Use of Independent Ensemble Averaging for DSMC solutions in unsteady flow.

Doug Auld
School of Aerospace, Mechanical and Mechatronic Engineering,
University of Sydney, 2006, NSW, Australia.

Abstract. Many gas flows at the nano- and micro-scale level show non-continuum behaviour that is best captured using rarefied flow methods such as Direct Simulation (Monte-Carlo) Method (DSMC). In these flow regimes unsteady effects such as flow bifurcation, oscillatory behaviour and disturbances due to acoustic wave propagation have a significant impact on flow properties. To accurately capture such behaviour, simulation techniques such as DSMC require ensemble averaging over samples of many millions of simulated molecules. This requirement means that accuracy limits are set by available memory capacity and CPU speeds of the computer systems used. In order to overcome the computational limits several different methods have been used to make use of parallel computer architectures or cluster computer systems. The current research is aimed at investigating the use of such independent ensemble averaging to reduce statistical scatter in solutions involving unsteady, time varying flow. Using a single initial solution at a given time, independent parallel solutions are run for nominated sample times.

INTRODUCTION

The Monte Carlo Direct Simulation Method [1] is a well-benchmarked technique for the simulation of rarefied flows or small-scale flow as high to medium Knudsen number. Gas flows in the nano- or micro-scale often exhibit properties which are governed both by the Reynolds number (Re) scale of the flow and by the Knudsen number (Kn). DSMC has thus been used as a technique to accurately simulate these small scale flows. The regime of interest is near-continuum with Kn number less than 0.01 and Reynolds numbers ranging between 10 and 1000. In many cases, these flows exhibit unsteady behaviour such as flow bifurcation, oscillatory behaviour and disturbance propagation due to acoustic waves. Also due to the relatively high number densities involved large ensemble averages over many millions of simulated molecules are needed to determine the trends in the normally statistically scattered results. The requirement on CPU memory and processing time is extremely large for this class of problem.

The use of parallel computer architecture arrangements in order to accommodate the large number of molecules is necessary. Previous solutions in this flow regime have been carried out using domain decomposition techniques such as multi-grid domain decomposition [2], [3] using PVM or MPI computer architectures to allow the fluid flow problem to be distributed across a group of clustered computers. This research [3] has found that medium level improvements in computational efficiency are obtainable. However, in these low Reynolds number flows, due to the dominance of the random thermal motion of molecules over average motion, there exists a large requirement for processor-to-processor communication. As molecules are continuously undergoing random motion from one domain to another, simulations with large numbers of sample molecules were found to be communication dominated hence computationally inefficient.

In this study the method of independent ensemble averaging was used in place of domain decomposition. Each computer processor carries out an independent solution of the flow field for a nominated period of time. The ensemble average of all solutions is then used to predict the current flow solution. For steady flow cases, this requires zero communication between processors and thus gives the best efficiency possible for a parallel processor algorithm. For the unsteady flow cases that are studied here, the solutions cannot be run truly independently as synchronisation between separate processor solutions is required in order track the true long term time dependent variation in the flow. The level of communication required for this synchronisation and its effect on the flow solutions is shown in the following sections of this paper.

FLOW GEOMETRY

Previous research using steady flow simulations [4] and ensemble averaging of independent statistical solutions for the case of an expanding jet has shown that DSMC is an accurate method for representing these flows. Flow field geometry is shown in fig. 1.

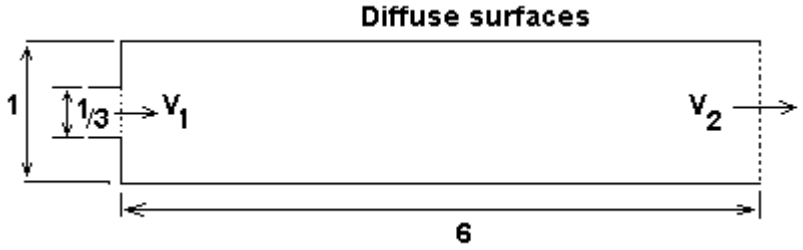

Figure 1. Flow field geometry.

The general properties of the flow used for this study are ideal air; density $5x10^{-5}$ kg/m3; static temperature 300K; inlet velocity 150 m/s; exit velocity 50 m/s based on area ratio 3 to 1. The flow is modelled as a two dimensional problem. The Knudsen number based on channel height is of the order 0.001 in order to model near-continuum flow. The Mach number of the gas is equal to or less than 0.5 in all parts of the flow field in order to minimise compressibility effects on the result. The Reynolds number based on channel height of the flow is nominally 500. This ensures that solution should be greater than the stability limit required for creation of oscillatory instabilities in this type of flow.

A benchmark DNS solution [4] for a two-dimensional jet flow in a channel, (fig 2) has been used previously to verify the accuracy of the DSMC solutions for the steady flow and incipient transition to instability for this flow.

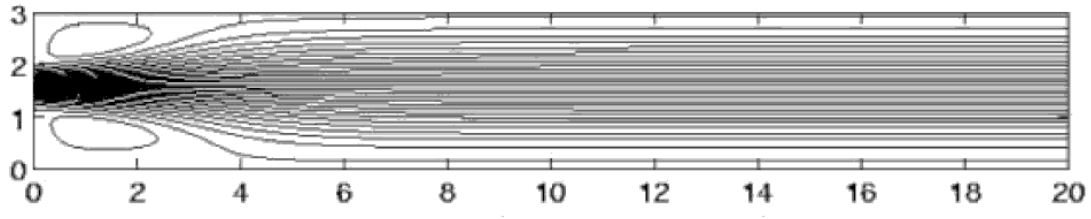

Figure 2. Benchmark steady flow field solutions from previous DNS calculation, ref 5.

DSMC PARALLEL CALCULATION ALGORITHM

The standard DSMC simulation procedure is well documented in ref 1. Modifications to the procedure to allow parallel calculation have been documented in refs [5], [6] and [7]. In these cases the aim has been to use domain decomposition to subdivide the flow field between processors using either structured or unstructured grids and so that a single program is running with multiple data sets distributed between CPUs. These techniques require high-speed communication between processors to account for the large number of molecules that randomly pass between the subdivided domains. Increasing parallelisation requires greater number of subdivisions and this increases the proportion of molecules that are moving from one processor memory to another during a particular time interval. Previous tests [2] done with simple parallel computer architecture have shown the communication time becomes significant after moderate number of subdivisions and then dominates so that no speed improvement is obtained with further sub division.

For steady flow solutions [7] an alternate method of parallelisation has been to run independent parallel solutions that can be used to give an accumulated sample after each separate processor has reached a steady state solution. This requires no communication between processor units and delivers the optimum performance in that there are no communication overheads.

A similar approach has been employed here but with the aim of modifying the procedure to allow for the solution of time varying flows. In these cases, the separate CPU solutions are not truly independent and time synchronisation is required so that each processor is correctly simulating the desired time interval of flow. This has been done by breaking the simulation sequence after a set number of sample intervals and requiring all processors to use just one of the current solutions as its seed for further calculation. The procedure is shown in the following diagrams. It is a two-program process with a master-sequencing program controlling independent DSMC simulations.

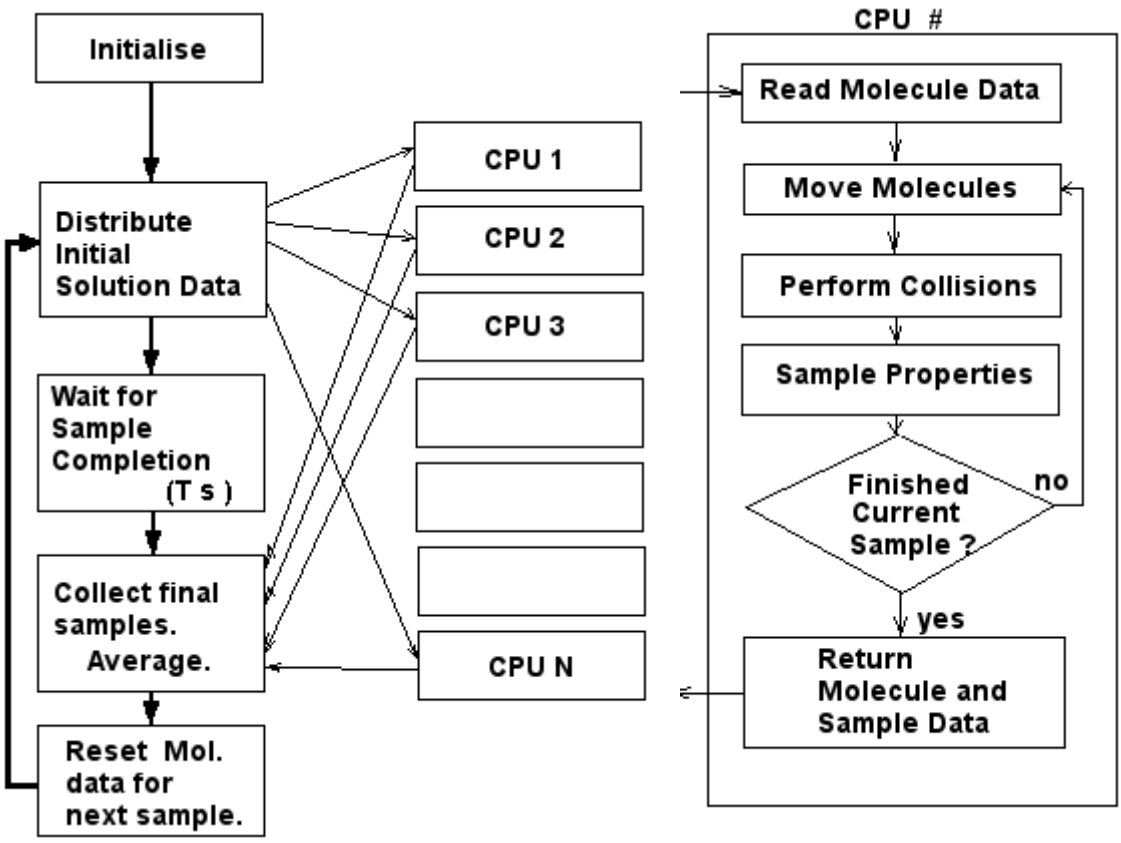

Figure 3. DSMC procedure for parallel ensemble averaging.

This procedure introduces a communications workload into the process. While all simulation CPUs run independent solutions, the distribution of initial solution and collection of current solution phases by the master program unit requires that molecule and geometry information be transferred. This overhead in this component has been investigated and was compared to the previous result were communication times for a domain decomposition solution were run.

In the cases used for result comparison the solutions were tested on identical CPU units. The hardware configuration is a cluster of ten Intel Pentium 4 CPU's each running at 2 Mhz and connected using a 100Mhz TCP/IP network. The machine operating environment is a Linux kernel system with software compiled using Intel Fortran compiler. All are standard components and no special configuration has been. The basis for the individual CPU simulations is the DS2V version 2 system provided by Prof. G. Bird. Only modifications have been made to the input-output data storage components to allow this standalone program to interface with the master control system. No modifications have been made to the DS2V DSMC simulation algorithm. In all simulation tests shown the number of simulated molecules for each CPU unit is 2.2 million. The resulting run information shown is obtained after a long duration simulation to ensure that any time zero initialisation effects have been damped from the flow.

Various values of ensemble sample time (Ts) have been used in the procedure in order to evaluate the effect on the solution.

SIMULATION RESULTS AND COMPARISONS

The following figures show sample outputs from the simulation process. Figure 4 shows Mach number contours for the established flow at a time of 4.03 secs. This image is based on an ensemble sample time of 800 millisecs in order to display the general nature of the flow. The oscillatory nature of the simulated flow is predominantly producing perturbations in vertical velocity along the length of the channel. Figure 5 (a to c) shows vertical velocity along the centreline of the channel for a set of ensemble sample times and CPU numbers.

Figure 4. General nature of flow field. Mach contours, time = 4.03 secs.

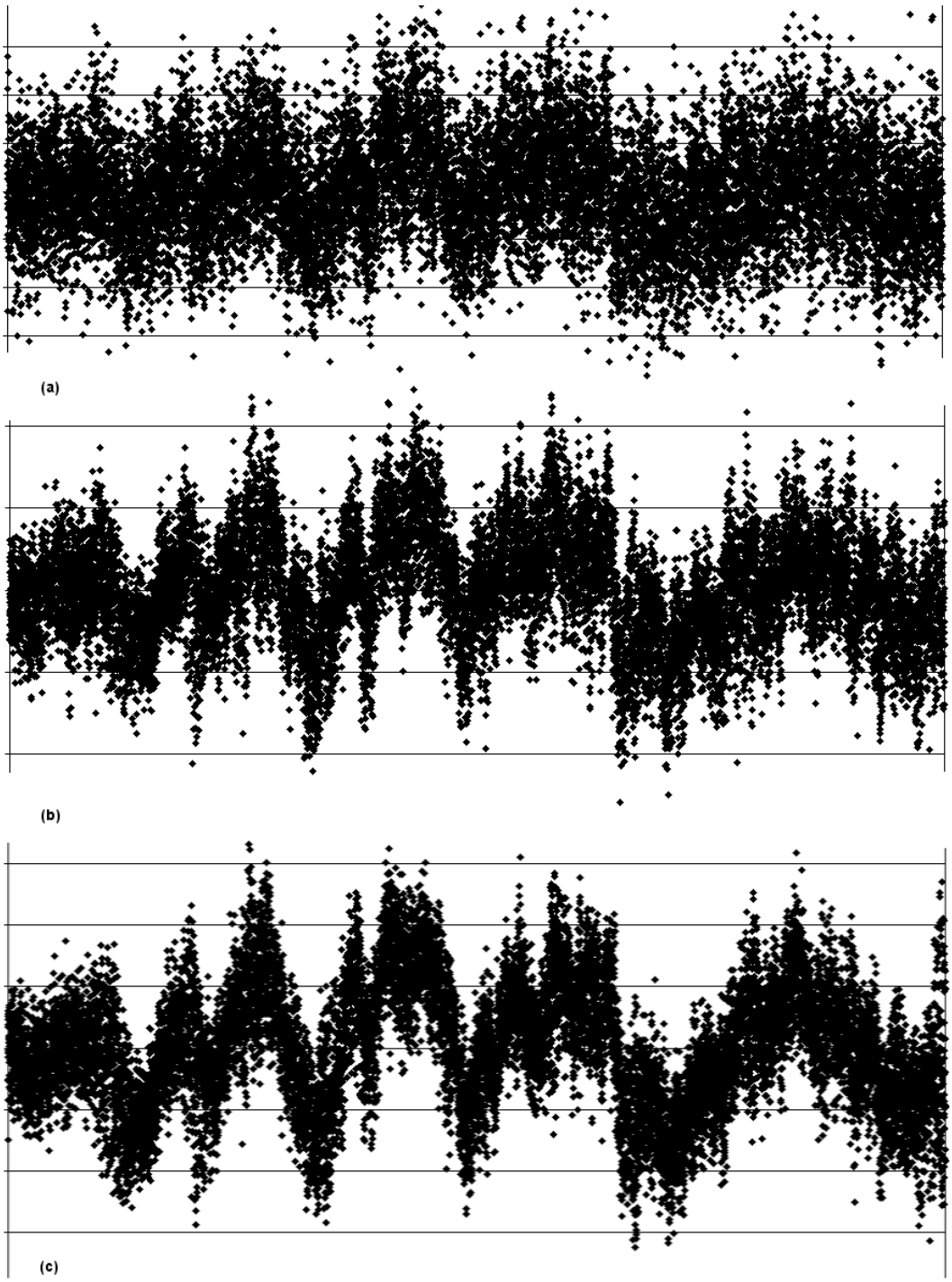

Figure 5. Vertical velocity samples. (a) ensemble time (Ts)= 7 millisecs, 1 CPU; (b) Ts = 7 millisecs, 10 CPU; (c) Ts = 14 millisecs, 10 CPU.

The communication time comparison between this current method and a domain decomposition method applied to the same hardware is shown in the following figure.

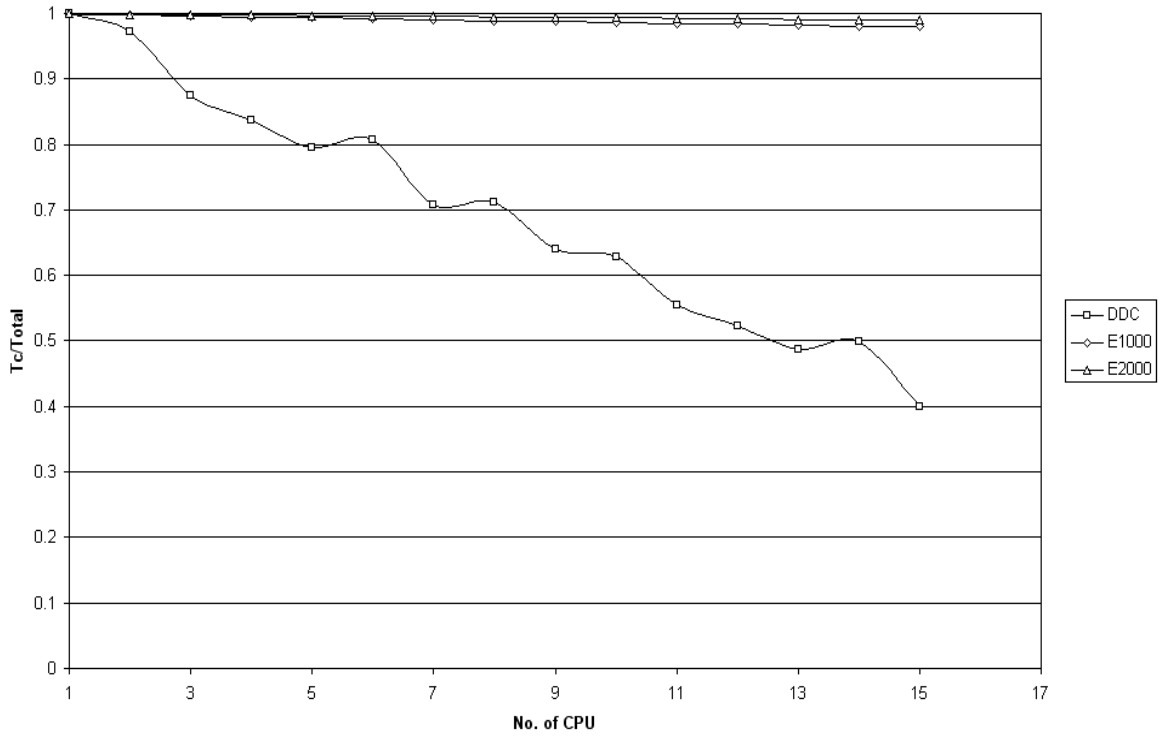

Figure 6. Communication efficiency of parallel processes. (Tc/Total) (Processor time/Total time), DDC – domain decomposition method; E1000 –Ensemble of 1000 samples; E2000 – Ensemble of 2000 samples.

As the ensemble average process for unsteady flow now requires interchange of molecule data between processors, this method does introduce a communication component into the overall process. However, simulation results indicate this is a significantly less time penalty than was involved in previous domain decomposition methods.

CONCLUSION

The use of independent ensemble averaging technique is able to be applied in cases were the flow is time dependent provided that a synchronisation process is incorporated. The optimisation of process is a topic for further investigation. Larger ensemble times and more sampling CPU's result in better definition of flow field properties. A limit on ensemble sample times is however imposed by the characteristic motion properties of the flow. As the propagation speed of the flow oscillations will be a function of the average speed in the channel, ensemble averages have been kept to approximately 10 millisecs, this still however produces a 5% to 10% variation in characteristic position while ensemble sampling takes place. In the case of high-speed flows the period chosen for ensemble averaging will need to be tailored to the flow minimum characteristic time.

Also from observation of the ensemble velocity distributions, it can be seen that the method introduces statistical fluctuations that are greater than expected in a normal distribution. This can be attributed to method of synchronisation where all following CPU solutions are seeded from one of the current CPU solutions. However the results produced by the methods multiple CPU solution algorithm produce results with greatly reduced statistical variation than an equivalent single CPU solution.

ACKNOWLEDGEMENTS

The author wishes to thank Prof. G. Bird for providing the DS2V DSMC code that is used as the core algorithm on each CPU and Prof S. Armfield for use of the University of Sydney, Fluid Mechanics Research Group BORG computer cluster to support this work.

REFERENCES

1. G A Bird, *Molecular Gas dynamics and the direct simulation of gas flows*. 1993.
2. D J Auld and W S Lan, *Simulation of lid-driven cavity flows by parallel DSMC method.* DSMC Theory Methods and Applications Conference Proceedings, Santa Fe, NM, Sept 2005
3. Cambridge, M.M.P., *MPI the complete reference.* 1998.
4. P L Morgan, D J Auld & S W Armfield , *A comparison of Eulerian and Lagrangian schemes for the simulation of an incompressible planar jet*. ANZIAM Journal (**45**) pp310-325 2004.
5. A Singh, and Y Zhao, Parallel unstructured dynamic grid direct Monte Carlo simulation of molecular gas dynamics and its applications. Journal of Scientific Computing, 2001. **16**(4): p. 553-568
6. J S Wu, and Y Y Lian*, Parallel three-dimensional direct simulation Monte Carlo method and its applications*. Computers and Fluids, 2003. **32**(8): p. 1133-1160.
7. D J Rader, M A Gallis, et al *DSMC Convergence Behaviour for Fourier Heat Flow*, DSMC Theory Methods and Applications Conference Proceedings, Santa Fe, NM, Sept 2005