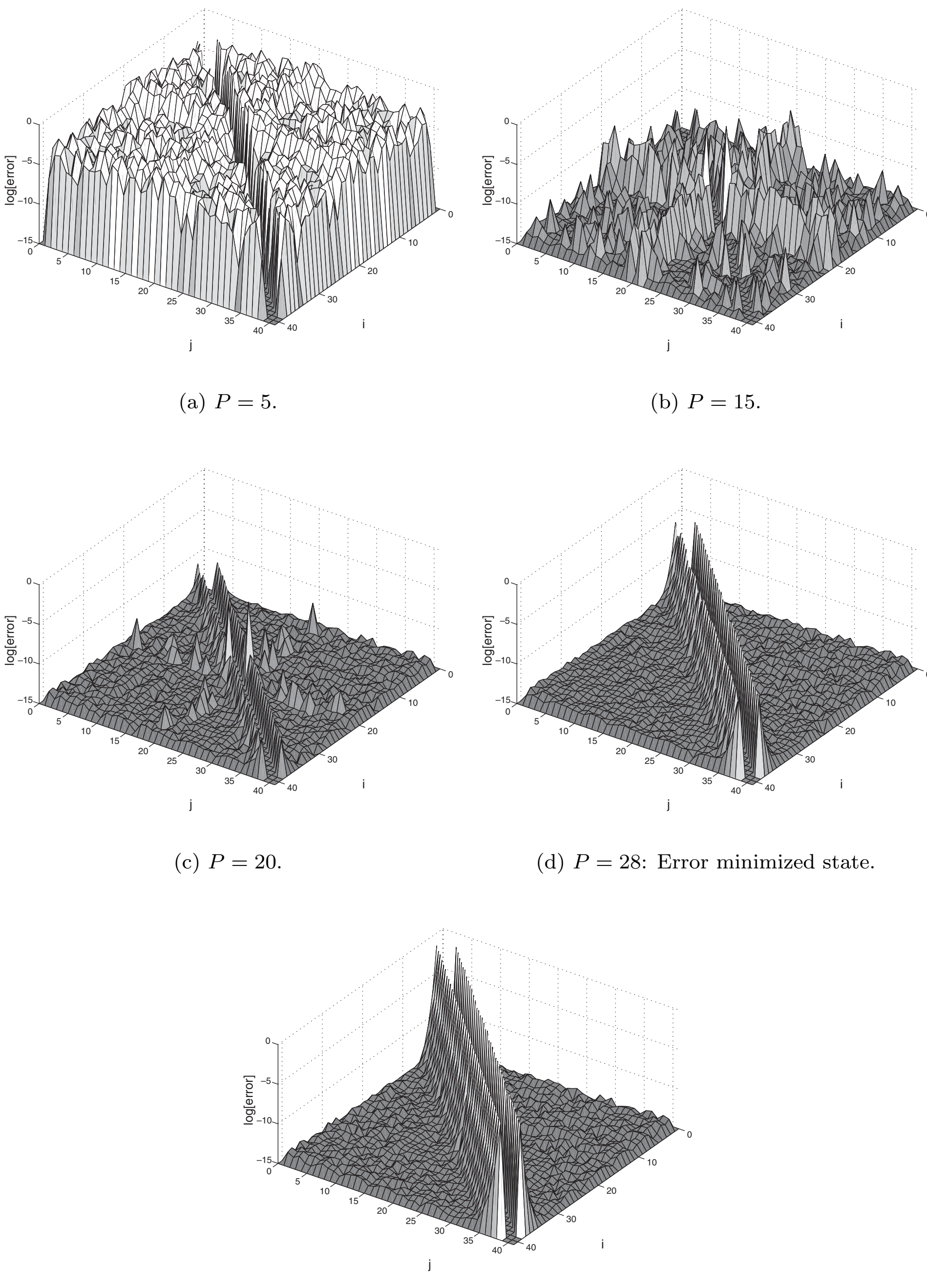
FIG. 5. *The slowest convergence line and the minimum error line for a* 1*-box buffer.*

tion of errors is illustrated in Figure 6. We can see the following error-convergence process:

(b-1) $P = 5$.
This error distribution is almost the same as the previous result (a-1) for the large buffer case except for the narrower diagonal region.

(b-2) $P = 15$.
To increase the truncation number, all the errors become smaller. The dominant error comes from the pair between boxes 19 and 21. The errors for their related boxes are also larger than those for other box pairs.

(b-3) $P = 20$.
Although the errors are still recognized for the worst-case interaction and some related pairs, the errors for almost all the interactions converge. However, we can recognize that other errors, which are not seen at the previous state for $P = 15$, come from all the nearest neighbor interactions. These errors depend on the divergence property of the Hankel functions in the translation operator.



(a) $P = 5$.

(b) $P = 15$.

(c) $P = 20$.

(d) $P = 28$: Error minimized state.

(e) $P = 35$.

FIG. 6. *Error-convergence process for a small buffer case. The worst-case interaction is assumed to be in the nearest box pair.*

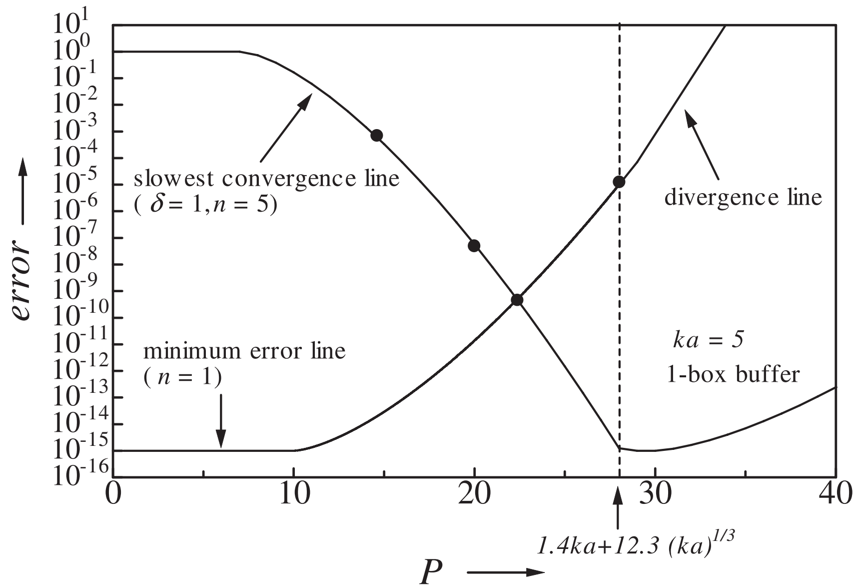


Fig. 7. *The slowest convergence line and the minimum error line for a* 1*-box buffer.*

(b-4) $P = 28$: Error minimized state.
Increasing the truncation number, the errors from the nearest neighbors become larger and the error from the worst-case interaction becomes smaller. Finally, these two errors approach the same level. This state can be defined as the error minimized state when the buffer size is small and the worst-case interaction is in the nearest box pair.

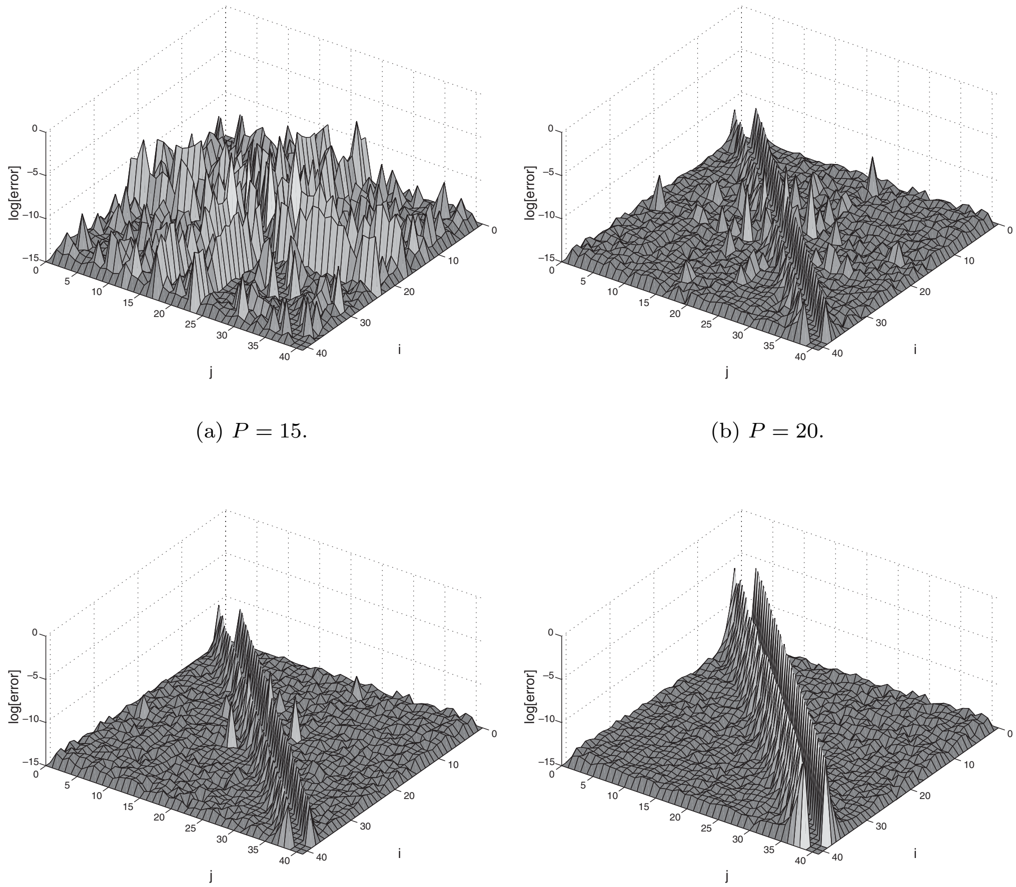
(b-5) $P = 35$.
If the truncation number is still increased, the error from all the nearest neighbor interactions tends to diverge.

3.2.2. Worst-case interaction not in a nearest box pair. We assume that the worst-case interaction ($\delta = 1$) is in the pair between boxes 17 and 23 ($n = 5$). Figure 7 shows the prediction of the slowest convergence line for the worst-case interaction and the minimum error line for the 1-box buffer case. The numerical experiment for a row of 40 contiguous boxes is shown in Figure 8.

(c-1) $P = 15$.
All the errors are in the process of converging. We can recognize that the dominant error comes from the worst-case interaction for boxes 17 and 23. The error for the related boxes is still large.

(c-2) $P = 20$.
The errors for the worst-case interaction and their related boxes are still in the process of converging. At this stage, we can recognize that another error comes from all the nearest neighbor interactions.

(c-3) $P = 22$: Error minimized state.
This state is determined by the convergence rate of the worst-case interaction and the divergence rate of the nearest neighbor interactions. When these two errors are of the same level, the total error from all the interaction pairs is



(a) $P = 15$.

(b) $P = 20$.

(c) $P = 22$: Error minimized state.

(d) $P = 28$: Optimal worst state.

FIG. 8. *Error convergence process for a small buffer case. The worst-case interaction is assumed not to be in the nearest box pair.*

minimized. To find this state, we need to know the information about the placement of the true worst-case interaction.

(c-4) $P = 28$: Optimal worst state.
The error from the nearest neighbor becomes larger and the error from the worst-case interaction completely converges. This state is the same as (b-4) for the worst-case interaction in the nearest box pair. If there is no information about the placement of the true worst-case interaction, we need to choose this number under the assumption that the worst-case interaction exists in the nearest box pair.

Next, we will investigate the error minimized state when the worst-case interaction ($\delta = 1$) consists of the farthest pair, that is, box 1 and box 40. There is a 38-box buffer between this worst pair. Figure 9 shows the error minimized state when $P = 22$. We can recognize that this state is determined by the worst-case interaction and the buffer size.

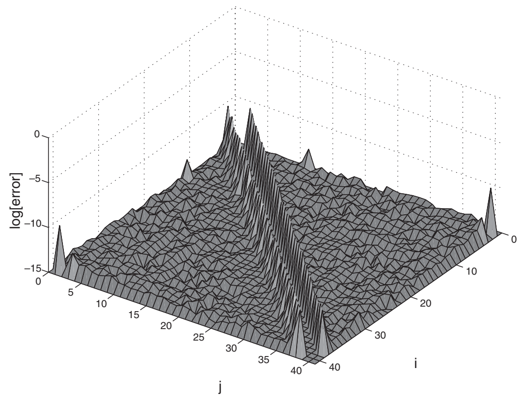


FIG. 9. *Error minimized state determined by the farthest box pair* ($P = 22$).

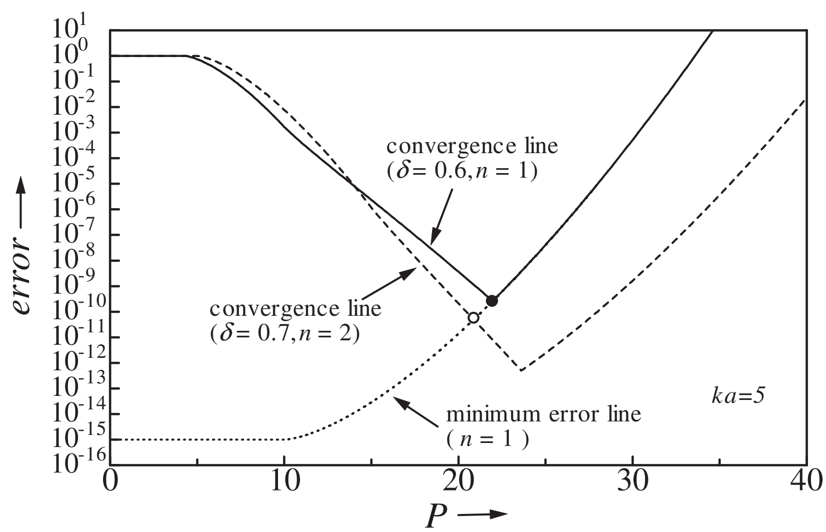


FIG. 10. *Slowest convergence line due to the interaction and the buffer size.*

4. Difficulties of error minimization. We have clarified that the information about the worst-case interaction is needed to minimize the truncation error for small buffer cases. However, finding the worst-case interaction and selecting the proper truncation number is not easy.

4.1. Worst-case interaction. We have used the term *worst-case interaction* without much elaboration. Generally speaking, it is difficult to recognize which interaction is the worst case. We investigate convergence processes of the interaction parameter $\delta = 0.6$ for a 1-box buffer pair and $\delta = 0.7$ for a 2-box buffer pair. Figure 10 shows the numerical result for these parameters. We can see that the error for the former case ($\delta = 0.6$, $n = 1$) is larger than that for the latter case ($\delta = 0.7$, $n = 2$) on the minimum error line. This example suggests that the worst-case interaction for small buffer cases is determined by the combination of the interaction parameter δ and the buffer size n.

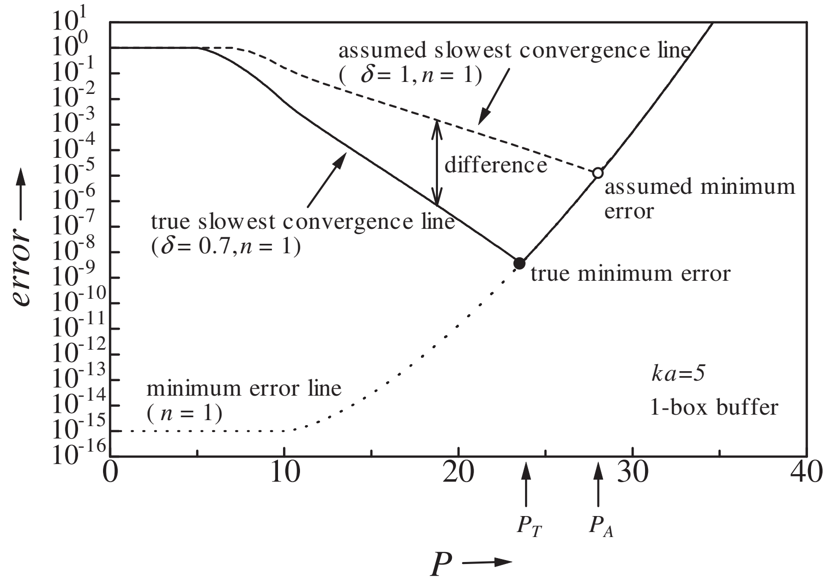


FIG. 11. *Difference of convergence processes between the assumed worst-case interaction and the true worst-case interaction for a* 1*-box buffer case.*

4.2. Truncation number. Even if we can find the worst-case interaction somehow, selecting the proper truncation number is still a difficult problem. Two examples are considered here.

We assume that the worst-case interaction is found for a 1-box buffer pair and its interaction parameter $\delta = 0.7$. Using this information, the convergence rate indicated by the solid line in Figure 11 is predicted. The error can be minimized only for the selection $P = P_T$ and becomes larger for P_A. In this case, the number P_T can be easily found considering the interaction parameter, such as $P_T := \sqrt{2}\delta ka + 12.3(\delta ka)^{1/3}$.

Next, we assume that a 2-box buffer pair has the worst-case interaction whose parameter $\delta = 0.7$. Figure 12 shows the prediction in this case. The same thing as in the previous example happens. We need to select P_T instead of P_A to minimize the error for the true worst-case interaction. However, this selection is not proper, since the error from any 1-box buffer pair becomes larger than the error from the worst-case interaction for this choice. Therefore, we need to select $P = P_{T_1}$ instead of P_T. This is another difficulty in minimizing the error. Since the selection depends on the placement of the worst-case interaction, other cases may arise.

5. Basic information. Before discussing algorithms to minimize the error, we will consider the information without searching for any interactions.

5.1. Optimal worst case. The given conditions without searching for any interactions are a box size ka and a buffer size $(n+1)ka$. The solid lines in Figure 13 correspond to the minimum error lines for the fixed box size. Since the worst case ($\delta = 1$) is assumed to exist, the truncation number should be selected as $P = \sqrt{2}ka+12.3(ka)^{1/3}$ to minimize the error. For this choice, the error for different box buffer cases is indicated by dots. The dominant error comes from the 1-box buffer pair and the error

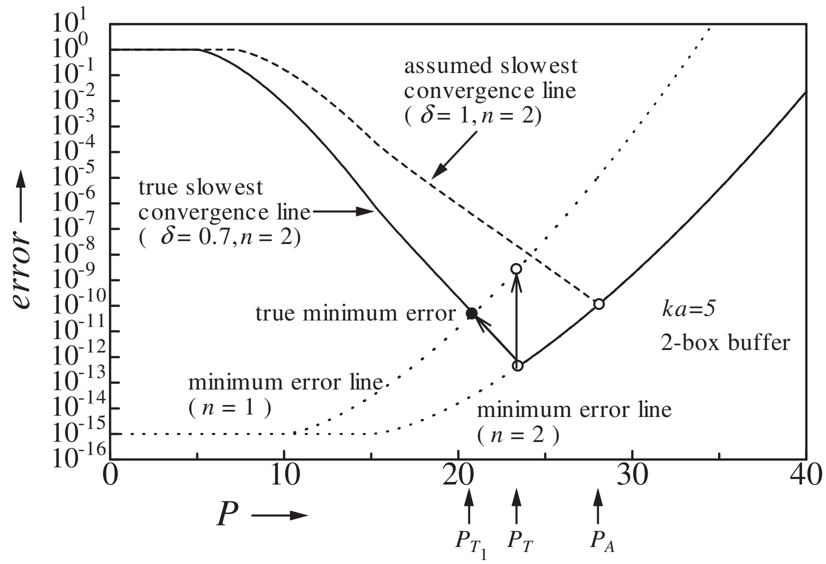


FIG. 12. *The way to find the proper truncation number to minimize the total error for a 2-box buffer case.*

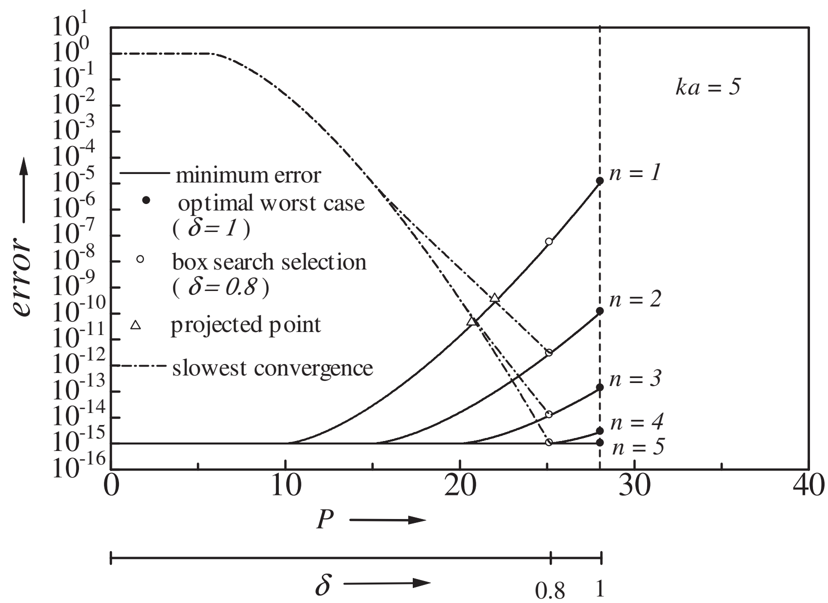


FIG. 13. *Choice of the truncation number for the optimal worst case, n-box search, and projection.*

from other boxes is smaller and negligible. The minimum error line converges to the value 10^{-15} for $0 < P < \sqrt{2}ka + 12.3(ka)^{1/3}$, when the number of buffer sizes satisfies $n > 0.4 + 12.3(ka)^{-2/3}$.

5.2. Box search. If we search elements in all the boxes individually and find the maximum distance $r_{\max}$ between an element and a box center, usually the error becomes smaller than that for the optimal worst case. Using the value $r_{\max}$, we can substitute the interaction parameter $\delta = r_{\max}/\sqrt{2}a$ (< 1), which is assumed to be $\delta = 0.8$ here. Compared with the optimal worst case, the errors for this choice (indicated by circles in Figure 13) become smaller. The dominant error still comes from the 1-box buffer pair.

5.3. Projection. If the truncation error is minimized, this value always should be found somewhere on the minimum error line for the smallest buffer case, that is, the 1-box buffer case for this example. After performing the box search, the largest interaction parameter is assumed to be $\delta = 0.8$. If this interaction truly exists for a 1-box buffer pair, the truncation number should be selected as $P \simeq \sqrt{2}\delta ka + 12.3(\delta ka)^{1/3}$ to minimize the error. However, we have not investigated all the interactions yet; there are various scenarios.

Let us assume that a 2-box buffer pair has the worst-case interaction whose parameter is $\delta = 0.8$. Although the error of this pair can be minimized for the same truncation number $P \simeq \sqrt{2}\delta ka + 12.3(\delta ka)^{1/3}$ as that for the 1-box buffer pair, this number cannot be selected. We should project this truncation number on the minimum error line for the 1-box buffer case indicated by triangles, as shown in Figure 13. To minimize the total error, we need to change the selection in this manner.

Next, we will discuss the way to find the projected point which is at the intersection between the minimum error line for the 1-box buffer case and the slowest convergence line for the worst-case interaction. If the interaction parameter δ is given, the convergence line for each buffer case can be predicted by the dashed-dotted line in Figure 13. The truncation number P at the projected point satisfies the following condition:

$$15 - \left[\frac{P-(n_a+1)ka}{1.8[(n_a+1)ka]^{1/3}}\right]^{3/2} = \left[\frac{P-\sqrt{2}\delta ka}{2.0(\delta ka)^{1/3}}\right]^{3/2} - \left[\frac{P-(n_b+1)ka}{1.8[(n_b+1)ka]^{1/3}}\right]^{3/2}, \tag{10}$$

where n_a is the smallest number of buffer boxes and n_b is the number of buffer boxes to find the slowest convergence line. Since the general solution does not exist for finding P, the value should be found approximately. Here, we will give an example for finding the intersection for the minimum error line for the 1-box buffer case and the slowest convergence line for the 2-box buffer case. The errors at five points M_1 to M_5, shown in Figure 14, can be predicted as follows:

(k-1) err_1 at M_1.
The error for the truncation number $P_1 = (n_b + 1)ka$ can be predicted by using the transformed excess bandwidth formula,

$$\log(1/\mathrm{err}_1) = \left[\frac{P_1-\sqrt{2}\delta ka}{2.0(\delta ka)^{1/3}}\right]^{3/2}. \tag{11}$$

(k-2) err_2 at M_2.
The error for the truncation number $P_2 = \sqrt{2}\delta ka + 12.3(\delta ka)^{1/3}$ can be predicted by using the new approach,

$$\log(1/\mathrm{err}_2) = 15 - \left[\frac{P_2-(n_b+1)ka}{1.8[(n_b+1)ka]^{1/3}}\right]^{3/2}. \tag{12}$$

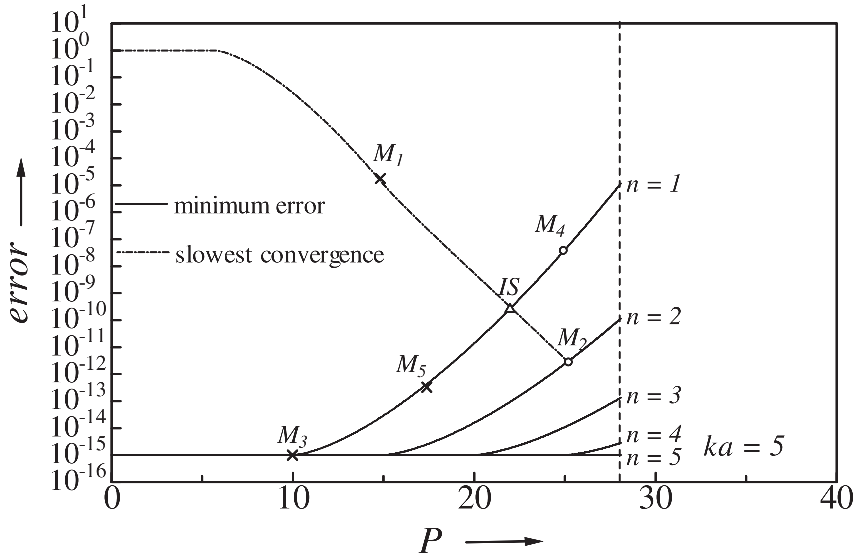


FIG. 14. *The way to find the projected points.*

Using this information, the segment $\overline{M_1M_2}$ can be approximated by the linear interpolation. Next, we approximate the minimum error line for the 1-box buffer case.

(k-3) err_3 at M_3.
We can predict the minimum error for $P_3 = (n_a + 1)ka$ by using the new approach,

$$\log(1/\,\mathrm{err}_3) = 15 - \left[\frac{P_3 - (n_a + 1)ka}{1.8[(n_a + 1)ka]^{1/3}}\right]^{3/2}. \tag{13}$$

(k-4) err_4 at M_4.
The error for the truncation number $P_4 = \sqrt{2}\delta ka + 12.3(\delta ka)^{1/3}$ can be predicted by using the new approach,

$$\log(1/\,\mathrm{err}_4) = 15 - \left[\frac{P_4 - (n_a + 1)ka}{1.8[(n_a + 1)ka]^{1/3}}\right]^{3/2}. \tag{14}$$

(k-5) err_5 at M_5.
Considering the truncation number P_3 and P_4, another number can be selected as $P_5 = 0.5(P_3 + P_4)$. For this number, the error can be predicted by

$$\log(1/\,\mathrm{err}_5) = 15 - \left[\frac{P_5 - (n_a + 1)ka}{1.8[(n_a + 1)ka]^{1/3}}\right]^{3/2}. \tag{15}$$

The segment $\overline{M_3M_4M_5}$ can be approximated by quadratic interpolation. The intersection point IS can be determined by using these approximated two segments. The same method can be applied to finding the projected points for the other buffer pairs. We do not need to continue this routine for all the buffer pairs, since the slowest convergence line becomes the same one for increasing the buffer box size.

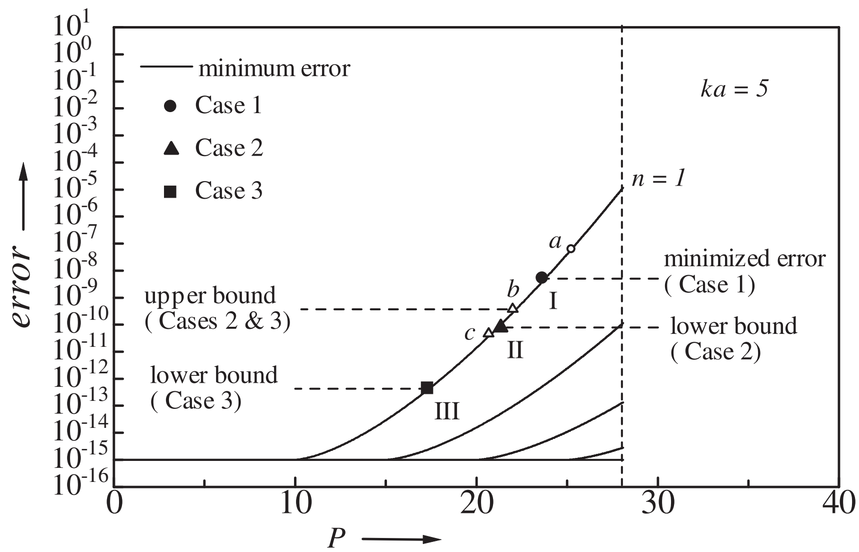


Fig. 15. *Buffer-based search for the* 1*-box buffer case: The worst error is found on Segment* I.

Using the projected points, we can divide the minimum error line for the 1-box buffer case into the following three segments (Figure 15):

(l-1) Segment I: between a and b.
If the minimized error is found on this segment, the worst-case interaction should exist for a 1-box buffer pair.

(l-2) Segment II: between b and c.
The worst-case interaction should be found for 1-box buffer or 2-box buffer pairs. The error for other box buffer pairs can be made smaller than this value.

(l-3) Segment III: the rest of the line.
If the minimized error is found in this segment, the worst-case interaction can exist for any box pairs.

6. Search algorithms. Considering the basic information, we propose search algorithms to minimize the truncation error and clarify their advantages and disadvantages.

6.1. Buffer-based search. The buffer-based search stands for searching the worst-case interaction in terms of changing the buffer size. Starting from all the 1-box buffer pairs, we will increase the buffer size until the worst-case interaction is found.

After searching all the interactions for 1-box buffer pairs, we can find the largest interaction parameter δ_1 among them. The truncation number to minimize this error can be selected as $P = \sqrt{2}\delta_1 ka + 12.3(\delta_1 ka)^{1/3}$. This error should be found somewhere on the minimum error line for the 1-box buffer case. The following three cases may arise, as shown in Figure 15:

(m-1) Case 1 (dot): error on Segment I.
If the error is found on Segment I, we do not need to search other box buffer pairs. This is the minimized error.

(m-2) Case 2 (triangle): error on Segment II.
The error can be considered as the lower bound. The upper bound is determined by the projected value for the 2-box buffer case. We need to search all the 2-box buffer pairs, and then the worst-case interaction can be determined.

(m-3) Case 3 (square): error on Segment III.
The error can be considered as the lower bound. The upper bound is the same as that for the projected one for the 2-box buffer case. We need to search all the 2-box buffer pairs.

If the error is found on Segment II or III, we need to continue the search for all the 2-box buffer pairs, and the largest interaction parameter for a 2-box buffer pair can be found such as δ_2. The projected value for $P = \sqrt{2}\delta_2 ka + 12.3(\delta_2 ka)^{1/3}$ can be found on Segment II or III. Considering the previous conditions, there are the following four possibilities:

(n-1) Case 2.1: error on Segment II.
In the previous step, the error is found on Segment II. The minimized error is determined by the larger value, this one or the error for Case 2. We can stop searching here.

(n-2) Case 2.2: error on Segment III.
In the previous step, the error is found on Segment II. The minimized error is the error for Case 2. We can stop searching here.

(n-3) Case 3.1: error on Segment II.
In the previous step, the error is found on Segment III. This value is the minimized error. We can stop searching here.

(n-4) Case 3.2: error on Segment III.
In the previous step, the error is found on Segment III. Compared with the error for Case 3, the larger one can be considered as the lower bound. The upper bound becomes the projected value for the 3-box buffer case.

Only for Case 3.2 do we need to continue the search for all the 3-box buffer pairs. After performing this, the projected value can be found in Segment III. Compared with the lower bound in the previous step, the larger one can be set as the new lower bound. If this value is not equal to the upper bound, we need to search all the 4-box buffer pairs. We can stop searching only when the projected value becomes the same as the upper bound or all the pairs are searched. This is the disadvantage.

6.2. Element-based search. The element-based search stands for searching the worst-case interaction in terms of distribution. When we perform the box search in section 5.2, it is possible to store the information about the distance of each element from the box center. Starting from the box which contains the longest distance element, we will search all the possible pairs. If we cannot find the pair for the worst-case interaction, the same search is performed for the box which contains the second longest distance element with other boxes.

This method is based on finding the largest interaction parameter δ. If a pair which contains these elements is found for a large buffer pair, we cannot conclude that this is the worst-case interaction, because there is a possibility that a smaller δ for a smaller buffer pair becomes the worst-case interaction, as shown in Figure 10. Therefore, all the possible pairs need to be searched in general. This is the disadvantage.

6.3. Combination search. To combine the advantages of the above two algorithms, we perform the buffer-based search first and then switch to the element-based search when the upper bound is fixed. In section 6.1, we have shown that the mini-

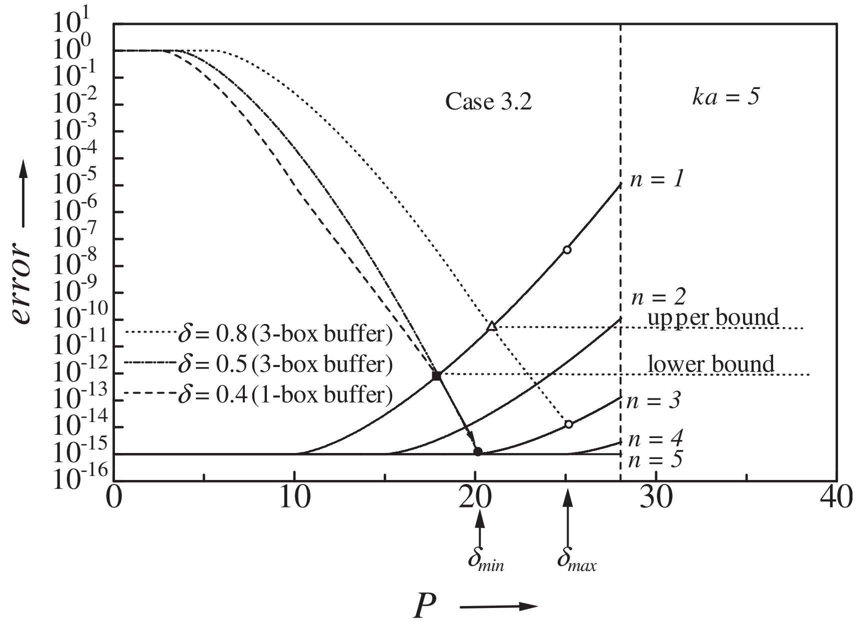


FIG. 16. *Definition of range of the interaction parameter:* $\delta_{\max}$ *and* $\delta_{\min}$.

mized error can be determined by using only the buffer-based search for Cases 1, 2, and 3.1. Here, we discuss the way to minimize the error for Case 3.2.

We perform the buffer-based search for the 1-box and 2-box buffer pairs. After finishing the 2-box buffer pair search, the upper bound and lower bound are fixed. As the next step, we project the lower bound, indicated by the square in Figure 16, on the minimum error line for the 3-box buffer case, since the interaction parameter for this value depends on the number of buffer boxes (see Figure 16). Therefore, we need to estimate the equivalent interaction parameter $\delta_{\min}$ for the lower bound in terms of the 3-box buffer case. The assumed largest interaction parameter $\delta_{\max}$ is the same as $\delta = 0.8$ in this case. Using these two values, the interaction parameter δ for the worst-case interaction is bound by $\delta_{\min} \leq \delta \leq \delta_{\max}$, as shown in Figure 16.

Next, we start the element-based search for the box that contains the longest distance element and all other boxes except 1-box and 2-box buffer cases. After this search, even if we do not find the interaction parameter which makes the lower bound larger, the upper bound always becomes smaller, which is replaced by the second largest element as the next search, since we exclude all the possibilities for $\delta_{\max} = 0.8$. The advantages of this method are as follows:

1. All the pairs for small box buffer cases are investigated first. The upper bound can be made smaller after each search.
2. When the upper bound fully converges to a certain value, the element-based search is applied. Again the upper bound always becomes smaller after each search.

6.4. Computational verification. Here, we verify the advantage of our error minimization techniques to compare absolute relative errors for various control methods. Errors are computed for randomly distributed elements inside the previ-

Table 1

The number of searches to reach the error minimized state by various methods. The absolute relative error is computed for randomly distributed elements inside 40 *contiguous square boxes. To vary the starting value of the pseudorandom process, three examples are considered for* $M = 1$ *and* 4*, where* M *is the number of elements in each box.* P *is the truncation number for the multipole expansion given by each method. The prediction means the estimated error and the error shows the true computational error.*

	Case 1				Case 2				Case 3			
Method	P	Prediction	Error	Number	P	Prediction	Error	Number	P	Prediction	Error	Number
Excess Bandwidth	28	1.0E-15	1.2E-05	0	28	1.0E-15	1.2E-05	0	28	1.0E-15	1.1E-05	0
New Approach	28	1.2E-05	1.2E-05	0	28	1.2E-05	1.2E-05	0	28	1.2E-05	1.1E-05	0
Direct Method	24	–	4.2E-09	20748	22	–	4.6E-10	20748	25	–	7.1E-08	20748
Buffer-Based Search	23	1.3E-09	4.6E-09	38	22	1.9E-10	4.6E-10	741	25	3.8E-08	7.1E-08	38
Element-Based Search	23	1.3E-09	4.6E-09	741	22	1.9E-10	4.6E-10	741	25	3.8E-08	7.1E-08	741
Combination Search	23	1.3E-09	4.6E-09	38	22	1.9E-10	4.6E-10	175	25	3.8E-08	7.1E-08	38

(a) $M = 1$.

	Case 1				Case 2				Case 3			
Method	P	Prediction	Error	Number	P	Prediction	Error	Number	P	Prediction	Error	Number
Excess Bandwidth	28	1.0E-15	1.2E-05	0	28	1.0E-15	1.2E-05	0	28	1.0E-15	1.2E-05	0
New Approach	28	1.2E-05	1.2E-05	0	28	1.2E-05	1.2E-05	0	28	1.2E-05	1.2E-05	0
Direct Method	25	–	6.0E-08	331968	24	–	7.8E-08	331968	26	–	1.7E-07	331968
Buffer-Based Search	25	9.7E-08	6.0E-08	608	25	9.5E-08	9.0E-08	608	26	2.7E-07	1.7E-07	608
Element-Based Search	25	9.7E-08	6.0E-08	11856	25	9.5E-08	9.0E-08	11856	26	2.7E-07	1.7E-07	11856
Combination Search	25	9.7E-08	6.0E-08	608	25	9.5E-08	9.0E-08	608	26	2.7E-07	1.7E-07	608

(b) $M = 4$.

ously assumed 40 contiguous square boxes and the maximum value is picked up. It is shown in the column "Error" in Table 1. To vary the starting value of the pseudorandom process, three examples are considered for $M = 1$ and 4, where M is the number of elements in each box. The number of elements N corresponds to $40M$.

The excess bandwidth stands for selecting the truncation number P by (5) under the condition $d_0 = 15$. The same truncation number is given by (8) for $\delta = 1$ in terms of the new approach. Considering the prediction, the new approach is always better. The number of searches to reach the error minimized state shown in the column "Number" is zero for both methods.

The direct method consists of computing errors of all the possible pairs for changing the truncation number P and selecting the minimized one. It is much smaller than the error given by the new approach. Although the error can be minimized by this method, the number of searches is proportional to $O(N^2)$.

The advantage of the proposed algorithms clearly can be seen in the table. They give the precise truncation number to reach the error minimized state and the excellent error prediction for all the cases. The combination search is the most efficient, and the number of searches is drastically reduced and proportional to $O(N)$. The buffer-based search is of the same cost except for Case 2 in Table 1(a).

7. Conclusions. We have studied the condition to minimize the truncation error of the multipole expansion. For large buffer cases, the truncation number is selected to control the worst-case interaction, which is assumed to exist. If the error is minimized for the assumed worst case, the total error from all the interaction pairs is also minimized. For small buffer cases, the truncation number is selected to control the worst-case interaction, which is assumed to be in the nearest box pair. However, the truncation number which minimizes the error for the assumed worst case usually does not minimize the total error from all the possible pairs.

Our computational results show that the combination search gives the excellent error prediction and precise truncation number to reach the error minimized state without searching through all the possible pairs. The cost is proportional to $O(N)$, where N is the number of elements. For many cases, the error can be minimized by using the buffer-based search with the same computational cost.

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 Trans. Antennas and Propagation, 35 (1993), pp. 7–12.

[3] C. C. Lu and W. C. Chew, *Fast algorithm for solving hybrid integral equation*, IEE Proceedings Part H: Microwaves, Antennas and Propagation, 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] S. Ohnuki and W. C. Chew, *Numerical accuracy of multipole expansion for 2D MLFMA*, IEEE Trans. Antennas and Propagation, 51 (2003), pp. 1883–1890.

[8] S. Ohnuki and W. C. Chew, *Truncation error analysis of multipole expansion*, SIAM J. Sci. Comput., 25 (2003), pp. 1293–1306.

[9] W. C. Chew, *Waves and Fields in Inhomogeneous Media*, Van Nostrand Reinhold, New York, 1990; reprinted by IEEE Press, New York, 1995.

[10] W. C. Chew, J. M. Jin, E. Michielssen, and J. M. Song, eds., *Fast and Efficient Algorithms in Computational Electromagnetics*, Artech House, Boston, 2001.

[11] W. J. Wiscombe, *Improved Mie scattering algorithms*, Appl. Optics, 19 (1980), pp. 1505–1509.

[12] O. M. Bucci and G. Franceschetti, *On the spatial bandwidth of scattered fields*, IEEE Trans. Antennas and Propagation, 35 (1987), pp. 1445–1455.

[13] V. Rokhlin, *Sparse Diagonal Forms for Translation Operations for the Helmholtz Equation in Two Dimensions*, Research Report YALEU/DCS/RR-1095, Dept. of Computer Science, Yale University, New Haven, CT, 1995.

[14] J. M. Song and W. C. Chew, *Error analysis for the truncation of multipole expansion of vector Green's functions*, IEEE Micro. Wireless Components Lett., 11 (2001), pp. 311–313.