

## Stochastic Estimation of the Frobenius Norm in the ACA Convergence Criterion

A. Heldring, E. Ubeda, and J. M. Rius

**Abstract**—The adaptive cross approximation (ACA) algorithm has been used in many fast Integral Equation solvers for electromagnetic Radiation and Scattering problems. It efficiently computes a low rank approximation to the interaction matrix between mutually distant parts of a scattering object. The ACA is an iterative algorithm that needs an accurate and efficient convergence criterion. The evaluation of this criterion may consume a considerable part of the computational resources. This communication presents an efficient new way to evaluate the convergence criterion, using a stochastic approach.

**Index Terms**—Adaptive cross approximation (ACA), computational electromagnetics, method of moments.

### I. INTRODUCTION

The adaptive cross approximation (ACA) algorithm for fast compression of rank deficient matrix sub-blocks was originally introduced for integral equations with “asymptotically smooth” (non-oscillatory) kernels [1]. In 2003, Zhao *et al.* [2] showed that the algorithm works well in practice for electromagnetic problems, even though the rigorous mathematical basis from [1] no longer holds. Since then, ACA has been used successfully in several studies of electromagnetic radiation and scattering problems, for example [3]–[5].

Typical usage of the ACA consists in hierarchically subdividing the problem impedance matrix and compressing all the off-diagonal blocks, of which there may be thousands, ranging in dimensions from a few hundred to tens of thousands. For a detailed description of the ACA we refer to the extensive literature, starting with [1] and [2]. In a nutshell: The ACA is an iterative method. At step  $k$  it picks a row and a column of the original block and recombines these to update the current approximation. The choice of which rows and columns to incorporate (pivoting) is necessarily based on incomplete information; the main benefit of the ACA, its efficiency, would be unacceptably compromised if all the elements of the original block had to be known.

An essential ingredient of the ACA algorithm is the convergence criterion; it should be as accurate as possible to minimize the number of iterations necessary for a prescribed relative error in the approximation. At the same time, it should be efficient; the computational effort involved should be small compared to the rest of the algorithm. In the ACA, the assumption is made that the absolute residual error at every iteration is of the order of magnitude of the latest update. With oscillating kernels, this is sometimes not fulfilled due to erroneous pivoting. In extreme cases, this may lead the algorithm to declare convergence long before it is actually reached. Some proposed strategies to avoid this problem are to restart the ACA [6] or apply some averaging [7].

Manuscript received September 19, 2014; revised December 19, 2014; accepted December 20, 2014. Date of publication December 29, 2014; date of current version February 27, 2015. This work was supported by the Spanish Ministry of Economy and Competitiveness under projects TEC2012-37582-C04-02, TEC2010-20841-C04-02 and by the European FEDER fund.

The authors are with the AntennaLab, Dept. of Signal Processing and Telecommunications, Universitat Politècnica de Catalunya, 08034 Barcelona, Spain (e-mail: heldring@tsc.upc.edu).

Color versions of one or more of the figures in this communication are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TAP.2014.2386306

These strategies solve the problem of premature convergence at little extra computational cost, yielding a reliable estimate of the order of magnitude of the residual absolute error.

However, as a measure of accuracy, we need the relative rather than the absolute error. Therefore, the above absolute error is compared to the Frobenius norm of the entire block. The latter, of course, can also only be estimated. In the original algorithm, the estimated Frobenius norm is updated with every iteration, in a way that does indeed converge to the true Frobenius norm for asymptotically smooth kernels. In a recent paper [7] we demonstrated that this is not true for oscillatory kernels. We also proposed an alternative procedure that does converge to the true norm.

Both the original, erroneous procedure and the one in [7] require a non negligible computational effort. The relative importance of this effort depends on the type of problem, the code and the platform; if the computation of the matrix elements is slow, then this will dominate the computation time. If, on the other hand, this part of the algorithm is highly optimized, then the linear algebra involved in estimating the norm may become the most expensive part.

In the example we presented in [7] and use again here in Section II, the time spent estimating the norm is more than 40% of the total time. This incentivized us to look for a more efficient procedure. In this communication we consider a stochastic approach, motivated by the following three considerations: (1) The absolute residual error of the ACA applied to oscillating kernels has an unavoidable uncertainty that is hard to quantify, but empirically it is on the order of 50% (illustrated in Section II). In consequence, an accuracy of 10% is more than sufficient for the Frobenius norm. (2) The Frobenius norm is proportional to the mean value of the squared norms of the matrix elements, which makes it particularly tractable by stochastic methods. (3) The Frobenius norm is independent of the phase of the matrix elements. The computational effort involved in a stochastic approach is therefore independent of the frequency.

### II. UNCERTAINTY IN THE ACA ABSOLUTE ERROR

The ACA algorithm declares convergence when the *relative error*

$$\varepsilon_{rel} = \frac{\|Z - UV^T\|_F}{\|Z\|_F} \quad (1)$$

in the ACA approximation  $UV^T$  of the (sub-) matrix  $Z$  is estimated to be below a chosen threshold value. The numerator of (1) is the *absolute error*  $\varepsilon_{abs}$ , and the denominator of (1) is the Frobenius norm of  $Z$ . Without detailing the procedure, which is extensively explained in the literature referenced above, in this section we illustrate the unavoidable uncertainty in the estimation of  $\varepsilon_{abs}$  with an example. The example is the same one that we used in [7]. It concerns the mutual impedance matrix at 300 MHz of two square perfectly conducting flat plates, with edge lengths equalling 5 m, facing each other at a distance of 10 m, modelled with the Electric Field Integral Equation, discretized into RWG basis functions on a triangular mesh. The plates are represented by 7400 basis functions each. Fig. 1 shows the relative error in the ACA estimation of  $\varepsilon_{abs}$

$$\Delta_{abs} = \frac{|\varepsilon_{abs} - \varepsilon_{abs}^{estim}|}{\varepsilon_{abs}} \quad (2)$$

as a function of the ACA iteration step. We see that this error, which gives rise to an equal relative error in (1), regularly exceeds a value of 50%. This error is caused by the incomplete knowledge of the elements

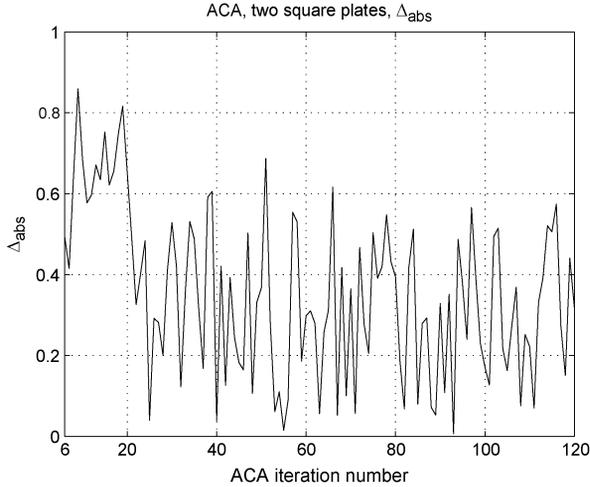


Fig. 1. The relative error committed by the ACA algorithm when it estimates the absolute difference between the ACA approximation and the true matrix.

of  $Z$  and therefore unavoidable. As a result, when the ACA terminates with a desired relative error of  $10^{-3}$  (a typical value), the real relative error is likely to be somewhere between  $0.7 \times 10^{-3}$  and  $1.5 \times 10^{-3}$ . This does not disqualify the ACA approximation, it just means that we only know the order of magnitude of the ACA residual error. It also implies that there is no need for high precision in the estimate of the Frobenius norm in the denominator of (1).

### III. STOCHASTIC FROBENIUS NORM ESTIMATION

From the definition of the Frobenius norm of a matrix  $Z \in \mathbb{C}^{m \times n}$ , with elements  $z$ , we have

$$\|Z\|_F = \sqrt{mn\mu_z}, \quad (3)$$

with  $\mu_z = \overline{|z|^2}$ , the mean value of the squared norms of the elements. Using standard statistical theory (e.g., [8]), if we pick  $N$  elements from  $Z$  by simple random sampling (SRS) [9], then the expected value of the *sample-mean*  $\mu_s = \overline{|z_s|^2}$ , equals  $E[\mu_s] = \mu_z$  and the expected value of the sample standard deviation

$$\sigma_s = \sqrt{\frac{1}{(N-1)} \sum_{i=1}^N (|z_i|^2 - \mu_s)^2} \quad (4)$$

is the true standard deviation of the squared norms of the elements  $E[\sigma_s] = \sigma_z$ . According to the Central Limit Theorem, for large enough  $N$ , the probability density distribution of the sample-mean will approach the normal distribution  $N(\mu_z, \sigma_z)$ , irrespective of the distribution of  $|z|^2$  [10]. Because we do not know either  $\mu_z$  or  $\sigma_z$  we must use Student's t-test [10] to establish the probability that the difference between the estimated mean  $\mu_s$  and the true mean  $\mu_z$  lies within a chosen range:

$$Pr \left[ \mu_s - t_{(\frac{\alpha}{2}, N-1)} \frac{\sigma_s}{\sqrt{N}} < \mu_z < \mu_s + t_{(\frac{\alpha}{2}, N-1)} \frac{\sigma_s}{\sqrt{N}} \right] = 1 - \alpha \quad (5)$$

where  $\alpha$  fixes the probability and  $t_{(\frac{\alpha}{2}, N-1)}$  is a parameter of the Student's t-distribution. Defining the relative error in the mean as

$$\Delta_\mu = \frac{|\mu_z - \mu_s|}{\mu_s}, \quad (6)$$

by rearranging the argument of the probability function in (5), we find that for a given choice of  $\alpha$  and  $N$ ,

$$\Delta_\mu < \frac{t_{(\frac{\alpha}{2}, N-1)} \sigma_s}{\sqrt{N} \mu_s} \quad (7)$$

with probability  $1 - \alpha$ . Finally, since the Frobenius norm is proportional to the square root of  $\mu$ , the relative error in the estimated Frobenius norm equals

$$\Delta_F = \frac{\Delta_\mu}{2}. \quad (8)$$

The objective of our procedure is to obtain, using a minimum number of samples, an estimate with a prescribed relative error. Accordingly, we start with a chosen initial number of samples  $N$  and iteratively keep adding samples and updating (8), until  $\Delta_F$  drops below a chosen maximum tolerance value  $\Delta_{tol}$  with probability  $1 - \alpha$ . Both the sample mean and the sample standard deviation can be incrementally updated for efficiency [11].  $t_{(\frac{\alpha}{2}, N-1)}$ ,  $\mu_s$  and  $\sigma_s$  all converge with growing  $N$  so the rate of convergence of the iterative process is globally proportional to  $1/\sqrt{N}$ . The actual convergence rate depends on the true spread of the sampled variable,  $\sigma_z/\mu_z$ . If the matrix elements assume a wide range of values, for example due to a very uneven geometry discretization or a wide range of interaction distances, more iterations are needed. As mentioned in the introduction, we can tolerate a relatively high error. This benefits a speedy convergence. On the other hand, we wish to virtually exclude the risk of an unacceptable error level, which calls for a very small value of  $\alpha$ .

Two representative numerical examples, presented in Section IV, give an indication of whether the iteration converges sufficiently fast for this procedure to be a competitive alternative in the framework of ACA applied to electromagnetic problems.

### IV. TWO NUMERICAL EXAMPLES

We have applied the procedure outlined in Section III to two scattering problems. The first example is the one with two square plates from Section II. The true mean value of the squared norms of the interaction matrix elements for this problem equals  $\mu_z = 2.9 \times 10^{-3}$  with a standard deviation of  $\sigma_z = 2.9 \times 10^{-3}$ , equal to the mean value. We chose the following values for the parameters in (7):  $\Delta_{tol} = 0.1$  and  $\alpha = 0.1\%$ . To test the validity of (7), we executed the procedure 100,000 times, and evaluated the true residual relative error in the Frobenius norm after convergence every time. In order to make sure the Central Limit Theorem applied, the initial number of samples was set to 100. The value of  $t_{(\frac{\alpha}{2}, N-1)}$  was fixed at  $t_{(0.0005, 99)} = 3.39$ . In reality the value of  $t$  decreases towards a value of 3.29 for infinite  $N$ , but we chose to maintain a "worst case" value. The distribution of the true relative errors obtained from each run is shown in Fig. 2. There were 110 values that did not comply with (7) (they had a relative error more than 10%), not too far off from the expected number of 100. More importantly, the largest true error was 0.1223, which would still not compromise the ACA accuracy. Indeed, the probability of a given

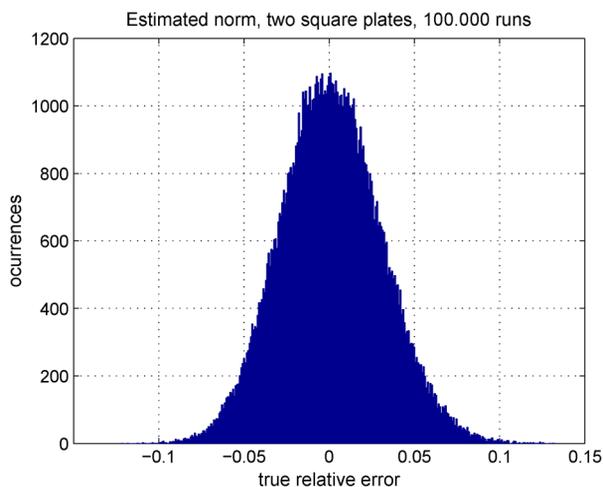


Fig. 2. Histogram of 100,000 estimates of the Frobenius norm of the interaction matrix of two square conducting plates.

error decreases fast with its size. As an illustration, with these same results, the probability of a relative error of 0.183 as predicted by (7) is one in  $10^8$ !

The number of samples needed was 284 on average, with a standard deviation of 25 and a maximum of 394. Comparing this with the computational effort involved in the ACA algorithm, of evaluating a full row and a full column of  $n = m = 7400$  elements each with every iteration plus algebraic operations of complexity  $k(n + m)$  at iteration  $k$ , shows that the operation count for estimating the Frobenius norm is insignificant.

In order to illustrate that the procedure is indeed frequency independent we raised the frequency from 300 MHz to 1.2 GHz and adapted the mesh accordingly, to 119,600 basis functions on each plate. Over 100,000 runs, the average number of samples needed for convergence was 285, with a standard deviation of 25, an insignificant difference with the previous result.

In the above example the interaction matrix was relatively homogeneous. In order to test the procedure under more adverse conditions we used a problem geometry consisting of two open cones, 5 m long, with a 1 m base diameter, aligned as depicted in Fig. 3, with a distance between the axes of 5 m. This problem geometry combines several difficult features: a high spread in mesh size (the largest triangular element surfaces, near the cone base, are a factor 2700 larger than the smallest ones, near the peak) and in relative orientation of the mesh elements. The simulation frequency was 300 MHz, but as mentioned above, the Frobenius norm estimation is independent of the frequency. The mean value of the squared norms of the interaction matrix elements for this problem equalled  $\mu_z = 6.0 \times 10^{-4}$  with a standard deviation of  $\sigma_z = 1.4 \times 10^{-3}$ . Again, we executed the estimation algorithm 100,000 times, using the same parameters in (7) as in the first example. The true relative error is distributed as in Fig. 4. This time there were 94 values that did not comply with (7). As expected, for this case many more iterations were necessary: 1622 on average, with a standard deviation of 207 and a maximum, among the 100,000 executions, of 3251 iterations. Nevertheless, even the maximum is still well below the operation count for a single ACA step. Note that the square root of the average number of iteration steps for the cones is 2.39 times that for the square plates. The true spread  $\sigma_z/\mu_z$  of the interaction matrix of the cones is 2.40 times that of the square plates. The two values are very close, as predicted by (7).

Finally, Table I shows the computation times for the ACA compression of the square plate interaction matrix at 300 MHz and 1.2 GHz.

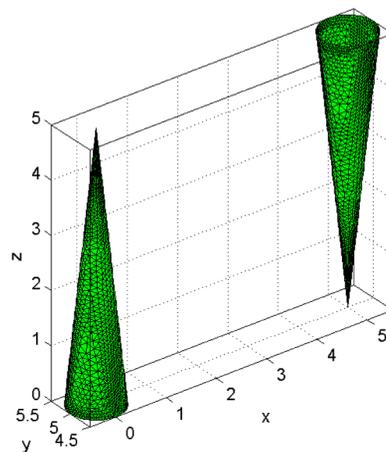


Fig. 3. Two conical surfaces.

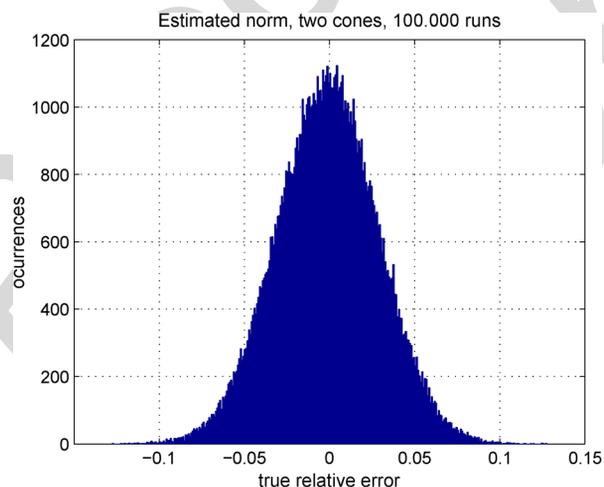


Fig. 4. Histogram of 100,000 estimates of the Frobenius norm of the interaction matrix of two conical conducting surfaces.

TABLE I  
COMPUTATION TIMES FOR ACA COMPRESSION OF TWO 5 m  $\times$  5 m SQUARE PLATES INTERACTION MATRICES

Computation time (seconds)	ACA	Frob. estim.	Total
300 MHz, 7400 unknowns per plate			
Deterministic method from [7]	2.11	1.64	3.75
Stochastic method, this paper	2.11	0.03	2.14
1.2 GHz, 119,600 unknowns per plate			
Deterministic method from [7]	795	710	1505
Stochastic method, this paper	795	0.03	795

Convergence was declared when the relative error in the ACA approximation was estimated at  $10^{-3}$ , which took  $k = 99$  steps for the lower frequency and  $k = 496$  steps for the higher frequency. The relative gain in computation time with the stochastic method is even more pronounced for the higher frequency, almost 50%. This is because for large  $k$ , the computation time of the ACA algorithm and that of the deterministic norm-estimation both scale with  $k^2(n + m)$  (see [7]) but the ACA algorithm also needs to compute the matrix elements, an operation that scales with  $k(n + m)$ . This latter term is proportionally more important for the lower frequency because  $k$  is smaller.

The average computation time of the stochastic procedure in the case of the two cones is 0.17 seconds which, compared with the timing for the plates (Table I), confirms that the computation time is proportional

to  $N$ . It should be stressed here that, as already mentioned in the introduction, these computation times depend on many problem-, implementation-, and hardware related factors. The presented examples were implemented in Matlab<sup>(R)</sup>, with a call to an external C++ routine for computing the matrix elements, on a PC with a Dual Intel Xeon<sup>(R)</sup> X5482 processor at 3.2 GHz.

#### V. CONCLUSION

An essential part of the ACA algorithm, the convergence criterion, involves estimating the Frobenius norm of the matrix to be approximated. The relative computational effort associated with this may be considerable if it is done in a conventional deterministic manner, in particular for problems with an oscillating kernel.

This communication proposes a stochastic approach to estimate the Frobenius norm and presents two numerical examples illustrating an important gain in computation time. The stochastic approach is shown to yield a manyfold reduction in operation count, with parameters chosen to reduce the probability of an unacceptable chance result virtually to zero. This resulted, in the author's implementation of the ACA, in a speed-up of more than 40%.

#### REFERENCES

- [1] M. Bebendorf, "Approximation of boundary element matrices," *Numer. Matem.*, vol. 86, no. 4, pp. 565–589, 2000.
- [2] K. Zhao, M. N. Vouvakis, and J.-F. Lee, "The adaptive cross approximation algorithm for accelerated method of moments computations of EMC problems," *IEEE Trans. Electromag. Compat.*, vol. 47, no. 4, pp. 763–773, Nov. 2005.
- [3] G. Hislop, S. Hay, and A. Hellicar, "Efficient sampling of electromagnetic fields via the adaptive cross approximation," *IEEE Trans. Antennas Propag.*, vol. 55, no. 12, pp. 3721–3725, Dec. 2007.
- [4] R. Maaskant, R. Mittra, and A. Tjihuis, "Fast analysis of large antenna arrays using the characteristic basis function method and the adaptive cross approximation algorithm," *IEEE Trans. Antennas Propag.*, vol. 56, no. 11, pp. 3440–3451, Nov. 2008.
- [5] A. Heldring, J. M. Tamayo, E. Ubeda, and J. M. Rius, "Accelerated direct solution of the method-of-moments linear system," *Proc. IEEE*, vol. 101, no. 2, pp. 364–371, Feb. 2013.
- [6] J. Laviada, R. Mittra, M. R. Pino, and F. Las-Heras, "On the convergence of the ACA," *Microwave Opt. Technol. Lett.*, vol. 51, no. 10, pp. 2458–2460, Oct. 2009.
- [7] A. Heldring, E. Ubeda, and J. M. Rius, "On the Convergence of the ACA Algorithm for Radiation and Scattering Problems," *IEEE Trans. Antennas Propag.*, vol. 62, no. 7, pp. 3806–3809, Jul. 2014.
- [8] J. R. Taylor, *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements*. Mill Valley, CA, USA: University Science Books, 1982.
- [9] D. S. Moore, G. P. McCabe, and B. A. Craig, *Introduction to the Practice of Statistics*, 6th ed. New York, NY, USA: W.H. Freeman, 2009.
- [10] Rachev, S. T. Stoyanov, S. V. Fabozzi, and J. Frank, *Advanced Stochastic Models, Risk Assessment, and Portfolio Optimization: The Ideal Risk, Uncertainty, and Performance Measures*. Hoboken, NJ, USA: Wiley, 2008.
- [11] D. E. Knuth, *Seminumerical Algorithms*, 3rd ed. Boston, MA, USA: Addison-Wesley, 1998, vol. 2, The Art of Computer Programming.