An efficient linear scaling algorithm for tight binding molecular dynamics

S. Goedecker
Cornell Theory Center, Ithaca,NY 14853-3801,USA

L. Colombo
Dipartimento di Fisica, Università di Milano
via Celoria 16, I-20113 Milano, Italy

(Received)

Abstract - A novel formulation for tight binding total energy calculations and tight binding molecular dynamics, which scales linearly with the size of the system, is presented. The linear complexity allows us to treat systems of very large size and the algorithm is already faster than the best implementation of classical diagonalization for sytems of 64 atoms. In addition, it is naturally parallelizable and it permits us therefore to perform molecular dynamics simulations of systems of unprecedented size. Finite electronic temperatures can also be taken into account. We illustrate this method by investigating structural and dynamical properties of solid and liquid carbon at different densities.

PACS numbers: 71.10+x, 71.20.Ad, 61.20.Ja, 71.25.Cx.

The tight binding method[1] (TBM) is one of the most widely used electronic structure methods. The method and its variants such as the extended Hückel method[2] have been applied to a wide range of systems ranging from transition metals to covalent materials. When a suitable repulsive potential is added in the total energy expression, TBM can also be used for tight binding molecular dynamics (TBMD) simulations[3–6]. Even though the TBM is not as accurate as ab initio density functional and Car-Parrinello molecular dynamics simulations[7] (CPMD) calculations it is a useful tool since the computational effort is significantly reduced in comparison with ab initio methods and it is therefore possible to treat bigger systems or to follow the evolution of the system over more time steps in a molecular dynamics simulation. Consequently, a large class of problems relevant to condensed matter physics which fall out of reach of CPMD can be studied by means of TBMD. In spite of this, the application of TBDM to *very large* systems (hundreds or thousands of atoms) was hindered because the conventional formulation of TBMD had a cubic scaling with respect to the number of atoms N_{at} in the system and because there were no algorithms which could take advantage of the powerful parallel computers. Depending on whether standard diagonalization techniques (SDT) or Car–Parrinello type fictitious Lagrangian dynamics are used, the cubic term arises either from the matrix diagonalization or the orthogonalization of the wave functions. This problem of the cubic scaling has been recognized by the computational physics community and several proposals have recently been put forward to overcome this bottleneck. They are either based on the concept of localized orbitals[8] or on the calculation of the density matrix[9].

In this paper we present a novel formulation of TB based methods, in which the computational workload scales linearly with N_{at}. In contrast to other approaches it is based on a projection step rather than on minimization. The projection approach requires much less interprocessor communication when it is implemented on a parallel machine than the current minimization approaches[10]. Furthermore it is free of numerically cumbersome local minima which are found in restricted minimization schemes. Our resulting computational framework is extremely efficient and can be applied to both TB total energy calculations and TBMD simulations of systems with unprecedented size.

Within the TBMD scheme the total energy E_{tot} of a given system is expressed as[3–6]

$$E_{tot} = \sum_{\alpha} \frac{p_{\alpha}^2}{2m_{\alpha}} + \sum_i \epsilon_i f(\frac{\epsilon_i - \mu}{kT}) + U_{rep} \, , \tag{1}$$

where ϵ_i are the single–particle energies obtained from the TB Hamiltonian H, $f(\frac{\epsilon_i-\mu}{kT})$ is the Fermi–Dirac distribution function, and U_{rep} is a suitable effective potential. During a conventional TBMD simulation most of the computational work is spent in calculating the eigenvalues $\{\epsilon_i\}$ and the eigenvectors $\{\Psi_i\}$ of H. They respectively enter in the band structure term

$$E_{bs} = \sum_i \epsilon_i f(\frac{\epsilon_i - \mu}{kT}) \tag{2}$$

and in the Hellman–Feynman (HF) force term which is the derivative of E_{bs} with respect to atomic displacements $\vec{r}_\alpha$

$$\vec{f}_\alpha = \sum_i < \Psi_i | \frac{\partial H}{\partial \vec{r}_\alpha} | \Psi_i > f(\frac{\epsilon_i - \mu}{kT}) \,. \tag{3}$$

By introducing the Fermi matrix $F_{\mu,T}$

$$F_{\mu,T} = f(\frac{H-\mu}{kT}) \tag{4}$$

these two quantities can be expressed as

$$E_{bs} = Trace[HF] = \sum_{l\alpha} \sum_{l'\alpha'} < \varphi_{l\alpha} | H | \varphi_{l'\alpha'} >< \varphi_{l'\alpha'} | F | \varphi_{l\alpha} > \tag{5}$$

$$\vec{f}_\alpha = Trace[\frac{\partial H}{\partial \vec{r}_\alpha} F] = \sum_{l\alpha} \sum_{l'\alpha'} < \varphi_{l\alpha} | \frac{\partial H}{\partial \vec{r}_\alpha} | \varphi_{l'\alpha'} >< \varphi_{l'\alpha'} | F | \varphi_{l\alpha} > \tag{6}$$

where $\varphi_{l\alpha}$ is the l–th TB basis set function (Löwdin orbital) centered on the α–th atom site. The above expressions for E_{bs} and $\vec{f}_\alpha$ follow readily from the fact that the trace is invariant under the unitary transformation from the Ψ_i–representation to the $\varphi_{l\alpha}$–representation, and that the Fermi matrix is diagonal in the basis set of the eigenfunctions Ψ_i. The evaluation of the traces is trivial and scales linearly with N_{at}. It remains now to be shown how the above defined Fermi matrix and the matrix products in eqs.(5,6) can be calculated in a numerically efficient way.

The Fermi distribution can be approximated in the interval spanned by the lowest ϵ_{min} and highest ϵ_{max} eigenvalue of H by a polynomial $p_F(x)$ of degree n_{pl}. Once this polynomial is known, each column of the Fermi matrix is calculated by applying the matrix polynomial $p_F(H)$ to $\varphi_{l\alpha}$. This operation requires a computer time proportional to $n_{pl} \times N_{at}$. Since the dimension of the Fermi matrix is proportional to N_{at}, the method as outlined above implies a computational effort proportional to $n_{pl} N_{at}^2$, i.e an $O(N_{at}^2)$ scaling with respect to the number of atoms.

However, in insulators and disordered metallic liquids an $O(N_{at})$ scheme can be obtained if one takes advantage of the fact that the elements of a column $F|\varphi_{l\alpha}>$ of the Fermi matrix (hereafter refered to as a *localized orbital* (LO)) decay exponentially as shown in Fig.1. To get a linear scheme, one simply zeroes each column element of LOs corresponding to those atoms falling outside a sufficiently large sphere centered at the α atom. We name this sphere the *localization volume*. The localized orbitals are not to be confused neither with TB basis functions nor with eigenfunctions Ψ_i. By making use of the cut off LO's, all the matrix–vector multiplications needed to compute the Fermi matrix and eqs.(5,6) have to be performed only in the localization volume which is much smaller than the volume of the simulation cell for large systems.

Ordered metals can in principle also be treated with a linear scaling if a nonzero electronic temperature is used. In this case the off diagonal elements of the density matrix again decay exponentially whereas they decay only algebraically at zero electronic temperature. However, unless one uses very high electronic temperatures the exponential decay is very slow compared to most insulators and the localization volume will be equal to the volume of the system even for rather large systems. One would therefore observe a quadratic scaling unless one treats gigantic systems, where one would again have linear scaling. Preliminary tests for this kind of metallic systems indicate that our method is faster than standard diagonalization for systems containing several hundred atoms.

The degree n_{pl} of the polynomial is of the order of $n_{pl} \approx (\epsilon_{max} - \epsilon_{min})/kT$. The method becomes therefore slower and slower as one goes to lower temperatures. In most applications, however, the relevant physical properties are only slightly affected by the electronic temperature and one can therefore chose an electronic temperature which is higher than the physical one in order to further improve the numerical efficiency. In the case of an insulator, it is not necessary to reproduce the Fermi distribution in the region of the band gap and one can find a polynomial of degree $n_{pl} \approx (\epsilon_{max} - \epsilon_{min})/\epsilon_{gap}$, where ϵ_{gap} is the energy gap, which gives good accuracy in the valence band region even at very low temperatures.

We have applied our method to solid and liquid carbon in the high (ρ=4.4 g/cm^3) and low density (ρ=2.0 g/cm^3) phase. Carbon is a system of interest in many field of science and its phase diagram is still a matter of discussion. Moreover, there are CPMD simulation[11,12] to compare with in order to establish the reliability of the present computational framework. The two–center TB hopping integrals required in eqs.(5,6) have been determined on a minimal basis set consisting of four sp^3 orbital for each atom[13] and scaled upon interatomic distance according the Harrison's rule.[1] The repulsive potential U_{rep} (see eq.(1)) has been fitted onto the cohesive energy curves of carbon in diamond and graphite structures and for the linear chain. Further details will be published in a following paper.

The large bandgap ($\approx$ 8eV) of diamond allows us to use a polynomial of rather low degree ($n_{pl} \approx 50$) and produces LOs which decay very rapidly (see Fig.1). We cut off the localized orbitals at a radius of $\approx$ 11 atomic units. With this choice of parameters, the evaluation of E_{bs} was correct to within less than one meV/atom when compared to a calculation using standard diagonalization technique (SDT). In a microcanonical simulation of 1000-atom crystal diamond the total energy was conserved to within 10^{-4} eV/atom at a temperature of 300K. In order to prove the high numerical efficiency of our method, we report in Figs.2 and 3 the timing results for one time step of a TBMD run. In Fig.2 we show that as soon as the volume of the simulation cell gets larger than the localization volume of LOs ($N_{at} \geq 64$) a nearly linear behaviour is observed. As a matter of fact, the method is faster than SDT even in the case of small 64 atom systems.

The corresponding timings with SDT (LAPACK library) are 5.4 sec for 64 atoms, 164 sec for 216 atoms, 54 min for 512 atoms and 253 min for 1000 atoms. In the case of a 1000 atom system, we reduced the execution time from more than 4 hours to less than 4 minutes. In addition, the memory requirements are also drastically reduced: In our implementation, the calculation of a 1000 atom system requires only 5MByte of memory. Since each localized orbital can be calculated independently from the others, our method is trivial to parallelize. The timings are shown in Fig.3. The computational speed on 25 IBM SP1 processors was over one giga FLOP. One TBMD time step on 1000 atoms, takes now roughly ten seconds on the parallel computer with our $O(N_{at})$ algorithm.

The simulation for carbon were carried out with a periodically repeated cubic cell containing 512 atoms. We used a time step as small as 0.5 10^{-15}s and the sampling was done during 2000 time steps after 4000 time steps of equilibration. In Fig.4 we report the particle–particle correlation function $g(r)$ as calculated at 6500 K for the high–density phase. Well defined peaks corresponding to the shells of neighbouring atoms can be observed up to a distance of $\approx$ 12 atomic units. The system does not display any tendency to melt as confirmed by the atomic mean square displacement. The excellent agreement to short range CPMD data[11] (dashed line) confirms the reliability of the present calculation. In Fig.5 we show the $g(r)$ for low density liquid carbon at 5000K. It is clearly seen that the crystal structure has been completely lost, apart from the short–range features which, again, are in rather good agreement to first–principle data (dashed line).[12] Moreover, the liquid structure has been found to be dominated by two–fold and three–fold coordinated atoms, as in Ref.[12]. The diffusive behaviour results into a self diffusion coefficient of 3.5×10^{-4} cm^2 s^{-1} as compared with 2.4×10^{-4} cm^2 s^{-1} from the CPMD simulation.[12]. When we applied our method to small cells, which are still accessible to SDT, the results were always quasi indistinguishable from the results obtained by SDT.

In conclusion, we have presented a new linear scaling method for tight–binding molecular dynamics. We have proved that the method is extremely efficient and easy to parallelize. We have presented practical applications to carbon in different phases and at different densities.

We are grateful to Prof. M. Teter (Cornell) for helpful discussions and kind hospitality. We acknowledge computational support by Swiss Scientific Computing Center (Manno, Switzerland) and by Cornell Theory Center (Ithaca, USA). One of us (S.G.) acknowledges financial support from Swiss and U.S. National Science Foundation.

References

1. "*Many-Atom Interactions in Solids*, edited by R. M. Nieminen, M. J. Puska, M. J. Manninen, Springer, Berlin (1989); J.A. Majewski and P. Vogl in "*The Structure of Binary Compounds*", North Holland, Amsterdam (1989), p. 287

2. T. A. Albright, J. K. Burdett, and M. H. Whangbo, "*Orbital Interactions in Chemistry*, John Wiley and Sons, New York (1985)

3. R. Virkkunen, K. Laasonen, and R.M. Nieminen, J. Phys. Condens. Matter **3**, 7455 (1991)

4. C.Z. Wang, C.T. Chan, and K.M. Ho, Phys. Rev. B**45**, 12227 (1992)

5. L. Goodwin, L. Skinner and A.J. Pettifor, Europhys. Lett. **9**, 701 (1989)

6. C. Molteni, L. Colombo, and L. Miglio, Europhys. Lett. **24**, 659 (1993)

7. R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985)

8. G. Galli and M. Parrinello, Phys. Rev. Lett. **69**, 3547 (1992) ; F. Mauri, G. Galli and R. Car, Phys. Rev. B**47**, 9973 (1993) ; P. Ordejon, D. Drabold, M. Grunbach and R. Martin, Phys. Rev. B**48**, 14646 (1993) ; W. Kohn, Chem. Phys. Lett. **208**, 167 (1993) ;

9. X.-P. Li, W. Nunes and D. Vanderbilt, Phys. Rev. B**47**,10891 (1993) ; M. S. Daw, Phys. Rev. B**47**, 10895 (1993) ; A. Alavi and D. Frenkel, J. Chem. Phys. **97**, 9249 (1992) ; W. Yang, Phys. Rev. Lett. **66**, 1438 (1991) ; M. Aoki, Phys. Rev. Lett. **71**, 1438 (1993)

10. F. Khan, Bulletin of the American Physical Society **39**, 776 (1994)

11. G. Galli, R.M. Martin, R. Car, and M. Parrinello, Science **250**, 1547 (1990)

12. G. Galli, R.M. Martin, R. Car, and M. Parrinello, Phys. Rev. Lett. **63**, 988 (1989)

13. D.J. Chadi and R.M. Martin, Solid State Commun. **19**, 643 (1976)

Figure captions

Fig. 1 - The decay of a localized orbital $< \varphi_{l=2s,\beta}|F|\varphi_{l=2s,\alpha} >$ versus the distance $|\vec{r}_\alpha - \vec{r}_\beta|$ for fixed α. Full circles denote the high density crystalline phase, empty circles the low density liquid phase.

Fig. 2 - CPU time on a IBM RS6000/550 workstation to calculate E_{bs} and HF forces with the present method for cells containing carbon atoms.

Fig. 3 - Execution speed as a function of the number of processors on a IBM SP1. The system considered is a 1000 carbon atom cell.

Fig. 4 - Particle–particle correlation function $g(r)$ for solid high density carbon at 6500K. The dashed lines represent the CPMD results of Ref.[7]

Fig. 5 - Particle–particle correlation function $g(r)$ for liquid low density carbon at 5000K. The dashed line represents the CPMD results of Ref.[7].

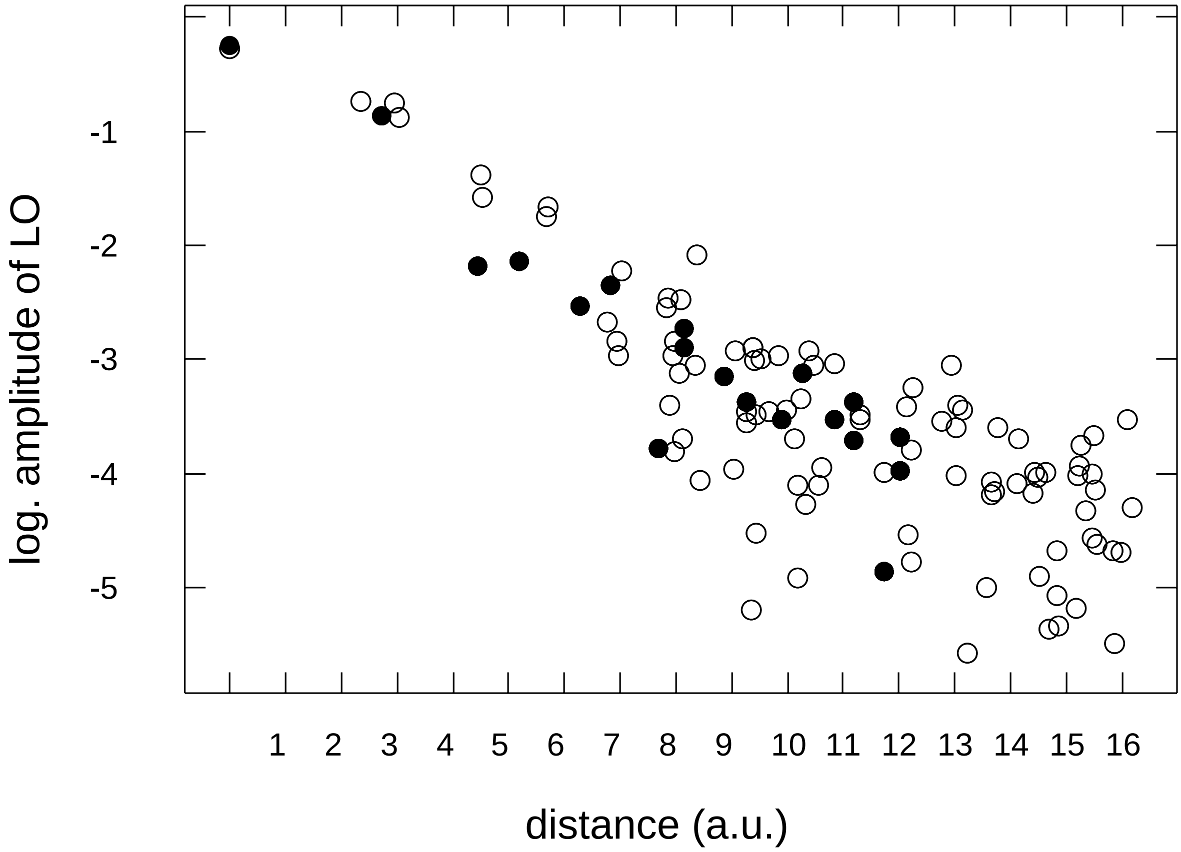

Figure 1: *The decay of a localized orbital* $<\varphi_{l=2s,\beta}|F|\varphi_{l=2s,\alpha}>$ *versus the distance* $|\vec{r}_\alpha - \vec{r}_\beta|$ *for fixed* α. *Full circles denote the high density crystalline phase, empty circles the low density liquid phase.*

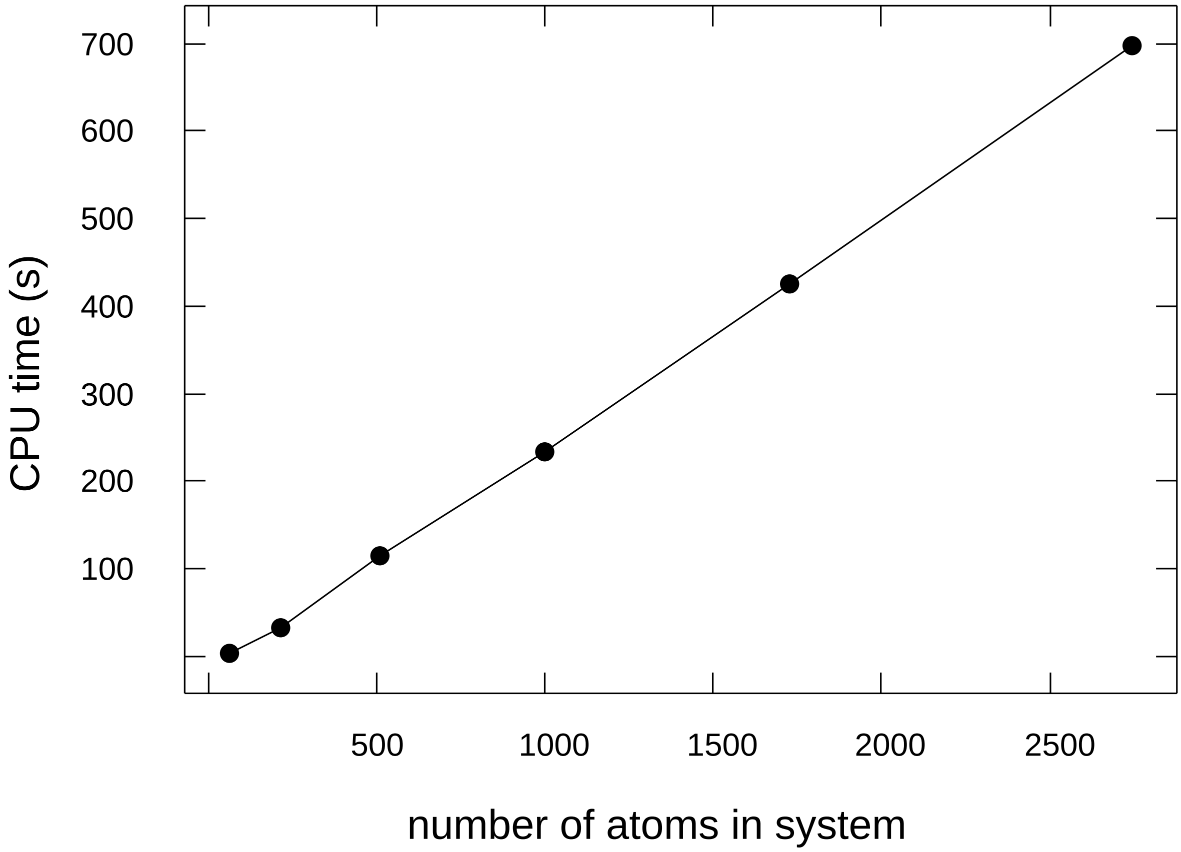

Figure 2: *CPU time on a IBM RS6000/550 workstation to calculate E_{bs} and HF forces for a cell of 1000 carbon atoms with the present method*

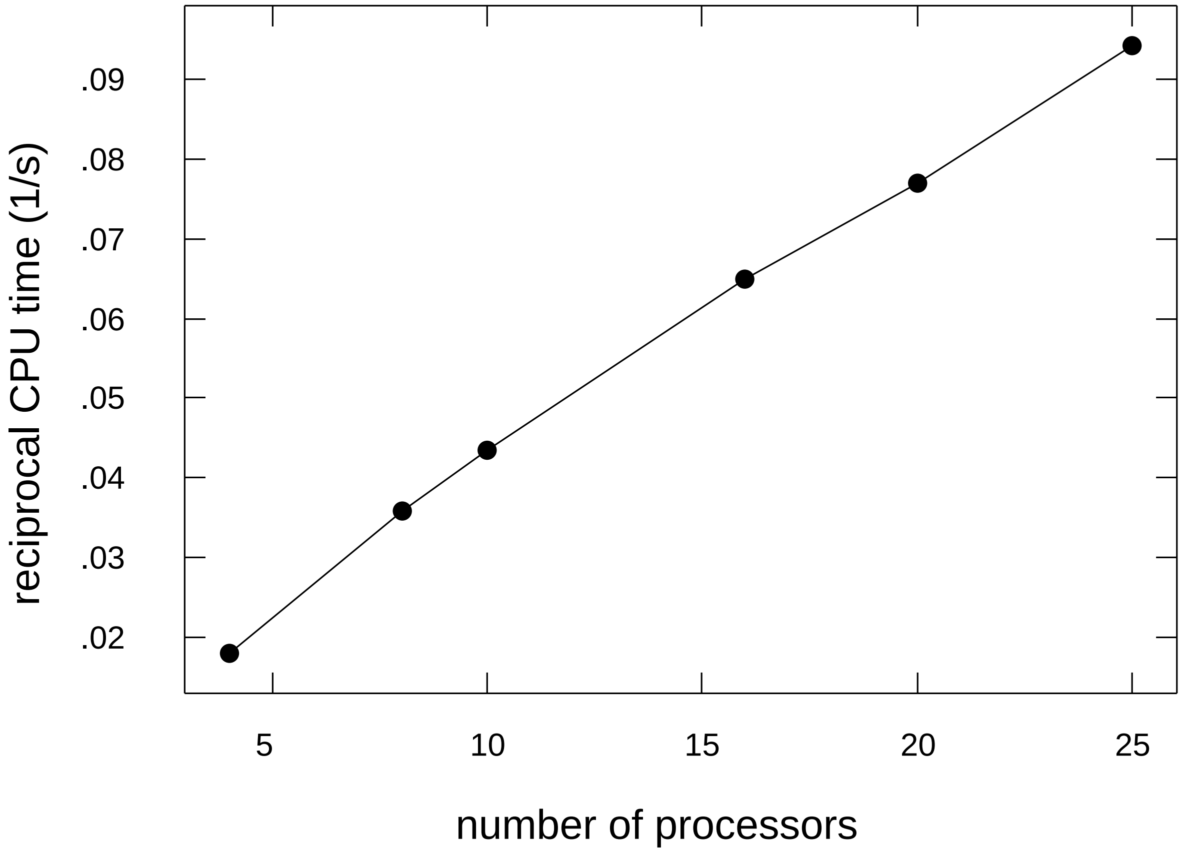

Figure 3: *Execution speed as a function of the number of processors on a IBM SP1. The system considered is a 1000 carbon atom cell*

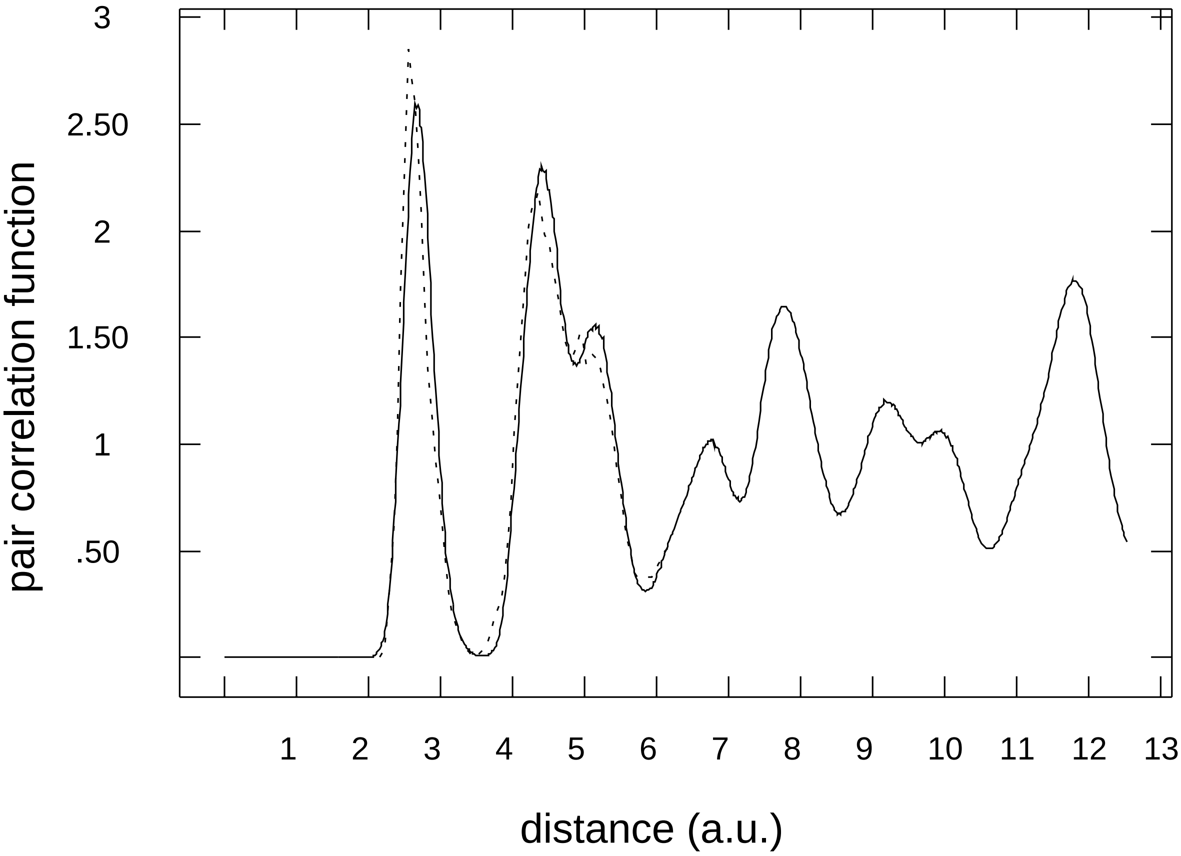

Figure 4: *Particle–particle correlation function $g(r)$ for solid high density carbon at 6500K. The dashed lines represent the CPMD results of Ref.[7]*

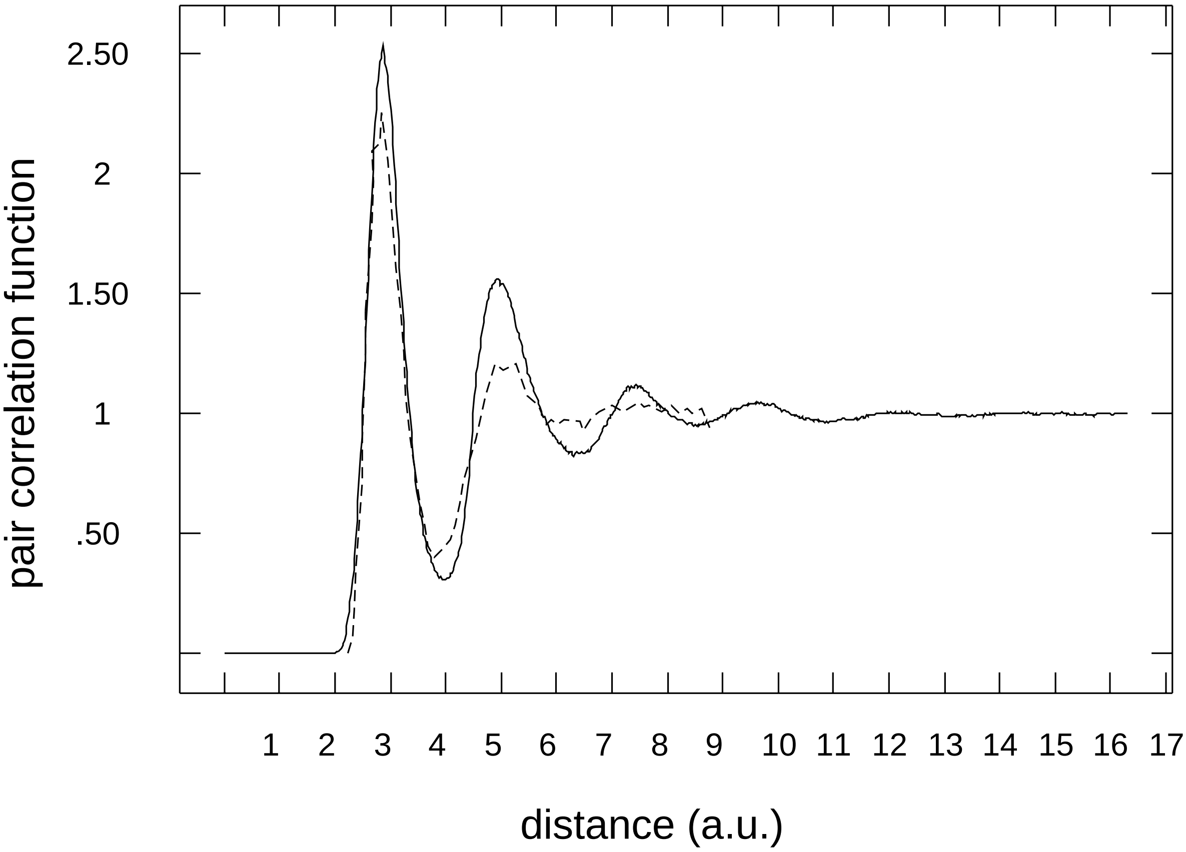

Figure 5: *Particle–particle correlation function $g(r)$ for liquid low density carbon at 5000K. The dashed line represents the CPMD results of Ref.[7].*