Daresbury Laboratory

INFORMATION QUARTERLY FOR COMPUTER SIMULATION OF CONDENSED PHASES

An informal Newsletter associated with Collaborative Computational Project No.5 on Molecular Dynamics, Monte Carlo & Lattice Simulations of Condensed Phases.

Number 40 February 1994

Contents

Editor: Dr. M. Leslie

SERC Daresbury Laboratory
Daresbury, Warrington WA4 4AD UK

IMPORTANT ANNOUNCMENT

RE-REGISTRATION OF USERS

All users who have been registered to receive the newsletter since 1990 or earlier are asked to re-register. A separate sheet is included with the mailing of this newsletter. If your copy does not include this form you need not do anything. You may re-register either

by returning the form to the address shown in the form.

You may also reply by electronic mail confirming that you wish to continue receiving the newsletter and noting any changes to the address.

Failure to re-register will result in your name being deleted from the mailing list.

General News

FUTURE MEETINGS

A summary table is given below, further details may be found inside. (CCP5 is not involved in the organization of the NATO ASI).

TOPIC	DATES	LOCATION
How to derive the interatomic potentials needed for simulation studies	4-5 July 1994	Oxford
The modelling of condensed phases using cellular automata	18-20 July 1994	Manchester
CCP5 ANNUAL MEETING:- ORDER IN LIQUIDS	5-7 September 1994	Sheffield
Observation and Prediction of Phase transitions in complex fluids (NATO ASI)	26 July - 5 August 1994	Varenna

CCP5 PROGRAM LIBRARY

There are two additions to the program library. MDMEGA, written by Dr. W. Smith at Daresbury, is a parallel domain decomposition molecular dynamics program. MOLDY, written by Dr. K. Refson of Oxford University, is a general purpose program for molecular dynamics simulations of condensed matter. It can handle rigid polyatomic molecules, atoms or ions or any mixture of these. It can use either the NVE ensemble or the NσT ensemble using the constant stress algorithm of Parrinello and Rahman. The program is written in "C". New additions to the library from our readers are always welcome.

CRAY NEWS

CCP5 participants are reminded that CCP5 has an annual allocation of Cray time at Rutherford Laboratory. This is available for the development of simulation programs which are of general use to the CCP5 community. Readers who wish to use some of this allocation should write to the CCP5

Secretary, Dr. M. Leslie, TCS Division, SERC Daresbury Laboratory, Daresbury, Warrington WA4 4AD.

The Science and Engineering Research Council has chosen a 256-node Cray T3D parallel supercomputer for use in "Grand Challenge" scientific applications. It will have a total of 256 processing elements, each of which is a DEC Alpha RISC processor with 64 MB of memory and 150 Mflops peak performance. Users will interact with the system through a single processor CrayY-MP4E system with 212 GB of disk and running ray's UNICOS operating system.

INTEL NEWS

CCP5 also has an annual allocation of time on the Intel IPSC/860 at Daresbury. If any CCP5 member wishes to make use of some of this time please contact M. Leslie at Daresbury.

CCP5 FUNDS FOR COLLABORATIONS

CCP5 can make available funds of up to £200 per annum for groups of two or more UK researchers wishing to undertake a collaborative project within the scientific area covered by CCP5. The funds are intended to cover travel and subsistence costs. Researchers who wish to apply for funds are requested to submit a brief proposal (about 1/2 a page) describing the intended work to Dr. M. Leslie, SERC Daresbury Laboratory, Daresbury, Warrington, Cheshire. Alternatively reply by Email to M.LESLIE@UK.AC.DL

CCP5 VISITORS PROGRAM

CCP5 organizes a visitors program which funds the visit to the UK of overseas collaborators. We would normally expect a visitor to visit three sites in the UK and give a lecture at each site. These lectures would be open to all members of CCP5 as well as members of the host university. The visit would normally last between one or two weeks. CCP5 would pay for the cost of travel to the UK and within the UK between universities. CCP5 would expect some contribution towards accommodation expenses at the host university to be met by the university. We will also consider longer collaborations or visits just one place if this can be justified by the nature of the work to be done. If you have an overseas collaborator who you would like to invite under this program, please make a request to Dr. M. Leslie, SERC Daresbury Laboratory, Daresbury, Warrington, Cheshire. UK Alternatively reply by Email to M.LESLIE@UK.AC.DL

JANET:	m.leslie@dl
INTERNET:	m.leslie@dl.ac.uk
EARN/BITNET:	m.leslie%dl.ac.uk@ukacrl

EMAIL MAILING LIST

We are trying to update our Email distribution list. At present we have Email addresses for only 50% of newsletter readers. I have sent Email concerning the electronic distribution of newsletters (see below) recently to the Email distribution list; if you did not receive a copy of this please send me Email so that I can add your Email address to the list. It is important to get this list up to date as there is a proposal to limit paper distribution of newsletters (See Below) and if you are not on the Email list we will not be able to notify you about new newsletters.

PROPOSED ELECTRONIC DISTRIBUTION OF CCP5 NEWSLETTERS

CCP5 is proposing that in future newsletters should be made available to the majority of our readers by electronic means, and that paper distribution by mail will be limited. This issue and the following one will be sent in the usual way by post, the issue due to be published in August 1994 will not be distributed to all readers on the mailing list unless individuals have made a case for us to do this. The reason for this is partly the cost of printing and mailing and partly because we feel that it will provide a faster service for readers overseas. It is proposed, when a new newsletter is available by anonymous ftp, that a short Email message should be broadcast from Daresbury to let our readers know. We would welcome any comments from readers, in particular:

1. Whether this form of distribution will cause readers difficulties as a result of unreliable or non-existent ftp services.
2. What form of distribution would be most convenient (would a postscript file be sufficient or would there be any demand for the LaTeXfiles).
3. Finally, any general experience of similar systems for distributing information.

HOW TO GET THIS NEWSLETTER BY FTP

1.	move to the desired directory on YOUR machine	
2.	type:	ftp 148.79.80.10
	or:	ftp gserv1.dl.ac.uk
3.	enter userid:	anonymous
4.	enter passwd:	*enter your name and site*
5.	change to ccp5.newsletters/40 directory:	cd ccp5.newsletters/40
6.	change to postscript or latex subdirectory	cd ps
	or	cd latex
7.	to get all the files in the directory :	mget *.*
8.		quit

In order to allow users to test the system, this newsletter is available by anonymous ftp or by sending an Email message to the Daresbury info-server, as described in the table. Please note that this is a trial service and details may change at a future date. The newsletter has been placed (in separate directories) both as a postscript file and as the source latex files. The latex directory has two figures in postscript in it. (Technical note:- The latex and postscript both use art11 style; the printed copy of the newsletter uses a slightly different page size to fit onto A4 paper).

REQUEST FOR CONTRIBUTIONS

Contributors to the current issue

Our thanks go to:

D. Frenkel Institute for atomic and Molecular Physics
G. Mooij Amsterdam
The Netherlands

K. Refson Department of Earth Sciences
University of Oxford
UK

The deadline for contributions for the next 2 newsletters will be **15 April 1994 and 15 July 1994** Readers are reminded that contributions are always welcome. Contributions may be sent by Email in LaTeX; this makes the task of collating the newsletter simpler for the editor.

MEETING AND WORKSHOP ANNOUNCEMENTS

The modelling of condensed phases using cellular automata

This meeting is being organized by the Statistical mechanics group of the Institute of Physics. For further information about the meeting write to the address below.

Dr. A. Masters
Department of Chemistry
University of Manchester
Oxford Road
Manchester M13 9PL
UK
Email mbdtsam@cms.mcc.ac.uk

CCP5 will be sponsoring the visit of Dr. A. Ladd (Lawrence Livermore Laboratory) to this meeting. He will also be visiting other universities in the UK as part of the CCP5 visitors program to give seminars and collaborate with CCP5 members. (Provisionally Oxford, Bristol and Cambridge during the week prior to the conference, further details will be in the next newsletter).

Observation and Prediction of Phase transitions in complex fluids (NATO ASI)

This meeting is part of the NATO ASI series; CCP5 is not involved in its organization. Topics include colloids, liquid crystals, polymer melts, membranes and bilayers. For further details contact

M. Baus
Faculté des sciences
C.P.231
Université Libre de Bruxelles
1050 BRUXELLES
Belgium
e-mail mbaus@ulb.ac.be

The closing date for applications is May 1st. 1994.

HOW TO DERIVE THE INTERATOMIC POTENTIALS NEEDED FOR SIMULATION STUDIES

A two day meeting to be held at Mansfield College Oxford from 12:00 on 4th July to 5:00 on 5th July 1994 Sponsored by CCP5

AIM: To encourage participants to air their results, experience and prejudices concerning the practice as well as the theory of how to derive useful atomistic modelling potentials. Both empirical fitting and ab initio techniques will be represented, as applied to bulk and surface studies of static and dynamic phenomena. Poster presentations are eagerly sought and short manuscripts accepted for the proceedings. The list of speakers includes:

David Cooper, Univ. Liverpool	A modern VB approach for interionic potentials
Julian Gale, Royal Institution	Future directions in empirical potential derivation
Mike Gillan, Univ. Keele	The computer modelling of oxide surfaces
Sally Price, UCL	Anisotropic atom-atom potentials for molecules
David Pettifor, Oxford Univ.	Bond order potentials for the atomistic simulation of covalent stytems
John Murrell, Univ. Sussex	Towards a general strategy for global interatomic potentials for elemental solids and clusters
Mike Payne, Cambridge Univ.	Ab initio data bases for testing empirical potentials
Paul Madden, Oxford Univ.	Realistic description of many-body polarisation effects in simulations of ionic systems

The cost for full board participants is £62 (including dinner on the 4th and lunch on both days). For those not needing overnight accommodation the cost is £42.

Due to the financial support of BIOSYM and Unilever plc. we are able to offer a full board reduced price of £30 for students. Those requiring further information please write to either:

R. W. Grimes
The Royal Institution
21 Albemarle St.
London W1X 4BS

A. H. Harker
Building 424.4
Harwell Laboratory
Didcot, Oxon. OX11 ORA

CCP5 ANNUAL MEETING

ORDER IN LIQUIDS

Sheffield Hallam University
5th to 7 September 1994

Preliminary Notice - Scope of Meeting

The Annual Meeting of CCP5 traditionally provides a forum for the the presentation and discussion of the results of molecular simulations across a wide spectrum of scientific problems. In addition to satisfying this more general aim, the special theme of this meeting will be Order in Liquids. Molecular liquids and mesophases exhibit a wide variety of long range and short range order. Many molecular liquids exhibit self assembly and phase transitions in the presence of appropriate thermodynamic fields. The use of computer simulations is now an important tool in investigating these materials and also the way in which the intermolecular potentials yield the observed macroscopic behaviour.

It is expected that the meeting will be broadly based but will include discussion of:- molecular liquids, polymers, lyotropic liquid crystals, thermotropic liquid crystals, self assembly, phase transitions, choice of potential, methods of extracting data from a simulation, methods for making comparison with experiment and methods for improving the efficiency of the simulation.

The meeting and associated accommodation will be in a pleasant campus environment close to the centre of Sheffield. Contributions from research students will be encouraged with a reduced fee. Further details and a call for papers will be circulated shortly. In order to be placed on the mailing list please contact:-

Prof C M Care
Materials Research Institute
Sheffield Hallam University
Pond Street
Sheffield
S1 1WB
Email: C.M.Care@shu.ac.uk
Fax: (0742) 533501

Dr D J Cleaver
Department of Applied Physics
Sheffield Hallam University
Pond Street
Sheffield
S1 1WB
Email: D.J.Cleaver@shu.ac.uk

The CCP5 Program Library W. Smith

CCP5 Program Library Conditions of Distribution

The CCP5 Program Library provides programs and documentation free of charge to academic centres upon application to Dr. W. Smith, TCS Division, S.E.R.C. Daresbury Laboratory, Daresbury, Warrington WA4 4AD, U.K.. Please supply a magnetic tape to receive the copies. Industrial and commercial applicants should enclose a £100 handling charge. No magnetic tape need be sent in this case. Listings of programs are available if required. Please note that use of inappropriate packing for magnetic tapes (e.g. padded bags) may result in them being returned without the required software. Please ensure that these forms of packaging are not used. A list of programs available is presented in the following pages.

All applicants will be required to sign an agreement not to exploit the programs for commercial purposes e.g. for resale or distribution as part or whole of a commercial product.

Readers should also note that we are authorised to supply the example programs originally published in the book "Computer Simulation of Liquids", by M.P. Allen and D.J. Tildesley (Clarendon Press, Oxford 1987). These are supplied in the same manner as the resident CCP5 programs. We are grateful to Mike Allen and Dominic Tildesley for their permission.

We should also like to remind our readers that we would welcome further contributions to the Program Library. The Library exists to provide support for the research efforts of everyone active in computer simulation and to this end we are always pleased to extend the range of software available. If any of our readers have any programs they would like to make available, please would they contact Dr. Smith.

Please Note: For copyright reasons we are not able to supply the programs CASCADE, SYMLAT, THBFIT,THBPHON and THBREL free of charge to Universities outside the United Kingdom.

Programs from the Book: "Computer Simulation of Liquids" by M.P. Allen and D. Tildesley, Clarendon Press, Oxford 1987.

These programs originally appeared on microfiche in the book "Computer Simulation of Liquids" by M. P. Allen and D. J. Tildesley, published by Oxford University Press, 1987. They are made freely available to members of CCP5, in the hope that they will be useful. The intention is to clarify points made in the text, rather than to provide a piece of code suitable for direct use in a research application. We ascribe no commercial value to the programs themselves. Although a few complete programs are provided, our aim has been to offer building blocks rather than black boxes. As far as we are aware, the programs work correctly, but we can accept no responsibility for the consequences of any errors, and would be grateful to hear from you if you find any. You should always check out a routine for your particular application. The programs contain some explanatory comments, and are written, in the main, in FORTRAN-77. One or two routines are written in BASIC, for use on microcomputers. In the absence of any universally agreed standard for BASIC, we have chosen a very rudimentary dialect. These programs have been run on an Acorn model B computer. Hopefully the translation of these programs into more sophisticated languages such as PASCAL or C should not be difficult.

M.P.Allen

CCP5 Program Library E-Mail and anonymous ftp Service

It is possible for CCP5 members to get copies of CCP5 programs through E-mail or anonymous ftp *automatically.*

Email

To access the library by E-mail send an E-mail message to **info-server@uk.ac.dl**. The contents of the E-mail message should be as shown in the table. (Note: the use of upper and lower case is significant - this is a unix system!)

```
request sources
topic index CCP/ccp5
topic CCP/ccp5/program-name
```

Where program-name is the name of the desired source code. A mail server will automatically process this message and return a copy of the source code to your e-mail address. Please note the following however:

The program source will be returned to you in **uue** format, which is a form of encoding most suitable for mail messages. It can easily be decoded on any unix system using the uudecode command. (Check your local unix *man* file for details). Also, to speed the transfer, the source will be split into files of 1200 records each, so expect two or three such files for the average CCP5 program. Once again, uudecode will help you to sort things out.

Readers who do not have unix facilities should include the following lines at the start of the above message.

```
line-limit: nnnnn
coding: off
```

Where nnnnn is the number of records in the source (in most cases 6000 should be enough). The program will be sent in plain FORTRAN as a single file. It may take a while to arrive, but be patient! Also beware in case your system mailer cannot handle messages of this size.

The programs from "Computer Simulation of Liquids" are also available. To access them use program-name F.01, for example.

anonymous ftp

The entire CCP5 program library is now accessible by anonymous FTP. The procedure is as shown in the table.

Readers must realize that the terms of use and distribution of the CCP5 programs that have applied hitherto will be maintained. Users should not redistribute or sell the programs, nor is any liability accepted for their use, either by SERC or the program authors. It is a requirement on the user that the programs be fully tested for their intended purpose. Any bugs found should be reported to the librarian, for the benefit of other users.

Lastly readers should realize that this means of transfer does not usually include any program documentation. So if you are unable to make sense of the programs, write for the documentation!

We are grateful to Mr. P. Griffiths of Daresbury's ITS Division for implementing this facility.

1.	move to the desired directory on YOUR machine	
2.	type:	ftp 148.79.80.10
	or:	ftp gserv1.dl.ac.uk
3.	enter userid:	anonymous
4.	enter passwd:	*enter your name and site*
5.	change to ccp5 directory:	cd ccp5
6.	to list the directory contents:	ls
7.	if desired change to lower directory e.g.	cd TEQUILA
8.	to get a compressed file (.Z)	binary
9.	to get a single file type:	get filename
10.	to get multiple files type:	mget *.*
11.		quit

Availability of the Allen/Tildesley example programs at Cornell

Appendix F of the Book "Computer Simulation of Liquids" by M. P. Allen and D. J. Tildesley describes a method whereby the example programs may be obtained from the statistical mechanics group FTP facility at Cornell. This facility is no longer operational as advertised, due to software and hardware changes. However, the programs are still available. To obtain them, please follow the procedure outlined here. The description below is taken from the HELP file that is distributed by the file server; to obtain the Allen/Tildesley example programs, simply use "ALLEN_TILDESLEY" as the package name (without the quotes, note underscore character _ not hyphen -).

STATMECH is a file distribution service for the Statistical Mechanics community that uses electronic mail facilities to deliver files. To communicate with STATMECH, send an EMAIL message to: statmech@cheme.tn.cornell.edu Commands are sent in the body of hte message you send to STATMECH (not in the subject line). Several commands may be sent at one time; just put one command per line. For each request you make, a transaction log is returned to you indicating

STATMECH commands:

SENDME package	Sends all parts of the specified package.
SENDME package.n	Sends part n of the specified package.
LIST [pattern]	Gives a brief description of all packages matching "pattern". If pattern is omitted, a description of all packages is sent.
HELP	Sends this help file.

the status of the request. The status report will indicate whether the request was successfully completed, and when the file was or will be sent. Large files are sent only during off-peak hours.

Problems, questions and comments about STATMECH service on this system should be directed to "statmech-mgr@cheme.tn.cornell.edu".
Steve Thompson, School of Chemical Engineering, Cornell University, Ithaca NY 14853 USA.

THE CCP5 PROGRAM LIBRARY.

Program	Type	Model	Algorithm	Properties	Authors
ADMIXT	MD	LJA MIX	LF	TH MSD RDF	W. Smith
CARLOS	MC	VS AQ		TH	B. Jonsson, S. Romano
CARLAN	DA	CARLOS structure analysis			B. Jonsson, S. Romano
CASCADE	LS	DIL	EM	TH STR	M. Leslie, W. Smith
CURDEN	DA	Current Density Correlations			W. Smith
DENCOR	DA	Density Correlations			W. Smith
HLJ1	MD	LJA	LF	TH MSD RDF	D. M. Heyes
HLJ2	MD	LJA	LF	TH MSD RDF VACF	D. M. Heyes
HLJ3	MD	LJA	LF LC	TH MSD RDF	D. M. Heyes
HLJ4	MD	LJA	LF CP CT	TH MSD RDF	D. M. Heyes
HLJ5	MD	LJA SF	LF	TH MSD RDF	D. M. Heyes
HLJ6	MD	LJA	TA	TH MSD RDF	D. M. Heyes
HMDIAT	MD	LJD	G5 Q4	TH MSD QC	S. M. Thompson
HSTOCH	MD SD	VS BA	LF CA	TH	W. F. van Gunsteren, D. M. Heyes
MCN	MC	LJA		TH	N. Corbin
MCLSU	MC	LJA		TH	C. P. Williams, S. Gupta
MCMOLDYN	MD MC	LJS FC AQ	LF QF G5 QS	TH RDF	A. Laaksonen
MCRPM	MC	RPE		TH RDF	D. M. Heyes
MDATOM	MD	LJA	G5	TH RDF MSD QC	S. M. Thompson
MDATOM	MD	LJA	LF	TH MSD RDF	D. Fincham
MDCSPC2P	PRMD	BHM	LF	TH STF RDF VACF MSD	W. Smith
MDCSPC4B	PRMD	BHM FC	G5 G4	TH STF RDF	W. Smith
MDDIAT	MD	LJD	LF CA	TH MSD	D. Fincham
MDDIATQ	MD	LJD PQ	LF CA	TH MSD	D. Fincham
MDIONS	MD	BHM	LF	TH MSD RDF STF	D. Fincham, N. Anastasiou
MDLIN	MD	LJL	G5 Q4	TH MSD QC	S. M. Thompson
MDLINQ	MD	LJL PQ	G5 Q4	TH MSD QC	S. M. Thompson
MD3DLJ_C	MD	LJA MIX	LF LC	TH MSD RDF	M. Bargiel, W. Dzwinel, J. Kitowski, J. Mościński
MDMANY	MD	LJS FC	LF QF	TH	D. Fincham, W. Smith
MDMIXT	MD	LJS MIX	LF QF	TH	W. Smith
MDMPOL	MD	LJS FC MIX	LF QF	TH	W. Smith, D. Fincham
MDNACL	MD	BHM	LF	TH MSD RDF	W. Smith

MDPOLY	MD	LJS	G5 Q4	TH MSD QC	S. M. Thompson
MDMEGA	MD	LJA		DD	W. Smith
MDMULP	MD	LJS PD PQ MIX	LF QF	TH	W. Smith
MDSGWP	MD	LJA SGWP	LF	TH VACF RDF QC	W. Smith, K. Singer
MDTETRA	MD	LJT	G5 Q4	TH MSD QC	S. M. Thompson
MDZOID	MD	GAU	LF QF	TH MSD RDF VACF	W. Smith
MOLDY	MD	BHM MIX	BE CP	RDF	K. Refson
NAMELIST	UT	Namelist emulation			K. Refson
NEMD	MD	LJA	SLLOD	Macintosh	D. J. Evans
NSCP3D	UT	Hard sphere packing			M. Bargieł, J. Mościński
PIMCLJ	PIMC	LJA	MC	TH RDF QC	K. Singer W. Smith
SCN	MC	LJA	RFD	TH	N. Corbin
SMFK	MC-SCF	Cylindrical Polyelec.			A. P. Lyubartsev
SLS_PRO	MD	Proteins	LF	TH RDF	A. Raine
SOTON_PAR	MD	LJA	LC	TH	M. R. S. Pinches
SURF	MD	BHM TF 2D	LF	TH RDF	D. M. Heyes
SYMLAT	LS	PIL	EM SYM	TH STR	Harwell
TEQUILA	GP				A. Wilton, F. Mueller-Plathe
THBFIT	LS	PIL	EM	Potential fitting	Harwell
THBPHON	LS	PIL 3B	EM	Phonon dispersion	Harwell
THBREL	LS	PIL	EM	TH STR	Harwell
XEDS	GP				D. Nikolow, W. Alda, J. Kitowski

Key

Program types

MD	Molecular dynamics
MC	Monte Carlo
PRMD	Parrinello-Rahman MD
LS	Lattice simulations
SD	Stochastic dynamics
DA	Data analysis
UT	Utility package
PIMC	Path Integral Monte Carlo
GP	Graphics program

Properties

TH	Thermodynamic properties.
MSD	Mean-square-displacement
RDF	Radial distribution function
STF	Structure factor
VACF	Velocity autocorrelation function
QC	Quantum corrections
STR	Lattice stresses

System models		Algorithm	
LJA	Lennard-Jones atoms	G5	Gear 5th order predictor-corrector
LJD	Lennard-Jones diatomic molecules	Q4	Quaternion plus 4th. order Gear P-C.
LJL	Lennard-Jones linear molecules	LF	Leapfrog (Verlet)
LJT	Lennard-Jones tetrahedral molecules	QF	Fincham Quaternion algorithm
LJS	Lennard-Jones site molecules	QS	Sonnenschein Quaternion algorithm
RPE	Restricted primitive electrolyte	BE	Beeman algorithm
BHM	Born-Huggins-Meyer ionics	LC	Link-cells MD algorithm
SGWP	Spherical gaussian wavepackets	CP	Constant pressure
TF	Tosi-Fumi ionics	CT	Constant temperature
VS	Variable site-site model	TA	Toxvaerd MD algorithm
BA	Bond angle model	CA	Constraint algorithm
PD	Point dipole model	EM	Energy minimisation
PQ	Point quadrupole model	SYM	Symmetry adapted algorithm
MIX	Mixtures of molecules	RFD	Rossky-Friedman-Doll algorithm
GAU	Gaussian molecule model	SLLOD	Thermostatted SLLOD equations
FC	Fractional charge model	DD	Domain decomposition parallel algorithm
PIL	Perfect ionic lattice model		
DIL	Defective ionic lattice model		
3B	3-body force model		
2D	Two dimensional simulation		
SF	Shifted force potential		
FC	Fractional charge model		
AQ	Aqueous solutions		

Programs from the Book "Computer Simulation of Liquids"

F.1	Periodic boundary conditions in various geometries
F.2	5-value Gear predictor-corrector algorithm
F.3	Low-storage MD programs using leapfrog Verlet algorithm
F.4	Velocity version of Verlet algorithm
F.5	Quaternion parameter predictor-corrector algorithm
F.6	Leapfrog algorithms for rotational motion
F.7	Constraint dynamics for a nonlinear triatomic molecule
F.8	Shake algorithm for constraint dynamics of a chain molecule
F.9	Rattle algorithm for constraint dynamics of a chain molecule
F.10	Hard sphere molecular dynamics program
F.11	Constant-NVT Monte Carlo for Lennard-Jones atoms
F.12	Constant-NPT Monte Carlo algorithm
F.13	The heart of a constant μVT Monte Carlo program
F.14	Algorithm to handle indices in constant μVT Monte Carlo
F.15	Routines to randomly rotate molecules
F.16	Hard dumb-bell Monte Carlo program
F.17	A simple Lennard-Jones force routine
F.18	Algorithm for avoiding the square root operation
F.19	The Verlet neighbour list
F.20	Routines to construct and use cell linked-list method

How to optimize Configurational Bias Monte Carlo?

Daan Frenkel and Germonda Mooij
FOM Institute for Atomic and Molecular Physics
Kruislaan 407
1098 SJ Amsterdam
The Netherlands

1 Introduction

Configurational-Bias Monte Carlo (CBMC) is a dynamic MC scheme that makes it possible to achieve large conformational changes in a single trial move that affects a large number of monomeric units of a chain molecule [1, 2, 3, 4]. The CBMC method is based on the Rosenbluth sampling scheme [5, 1, 2] for lattice systems. In this scheme, the molecular conformation is built up step-by-step, in such a way that, at every stage, the next monomeric unit is preferentially added in a direction that has a large Boltzmann weight. This increases the probability of generating a trial conformation that has no hard-core overlaps. As explained below, the probability of acceptance of the trial conformation is given by the ratio of the 'Rosenbluth weights' of the new and the old conformations. Whereas the original Rosenbluth scheme was devised for polymers on a lattice, the CBMC scheme will also work for chain molecules in continuous space. Unlike the reptation algorithm [6], CBMC can be used in the simulation of grafted chains and ring polymers. Recently, the CBMC method has been integrated with the Gibbs-ensemble technique to simulate liquid-vapour and fluid-fluid phase equilibria of chain molecules [7]. In Gibbs-ensemble simulations of phase coexistence, simulations of the two coexisting phases (e.g. liquid and vapor) are carried out in parallel. In addition to MC trial moves of the molecules within either system, we also allow the two systems to exchange volume and mass. CBMC trial moves are used to swap chain molecules between the two systems. Clearly this requires complete regrowth of the entire chain. For long chains this becomes expensive and, at present, Gibbs-ensemble simulations are limited to chain molecules with less than 50 carbon atoms [8]. For simple CBMC sampling the situation is less serious, because one can choose not to regrow the entire chain but only part thereof. In the limit that only one monomeric unit is regrown, CBMC reduces to the reptation algorithm, but in general it will be advantageous to regrow a larger number of monomeric units. Of course, the computational cost per trial move is higher for CBMC than for reptation and hence it becomes important to be able to construct the most efficient MC move for a given system.

The efficiency of the Rosenbluth sampling technique depends on the choice of a set of parameters, namely the number of trial insertions for a given segment i, k_i. As described in the next section, k_i can, in principle, be chosen freely. However, the choice of k_i affects the efficiency of the sampling scheme. The aim of the present contribution is to show that the choice of the parameters in a CBMC simulation is not a question of black magic. Rather, as we shall show, there are systematic techniques to optimize the algorithm. In particular, we show how the efficiency of a CBMC program can be optimized with respect to k_i. Although we apply our analysis to the CBMC scheme, it is in fact much more general, and can be used to optimize the efficiency of any MC trial move that can be decomposed into a sequence of elementary steps. Before discussing the optimization of CBMC sampling, we briefly review the basic idea behind the method.

2 Configurational-Bias Monte Carlo

2.1 Rosenbluth sampling

The Configurational-Bias Monte Carlo scheme for continuously deformable chain molecules [3], is based on Rosenbluth sampling [5, 1, 2] for lattice systems. Chain configurations are generated by successive insertion of the bonded segments of the chain. When the positions of the segments are chosen at random, it is very likely, that one of the segments will overlap with another particle in the fluid, which results in rejection of the trial move. The Rosenbluth sampling scheme increases the insertion probability by looking one step ahead. On lattices, the availability (i.e. the Boltzmann factor) of all sites adjacent to the previous segment can be tested. In continuous space, there are in principle an infinite number of positions that should be tested (e.g. in the case of a chain molecule with rigid bonds, all points on the surface of a sphere with a radius equal to the bond length). Of course, it is not feasible to scan an infinite number of possibilities. Fortunately, however, it turns out that it is possible to construct a correct Monte Carlo scheme for off-lattice models in which only a finite number of trial segments (k), is selected either at random or, more generally, drawn from the distribution of bond-lengths and bond-angles of the 'ideal' chain molecule. ¿From here on, the procedure is the same for lattices and continuous space systems. For each of the trial positions, we compute the Boltzmann factor associated with the non-bonded interactions (more precisely, the contributions of all those interactions that have not yet been accounted for in the generation of the trial positions). One of these trial positions is then selected with a probability proportional to its Boltzmann factor. In this way, regions of high potential energy, such as the hard core of another particle, are avoided and configurations with a non-vanishing Boltzmann weight are generated. To correct for the bias introduced by this very non-random sampling procedure, a weight has to be assigned to each conformation, Γ, called the Rosenbluth weight W_Γ [5]. The contribution of each i th segment to this Rosenbluth weight is equal to the average of the Boltzmann factors of the trial positions for this segment:

$$W_{\Gamma_i} = \frac{1}{k_i} \sum_{j=1} k_i \mathrm{e}\ -\beta U\ _{\mathrm{nb}\Gamma_{ij}}, \tag{1}$$

where $\beta = 1/k_B T$ and $U\ _{\mathrm{nb}\Gamma_{ij}}$ is the non-bonded energy of the j th trial direction for the i th segment. The Rosenbluth weight of the total configuration Γ, is the product of the weights of the individual segments, including the Boltzmann factor of the energy of the first segment, U_{Γ_0}:

$$W_\Gamma = \mathrm{e}\ -\beta U_{\Gamma_0} \prod_{i=1} \ell W_{\Gamma_i}, \tag{2}$$

where ℓ is the chain length. In the original Rosenbluth scheme, every chain conformation Γ was given a statistical weight proportional to W_Γ. However, as explained in ref. [9], this approach fails when the largest contribution to the equilibrium properties of a chain molecule come from conformations that have a large Rosenbluth weight W, but a very small probability $P(W)$ of being generated in the Rosenbluth sampling scheme. The Configurational-Bias MC scheme that we discuss below, was designed to avoid this problem.

2.2 CBMC: 'Dynamic' Rosenbluth sampling

The Configurational-Bias Monte Carlo method is a sampling scheme that employs the Rosenbluth method (extended to continuously deformable molecules [3]) to generate trial conformations. However, it does not suffer from the sampling problem of the original Rosenbluth scheme, because all chain conformations are generated with the correct statistical weight: hence, all averages obtained with the CBMC method are *unweighted* averages over MC configurations, and the problems that are associated with the re-weighting in the original Rosenbluth scheme, disappear.

The CBMC procedure for generating a new conformation of a chain is as follows. First, a chain is chosen at random. Next, a trial conformation for this chain is generated by means of the Rosenbluth sampling scheme and a Rosenbluth weight for this new conformation is calculated. Next, we should decide if we accept the proposed 'move'. To this end, we must compare the Rosenbluth weight W_{new} of the trial conformation with W_{old}, the weight of the old conformation. In fact, the computation of the latter quantity is a bit subtle. In case of a lattice system it is obvious what the trial directions for the old conformation are, and hence its Rosenbluth weight can be evaluated unambiguously. In contrast, for continuously deformable chains the trial directions are chosen at random for every new conformation, and it is not immediately obvious what choice should be made for the calculation of the Rosenbluth weight of the old conformation. As shown in ref. [3], it can be proven that the following simple procedure satisfies detailed balance, and thereby fulfills a sufficient condition to ensure that all chain conformations are generated with a probability proportional to their Boltzmann weight: around every segment i of the old chain, $k_i - 1$ trial directions are drawn from the same probability distribution as the one from which the directions for the trial conformation are chosen. The old Rosenbluth weight is calculated, by treating the $k_i - 1$ trial directions *plus the direction in which the segment of the old chain is situated*, as the set of 'trial' directions for the existing conformation. Finally, we compute the ratio of the Rosenbluth weights of the new and the old conformations. We use a Metropolis-like criterion to decide on the acceptance of the trial move, i.e. the trial move is accepted with a probability P_{acc},

$$P_{\mathrm{acc}} = \mathrm{Min}(1, \frac{W_{\mathrm{new}}}{W_{\mathrm{old}}}). \tag{3}$$

The procedure sketched above is valid for a complete regrowth of the chain, but it is also possible to regrow only part of a chain, i.e. to cut a chain at a (randomly chosen) point and regrow the cut part of the chain either at the same site or at the other end of the molecule. Clearly, if only one segment is regrown and only one trial direction is used, CBMC reduces to the reptation algorithm (at least, for linear homo-polymers). It should be stressed that there are many possible ways to generate a trial conformation. For instance, one can generalize the 'pivot' algorithm [10]. In the pivot algorithm a new conformation is generated by rotating a molecule over a random angle around a randomly selected 'pivot' segment. The pivot algorithm is very efficient for isolated chains, but becomes inefficient for molecules in dense media. However, with CBMC, one can introduce a larger number of pivots in a chain molecule, in such a way that the acceptance of the trial moves is enhanced (at the expense of additional computation). Of course, when CBMC is combined with Grand Canonical and Gibbs-ensemble MC simulations, where entire molecules are exchanged, it is necessary to include moves that attempt to (re)grow chains completely.

One choice remains to be made before applying the Rosenbluth sampling scheme for continuously deformable chain molecules to CBMC and chemical potential calculations, namely the choice for the number of trial directions k_i. Too many trial directions increase the cost of a simulation cycle, but too few trial directions lower the acceptance rate, and increase the simulation length. Clearly, we wish to have simple guidelines that allow us to select k_i for every segment such that it optimizes the efficiency of the simulation. In the following section we show how the optimal values for the set $\{k_i\}$ and the maximum efficiency achievable can be estimated.

3 Efficiency of Configurational-Bias Monte Carlo

In order for the Rosenbluth sampling scheme to work, it is essential to generate, on average, at least one trial position that has a non-negligible Boltzmann weight for every segment. If all trial positions have a small Boltzmann weight, the Rosenbluth weight of the new conformation is virtually zero, while the Rosenbluth weight of the existing conformation is necessarily finite, and the trial move will be rejected. The probability of finding at least one trial position with a non-negligible Boltzmann weight, depends on the choice for the value of k_i, i.e. the number of trial directions that are scanned when looking for an acceptable position of the next, i th, segment. In discussing the efficiency of the CBMC scheme, it is convenient to consider monomeric units with a hard repulsive core because in that case the Boltzmann weight associated with conformations that have hard-core overlaps is strictly zero. However, the general conclusions carry over to systems with continuous intermolecular interaction potentials.

Two trends determine this choice for optimal k_i-values, $k_{i\ \mathrm{opt}}$. On the one hand, the probability of a successful chain insertion grows with increasing k_i. There is an upper limit to that, because when the space to insert another segment is simply not available, there is no point in generating more and more trial directions. Moreover, the computational cost also rises with increasing k_i. The optimal choice for k_i depends on density, temperature and the nature of the intermolecular interactions. For instance, at high densities a larger number of trial directions is needed to regrow a given number of segments than at low densities. It can also be expected that $k_{i\ \mathrm{opt}}$ varies along a chain. After successful insertion of part of the chain, a larger number of trial directions should be chosen for the next segment, in order to minimize the probability that we waste the computational effort that has already been invested in this trial move.

Below, we show how we can arrive at an estimate of the optimal values $k_{i\ \mathrm{opt}}$. To do so, we should first define what we mean by the 'efficiency' of a given CBMC trial move. Loosely speaking, we expect the efficiency to be proportional to the probability that a given trial conformation is successfully generated and inversely proportional to the computational cost of that trial move. For a chain of ℓ segments

$$\mathrm{Eff}(\ell) = \frac{\langle P(\ell)\rangle}{\langle \mathrm{Cost}(\ell)\rangle}, \tag{4}$$

where $\langle P(\ell)\rangle$ is the probability to find for every segment at least one trial direction with a non-negligible Boltzmann weight, in which case the chain can be inserted successfully. $\langle \mathrm{Cost}(\ell)\rangle$ is the average cost for trying to insert the chain, measured in the number of times the energy of a trial direction is calculated. The extra cost for trying to insert a chain which is one segment longer, depends linearly on the number of trial directions and on the probability to insert ℓ segments successfully. So, the average cost for one trial insertion of a chain of length $\ell + 1$ is given by

$$\langle \mathrm{Cost}(\ell+1)\rangle = \langle \mathrm{Cost}(\ell)\rangle + 2k_{\ell+1} \times \langle P(\ell)\rangle\,. \tag{5}$$

where we have introduced, as our unit of computational cost, the amount of computation needed to compute the energy for one trial segment. In the computational cost of a trial move in the CBMC scheme, we have included the cost of the energy calculations for the $k_{\ell+1}$ 'trial' directions of the old conformation, needed to compute the 'old' Rosenbluth weight W_{old}. The probability to find at least one acceptable position for the $\ell+1$ th segment, $\langle P_{\mathrm{add}}(k_{\ell+1})\rangle$, also increases with $k_{\ell+1}$. If we assume that subsequent insertions of segments are independent, $\langle P(\ell+1)\rangle$ is given by

$$\langle P(\ell+1)\rangle = \langle P(\ell)\rangle \times \langle P_{\mathrm{add}}(k_{\ell+1})\rangle\,. \tag{6}$$

Equations 5 and 6 can be combined with equation 4 to yield the following very simple recursive relation

$$\frac{\mathrm{Eff}(\ell+1)}{\mathrm{Eff}(\ell)} = \frac{\langle P_{\mathrm{add}}(k_{\ell+1})\rangle}{1 + 2k_{\ell+1} \times \mathrm{Eff}(\ell)}. \tag{7}$$

Together with the 'boundary' condition $\mathrm{Eff}(\ell = 1)$, equation 7 allows us to compute the efficiency of a trial move for a given set of k_i-values. The values of the set $\{k_i\}$ affect both the numerator and the denominator of equation 7. Our aim is to vary all k_i-values until the optimum efficiency is reached.

The computational cost of the insertion of the first monomer of the chain is zero if we simply start regrowing part of an existing chain. However, if we must successfully insert one monomer before we can continue growing the rest of the chain, then the computational cost of the first insertion is non-negligible and this, in turn, will affect (increase) the optimal values for all subsequent k_i's. In addition to Eff(1), we must know $\langle P_{\mathrm{add}}(k_{\ell+1})\rangle$ for all ℓ. $\langle P_{\mathrm{add}}(k_{\ell+1})\rangle$ can be determined numerically by calculating

$$\langle P_{\mathrm{add}}(k_{\ell+1})\rangle = 1 - \langle (1 - P_{\mathrm{add}}(1))\ k_{\ell+1}\rangle. \tag{8}$$

In words: the probability to generate at least one acceptable trial segment is equal to one minus the probability that not a single acceptable trial segment is generated in $k_{\ell+1}$ attempts. In equation 8, $P_{\mathrm{add}}(1)$ is the probability that the insertion of a single trial segment will be successful. It should be noted that this probability is a fluctuating quantity: the angular brackets in equation 8 denote averaging over the equilibrium configurations of the fluid. Of course, we can make a crude estimate of $\langle P_{\mathrm{add}}(k_{\ell+1})\rangle$ by ignoring all fluctuations, in which case we get the 'mean-field' estimate

$$\langle P_{\mathrm{add}}(k_{\ell+1})\rangle = 1 - (1 - \langle P_{\mathrm{add}}(1)\rangle)\ k_{\ell+1}. \tag{9}$$

Although equation 9 is useful for order-of-magnitude estimates, we shall not use it in what follows. Rather, we shall compute $\langle P_{\mathrm{add}}(k_{\ell+1})\rangle$ by simulation. In stead of computing $\langle P_{\mathrm{add}}(k_{\ell+1})\rangle$ for all ℓ, we measured it for $\ell \leq 2$, and assume that for $\ell > 2$, the values for $\ell = 2$ can be used as an estimate. We verified this assumption under various conditions by calculating $\langle P_{\mathrm{add}}(k_{\ell+1})\rangle$ for all ℓ and we found no significant difference in the answers.

The procedure described above allows us to determine numerically the values for the set $\{k_i\}$ that maximize equation 7, and thereby the efficiency to generate an acceptable trial conformation for a chain in a CBMC move. Thus far we have ignored the fact that this trial conformation, although acceptable in principle, may be rejected in practice. As stated before (equation 3), the overall acceptance probability is determined by the ratio of the new and the old Rosenbluth weights: $W_{\mathrm{new}}/W_{\mathrm{old}}$. Below we shall show that this further attrition of trial conformation also affects the efficiency of the CBMC scheme.

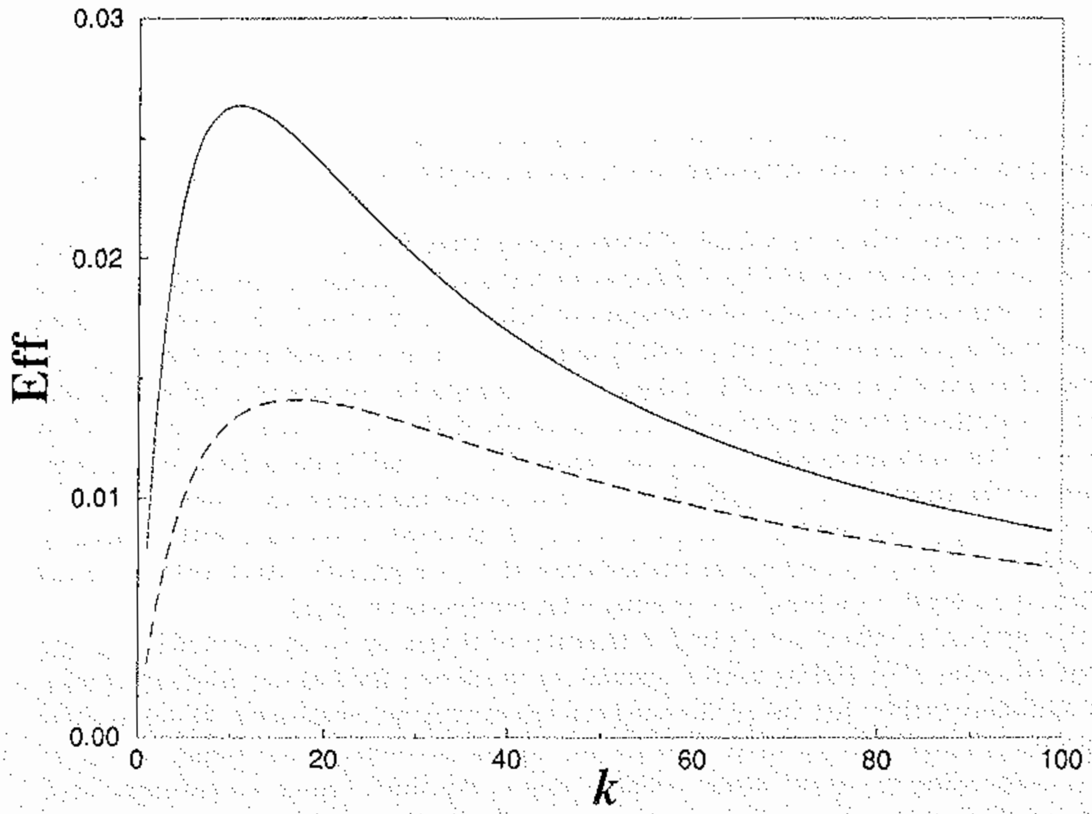

Figure 1: The efficiency, as defined by equation 7, for inserting a hard dimer (———) and a fully flexible trimer of hard spheres (– – –) into a fluid of hard spheres at several densities $\rho\sigma$ 3, over a range of k-values.

ℓ			
	0.3	0.4	0.5
2	5	10	29
3	9	18	55
4	12	27	86
5	15	35	>100
6	18	43	>100
7	20	51	>100
8	22	59	>100
9	25	66	>100
10	27	73	>100
11	29	80	>100
12	30	87	>100

Table 1: Optimal k_ℓ-values, for insertion of a chain by Rosenbluth sampling in the Configurational-Bias Monte Carlo scheme.

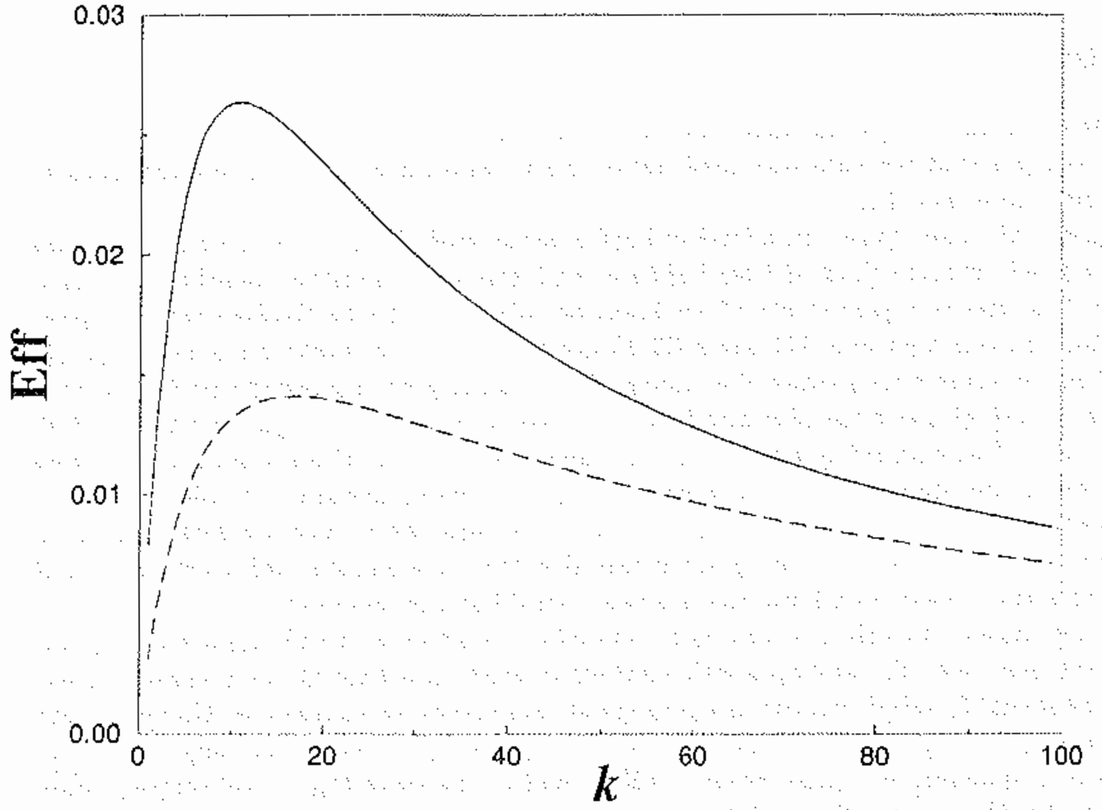

Figure 1: The efficiency, as defined by equation 7, for inserting a hard dimer (——) and a fully flexible trimer of hard spheres (– – –) into a fluid of hard spheres at several densities $\rho\sigma$ 3, over a range of k-values.

ℓ			
	0.3	0.4	0.5
2	5	10	29
3	9	18	55
4	12	27	86
5	15	35	> 100
6	18	43	> 100
7	20	51	> 100
8	22	59	> 100
9	25	66	> 100
10	27	73	> 100
11	29	80	> 100
12	30	87	> 100

Table 1: Optimal k_ℓ-values, for insertion of a chain by Rosenbluth sampling in the Configurational-Bias Monte Carlo scheme.

4 Results

As an example we studied a system with only hard core interactions, but it should be noted that the efficiency analysis that we have presented above can be applied with minor modifications to systems with soft potentials [11].

Let us consider the case where, in a fluid of hard spheres with diameter σ at number density $\rho\sigma$ 3, we insert a fully flexible chain of ℓ hard spheres with the same diameter, attached at a fixed bond length σ. The insertion probability of one segment (which for this particular system is given by the Carnahan-Starling equation [12]) gives Eff(1) and by inserting a second segment $\langle P_{add}(k_2)\rangle$ is calculated from equation 8 for a range of k_2-values. The efficiency for successfully adding another segment, Eff(2), is calculated from equation 7, and the result is shown in Figure 1 for a fluid at density $\rho\sigma$ 3 = 0.4. The maximum determines the value of k_{opt_2}. Eff(3) for a fluid at the same density is plotted in the same Figure, which shows a shift of the maximum to a value for k_{opt_3} which is higher than k_{opt_2}. As already mentioned, k_{opt_ℓ} is expected to increase with ℓ, because more and more effort is invested previously in the insertion of $\ell - 1$ segments, which will be wasted if all the trial directions result in a hard core overlap with spheres in the fluid.

In Table 1 the optimal k_ℓ-values are listed for insertion of chains up to 12 segments long into a fluid at various densities. For adding a fifth segment or more in a fluid at the highest density, $\rho\sigma$ 3 = 0.5, the optimal k_i-values fell out of the range of values that we considered. However, here the efficiency is already close to its optimal value for the highest k_i-values in our range. In Figure 2 we show the corresponding maximal values of Eff(ℓ), and in the same Figure we compare these efficiencies with the efficiencies of a random insertion, i.e. the limit $k_\ell = 1$ for all ℓ. It shows a considerable increase of efficiency using CBMC, and much longer chain lengths are feasible. In the same figure, we also indicate the effect of the attrition of acceptable trial conformations due to the acceptance criterion (equation 3). The decrease is estimated as $\langle W_{new}/W_{old}\rangle$, where W_{new} is now only averaged over chains already inserted successfully (i.e. only the suitable trial conformations are considered). It is possible to give a rough estimate of the maximum chain length that can be reached: if we impose that in a typical MC run we wish to limit the number of evaluations of the potential energy of a segment to a value of 10 8 (i.e. something that can be achieved in a reasonable amount of time on most workstations).

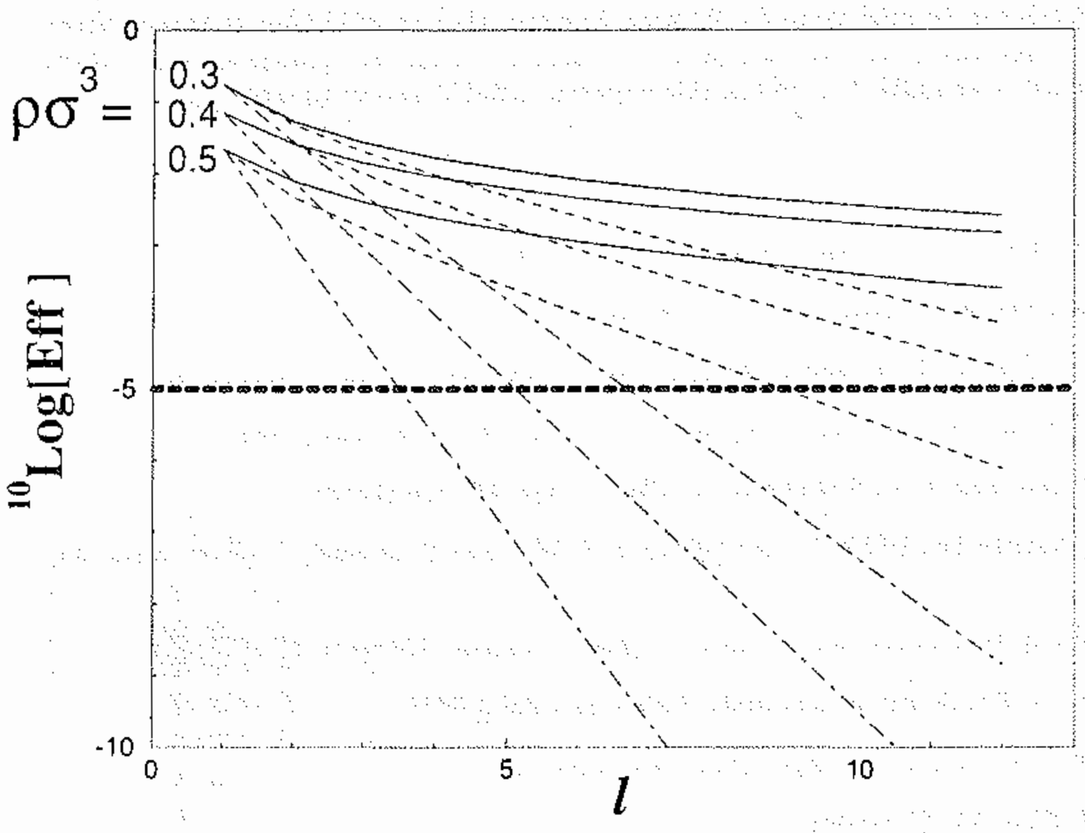

Figure 2: The efficiency (equation 7) for inserting a fully flexible chain of l hard spheres into a fluid of hard spheres at several densities $\rho\sigma$ 3. Both the efficiency of a random insertion $(- \cdot - \cdot -)$, i.e. $k_\ell = 1$ for all ℓ, and the maximal efficiency (———), obtained by choosing the optimal k-values, are shown. In the same Figure we show the efficiency for acceptance of a CBMC move $(- - -)$ by the acceptance criterion 3. The dashed horizontal line shows the minimal efficiency needed for a simulation of typical length.

Moreover, we assume that, in order to sample configuration space effectively, we need at least 10^3 successful chain insertions. This implies that the minimal efficiency needed is of the order of 10^{-5}. Figure 2 shows, that random insertion does not fullfil this requirement for chains longer than three segments at $\rho\sigma^3 = 0.5$, five segments at $\rho\sigma^3 = 0.4$ or seven segments at $\rho\sigma^3 = 0.3$. The CBMC scheme can be used at least up to $\ell = 12$ for $\rho\sigma^3$=0.3 and 0.4, and at the higher density $\rho\sigma^3 = 0.5$ it can be used up to $\ell = 9$. Note, however, that at still higher densities, where trhe probability of successful insertion of a monomer becomes small, the optimal value of ℓ shifts to lower values until, eventually, ℓ=1. When this happens, CBMC reduces to reptation.

References

[1] J. Harris and S.A. Rice, *J. Phys. Chem.* **88**, 1298 (1988)

[2] J.I. Siepmann and D. Frenkel, *Mol. Phys.* **75**, 59 (1992)

[3] D. Frenkel, G.C.A.M. Mooij, and B. Smit, *J. Phys.: Condens. Matter* **4**, 3053 (1992)

[4] J. J. de Pablo and M. Laso and U. W, Suter, J. Chem. Phys. **96**, 2394 (1992).

[5] M.N. Rosenbluth and A.W. Rosenbluth, *J. Chem. Phys.* **23**, 356 (1955)

[6] A.K. Kron, Polymer Sci. USSR 7, 1361 (1965), F.T. Wall and F. Mandel, *J. Chem. Phys.* **63**, 4529 (1975)

[7] G.C.A.M. Mooij, D. Frenkel, and B. Smit, *J. Phys.: Condens. Matter* **4**, L255 (1992)

[8] J.I Siepmann, S. Karaborni, and B. Smit, *Nature* **365** 330 (1993)

[9] J. Batoulis and K. Kremer, *J. Phys. Chem.* **90**, 2370 (1989)

[10] M. Lal, *Mol. Phys.* **17**, 57 (1969)

[11] D. Frenkel "Advanced Monte Carlo Techniques" in the Proceedings of the NATO Advanced Study Institute on Computer Simulation in Chemical Systems (M.P. Allen and D.J. Tildesley, eds), Alghero (1992)

[12] N.F. Carnahan and K.E. Starling, *J. Chem. Phys* **51**, 635 (1969)

A GENERAL-PURPOSE MD CODE

Keith Refson

Moldy is a general-purpose molecular dynamics simulation program which I wrote initially for my own research. It is sufficiently flexible that it ought to be useful for a wide range of simulation calculations of atomic, ionic and molecular systems and I am therefore offering it to the CCP5 community.

The design objectives were for a single program which is able to simulate a wide range of systems without arbitrary restrictions on the number or form of constituents. The system is specified at run time with a description file so there is no need to recompile when changing systems. Another design objective was that the program should handle much more of the bookkeeping than is traditionally done, especially with regard to keeping track of consistency of parameters, restart files, output trajectories and so forth. You don't have to worry about array sizes and limits because this is all handled automatically. It ought to be easier to concentrate on the science by making starting and keeping track of a new simulation a simpler process and helping to eliminate many of the frustrating and time-wasting mistakes that every simulator is so familiar with.

The program can handle any mixture of atoms or polyatomic molecules (linear or otherwise) of any size within the rigid-molecule approximation. There are no limits on the number of atoms in a molecule, the number of molecular species or number of molecules. The system can be in the liquid or solid state, with MD cells of arbitrary dimensions and angles and the simulation may be conducted at either constant volume or constant stress using the Parrinello-Rahman algorithm. Interactions are by pair-potentials (based at atomic sites in the case of molecules) with or without coulombic interactions. Most common forms of potential functions are supported (Lennard-Jones, Buckingham, Born-Meyer, MCY) and the program is designed to make it very easy to add others. Short-ranged forces are handled using the link-cell method and the long-ranged coulombic forces by the Ewald sum. Therefore the program ought to be suitable for simulating very large systems.

There are several features which are slightly novel. First, Moldy does not use the usual "minimum-image" convention, but instead includes interactions between a molecule and ALL of its periodic images that lie inside the cut-off radius. This is more strictly correct and just as easy to implement as minimum-image because of the link-cell algorithm. Second, Moldy incorporates a method of generating initial configurations for liquid systems called a "skew start". This can reliably generate a configuration which is partially ordered but avoids molecular overlap. Finally, there is a capability for defining a "framework" which is a rigid super-molecule permeating all of space. This may be used to model rigid surfaces or zeolite-like cages, for example.

One other aspect of moldy which might be unfamiliar is that is is written in C rather than FORTRAN. Fortran does not have the flexibility of dynamic memory allocation to allow the automatic sizing of the arrays which Moldy needs. This ought to present no problems as C compilers are just as or more common than FORTRAN ones. The program is highly portable and has been optimised for both vector supercomputers (cray and convex), but also runs fast on modern unix workstations and even PCs. There is a parallel version for shared-memory parallel-processors including explicitly Cray and Convex machines. A port to distributed memory parallel architectures ought to be straightforward, though the current parallelization strategy will not scale well to very large numbers of processors.

The program incorporates radial distribution function calculations and running accumulation of many of the usual thermodynamic averages. Any more sophisticated analysis can be performed by storing configurational data throughout the run for later analysis. There are flexible facilities for doing this. In addition to the main program there are utilities for manipulating dump datasets and an interface to a molecular graphics module for AVS.

Given the above claims for generality and flexibility it is, only fair to mention the current limitations of the program. Only pair potentials are supported at the moment. New forms of potential function are easily added, but bond-bending or 3-body forces or shell models will take rather more work. The program treats molecules as rigid bodies using the quaternion algorithms, and no flexibility or other constraints are allowed. There is also no support at present for a thermostat - temperature is controlled by velocity scaling techniques. Since the source code is freely available I hope that others with a need for these facilities will be able to add modules and extend the capabilities.

Moldy differs in functionality and strategy from the CCP5 project, DL_POLY. While DL_POLY is implemented as a series of modules to be bolted together, Moldy is a single program which is configured at run-time by specification files. This does lead to less choice, for example in matters of boundary conditions and treatment of long-range forces, but makes starting a new simulation substantially easier and less prone to error. Moldy is aimed at systems of small molecules for which the rigid-molecule approximation is useful. It therefore supports massless sites and implements interaction cutoffs using a molecular rather than a site criterion. The "feature-lists" differ - Moldy implements the constant-stress ensemble for the study of solid-state phase transitions, whereas DL_POLY offers thermostats - but we will no doubt see the holes filled as both programs develop. Moldy shares with DL_POLY the principle of giving the user control over the source code, and is designed in a modular fashion using the principles of structured programming to encourage extensions to be added as needed. Finally, moldy is available now to anyone in CCP5 or otherwise who wishes to use it.

The source code may be obtained from the CCP5 program library in the usual way, and also directly by anonymous file transfer from Oxford. Connect to "earth.ox.ac.uk" using "ftp", with an account name of "anonymous" and your email address as password. The relevant files are all in the "/pub" directory and are

- moldy-2.6.tar.Z — The Unix distribution (also for MSDOS)
- moldy-2.6.com — The VMS distribution
- moldy-manual.ps.Z — The Manual in PostScript form. Note that the distribution files already contain the LaTeX source.

(The current release is 2.6; I intend to keep to this naming scheme and just keep the highest release on the ftp server.)

Please note that moldy is copyrighted and distributed under the GNU public license which is designed to encourage its distribution and modification. This is to ensure that the source code of moldy and any improvements made to it remain freely available. I would like to encourage anyone who improves the program to return the changes to me so they can be made incorporated into future releases for the benefit of all.

I am also keeping a list of email addresses of anyone who uses the program for notification of updates, bugs and so forth. Please notify me if you would like to be added to this list, preferably by email to Keith.Refson@earth.ox.ac.uk.