



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Coupling Molecular Dynamics and Direct Simulation Monte Carlo using a general and high-performance code coupling library

Citation for published version:

Longshaw, SM, Pillai, R, Gibelli, L, Emerson, DR & Lockerby, DA 2020, 'Coupling Molecular Dynamics and Direct Simulation Monte Carlo using a general and high-performance code coupling library', *Computers and Fluids*, vol. 213, 104726. <https://doi.org/10.1016/j.compfluid.2020.104726>

Digital Object Identifier (DOI):

[10.1016/j.compfluid.2020.104726](https://doi.org/10.1016/j.compfluid.2020.104726)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Publisher's PDF, also known as Version of record

Published In:

Computers and Fluids

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.





Coupling Molecular Dynamics and Direct Simulation Monte Carlo using a general and high-performance code coupling library



S.M. Longshaw^{a,*}, R. Pillai^b, L. Gibelli^b, D.R. Emerson^a, D.A. Lockerby^c

^aScientific Computing Department, UKRI-STFC, Daresbury Laboratory, Warrington, UK

^bSchool of Engineering, Edinburgh University, Edinburgh, UK

^cSchool of Engineering, Warwick University, Warwick, UK

ARTICLE INFO

Article history:

Received 31 January 2020

Revised 1 July 2020

Accepted 24 August 2020

Available online 18 September 2020

Keywords:

Molecular dynamics

Direct simulation Monte Carlo

Coupling

Supercomputing

ABSTRACT

A domain-decomposed method to simultaneously couple the classical Molecular Dynamics (MD) and Direct Simulation Monte Carlo (DSMC) methods is proposed. This approach utilises the MPI-based general coupling library, the Multiscale Universal Interface. The method provides a direct coupling strategy and utilises two OpenFOAM based solvers, mdFoam+ and dsmcFoam+, enabling scenarios where both solvers assume one discrete particle is equal to one molecule or atom. The ultimate goal of this work is to enable complex multi-scale simulations involving micro, meso and macroscopic elements, as found with problems like evaporation.

Results are presented to show the fundamental capabilities of the method in terms of mass and kinetic energy conservation between simulation regions handled by the different solvers. We demonstrate the capability of the method by deploying onto a large supercomputing resource, with attention paid to the scalability for a canonical NVT ensemble (a constant number of atoms N , constant volume V and constant temperature T) of Argon atoms. The results show that the method performs as expected in terms of mass conservation and the solution is also shown to scale reasonably on a supercomputing resource, within the known performance limits of the coupled codes. The wider future of this work is also considered, with focus placed on the next steps to expand the capabilities of the methodology to allow for indirect coupling (where the coarse-graining capability of the DSMC method is used), as well as how this will then fit into a larger coupled framework to allow a complete micro-meso-macro approach to be tackled.

© 2020 The Authors. Published by Elsevier Ltd.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

1. Introduction

Many fundamental physical processes are intrinsically multi-scale in nature when they are considered from a modelling and simulation perspective. This is due to the fact that they typically involve phenomena that can only be described by constitutive laws which are valid at the meso- or micro-scale and therefore cannot be modelled using just macroscopic methods. A common approach to resolving this is to model the majority of the process macroscopically and then incorporate mesoscopic or microscopic effects by way of generalised parameterisation.

An example of this is the typical macroscopic approach taken to modelling evaporation, where it is necessary to consider a phase-change problem in which liquid transforms to gas. In this situa-

tion, the concept of transformation from liquid to gas is captured in the form of an interface and the mass flux through this defined according to macroscopic descriptions by way of coefficients for evaporation and condensation [1–3]. The problem with this approach is that where liquid transforms to gas is not really a distinct and immediate change in state. Rather, a Knudsen layer exists within the gas near to the liquid that can only be modelled using microscopic or mesoscopic approaches. A layer also exists between this and the liquid phase that can only be treated using microscopic approaches, see Fig. 1 for a full depiction. While parameterising the micro-meso part of this complex macro-meso-macro system allows for an approximate continuum solution to the problem [4,5], to go beyond this and fully understand the physics of the phenomena needs the system to be considered using a resolved multi-scale approach or using a method appropriate to the lowest length- and time-scales present in the system. Examples of simulating evaporation using just classical Molecular Dynamics (MD) exist [6] but this is notably restrictive in terms of

* Corresponding author.

E-mail address: stephen.longshaw@stfc.ac.uk (S.M. Longshaw).

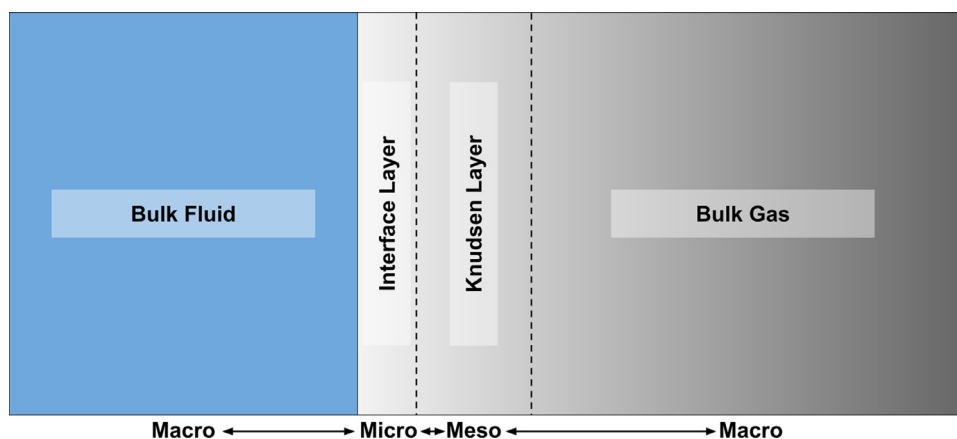


Fig. 1. The macroscopic-microscopic-mesoscopic-macroscopic nature of problems like evaporation.

problem size that can be captured, due to the computationally intensive nature of MD.

It is modelling the process of evaporation that has driven the need for the work presented here. A method is proposed, along with associated software, that provides a fully-coupled (and general) simulation environment for problems where classical MD and Direct Simulation Monte Carlo (DSMC) can be used simultaneously to consider a problem. This is the first part of a larger overall piece of work which will next look at how to allow for coarse-graining with the DSMC method and then allow for simultaneous coupling between MD and Computational Fluid Dynamics (CFD) as well as DSMC and CFD in order to complete the full tri-scale modelling capability demanded by an evaporation problem.

This work presents a coupled software solution that uses the MD solver mdFoam+ [7] and the DSMC solver dsmcFoam+ [8] (both based in OpenFOAM [9]), this is available via public repository from [10], along with all test cases presented. The coupling is achieved using the general coupling library, the Multiscale Universal Interface (MUI) [11], which is a lightweight and header-only library written in C++.

The MUI library provides an ability to create one or more coupling *interfaces* between solvers that allow for time-associated data to be transferred using only Message Passing Interface (MPI) communications. It also provides spatial and temporal interpolation functionality to extract data from the interface in an appropriate format. The library is designed specifically to work within computing environments that support the MPI multiple programs, multiple data (MPMD) standard, it is also designed to integrate with codes that are already parallelised using MPI (or any other parallelisation or acceleration method). MUI has been designed to operate on supercomputers, therefore it pays particular attention to minimising inter-process communication, this is especially important when considering coupled solutions where each solver is MPI parallelised and the problem highly decomposed, as is the case here.

In this paper we present the algorithmic design of the implemented coupling methodology and then show that the approach reliably captures properties such as mass and energy conservation. The computational performance of the solution is considered on the UK's national supercomputing service, ARCHER, and next steps for this work are highlighted that look toward the creation of a fully-coupled tri-scale simulation methodology, including coupling the MD and DSMC solvers to macroscopic CFD [12–14].

2. Background

Efforts to simultaneously couple MD and DSMC using a domain-decomposed approach are evident over the past two decades, with

approaches ranging from direct coupling strategies [15,16] that allow DSMC to be used as a direct replacement for MD (or similar methods) in order to facilitate the simulation of significantly larger problem sizes (due to the inherent lower computational demand of the method compared to MD), through to strategies that allow DSMC to be used as a coarse-grained method alongside MD [17–19]. Notably these strategies are distinct; the first treats MD molecules (or atoms) and DSMC parcels (or particles) as functionally equivalent; the second relies on collecting statistical data from each domain that enables an interface-based coupling where length and time scales can be different. This work presents a capability based on the first of these approaches, with a view towards extending its design in the future to enable the second.

The types of problems that can be tackled using a coupled MD-DSMC approach are varied. They range from fundamental phenomena, like evaporation (as discussed in Section 1), through to highly multi-scale problems, like re-entry of satellites. This particular problem has already been considered using DSMC [20] but would benefit from an ability to incorporate MD to better resolve aspects of the physics, such as fluid-wall interactions and adsorbed gas layers. Whether the use of a direct or indirect coupling method is most appropriate depends on the problem being considered. In cases where the length and time-scales of the physical domains being simulated by MD and DSMC are functionally equivalent (i.e. of a similar order of magnitude) then a direct approach is likely the best as this eliminates the need to average quantities and therefore provides as exact a solution as possible. Where domains are significantly different then an indirect approach is needed that allows the DSMC simulated domain to be coarse-grained.

2.1. Molecular dynamics

Molecular Dynamics (MD) is a method used to describe the application of the classical Newtonian equations of motion to atoms or molecules, henceforth referral will only be made to *molecules* for brevity, however this should be treated interchangeably with *atoms* or *molecules* dependant on the simulation type.

Originally developed as a tool to simulate noble gases using hard-spheres, MD is now routinely applied to accurately study the motion of complex molecules in fields as disparate as materials science and biophysics [21]. In MD, the equation of motion for each molecule is calculated using Newton's second law:

$$\mathbf{F}_i = m_i \mathbf{a}_i = m_i \frac{d^2 \mathbf{r}_i}{dt^2} = \sum_{j \neq i}^N \left(-\frac{d\phi_{ij}(r_{ij})}{dr} \right) \frac{\mathbf{r}_{ij}}{r_{ij}}, \quad (1)$$

where \mathbf{F}_i and \mathbf{a}_i are the force acting on and acceleration of molecule i of mass m_i , respectively, and N is the total number

of molecules. \mathbf{r}_{ij} , r_{ij} , and $\phi_{ij}(r_{ij})$ are the position vector, distance and inter-molecular potential between the i th and j th molecule, respectively. The choice of $\phi_{ij}(r_{ij})$ in Eq. (1) determines the inter-molecular interactions, which in this work is specified by the 12-6 Lennard-Jones potential, given as:

$$\phi_{ij}(r_{ij}) = 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right], \quad (2)$$

where σ is the distance at which the inter-molecular potential is zero, and ϵ is the depth of the potential well. The r^{12} term in Eq. (2) is short-ranged and repulsive, and models Pauli repulsion resulting from the overlap of electron orbitals, while the r^6 term in Eq. (2) is long-ranged and attractive, modelling van der Waals attraction. The choice of σ and ϵ is determined by the fluid being modelled. In this work, the velocity-Verlet algorithm was used to integrate the equations of motion for all molecules. For computational efficiency, all forces are truncated at an appropriate radial distance (typically on the order of a few σ) from every molecule.

As the molecules move, collide, and advance in space and time, continuous molecular trajectories are produced. Using basic information about the molecule's motion (e.g. relative locations, velocities), measurable material properties (e.g. temperature, pressure) can be obtained using well-understood relations. Time-averaging can be used for steady flows to obtain accurate results for a system comprising of relatively small numbers of molecules. In this work, mdFoam+ [7], an MPI parallelised solver implemented entirely within the OpenFOAM [9] software framework, is used for all MD calculations.

2.2. Direct simulation Monte Carlo

The Direct Simulation Monte Carlo (DSMC) method is the most widely used method for simulating dilute gases [22]. It was initially introduced as the direct coding of a stochastic process which mimics the dynamics of a monatomic ideal gas but, later, it was proven that it provides solutions converging (within a suitable time limit), to that of the hard-sphere Boltzmann equation [23].

In DSMC simulations the gas is represented by a number of computational point-like particles. For general flow-fields, each particle can either represent multiples of real molecules or, when appropriate, can have a one-to-one mapping.

The simulation domain is divided into a mesh of cells whose size is smaller than the particles' mean free path, namely the average distance travelled by particles between two consecutive collisions. The motion of particles is decoupled from collisions by a fractional-step method, where the time step is shorter than the local mean free time between collisions. The particles are first translated as if they do not interact with each other. In this free-streaming sub-step, boundary conditions are taken into account. Then, the particles are sorted into cells and collisions are evaluated based on stochastic rules which, in effect, correspond to the Monte Carlo evaluation of the collision integral of the Boltzmann equation.

The mesh of cells is also used to evaluate the macroscopic properties of the ensemble. These are obtained through weighted averages of the particles' properties. As with MD, in steady flow conditions the averaging time can be long enough to obtain accurate results by a relatively small number of particles. However, for time-dependent problems the possibility of time averaging is lost or reduced and acceptable accuracy can be achieved only by increasing the number of particles or superposing several flow snapshots obtained from statistically independent simulations of the same flow. This work uses dsmcFoam+ [8], which, like mdFoam+, is an MPI parallelised solver implemented entirely within the OpenFOAM [9] software framework.

2.3. Coupling in HPC environments

This work uses the general coupling library MUI [11], as its basis. Generalised code coupling has received a number of notable contributions over the past decade [24,25] but, more recently, solutions designed specifically for use in high performance computing (HPC) environments have emerged [26] and MUI fits into this latter category.

The MUI library has been selected for this work for a number of key reasons; first and foremost it is designed to allow easy coupling between different methodologies. It does this by generalising domain-specific representations, such as a mesh, into a cloud of points and then providing spatial sampling techniques to interpolate these back into an appropriate representation (see Fig. 2). It has also been designed to handle the challenges of different time-stepping scales when considering multi-scale modelling approaches and therefore allows one solver to operate at a different time-stepping frequency to the other, while providing temporal interpolation methods to allow an appropriate value to be used in the domain with the lowest frequency. These capabilities make MUI ideally suited to create a direct coupling between an MD and DSMC solver, both of which are particle-based in nature, and as an indirect method is developed, the ability to interpolate in space and time will also prove useful.

A notable difference between MUI and other similar libraries is that its design is entirely based in the MPI MPMD paradigm, which has been part of the MPI standard for approximately five years and is implemented in all major distributions. Other similar coupling libraries rely on different data transport mechanisms, such as direct communication using the TCP/IP protocol or perhaps files stored on disk. The key issues with these approaches revolve around system portability and performance. In cases where anything other than MPI or files are used to communicate, then the majority of HPC systems do not allow this within standard security policy. Where files are the primary data transfer mechanism then this can introduce a significant overhead in large-scale distributed computing environments, especially where non-parallel file types are used.

Another key aspect of deploying coupled solutions in HPC environments is that often it is desirable to couple solvers designed to operate in a parallel manner, typically facilitated by MPI. However, as the architecture of the modern supercomputer changes as we move towards exascale systems, codes are likely to be both parallelised using MPI for use across distributed computing resources (nodes) and also accelerated to use local resources per node (i.e. GPUs or large numbers of CPU threads). While MUI relies on MPI as its communication platform, it does so in a way that ensures it is able to connect solvers already parallelised or accelerated. It also considers how best to reduce the computational overhead introduced by adding MPI message passing between solvers. This is achieved using its *Smart Send* capability, which allows for each MPI rank in a parallel run to be matched to only those that are needed to ensure a coupling interface works as expected. This functionality is described in more detail in [11] and depicted in Fig. 3. The efficacy of this method to reduce unnecessary MPI communication is entirely dependant on the nature of the two domains being coupled and how different regions of each domain overlap in terms of where data is being sent and where it is received.

3. MD-DSMC direct coupling method

The method presented produces a direct-coupling capability between mdFoam+ and dsmcFoam+ in which one molecule in the MD domain is represented by one particle in the DSMC domain (and vice versa). Both methods presented use a similar explicit fixed time-stepping mechanism (please refer to [7,8] for more de-

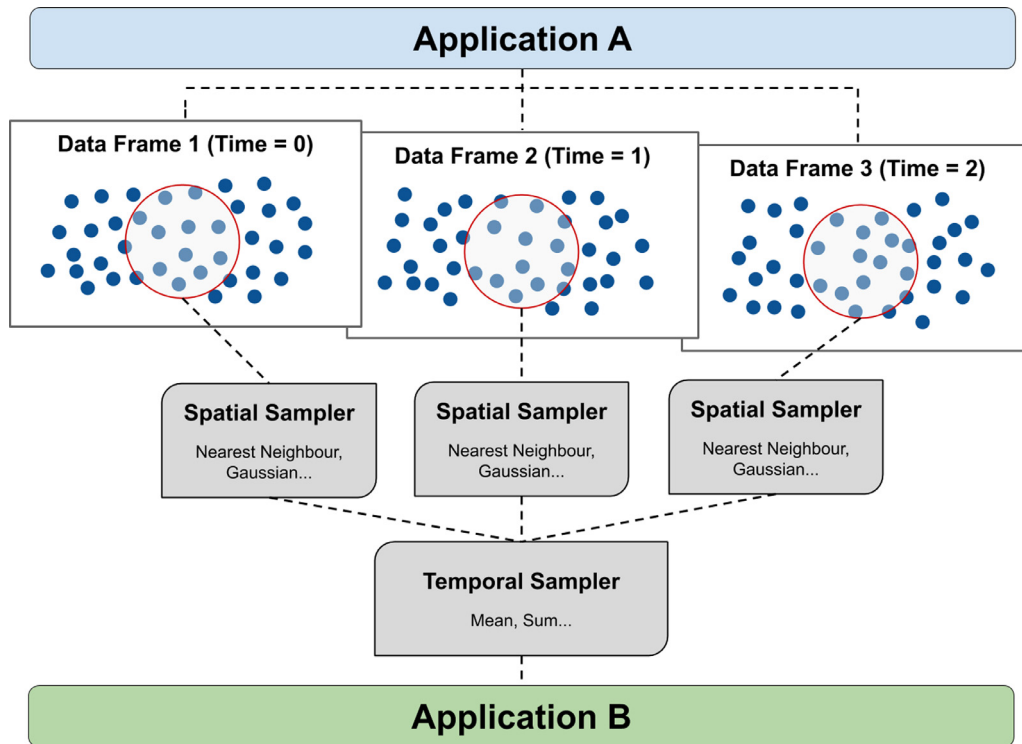


Fig. 2. Use of spatial and temporal data filters to allow generalised domain-specific values to be transferred between applications to enable coupling [11].

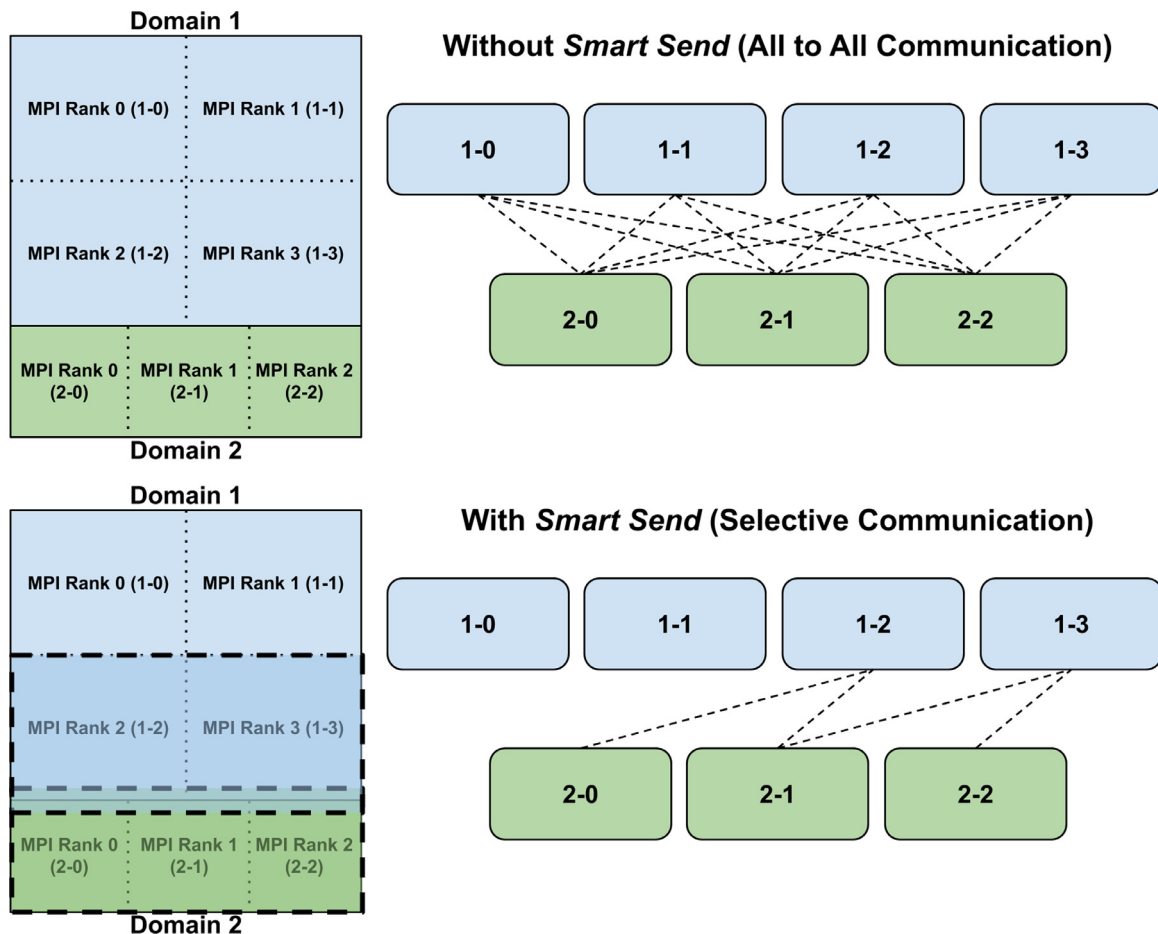


Fig. 3. MUI's communication reducing algorithm *Smart Send*. This prevents unnecessary data exchange between MPI ranks of different coupled solvers based on a geometric description of how each domain overlaps in terms of the regions in which it intends to send or receive data [11].

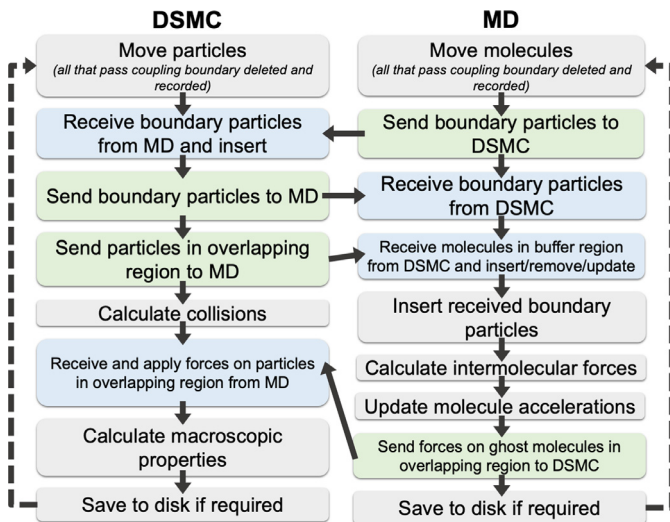


Fig. 4. Algorithmic design of the direct MD to DSMC coupling methodology between mdFoam+ and dsmcFoam+ using the MUI coupling library. Not shown is an initialisation step to ensure both the MD and DSMC domains are contiguous and the overlapping region is populated before the main calculation loop shown is started.

tails on the algorithms implemented in each solver) and it is assumed that coupled simulations performed using this method will operate at identical time-stepping frequencies. While the method shown is implemented using the two aforementioned OpenFOAM based solvers, it is fundamentally applicable for use between other, similarly capable MD or DSMC solvers. According to the definitions of Keyes et al. [27] this method is a two-way loose coupling with strongly coupled physics. The method is loose because neither solver is dependant on the other in order to operate and a distinct interface is used to facilitate data exchange, while the physics of the solution is strongly coupled because it relies on a significant overlap of solution domains, with both sides of the two-way coupling providing interaction with the other.

The complete framework is described in Fig. 4, with the overall cycle for each solver shown alongside the new steps needed to facilitate the coupling, as well as the direction data is transferred between them. An important point to note is that, as this is a direct coupling mechanism, there is no need for the use of spatial or temporal sampling, values are transferred directly. Therefore, MUI is used here as a data transport layer, with the expectation that future versions of the framework will use its sampling capability to enable a non-direct method.

From a computational algorithm perspective the MD and DSMC method are similar, both typically use an explicit time-stepping scheme and both generally conform to a regime in which particles are moved in space and time according to underlying forces. However, there is a significant difference between the methods which means they need to be considered quite differently, namely DSMC is stochastic whilst MD relies on Newtonian mechanics. MD directly represents a cohort of molecules and allows them to interact with either short or long-range force fields and electrostatic forces, whilst DSMC statistically represents the general behaviour of such a cohort using randomised interactions between particles that exist in the same computational cell of a mesh. This method allows DSMC to recover macroscopic properties of phenomena using the underlying stochastic molecular model as its basis, meaning it is possible for a single particle to represent many molecules. When combined, these two effects mean that DSMC is a significantly less computationally demanding problem than MD and, perhaps just as importantly, the fact it localises interactions between particles to a

single cell also means that it scales better in a parallel computing sense.

However, an important point is that DSMC is primarily used to capture the physics of rarefied gas flows (i.e. when the Knudsen number (Kn) of the flow, which is the ratio of the molecular mean free path length, relative to a representative length scale, is order 1 or larger), therefore there are many types of problem where it simply is not suited. MD has no such limitation, with the only caveat being the requirement that a suitable force field exists to capture the type of molecules being simulated. MD is orders of magnitude more computationally expensive than DSMC for similar problem sizes, therefore the goal of using a coupled MD-DSMC approach is to find a problem where DSMC can capture the physics of a rarefied gas, while the MD can be used to capture physics outside of the scope of DSMC, but not within the capabilities of alternative, less computationally demanding methods. The problem of evaporation (as defined in Section 1) is one such example. Notably, even when DSMC is used on a basis of one particle representing one molecule, it is still less computationally demanding than MD for the same problem due to its limited and simple inter-particle collision schemes when compared to even the short range force field calculations of MD.

The method presented uses a region where the MD domain overlaps the DSMC, in this region DSMC particles are directly represented in the MD simulation by a special type of molecule that is referred to here as a *ghost*. A ghost carries the same physical properties as the one it is representing in the DSMC domain but exists within the MD domain. It differs from a normal (i.e. non-ghost) molecule as its position is not updated due to force field calculations, but according to the position of the corresponding DSMC particle, molecules are added (or deleted) in the MD domain as the simulation progresses. Normal molecules interact with ghost molecules according to standard force field calculations and feedback is provided to the corresponding DSMC particles in the form of a returned external force, ghost molecules do not interact with each other in the MD domain. The size of the overlapping region is chosen on a per-case basis. However, a general rule is that its size should be minimised to reduce the number of ghost molecules needed, while still maintaining a full region of support for the force field calculations of the normal molecules in the MD domain. Typically a value slightly greater than the cut-off radius of the forcefield used has been found to be sufficient.

In order to facilitate mass transfer between the two domains, it is possible for a particle to pass from the DSMC domain completely and be copied into the MD domain (and vice versa). This is done by introducing a coupling boundary in each domain which represents a fixed point in space where, should a particle (DSMC), or molecule (MD) pass, it is deleted from its respective domain and added to the corresponding one. Conceptually this was a simple addition to both mdFoam+ and dsmcFoam+ as both use OpenFOAM as their underlying framework and therefore both use a typical mesh structure for their cell-linked list neighbour searching algorithms. As OpenFOAM is primarily a CFD codebase, the concept of boundaries is intrinsic to its framework, therefore adding a new type for use with these two solvers was straight-forward. One significant caveat however stems from the fact that DSMC is a stochastic method and MD is not.

This difference in nature means that a level of separation between molecules can be expected in an MD simulation, which is governed by the chosen force field and its parameters. However, as DSMC does not use a force field calculation to move its particles, rather it uses a randomly selective interaction regime, this means that it is possible for particles to exist closer to each other than is allowable in an equivalent MD simulation without any specific consequences. This is expected behaviour for DSMC but can introduce an issue when directly transferring DSMC particles as

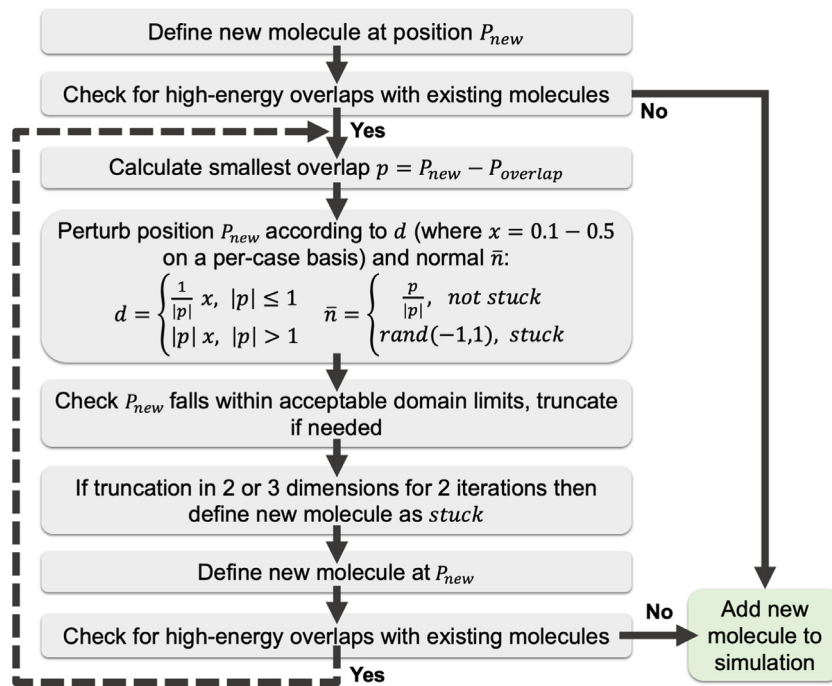


Fig. 5. Design of the algorithm to correct potential high-energy overlap events that can occur when inserting molecules into the MD domain at the position they exited the DSMC domain.

molecules into the MD domain, namely a new molecule may be instantaneously created closer to another than the chosen force field allows for, resulting in a significant (and spurious) release of energy into the MD simulation, as noted by Nedeia et al. [15]. In order to overcome this, before any new molecule is added to the MD domain due to a particle passing the DSMC coupling boundary, an overlapping energy check is performed on molecules that already exist within the force field cut-off radius around it. If placement at the initial position would result in a high-energy event occurring then an iterative scheme is used to find a new viable location as close as possible to the original. This algorithm is described in Fig. 5 and applied whenever a new molecule is inserted into the MD domain at P_{new} due to a particle passing a coupling boundary in the DSMC domain. The exact location of particles in a DSMC simulation does not have the same meaning as that of molecules in an MD simulation due to its stochastic nature, therefore the use of a different location for the creation of a molecule is acceptable as mass conservation is still preserved. Minimisation of how far the new molecule is perturbed is important to ensure mass is conserved in approximately the same location of the simulation domain.

4. Coupled results

There are two key aspects to evaluate the success of this coupled implementation. The first is to show that it reproduces the physics expected and the second is that it works in an HPC environment. The first of these is explored in this section and the second in Section 5.

In order to understand that the presented coupling mechanism captures mass transport between domains, a simple canonical NVT ensemble of Argon atoms was used. In this, mdFoam+ was used to simulate a central slice of the overall domain, with dsmcFoam+ used for the remaining outer regions (see Fig. 6). The outer boundaries for each domain were periodic, while the inner two (DSMC-MD and MD-DSMC) were of the new coupling type. To avoid the need to couple dsmcFoam+ with itself, the same instance was

used to simulate both outer regions, therefore they can be considered as a single simulation domain. The presented example is 250nm in length and 50nm in width and depth. It assumes an initial temperature of 300K and aims to represent Argon gas, therefore a density of $\rho_M = 1.603\text{kg/m}^3$ is used. This resulted in approximately 3000 molecules in the MD domain and 12,000 particles in the DSMC domain. The ensemble was simulated for a total of 200,000 time-steps to a final time of 1ns, resulting in a time-step size of $\Delta t = 5 \times 10^{-6}\text{ns}$. The DSMC simulation used a typical no time counter collision selection method and a Larsen Borgnakke variable hard sphere collision model, while this is unnecessarily complex for simulating Argon gas, it effectively simplifies to a classical hard sphere model in this case and shows capability for more complex future systems. The MD simulation used a standard 6-2 Lennard-Jones force field along with typical coefficients for Argon of $\sigma = 3.405 \times 10^{-10}$ and $\epsilon = 1.654 \times 10^{-21}$, a cut-off radius of 3.5σ was used.

Initialisation of each domain was achieved using the respective pre-processing tools for each solver, these produce layouts appropriate to their method, therefore an initial discrepancy between each domain in terms of important average values, such as kinetic energy, was unavoidable. In order to correct this and ensure the overall coupled domain started from a homogeneous state, an instantaneous method was used (an alternative would have been to simulate an equilibrated initial state by imposing a thermostat on one of the two domains in isolation). To achieve this, during the initialisation step for mdFoam+ and dsmcFoam+ in Fig. 5, the temperature of the MD (T_{md}) and DSMC (T_{dsmc}) domains are calculated and a velocity (U) re-scaling applied to the MD domain using a Gaussian thermostat according to

$$\tilde{U} = U \sqrt{\frac{T_{dsmc}}{T_{md}}} \quad (3)$$

The average linear kinetic energy per atom over the whole coupled ensemble is shown in Fig. 7, alongside the values for only the MD and DSMC domains. The fluctuations for the MD and DSMC domains are wholly expected as the overall number of atoms and

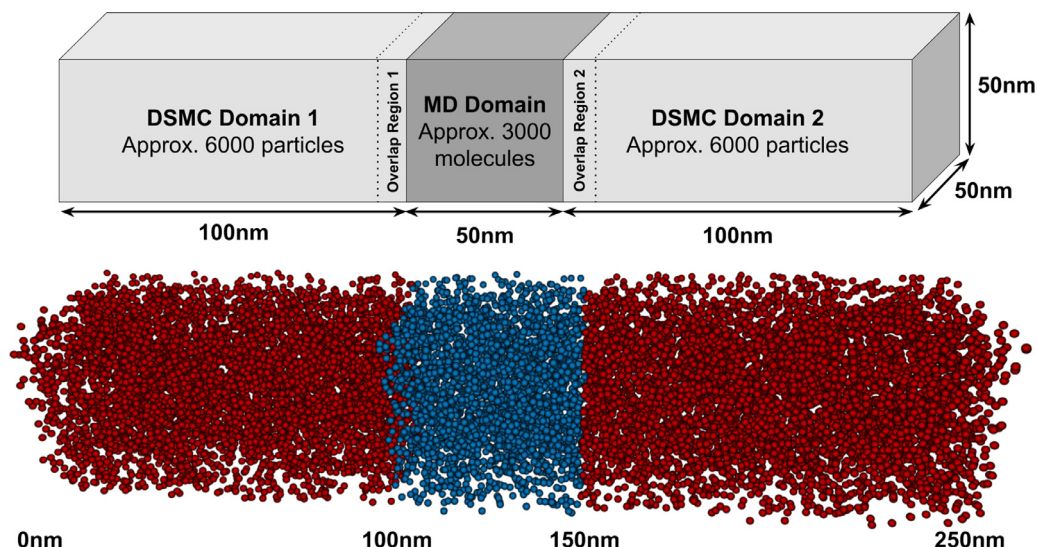


Fig. 6. A canonical NVT ensemble used to fundamentally test the direct MD-DSMC coupling method. The top of the image shows a schematic of the domain, while the bottom shows the resulting initial atomic layout.

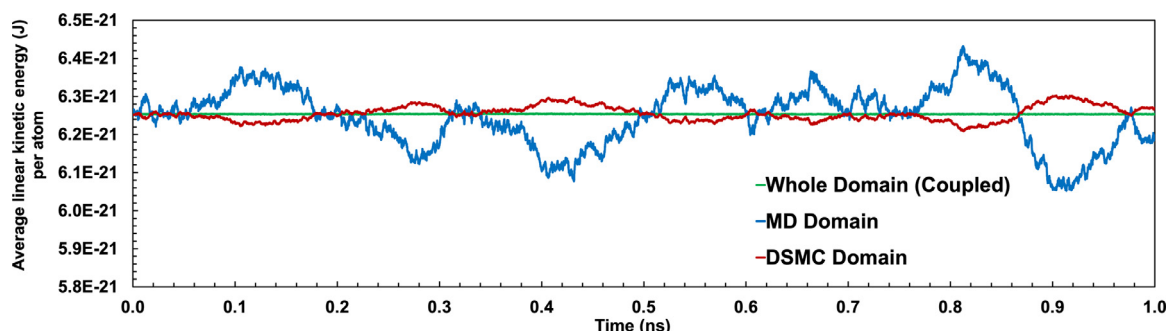


Fig. 7. Graph showing the average linear kinetic energy per atom for the entire coupled domain, the MD domain only and the DSMC domain only.

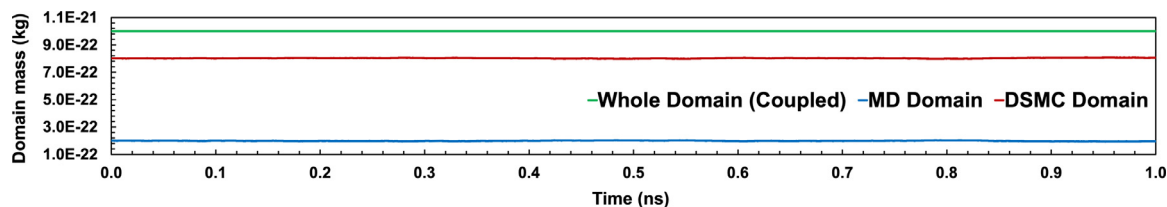


Fig. 8. Graph showing the total mass for the entire coupled domain, the MD domain only and the DSMC domain only.

particles in each also fluctuates by a small amount and demonstrate that energy transference between domains is occurring as expected. Notably the fluctuations of the MD domain are larger than that of the DSMC domain, this is simply because there are more DSMC particles than MD atoms. Importantly, the value for the coupled ensemble is as expected, demonstrating that linear kinetic energy is conserved in the coupled solution. In Fig. 8 the total mass of all atoms and particles in the MD, DSMC and coupled domains is presented, where it can be seen that the mass of the system is conserved.

To understand whether the coupled solution displays any unexpected behaviour in terms of density distribution, in particular around the areas where the domains overlap, a density profile of the whole coupled domain was produced. This can be seen in Fig. 9 as a matrix of histograms at six points in time during the complete 1 ns simulation. There are no notable anomalies in the distribution of molecules (in the MD domain) or particles (in the DSMC domain), with changing distributions in each of the 100 bins

within the levels expected for this kind of simulation. In particular, the bins overlapping the coupled regions (highlighted in green) do not show any particular bias and areas directly before and after each of the coupled region fluctuate in terms of their atomic density reasonably.

A comparison of the performance of the coupled solution against an MD and DSMC variant of the same simulation was also made using mdFoam+ and dsmcFoam+. All three simulations were run using the same computing hardware (Intel Core-i7 4770K providing 4 physical and 8 logical cores), using 4 MPI ranks for the MD and DSMC simulations and 4 MPI ranks for the MD domain plus 1 MPI rank for the DSMC domain for the coupled simulation. The results are shown in Fig. 10, where it can be seen that per time-step, the DSMC simulation was approximately 66× faster than the MD simulation time of 0.4109s at 0.0062s. The coupled variant was approximately 3× faster than the MD simulation at 0.1298s. The expected speed-up is case dependent, in this instance two thirds of the overall problem was simulated using DSMC, how-

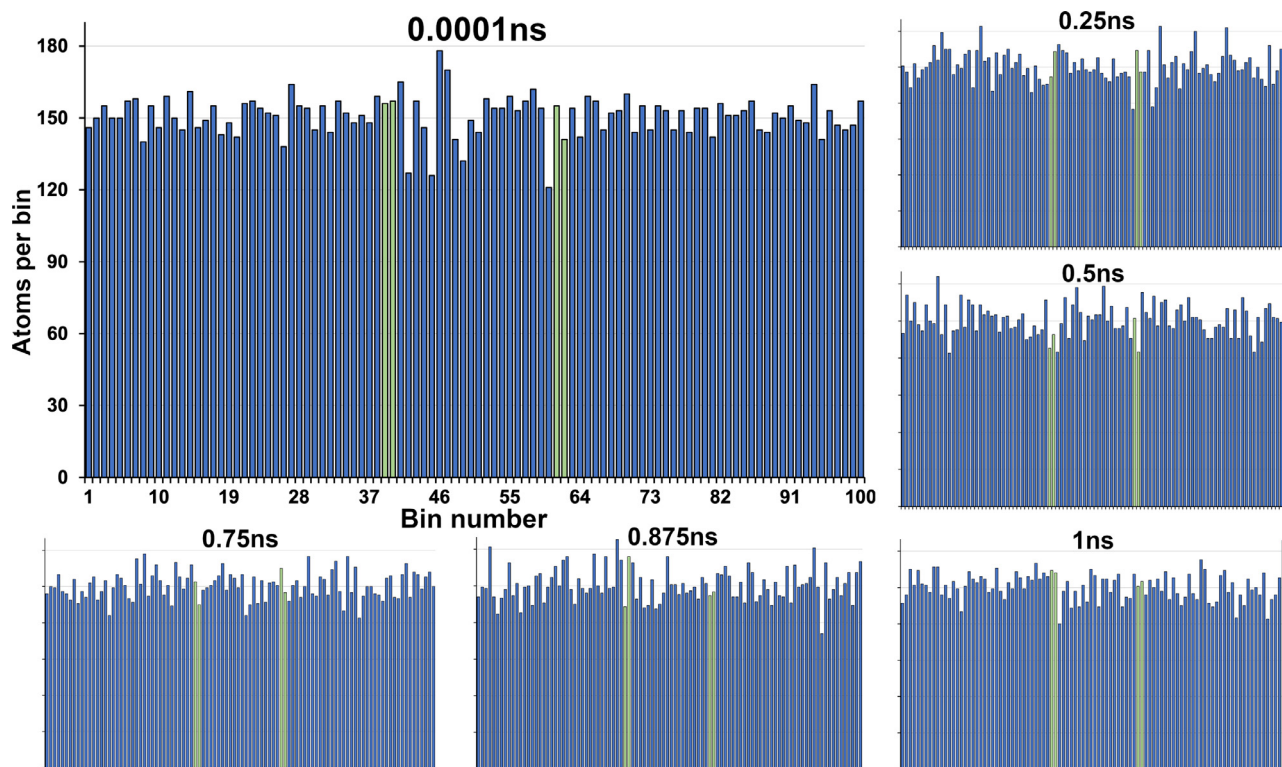


Fig. 9. Matrix of histograms showing the density profile of the coupled NVT solution according to 100 bins in the x direction, at six points in time through the entire 1ns of simulated time. Bins that intersect the overlapping coupled regions are highlighted in green. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

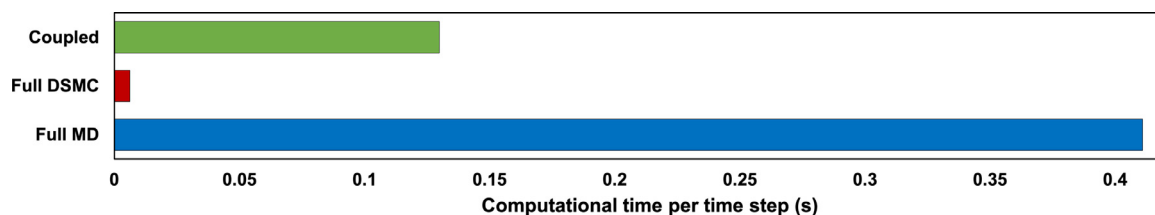


Fig. 10. Plot of the wall-clock time, per time-step of simulation, for the presented coupled canonical NVT ensemble, alongside the times for a full DSMC and MD version of the same problem.

ever if it were reduced or increased then this would change accordingly. It is clear, however, that the use of this coupled implementation can enable computational performance improvement when compared to using just MD.

5. Coupled performance

The two codes being coupled, mdFoam+ and dsmcFoam+, are both built using the OpenFOAM framework. This means that they both inherit its ability to use domain-decomposed MPI based parallelism. While the overall parallel scalability of each code is not best-in-class, this is explored and explained in each solvers respective publication [7,8] and is accepted as a baseline for this work. Therefore, while each may be outperformed by alternative codes, they both have advantages in other areas, such as extensibility and applicability to problems with arbitrarily complex boundaries. This work does not consider whether the two chosen codes perform well compared to alternatives, rather the overall performance of the coupled system is considered with the expected capabilities of each code taken into consideration.

For similar problem sizes and types dsmcFoam+ is faster than mdFoam+ (please see [7,8] for specifics), therefore the MD calculations are to be considered the bottleneck in this case. In this

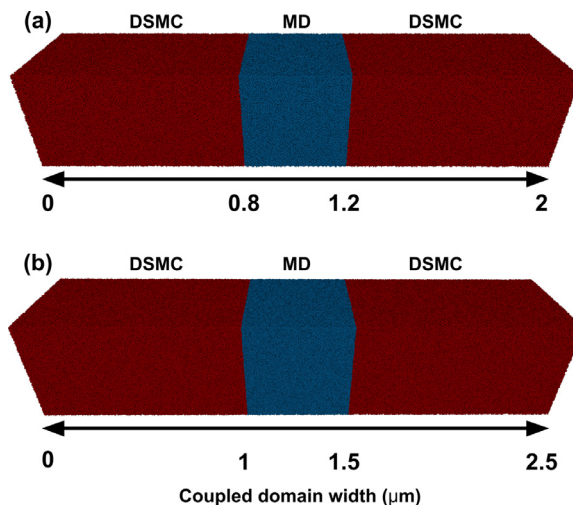


Fig. 11. Two large canonical NVT ensemble of Argon atoms used to test the parallel scaling of the presented direct MD-DSMC coupling method. The small (a) and large (b) ensembles are both shown, with red representing DSMC particles and blue MD molecules. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

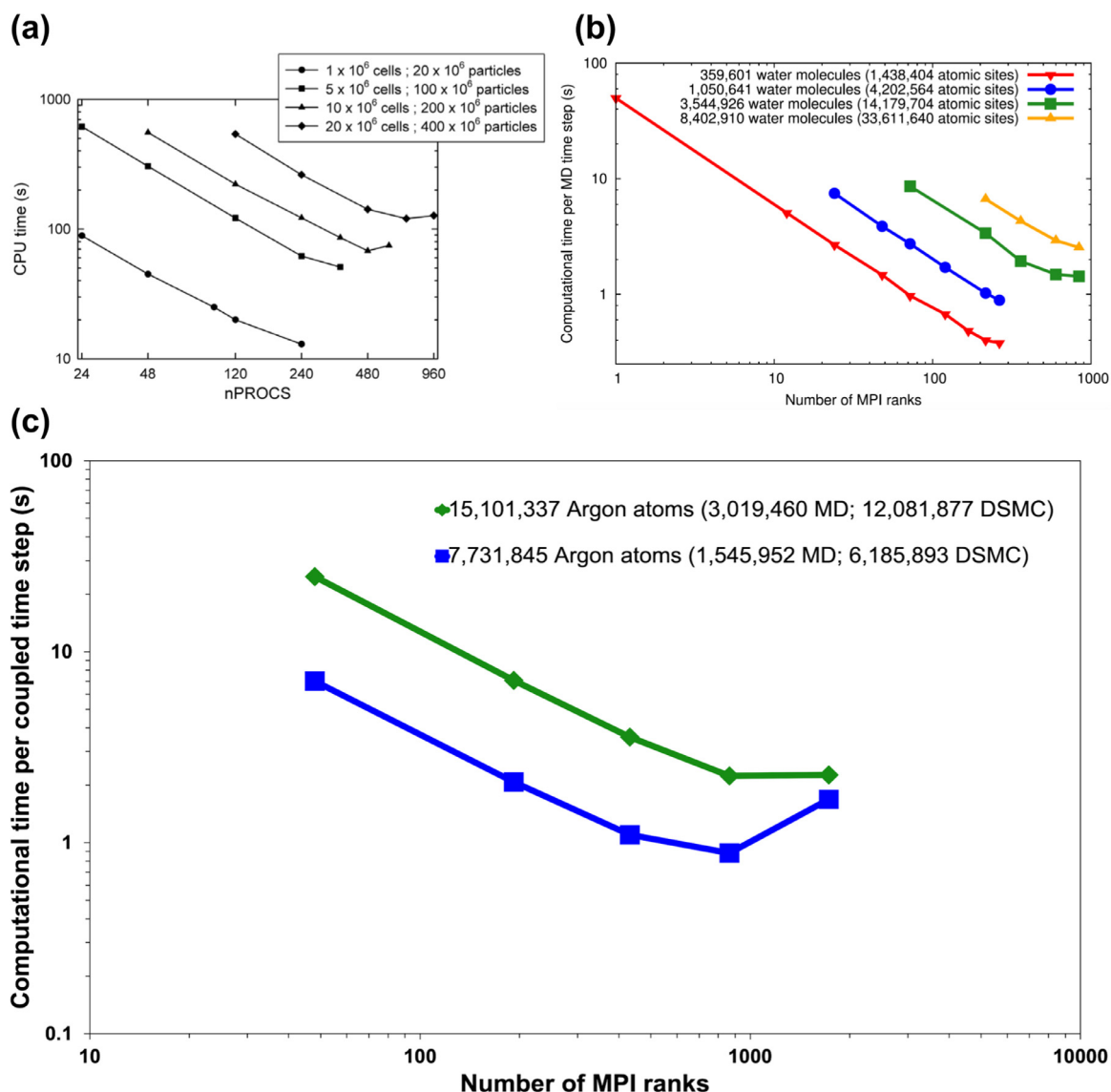


Fig. 12. Graphs showing the time per computational time step taken by the dsmcFoam+ (a) [8] and mdFoam+ (b) [7] solvers for various problem sizes, alongside results for two canonical NVT cases solved using the presented coupled method (c). All results obtained using the UK’s national supercomputing service ARCHER. Results are shown on a log-log scale to ensure consistency with graphs (a) and (b) from previous publications.

example the MD simulation is approximately four times smaller than the DSMC. However, the scalability of the DSMC algorithm is often close to linear (dependant on chosen underlying mesh cell sizes and frequency of particle collisions) while MD algorithms typically are not. It is therefore necessary to consider two aspects when working out how to best load balance available computing resources against each solver. The first is the size of the problem and the second is the computational demand from each solver. The first of these considerations may only be a variable in some scenarios, whilst others based on real-world problems may set this by definition, meaning only control over computational resource allocation is maintained. As the best way to load balance complex coupled simulations like this one is a topic in its own right and beyond the scope of this particular piece of work, a resource split of 50% of the total processors was allocated to each solver, in each case presented the number of DSMC particles is approximately 4× that of the MD atoms.

Tests were performed on the UK’s national supercomputing service ARCHER, a Cray XC30 machine with 4920 compute nodes, each providing 64GB of RAM and two 12 core Intel Ivy Bridge

CPUs. Past work has shown that mdFoam+ scales reasonably to approximately 300 processors with problem sizes of the order of 1 million molecules [7]. In order to explore the performance of this coupled solution, the same basic canonical NVT ensemble used in Section 4 was increased in size twice so a strong-scaling analysis could be made at two different problem sizes. The number of MD molecules and DSMC particles were therefore increased to approximately 1,500,000 and 6,000,000 respectively for a smaller case and 3,000,000 and 12,000,000 for a larger case. This produced an ensemble 2µm in length and 0.4µm in width and depth for the smaller case and 2.5µm in length and 0.5µm in width and depth for the larger case. Performance on between 2 (48 CPUs) and 72 (1728 CPUs) nodes was explored and results for the wall-clock time for a single iteration (averaged over 20 steps) can be seen in Table 1 and plotted in Fig. 12 in graph (c), alongside published scalability plots for mdFoam+ and dsmcFoam+ for reference. Visualisations of the two ensembles are shown in Fig. 11.

The scalability of the smaller test case starts to reduce between 18 and 36 nodes and then stops between 36 and 72, with overall performance worse at the higher node count. For the larger case,

Table 1

Wall-clock timings for between 2 and 72 nodes of the ARCHER supercomputer for two cases using the MD-DSMC coupled implementation. The processor resource split is shown, alongside recorded results of time per time-step for each case. The small case included approximately 7.5 million Argon atoms and the large case approximately 15 million.

| Nodes | Processors (MD/DSMC) | Wall-clock time (s) small | Wall-clock time (s) large |
|-------|----------------------|---------------------------|---------------------------|
| 2 | 24/24 | 7.02 | 24.72 |
| 8 | 96/96 | 2.08 | 7.08 |
| 18 | 216/216 | 1.1 | 3.57 |
| 36 | 432/432 | 0.88 | 2.24 |
| 72 | 864/864 | 1.69 | 2.26 |

reasonable scalability continues up to 36 nodes and then stops, with overall performance effectively identical at 72 nodes. This level of performance is inline with expectations given prior scalability studies of the two solvers, in particular mdFoam+ (please see graph (b) in Fig. 12) is the likely bottleneck, with the results achieved matching well with those obtained using the solver on its own. However, in this instance it was possible to simulate approximately 4× more Argon atoms in roughly the same computational time.

Generally speaking, it can be seen that the coupled solution does scale in an HPC environment, within the confines of each code used and taking factors such as problem size into account. This shows that the use of the MUI coupling library successfully implements the direct MD-DSMC method for parallel execution. It is also worth noting that while this implementation uses the mdFoam+ and dsmcFoam+ solvers, which are highly flexible in their design but notably slower than other similar codes computationally, the general methodology presented could be transferred to different MD or DSMC solvers in order to improve overall solution performance.

6. Conclusions and future work

This work has presented the first part of a larger piece of work, ultimately with the aim of producing a general simulation framework for complex problems that traverse scales all the way from microscopic to macroscopic such as studying evaporative processes. This initial work has concentrated on directly coupling a classical Molecular Dynamics (MD) solver with a Direct Simulation Monte Carlo (DSMC) solver using a general coupling library called the Multiscale Universal Interface (MUI). This method facilitates coupled simulation cases where molecules are directly considered in the DSMC domain and both domains utilise an equal time-step frequency. It is best suited for use cases where the problem size provides domains that are approximately equal in term of their order of magnitude but capturing the whole domain using MD directly is prohibitive due to computational cost, or where it is impossible to capture required physics using one method alone.

The method works as expected in terms of satisfying mass and kinetic energy conservation and is also shown to be capable when used in distributed high-performance computing environments. More work still needs to be done to better explore how to optimally load-balance problems and to understand what size of problem is needed to use a set amount of computing resource. The method also now needs to be applied to more complex physical problems in order to better understand its applicability to the wider problem types this work is aimed at. As both solvers coupled are already well-established and developed, with both providing a general simulation capability for their respective methods, this means that the resultant coupled method is also general in nature. In particular, both codes are able to handle the use of geometries of arbitrary complexity to define boundaries and domain shape, in turn this means this coupled solution is also able to consider problems with these attributes.

A next step for this work is to consider how to expand the presented methodology to handle indirect coupling, whilst maintaining the same HPC coupling framework. Specifically, this is needed when the DSMC portion of the coupled problem uses the method's coarse-graining capability, meaning a direct transference of mass by way of a molecule to particle exchange is no longer possible. This capability will enable the consideration of problems where the length and time-scales of the DSMC domain can be significantly different to that of the MD domain but careful consideration will have to be made to ensure conservation. Once this is complete it is then necessary to couple MD directly to macroscopic CFD and also to DSMC. The eventual goal is to produce a general and useful coupled framework using open-source software that is able to handle complex multi-scale problems, such as evaporation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

S.M. Longshaw: Conceptualization, Methodology, Software, Validation, Resources, Writing - original draft, Writing - review & editing, Visualization. **R. Pillai:** Conceptualization, Methodology, Validation, Writing - review & editing. **L. Gibelli:** Conceptualization, Methodology, Validation, Writing - review & editing. **D.R. Emerson:** Writing - review & editing, Supervision, Funding acquisition. **D.A. Lockerby:** Writing - review & editing, Project administration, Funding acquisition.

Acknowledgements

This work was supported by the UK Research and Innovation (UKRI) Engineering and Physical Sciences Council (EPSRC) under the programme grant EP/N016602/1. This work also used the ARCHER UK National Supercomputing Service.

References

- [1] Kunz R, Boger D, Stinebring D, Chyczewski LJ, Gibeling H, Venkateswaran S, et al. A preconditioned implicit method for two-phase flows with application to cavitation prediction. *Comput Fluids* 2000;29(8):849–75. doi:10.1016/S0045-7930(99)00039-0.
- [2] Kharangate C, Mudawar I. Review of computational studies on boiling and condensation. *Int J Heat Mass Transf* 2017;108:1164–96. doi:10.1016/j.ijheatmasstransfer.2016.12.065.
- [3] S NK, Premachandran B. A coupled level set and volume of fluid method on unstructured grids for the direct numerical simulations of two-phase flows including phase change. *Int J Heat Mass Transf* 2018;122:182–203. doi:10.1016/j.ijheatmasstransfer.2018.01.091.
- [4] S D, X J, C Q. Modeling of the evaporation and condensation phase-change problems with fluent. *Numer Heat Transf Part B* 2014;66(4):326–42. doi:10.1080/10407790.2014.915681.
- [5] S H, Z W, L Y, Y Z, M A. Simulation of the vapor-liquid two-phase flow of evaporation and condensation. *Int J Heat Technol* 2016;34(4):663–70. doi:10.18280/ijht.340416.

- [6] Meland R, Frezzotti A, Ytremus T, Hafskjold B. Nonequilibrium molecular-dynamics simulation of net evaporation and net condensation, and evaluation of the gas-kinetic boundary condition at the interphase. *Phys Fluids* 2004;16(2):223–43. doi:10.1063/1.1630797.
- [7] Longshaw S, MK B, Ramisetti S, Zhang J, Lockerby D, DR E, et al. mdFoam+: advanced molecular dynamics in OpenFOAM. *Comput Phys Commun* 2018;224:1–21. doi:10.1016/j.cpc.2017.09.029.
- [8] White C, Borg M, Scanlon T, Longshaw S, John B, Emerson D, et al. dsmcFoam+: an OpenFOAM based direct simulation Monte Carlo solver. *Comput Phys Commun* 2018;224:22–43. doi:10.1016/j.cpc.2017.09.030.
- [9] Weller HG, Tabor G, Jasak H, Fureby C. A tensorial approach to computational continuum mechanics using object-oriented techniques. *Comput Phys* 1998;12(6):620–31. doi:10.1063/1.168744.
- [10] OpenFOAM-2.4.0-MNF. 2020. <https://github.com/MicroNanoFlows/10.5281/zenodo.3596736>.
- [11] Tang Y, Kudo S, Bian X, Li Z, Karniadakis G. Multiscale universal interface: a concurrent framework for coupling heterogeneous solvers. *J Comput Phys* 2015;297(C):13–31. doi:10.1016/j.jcp.2015.05.004.
- [12] Smith E. On the coupling of molecular dynamics to continuum computational fluid dynamics. *The school of Mechanical Engineering*; 2013.
- [13] Barbante P, Frezzotti A. A comparison of models for the evaporation of the Lennard-Jones fluid. *Eur J Mech-B/Fluids* 2017;64:69–80. doi:10.1103/PhysRevE.57.7259.
- [14] Li J, Liao D, Yip S. Coupling continuum to molecular-dynamics simulation: reflecting particle method and the field estimator. *Phys Rev E* 1998;57(6):7259. doi:10.1103/PhysRevE.57.7259.
- [15] Nedeá S, Frijns A, van Steenhoven A, Markvoort A, Hilbers P. Hybrid method coupling molecular dynamics and Monte Carlo simulations to study the properties of gases in microchannels and nanochannels. *Phys Rev E* 2005;72(1):016705. doi:10.1103/PhysRevE.72.016705.
- [16] Gu K, Watkins CB, Koplik J. Atomistic hybrid DSMC/NEMD method for nonequilibrium multiscale simulations. *J Comput Phys* 2010;229(5):1381–400. doi:10.1016/j.jcp.2009.10.035.
- [17] Zeifman MI, Garrison BJ, Zhigilei LV. Direct simulation Monte Carlo calculation: strategies for using complex initial conditions. *Mater Res Soc Symp Proc* 2002;731:33–8.
- [18] Liang T, Ye W. An efficient hybrid DSMC/MD algorithm for accurate modeling of micro gas flows. *Commun Comput Phys* 2014;15(1):246–64. doi:10.4208/cicp.141112.160513a.
- [19] Watvisave D, Puranik B, Bhandarkar U. A hybrid MD-DSMC coupling method to investigate flow characteristics of micro-devices. *J Comput Phys* 2015;302:603–17. doi:10.1016/j.jcp.2015.09.012.
- [20] Scanlon T, White C, Borg M, Palharini R, Farbar E, Boyd I, et al. Open-source direct simulation Monte Carlo chemistry modelling for hypersonic flows. *AIAA J* 2015;53(6):1–11. doi:10.2514/1.J053370.
- [21] Allen MP, Tildesley DJ. *Computer simulation of liquids*. second ed. Oxford: Oxford University Press; 2017.
- [22] Bird GA. *Molecular gas dynamics and the direct simulation of gas flows*. Oxford: Oxford University Press; 1994.
- [23] Wagner W. A convergence proof for Bird's direct simulation Monte Carlo method for the Boltzmann equation. *J Stat Phys* 1992;66:1011–44.
- [24] Borgdorff J, Mamonski M, Bosak B, Groen D, Belgacem M, Kurowski K, et al. Multiscale computing with the multiscale modeling library and runtime environment. *Procedia Comput Sci* 2013;18:1097–105. doi:10.1016/j.procs.2013.05.275.
- [25] Joppich W, Kürschner M. MpCCI - a tool for the simulation of coupled applications. *Concurr Comput Pract Exper* 2006;18(2):183–92. doi:10.1002/cpe.913.
- [26] Bungartz H-J, Lindner F, Gatzhammer B, Mehl M, Scheufele K, Shukaev A, et al. preCICE - a fully parallel library for multi-physics surface coupling. *Comput Fluids* 2016;141:250–8. doi:10.1016/j.compfluid.2016.04.003.
- [27] Keyes D, McInnes L, Woodward C, Gropp W, Myra E, Pernice M, et al. Multiphysics simulations: challenges and opportunities. *Int J High Perform Comput Appl* 2013;27(1):4–83. doi:10.1177/1094342012468181.