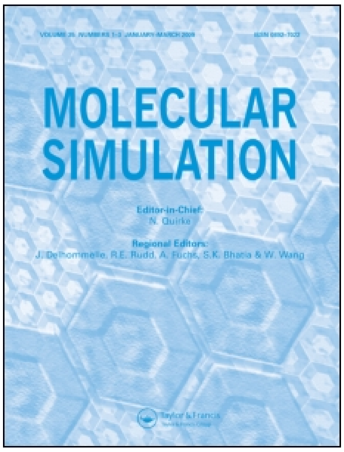

**Molecular Simulation**
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

# Fast Version of the Structure Adapted Multipole Method-Efficient Calculation of Electrostatic Forces in Protein Dynamics

Christoph Niedermeier[a]; Paul Tavan[a]
[a] Institut für Medizinische Optik, Theoretische Biophysik Ludwig-Maximilians-Universität München Theresienstraße 37, München, Germany

**To cite this Article** Niedermeier, Christoph and Tavan, Paul(1996) 'Fast Version of the Structure Adapted Multipole Method-Efficient Calculation of Electrostatic Forces in Protein Dynamics', Molecular Simulation, 17: 1, 57 — 66
**To link to this Article: DOI:** 10.1080/08927029608024094
**URL:** http://dx.doi.org/10.1080/08927029608024094

*Molecular Simulation*, 1996, Vol. 17, pp. 57–66
Reprints available directly from the publisher
Photocopying permitted by license only

# FAST VERSION OF THE STRUCTURE ADAPTED MULTIPOLE METHOD-EFFICIENT CALCULATION OF ELECTROSTATIC FORCES IN PROTEIN DYNAMICS

CHRISTOPH NIEDERMEIER and PAUL TAVAN[1]

*Institut für Medizinische Optik, Theoretische Biophysik Ludwig-Maximilians-Universität München Theresienstraße 37, D-80333 München, Germany*

*(Received September 1995; accepted September 1995)*

Traditional algorithms for evaluation of electrostatic forces in molecular dynamics simulations of proteins require computational work of the order $O(N^2)$ for a system of $N$ atoms. Truncation methods, which try to avoid that effort, entail intolerably large errors in forces, energies and other observables. Hierarchical multipole methods allow rapid and correct evaluation of electrostatic forces. We describe an improved version of the *Structure Adapted Multipole Method* (SAMM) which not only scales with $O(N)$ but also works as fast as truncation algorithms. Unlike the latter it manages to reproduce electrostatic forces with an error below 1%.

KEY WORDS: Molecular dynamics, protein dynamics, electrostatic forces, hierarchical multipole methods, structure adaptivity.

## 1 INTRODUCTION

Molecular dynamics (MD) simulations of proteins currently are restricted to relatively small systems of some thousand atoms and short simulation periods of a few nanoseconds. A description of biologically important processes which occur on much longer time scales requires the development of correspondingly more efficient algorithms. Of particular importance is the computation of long-range electrostatic interactions because, when carried out exactly that computation scales with $O(N^2)$. Commonly used truncation methods scale linearly with $N$ but lead to considerable errors in forces, energies and other observable quantities. These errors are too large for acceptable descriptions of proteins [1–3].

In contrast, hierarchical multipole methods allow a much more accurate computation of electrostatic forces with an algorithmic complexity of $O(N \log N)$ or even $O(N)$. In particular, the *Fast Multipole Method* (FMM) proposed by Greengard and Rohklin [4, 5] allows to calculate forces at arbitrary numerical accuracy. FMM uses a hierarchy of cubic grids for decomposing space at different levels of resolution. For each cell, a multipole expansion is used to represent the electrostatic force field

---

[1] To whom reprint requests should be addressed.

arising from the charges located within the cell. Use of a sophisticated scheme of hierarchical Taylor expansions for calculating forces from the multipole expansions enables the aforementioned algorithmic complexity of $O(N)$. Unfortunately, as compared to truncation methods considerable computational overhead is associated with FMM [6–8].

It is the purpose of this contribution to show how FMM strategies can be employed for the design of a method that allows computation of electrostatic forces much more accurate but as efficient as truncation methods. Instead of decomposing a system rigidly into a hierarchy of cubic cells, that methods takes advantage of intrinsic dynamical and structural properties of proteins for the decomposition. The improved method presented below scales with $O(N)$ and extends the *Structure Adapted Multipole Method* (SAMM) [9] which exibited an $O(N \log N)$ scaling behavior.

For a description of our new methods, we first summarize the basic features of SAMM. After that we explain how SAMM and FMM are merged into an enhanced algorithm whose accuracy and computational efficiency we compare with that of other methods.

## 2 SAMM ALGORITHM

Proteins are characterized by a hierarchy of structural properties. SAMM takes advantage of these properties in order to achieve an efficient hierarchical description of the electrostatic force field. In the following, we will briefly explain which structural properties of proteins are exploited for the purpose and how electrostatic forces are calculated. For closer details see Refs. [9, 10].

The basic idea of SAMM–and hierarchical multipole methods in general–is to approximately calculate the electrostatic force field arising from more distant parts of a system by multipole expansions of their charge distribution. At a chosen level of accuracy, according to the convergence criterion for multipole expansions the scales for coarse graining of the system may be chosen to increase with the distance between its interacting parts. The electrostatic potential acting on a particle at $\mathbf{r}_i$ can be expressed as

$$\Phi(\mathbf{r}_i) = \sum_{k=0}^{n} \Phi_k(\mathbf{r}_i), \tag{1}$$

where $\Phi_k(\mathbf{r}_i)$ denotes the electrostatic potential arising from parts of the system in a distance range $[R_k, R_{k+1}]$. These distance ranges correspond to levels $k$ of a structural hierarchy. As far as traditional multipole algorithms are concerned, that hierarchy is defined somewhat arbitrarily by a series of cubic grids. SAMM however employs a structure adapted hierarchical scheme as described in Figure 1. Groups of particles at different levels of the hierarchy are called *objects*. The left and central columns of the Figure show the object hierarchy and its definition by a bottom-up strategy. The lowest level of the hierarchy (level 0) corresponds to interactions between partially charged atoms at small distances. Level 1 corresponds to interactions

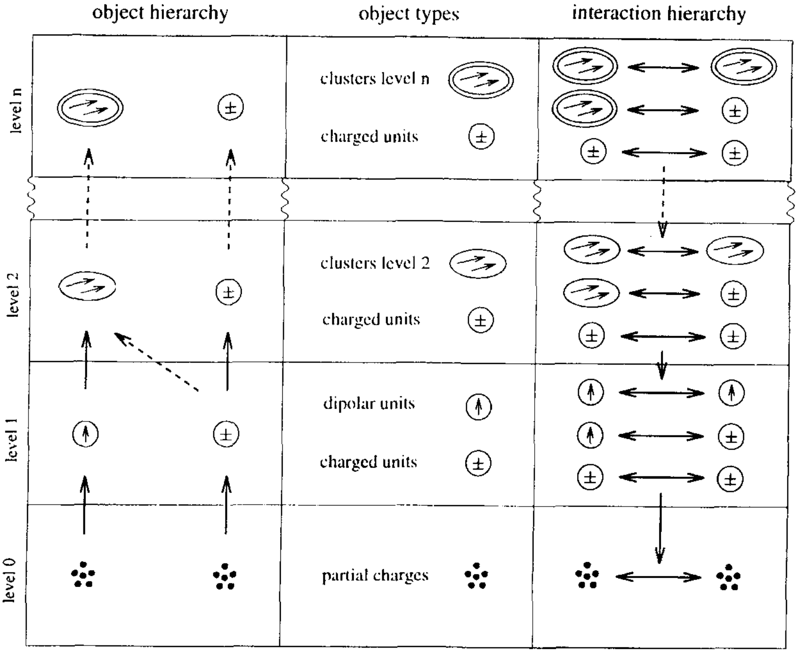

**Figure 1** Structure adapted hierarchical description of polar biological macromolecules.

between small chemical building blocks, so-called *structural units*. SAMM distinguishes between three types of structural units: charged, dipolar and neutral. Whereas the first two types are relevant for electrostatic interactions, neutral units exclusively consist of uncharged atoms which are subjected only to binding and van der Waals interactions. Therefore they are not shown in the Figure. Further below we will explain how these types are defined for protein-water systems. The objects on level 2 are either charged units or *clusters* of dipolar units. In the current implementation of SAMM charged units are not clustered because in proteins their number is small compared to the number of dipolar units. Clusters of dipolar units are formed by using self-organizing and adaptive vector quantization techniques (see Ref. [10]). Currently, the algorithm is restricted to three levels ($n = 2$) but more levels may be introduced [12] as indicated in Figure 1. The right column of the Figure shows how interaction lists are constructed. Objects on level $k$ interact with each other if their center-to-center distance is larger than $R_k$. Otherwise, they are split up into subobjects defined on the next lower level. By using a top-down strategy which starts at the topmost level interaction lists for all levels are obtained. Based on these interaction lists we constructed two alternative methods for evaluation of electrostatic forces. The previous method proposed by us in [9] will be sketched below. The improved methods adopted from FMM will be described in the next section.

For protein-water systems, the three types of structural units mentioned above are defined as follows: (i) water molecules and peptide groups form dipolar units; (ii) charged amino acid side chains as well as N— and C— terminal groups of the polypeptide chain form charged units; (iii) uncharged amino acid side chains form either neutral or dipolar units. Other groups as chromophores or lipids can be represented similarly (see Ref. [10]). Figure 2 illustrates this procedure for a peptide sequence of three amino acids. In the Figure, only those polar hydrogens that are explicitly represented in MD simulations are shown. The CHARMM parameter set PARAM19 is used for assignment of partial charges to atoms [11].

The electrostatic field of charged and dipolar objects is represented in a simple way: Only the lowest non-vanishing multipole moments are taken into account whereas all higher moments are neglected. Accordingly, charged objects are modeled by charges, dipolar objects by dipoles. Since in contrast to all higher multipole moments the lowest non-vanishing multipole moments are independent of the reference point, that representation can be optimized by selecting the reference point such that the contribution of the lowest neglected multipole moment is minimized [9]. For a charged unit that criterion simply yields the center-of-charge

$$\mathbf{r}_c = \frac{\sum_i q_i \mathbf{r}_i}{\sum_i q_i} \tag{2}$$

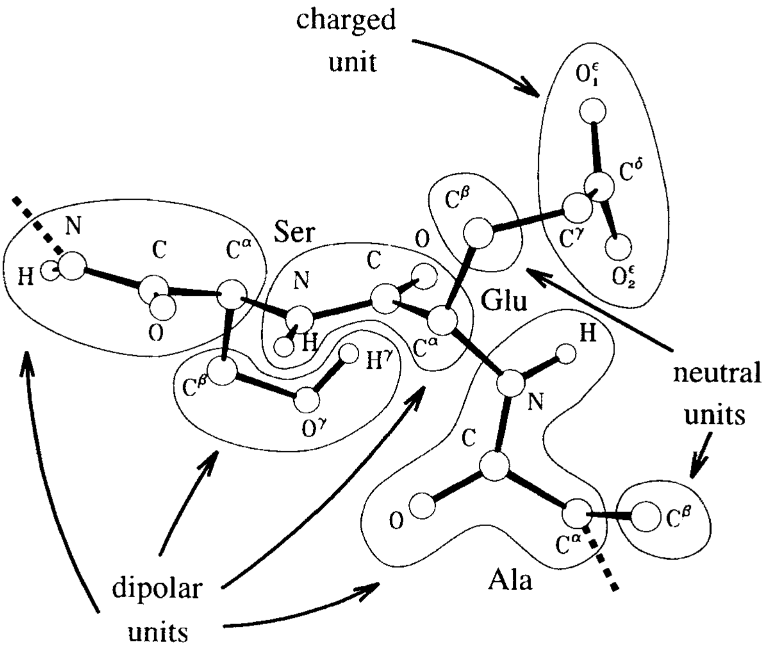

Figure 2 Structural units for the amino acid sequence X-Ala-Glu-Ser-Y.

as its reference point, where $\mathbf{r}_i$ denotes the position and $q_i$ the partial charge of atom $i$, and the dipole moment vanishes. For a dipolar unit one finds for the optimal reference point

$$\mathbf{r}_d = \frac{1}{3p^2}\bar{\mathrm{Q}}^0\mathbf{p} - \frac{1}{12p^4}(\mathbf{p}\bar{\mathrm{Q}}^0\mathbf{p})\mathbf{p}, \tag{3}$$

with $\mathbf{p} = \sum_i q_i \mathbf{r}_i$ and $Q^0_{\alpha\beta} = \sum_i q_i (3r_{i\alpha} r_{i\beta} - r_i^2 \delta_{\alpha\beta})$. As a result the largest principal component of the quadrupole tensor vanishes, *i.e.* it becomes "flat".

Evaluation of electrostatic forces follows a pseudo-particle scheme adopted from stellar dynamics whose scaling behavior is $O(N \log N)$. The scheme is illustrated in Figure 3. Each partial charge in the system interacts with partial charges closer than $R_1$, with structural units in the range $[R_1, R_2]$, with aggregates and charged units (level 2) etc. The interaction partners at a certain level are determined by treating each pair of objects A and B in the interaction list of that level (see Figure 1) as follows: Partial charges contained in object A interact with the charge or dipole moment of object B and vice versa. Thus, partial charges directly interact with objects of the hierarchy; these objects are looked upon as "pseudo-particles" carrying either charges or dipole moments.

## 3 IMPROVEMENTS

A faster method for computation of electrostatic forces can be adopted from FMM [5]. That method employs local descriptions of the electrostatic field by means of Taylor expansions. At each level of the hierarchy, such *local expansions* are determined for all objects where the centers of the objects are used as reference points.

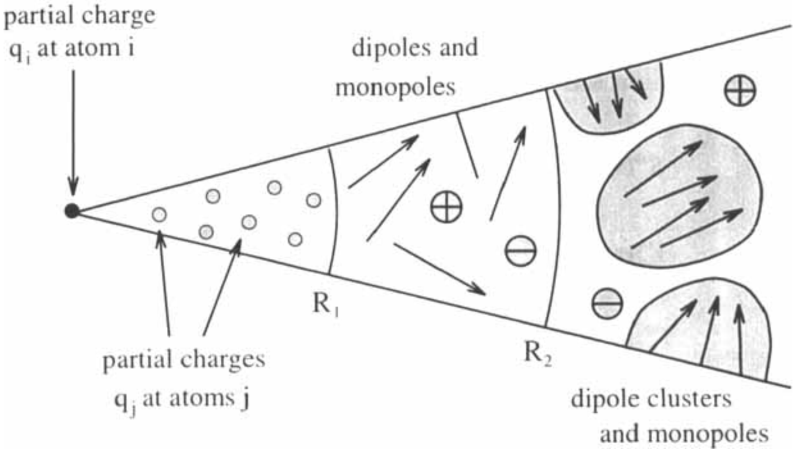

**Figure 3** Pseudo-particle scheme for evaluation of electrostatic forces.

Interactions between objects are evaluated by transforming the multipole expansion referring to an object into a local expansion at the other object. As compared to the pseudo-particle method mentioned above such a treatment of interactions contains an additional approximation and, therefore, is expected to be less accurate. On the other hand, local expansions can be determined by a hierarchical strategy that enables the desired overall $O(N)$ scaling behavior.

Local Taylor expansions of the electrostatic potential are determined with respect to the optimized reference points $\mathbf{x}$ of the respective objects. We found that a local expansion up to the second order in the relative position vector $\rho$ is sufficiently accurate for our purposes. Correspondingly, the Taylor expansion of the potential at position $\mathbf{x}+\rho$ reads

$$\tilde{\Phi}(\mathbf{x}+\rho)=S^{\Phi}(\mathbf{x})+\sum_{\alpha=1}^{3} V_{\alpha}^{\Phi}(\mathbf{x})\rho_{\alpha}+\frac{1}{2}\sum_{\alpha=1}^{3}\sum_{\beta=1}^{3} T_{\alpha\beta}^{\Phi}(\mathbf{x})\rho_{\alpha}\rho_{\beta}, \tag{4}$$

where $\rho_\alpha$ denote the components of $\rho$. Corresponding expansions of the cartesian components of the force $\tilde{F}_\gamma(\mathbf{x}+\rho)$ are obtained by computing the gradient of $\tilde{\Phi}(\mathbf{x}+\rho)$ with respect to $\mathbf{x}$. We have to determine these expansions only for interactions with charges and dipoles. The computational effort associated with this task can be reduced by exploiting relations between the coefficients of the expansion of the potential and those of the force components as well as symmetry relations (see Ref. [10]).

For interaction with a charge $Q$ the coefficients of the potential read

$$S^{\Phi}(\mathbf{x})=\frac{Q}{|\mathbf{x}|} \tag{5}$$

$$V_{\alpha}^{\Phi}(\mathbf{x})=-\frac{Q}{|\mathbf{x}|^{3}}x_{\alpha} \tag{6}$$

$$T_{\alpha\beta}^{\Phi}(\mathbf{x})=\frac{3Q}{|\mathbf{x}|^{5}}x_{\alpha}x_{\beta}-\frac{Q}{|\mathbf{x}|^{3}}\delta_{\alpha\beta} \tag{7}$$

whereas for interaction with a dipole $\mathbf{p}$ we obtain

$$S^{\Phi}(\mathbf{x})=\frac{\mathbf{px}}{|\mathbf{x}|^{3}} \tag{8}$$

$$V_{\alpha}^{\Phi}(\mathbf{x})=-\frac{3\mathbf{px}}{|\mathbf{x}|^{5}}x_{\alpha}+\frac{1}{|\mathbf{x}|^{3}}p_{\alpha} \tag{9}$$

$$T_{\alpha\beta}^{\Phi}(\mathbf{x})=\frac{15\mathbf{px}}{|\mathbf{x}|^{7}}x_{\alpha}x_{\beta}-\frac{3}{|\mathbf{x}|^{5}}(p_{\alpha}x_{\beta}+p_{\beta}x_{\alpha})-\frac{3\mathbf{px}}{|\mathbf{x}|^{5}}\delta_{\alpha\beta} \tag{10}$$

Corresponding coefficients for the force can be easily derived.

Figure 4 explains the hierarchical strategy of force computation based on local expansions. For the sake of simplicity, we discuss the case $n=2$, *i.e.* a hierarchy with three levels. At each level $k>0$, the interactions specified by the corresponding

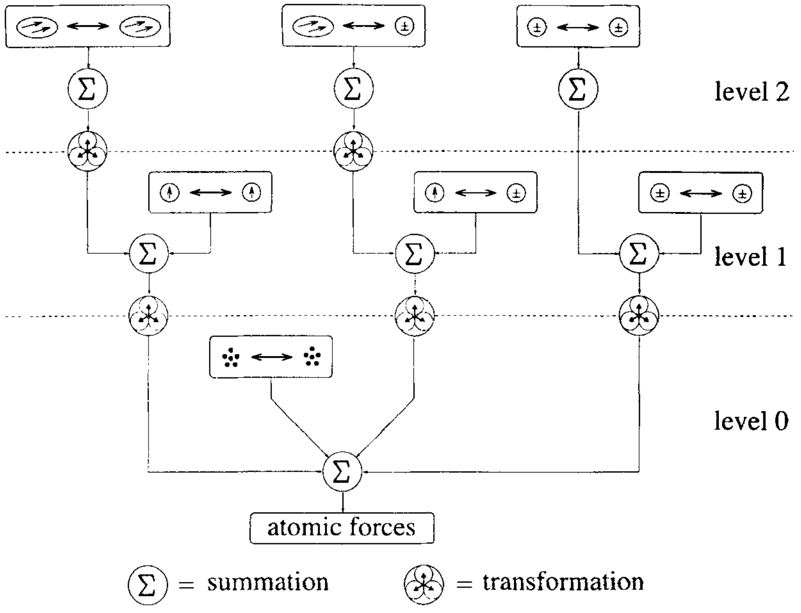

**Figure 4** Fast strategy for evaluation of forces demonstrated for a three-level hierarchy.

interaction list are converted into local expansions. Those expansions are added up for each object (symbol "summation"). Additionally, each object inherits a contribution from the next higher level which is obtained by transforming the local expansion of its parent object (symbol "transformation"). The transformation consists in a shift of the reference point of the expansion from the center $\mathbf{x}$ of the parent object to the center $\mathbf{x}'$ of the child. The transformation rules for the coefficients of the potential are given by

$$S'^{\Phi}(\mathbf{x}') = S^{\Phi}(\mathbf{x}) + \sum_{\alpha=1}^{3} V_{\alpha}^{\Phi}(\mathbf{x})\Delta x_{\alpha} + \frac{1}{2}\sum_{\alpha=1}^{3}\sum_{\beta=1}^{3} T_{\alpha\beta}^{\Phi}(\mathbf{x})\Delta x_{\alpha}\Delta x_{\beta}, \tag{11}$$

$$V'^{\Phi}_{\alpha}(\mathbf{x}') = V_{\alpha}^{\Phi}(\mathbf{x}) + \sum_{\beta=1}^{3} T_{\alpha\beta}^{\Phi}(\mathbf{x})\Delta x_{\beta}, \tag{12}$$

$$T'^{\Phi}_{\alpha\beta}(\mathbf{x}') = T_{\alpha\beta}^{\Phi}(\mathbf{x}), \tag{13}$$

where $\Delta x_{\alpha}$ denote the components of the shift vector $\Delta\mathbf{x} = \mathbf{x}' - \mathbf{x}$. The components of the force transform accordingly (see Ref. [10]). Local expansions at charged units do not have to be transformed (see right part of Figure 4) since these units are not clustered. The top-down procedure for determination of local expansions stops at level 1. There, long-range electrostatic interactions are computed by evaluating the local expansions at the positions of individual particles (also marked with symbol "transformation" at the transition to level 0). Short-range contributions are computed by evaluating pair interactions with nearby particles as specified by the interaction list for level 0. By adding up both contributions, total electrostatic energies and forces are obtained.

Efficiency and accuracy of the algorithm described above were investigated by performing several MD simulations of the membrane protein bacteriorhodopsin (BR) which consists of about 2400 particles, 1800 of them charged. MD simulations were carried out over 10000 time steps of 1fs. We compared the following algorithms for computation of electrostatic interactions:

| | |
|---|---|
| REF: | direct evaluation of all pair interactions |
| CUT: | truncation algorithm with cutoff radius 10 Å |
| SAMM-1: | SAMM with pseudo-particle method ($R_1 = 10\,\text{Å}, R_2 = 16\,\text{Å}$) |
| SAMM-2: | SAMM with fast multipole method ($R_1 = 10\,\text{Å}, R_2 = 16\,\text{Å}$) |

Beside electrostatic interactions van der Waals interactions are treated differently as well: REF and CUT compute van der Waals interactions along with electrostatic interactions whereas the SAMM algorithms consider van der Waals interactions separately by a truncation algorithm within a cutoff distance of 8 Å

We investigate accuracy and computational efficiency of methods CUT, SAMM-1 and SAMM-2 by comparisons with REF. To assess the numerical error of energy contributions we consider the rms error of all atomic energies

$$\varepsilon = \left( \frac{\sum_{i=1}^{N} |\tilde{E}_i - E_i|}{\sum_{i=1}^{N} |E_i|^2} \right)^{1/2}, \tag{14}$$

where $E_i$ denotes the energy contribution of atom $i$ as determined by REF, $\tilde{E}_i$ the approximate value and $N$ the number of particles. Analogously, we use

$$\eta = \left( \frac{\sum_{i=1}^{N} |\tilde{\mathbf{F}}_i - \mathbf{F}_i|}{\sum_{i=1}^{N} |\mathbf{F}_i|^2} \right)^{1/2} \tag{15}$$

for the investigation of forces. Computational efficiency is measured by $\tau = \Delta T / \Delta \tilde{T}$ where $\Delta T$ denotes the average computer time per MD time step for simulation REF and $\Delta \tilde{T}$ the corresponding computer time for the method under consideration.

Table 1 shows computational efficiencies[1] as well as rms errors of various contributions to atomic forces and energies. The error values were obtained by taking time averages of rms errors determined every 10 time steps. The data show that SAMM-2 produces only slightly larger errors than SAMM-1 whereas the errrors produced by CUT are larger by more than a factor 15. Only van der Waals forces and energies are represented more accurately by CUT because the cutoff distance is larger in that case. Furthermore the table shows the disproportion between the accuracy of van der Waals forces and electrostatic forces produced by CUT whereas

---

[1] Benchmarking runs were performed on a HP 735 workstation; other machines may produce slightly different results.

**Table 1** Computational efficiency and time-averaged errors of force and energy contributions. "ES" denotes electrostatic contributions, "vdW" van der Waals contributions and "total" total forces.

| method | error in % | | | | | | efficiency |
|---|---|---|---|---|---|---|---|
| | forces | | | energies | | | |
| | vdW | ES | total | vdW | ES | vdW + ES | |
| CUT | 0.175 | 12.9 | 3.68 | 5.73 | 20.6 | 20.1 | 8.11 |
| SAMM-1 | 0.562 | 0.804 | 0.256 | 13.5 | 1.04 | 1.51 | 6.24 |
| SAMM-2 | 0.560 | 0.876 | 0.276 | 13.5 | 1.09 | 1.54 | 8.09 |

SAMM exhibits no such disproportion. The two versions of the SAMM algorithm differ clearly in computational efficiency. SAMM-1 is about 30% slower than SAMM-2 which is as fast as CUT.

## 4 CONCLUSIONS

We demonstrated in this paper that the computational efficiency of SAMM can be increased by combining the method with an algorithmic strategy adopted from FMM. Truncation of Taylor expansions after the second order term produces errors which are only slightly larger than those produced by the previous version of SAMM. However, the computational efficiency of the method is increased considerably. The new version of SAMM is now as fast as truncation methods with a cutoff distance of 10 Å but manages to reduce errors of electrostatic forces by more than a factor of 15 to somewhat less than 1%.The rms error of the total forces turns out to be as low as 0.25%. This represents a considerable progress in comparison with truncation methods. However, whether SAMM is suited for replacing truncation algorithms not only depends on its accuracy but also on its influence on dynamical properties of the simulation system. Due to the approximations employed artificial noise forces arise which could spoil the quality of the simulation in spite of the improved accuracy of force calculation. We will address this problem and its solution in a forthcoming publication [10, 13].

### *Acknowledgements*

C. N. is grateful to the Boehringer Ingelheim Fonds for a fellowship; the study has been supported by the Deutsche Forschungsgemeinschaft (SFB 143/C1).

### *References*

[1] L. Charles, Brooks III, B. Montgomery Pettitt, and Martin Karplus. Structural and energetic effects of truncating long ranged interactions in ionic and polar fluids. *J. Chem. Phys.*, **83**, 5897–5908 (1985).

[2] R. J. Loncharich and B. R. Brooks. The effects of truncating long-range forces on protein dynamics. *Proteins*, **6**, 32–45 (1989).

[3] Helmut Grubmüller and Paul Tavan. Efficient algorithms for molecular dynamics simulations of proteins: How good are they? (submitted).

[4] L. Greengard and V. Rohklin. A fast algorithm for particle simulations. *J. Comp. Phys.*, **73**, 325–348 (1987).

[5] L. Greengard and V. Rohklin. Rapid evaluation of potential fields in three dimensions. Research Report RR-515, YALEU/DCS, Yale University, New Haven (1987).

[6] James F. Leathrum and John A. Board. The parallel fast multipole algorithm in three dimensions. Technical report, Dept. of Electrical Engineering, Duke University, Durham (1992).

[7] K. E. Schmidt and M. A. Lee. Implementing the fast multipole method in three dimensions. *J. Stat. Phys.*, **63**, 1223–1235 (1991).

[8] H. -Q. Ding, N. Karasawa and W. A. Goddard III. Atomic level simulations on a million particles: The cell multipole method for Coulomb and London nonbonded interactions. *J. Chem. Phys.*, **97**, 4309–4315 (1992).

[9] C. Niedermeier and P. Tavan. A structure adapted multipole method for electrostatic interactions in protein dynamics. *J. Chem. Phys.*, **101**, 734–748 (1994).

[10] C. Niedermeier. *Effiziente Modellierung elektrostatischer Wechselwirkungen in Proteindynamikrechnungen.* Dissertation, Ludwig-Maximilians-Universität München, Munich, December (1994).

[11] Bernhard R. Brooks, Robert E. Bruccoleri, Barry D. Olafson, David J. States, S. Swaminathan and Martin Karplus. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comp. Chem.*, **4**, 187–217 (1983).

[12] M. Eichinger. Paralleler schneller Multipolalgorithmus mit Mehrschrittverfahren für Molekulardynamiksimulationen. Diplomarbeit, Ludwig-Maximilians-Universität München, Munich, April (1995).

[13] C. Niedermeier and P. Tavan. The structure adapted multipole method–refinements for suppression of algorithmic artifacts in molecular dynamics simulations. (in preparation).