Modelling solutions of biomolecules at atomistic and continuum representation at the same time: hybrid MD/hydrodynamics framework

Sergey Karabasov, Dmitry Nerukh, Arturs Scukins,
Evgen Pavlov, Anton Markesteijn,
Pardis Tabaee, Makoto Taiji

Queen Mary University of London, Aston University,
Cambridge University, RIKEN

Multiscale Modelling of
Condensed Phase and Biological Systems, 8 Jan 2014

The project

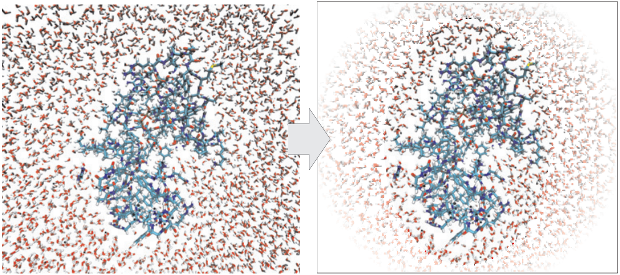
Title: ‘Using next generation computers and algorithms for modelling the dynamics of large biomolecular systems’

Consortium: 5 groups (Japan, UK, Russia)

Funding: G8 Research Councils Initiative on Multilateral Research Funding - Exascale Computing

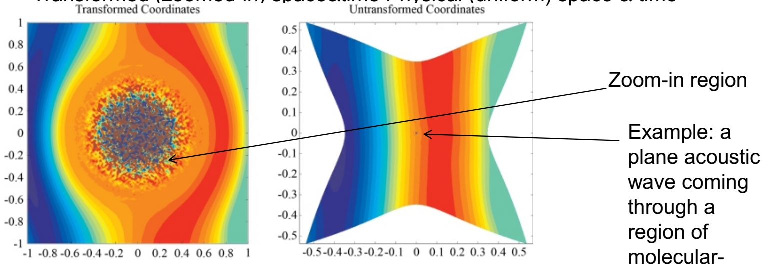
Hybrid MD/HD: motivation

- **Motivation**: multiphysics/multiscale (speed up and data reduction)
- **Examples**:
  - bridging atomistic times and microfluidic mixer times (9 orders of magnitude difference),
  - the effects of viscosity and hydrodynamic shear on protein folding.



multiphysics/multiscale

Transformed (zoomed-in) space&time Physical (uniform) space & time

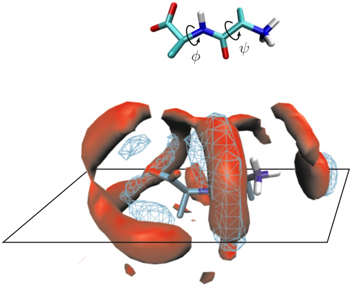


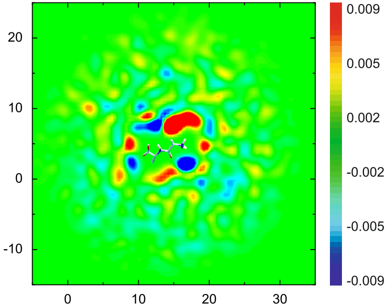
Example: a plane acoustic wave coming through a region of molecular-scale fluctuations

AP Markesteijn and SA Karabasov, *J. Comput. Phys.*, **258**, 137 (2014)

Fluctuations in biomolecular systems

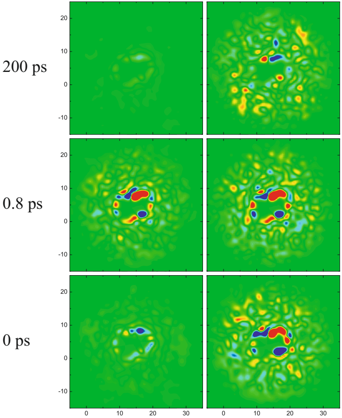
Water density around dialanine zwitterion





Question: to what extend the dynamics of water density (fluctuations) is connected with the dynamics of the peptide (quantified by its dihedral angles ϕ and ψ)?

Water density correlation with the peptide motion



Answer: not only connected but strongly correlated at very specific periods, when the conformational transitions occur

D Nerukh and S Karabasov, *J. Phys. Chem. Lett.*, **4**, 815 (2013)

The fundamentals: hydrodynamics

Continuous representation (hydrodynamics)

- All started with macroscopic thermodynamical quantities: the properties of the system **as a whole**, the largest possible scale.
- Describing the system at smaller scales: the properties become **fields** changing in **time**:

$$\rho(\mathbf{x}, t), \mathbf{u}(\mathbf{x}, t), T(\mathbf{x}, t).$$

- $\mathbf{x}$ is the Euclidean 3D space.
- The equations of motion are the FD equations.
- The solution is the values of the fields at each location in space at every instant of time: $\rho(\mathbf{x}, t), \mathbf{u}(\mathbf{x}, t), T(\mathbf{x}, t)$.

The fundamentals: atomistic

Atomistic representation

- The variables are the positions and momenta of the point masses, the atoms:

$$\{\mathbf{q}_1, \ldots, \mathbf{q}_N, \mathbf{p}_1, \ldots, \mathbf{p}_N\}$$

- The space is the $6N$-dimentional phase space.
- The atoms interact through empirically (in MD) defined Hamiltonian $H(\mathbf{q}, \mathbf{p})$
- The equations of motion describing $\mathbf{q}(t), \mathbf{p}(t)$ are the Hamilton equations

$$\frac{\mathrm{d}q_i(t)}{\mathrm{d}t} = \frac{\partial H(\mathbf{q}, \mathbf{p})}{\partial p_i}, \frac{\mathrm{d}p_i(t)}{\mathrm{d}t} = -\frac{\partial H(\mathbf{q}, \mathbf{p})}{\partial q_i}.$$

- The solution is the molecular trajectory: the values of the coordinates and momenta at every moment of time:

$$\mathbf{q}(t), \mathbf{p}(t).$$

Connecting the representations

Calculating the continuous density:

$$\rho_q(\mathbf{q};\mathbf{x},t) = \sum_{j=1}^{N} m\delta(\mathbf{q}_i(t) - \mathbf{x})$$

It is a *function of the molecular coordinates* (phase space variable), which also parametrically depends on $\mathbf{x}$ and t.

Connecting the representations

How the measurement is done: a probe of volume Δx is placed at the point $\mathbf{x}$ for a period of time Δt at time t.

The 'true' (measured) value of $\rho(\mathbf{x}, t)$ is obtained by overaging $\rho_q(\mathbf{q}; \mathbf{x}, t)$ over Δx and Δt.

$$\rho(\mathbf{x}, t) = \langle \rho_q(\mathbf{q}; \mathbf{x}, t) \rangle_{\Delta x, \Delta t}$$

Connecting the representations

This is MD→HD transformation.

HD→MD - ???

Describing fluctuations of the continuum

The Landau-Lifshitz Fluctuating Hydrodynamics (LL-FH) equations are a generalisation of the deterministic Navier-Stokes (NS) equations:

$$\frac{\partial \rho}{\partial t} + \nabla(\rho \mathbf{u}) = 0,$$
$$\frac{\partial \rho u_i}{\partial t} + \nabla(\rho u_i \mathbf{u}) = \nabla_j \left(\Pi_{ij} + \tilde{\Pi}_{ij}\right),$$
$$\frac{\partial \rho E}{\partial t} + \nabla(\rho E \mathbf{u}) = \nabla_j \left[\left(\Pi_{ij} + \tilde{\Pi}_{ij}\right) \cdot u_i\right] + \nabla(\mathbf{q} + \tilde{\mathbf{q}}).$$

Describing fluctuations of the continuum

The stress tensor consists of a deterministic part

$$\Pi_{ij} = -\left(p - \eta_V \nabla \mathbf{u}\right)\delta_{ij} + \eta\left(\partial_i u_j + \partial_j u_i - 2D^{-1}\nabla\mathbf{u}\cdot\delta_{ij}\right)$$

and a stochastic part, a random Gaussian matrix with zero mean and the covariance

$$\langle \tilde{\Pi}_{ij}(r_1,t_1)\cdot\tilde{\Pi}_{kl}(r_2,t_2)\rangle = 2k_BT\left[\eta\left(\delta_{ij}\delta_{ik} + \delta_{ik}\delta_{jl}\right) + \left(\eta_V - \frac{2}{3}\eta\right)\delta_{ij}\delta_{jk}\right]\delta(r_1 - r_2)\delta(t_1 - t_2).$$

This form of correlations follows from the fluctuation-dissipation theorem, which relates the thermal fluctuations to temperature.

Describing fluctuations of the continuum

The heat flow is also a sum of the averaged flow

$$q_i = \kappa \cdot \partial_i T$$

and a stochastic component with zero mean and the covariance

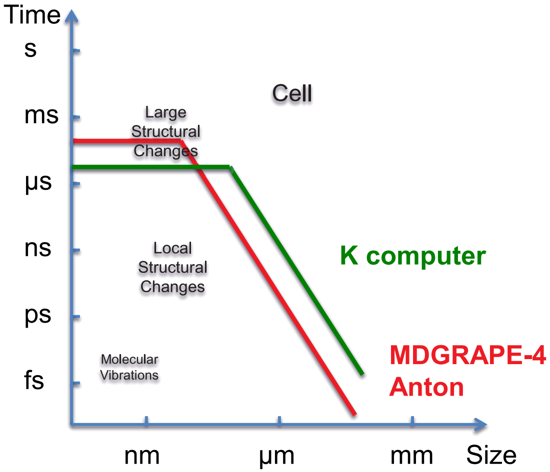
$$\langle \tilde{q}_i(r_1, t_1) \cdot \tilde{q}_j(r_2, t_2) \rangle = 2k_B \kappa T^2 \delta_{ij} \delta(r_1 - r_2) \delta(t_1 - t_2)$$

Describing fluctuations of the continuum

$$\tilde{\Pi}_{ij} = \sqrt{\frac{2k_B T}{\Delta x \Delta t}} \left(\sqrt{2\eta} \cdot \mathbf{G}^s_{ij} + \sqrt{D\eta_V} \frac{tr[\mathbf{G}]}{D} \mathbf{E}_{ij} \right)$$

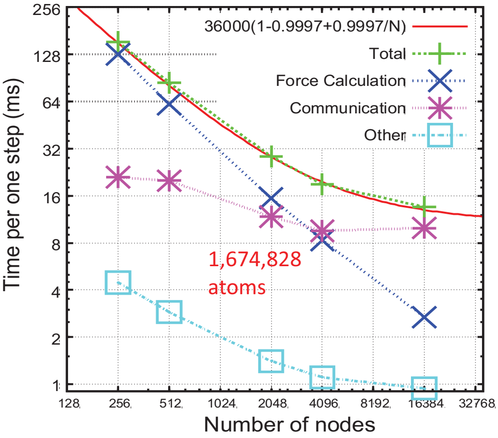
$$\tilde{q}_i = \sqrt{\frac{2k_B \kappa T^2}{\Delta x \Delta t}} \mathbf{G}_i$$

Biomolecular scales



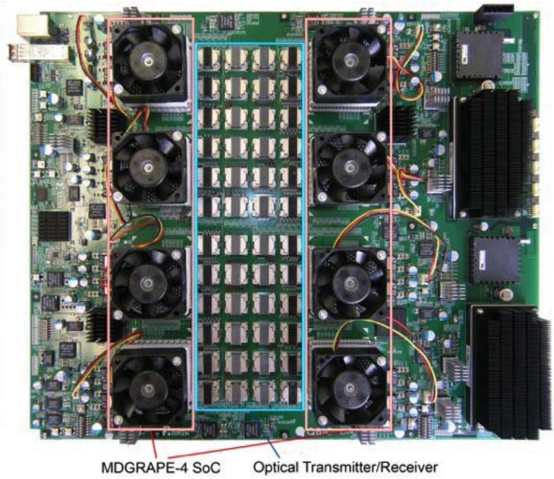
Quick digression: hardware

RIKEN **K-computer**: 10Pflops, 640000 nodes
scaling:

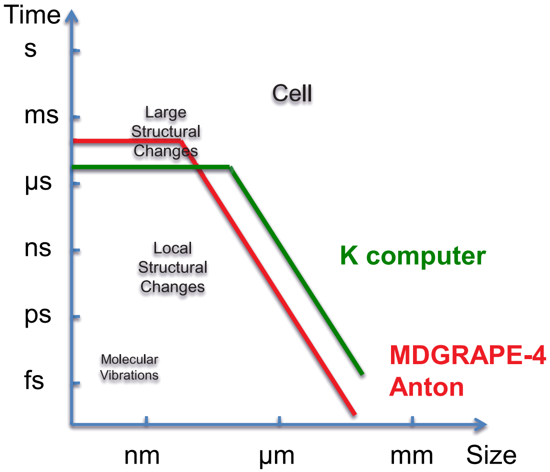


Hardware

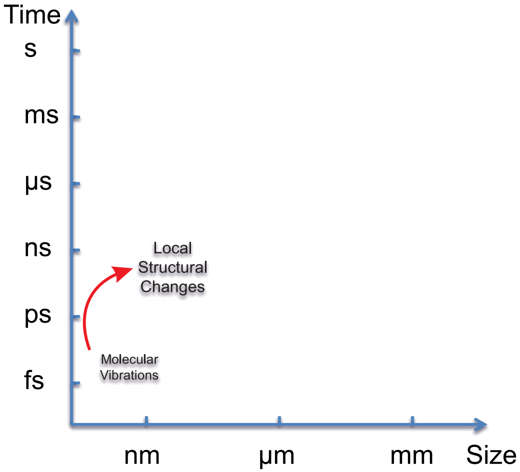
MDGRAPE-4: 100 ps per day for a 100K atoms system



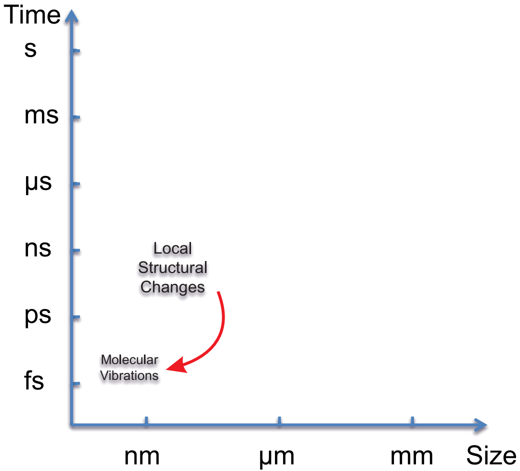
Biomolecular scales



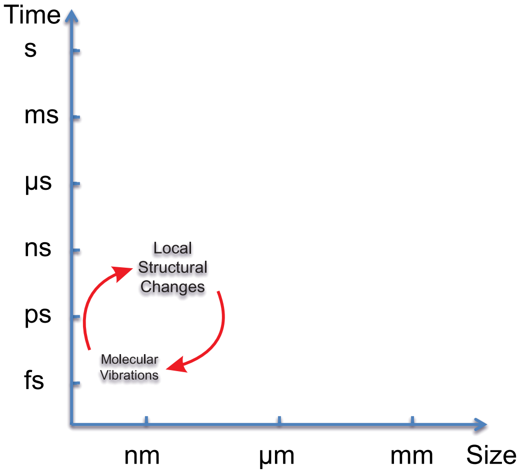
Scales coupling: acyclic 'bottom-up'



Scales coupling: acyclic ‘top-down’



Scales coupling: cyclic



An example of acyclic ‘top-down’ approach

O’Connell Thompson (1995):

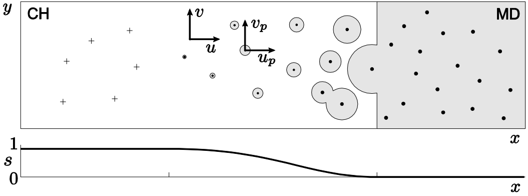
$$\frac{dx_{ip}}{dt} = u_{ip} + s\left(\frac{\sum_N m_{ip}}{M_{CFD}} u_i - \frac{\sum_N m_{ip} u_{ip}}{\sum_N m_{ip}}\right),$$

$$\frac{d}{dt} u_{ip}^{Newton} = \frac{F_{ip}}{m}$$

- CFD = Deterministic N-S model
- Application of repulsive barrier to retain particles
- Not necessarily conserves macroscopic momentum balance
- Not fully coupled (no feedback from MD to CFD)

see our poster for our version of ‘top-down’ coupling

Our framework



- The end domains HD and MD are described by purely hydrodynamic and purely Newtonian equations of motion respectively.
- In the hybrid domain the fluid consists of two "phases":
  - HD phase is a continuum water with volume fraction $s = \frac{V_1}{V}$,
  - MD phase is a phase that incorporates atoms, its volume fraction is $(1 - s)$.
- The parameter $s = s(x)$ is the function of space coordinates, such that $s = 1$ in the HD domain, $s = 0$ in the MD domain.

A Markesteijn, S Karabasov, A Scukins, D Nerukh, V Glotov, and V Goloviznin, submitted to *Phil. Trans. R. Soc. A*

Mass conservation

For HD phase:

$$\frac{\partial}{\partial t}(s\rho) + \frac{\partial}{\partial x_i}(u_i s\rho) = J,$$

For MD phase:

$$\frac{\partial}{\partial t}\left((1-s)\sum_{p=1,N(t)}\rho_p\right) + \frac{\partial}{\partial x_i}\left((1-s)\sum_{p=1,N(t)}\rho_p u_{ip}\right) = -J,$$

where $\rho_p = m_p/V$ is the density of MD particles and J is the birth/death rate due to the coupling between the phases.

$\tilde{\rho} = s\rho + (1-s)\sum_{p=1,N(t)}\rho_p$

Conservation of momentum

For HD phase:

$$\frac{\partial}{\partial t}(s u_i \rho) + \frac{\partial}{\partial x_j}(u_j u_i s \rho) = s F_i + J_2,$$

where J_2 is the HD-MD interaction force and F_i is the hydrodynamic force.

For MD phase:

$$\frac{\partial}{\partial t}\left((1-s)\sum_{p=1,N(t)} u_{i,p}\rho_p\right) + \frac{\partial}{\partial x_j}\left((1-s)\sum_{p=1,N(t)} \rho_p u_{i,p} u_{j,p}\right) = (1-s)\sum_{p=1,N(t)} F_{i,p} - J_2$$

$$\tilde{\rho}\tilde{u}_j = \left[s\rho u_j + (1-s)\sum_{p=1,N(t)} \rho_p u_{jp}\right]$$

Restricted dynamics

The deviations of $\tilde{\rho}$ are driven towards the correct value $\sum_{p=1,N(t)} \rho_p$:

$$\frac{D}{Dt_0}\left(\tilde{\rho} - \sum_{p=1,N(t)} \rho_p\right) = L^{(\rho)} \cdot \left(\tilde{\rho} - \sum_{p=1,N(t)} \rho_p\right),$$

where $\frac{D}{Dt_0} \cdot = \frac{\partial}{\partial t} \cdot + \nabla(\mathbf{u} \cdot)$,

and similarly for $\tilde{u}_j \tilde{\rho}$:

$$\frac{D}{Dt_0}\left(\tilde{u}_j \tilde{\rho} - \sum_{p=1,N(t)} u_{jp}\rho_p\right) = L^{(u)} \cdot \left(\tilde{u}_j \tilde{\rho} - \sum_{p=1,N(t)} u_{jp}\rho_p\right) + s\nabla_j\left(\Pi_{ij} + \tilde{\Pi}_{ij}\right).$$

Restricted dynamics

$\tilde{\rho}$ is diffused towards $\sum_{p=1,N(t)} \rho_p$:

$$L^{(\rho)}.\left(\tilde{\rho} - \sum_{p=1,N(t)} \rho_p\right) = \frac{\partial}{\partial x_i}\left(s(1-s)\alpha\frac{\partial}{\partial x_i}\left(\tilde{\rho} - \sum_{p=1,N(t)} \rho_p\right)\right).$$

$\tilde{u}_j\tilde{\rho}$ is diffused towards $\sum_{p=1,N(t)} u_{jp}\rho_p$:

$$L^{(u)}.\left(\tilde{u}_j\tilde{\rho} - \sum_{p=1,N(t)} u_{jp}\rho_p\right) =$$

$$\frac{\partial}{\partial x_i}\left(s(1-s)\beta\frac{\partial}{\partial x_i}\left(\tilde{u}_j\tilde{\rho} - \sum_{p=1,N(t)} u_{jp}\rho_p\right)\right).$$

The sources J and J_2

From these constrains the sources J and J_2 can be found:

$$J = s\frac{\partial}{\partial t}\sum_{p=1,N(t)}\rho_p + \frac{\partial}{\partial x_i}\left(su_i\sum_{p=1,N(t)}\rho_p\right) + \frac{\partial}{\partial x_i}\left(s(1-s)\alpha\frac{\partial}{\partial x_i}\left(\tilde{\rho}-\sum_{p=1,N(t)}\rho_p\right)\right),$$

$$J_2 = s\frac{\partial}{\partial t}\sum_{p=1,N(t)}\rho_p u_{jp} + \frac{\partial}{\partial x_i}\left(su_i\sum_{p=1,N(t)}\rho_p u_{jp}\right) - sF_j + \frac{\partial}{\partial x_i}\left(s(1-s)\beta\frac{\partial}{\partial x_i}\left(\tilde{u}_j\tilde{\rho}-\sum_{p=1,N(t)}u_{jp}\rho_p\right)\right),$$

Modified MD equations

For known J and J_2 MD equations are modified to preserve macroscopic conservation laws:

$$\frac{dx_{ip}}{dt} = u_{ip} + s(u_i - u_{ip}) + s(1-s)\alpha \frac{\partial}{\partial x_i}\left(\tilde{\rho} - \sum_{p=1,N(t)} \rho_p\right)\frac{1}{\rho_p N(t)},$$

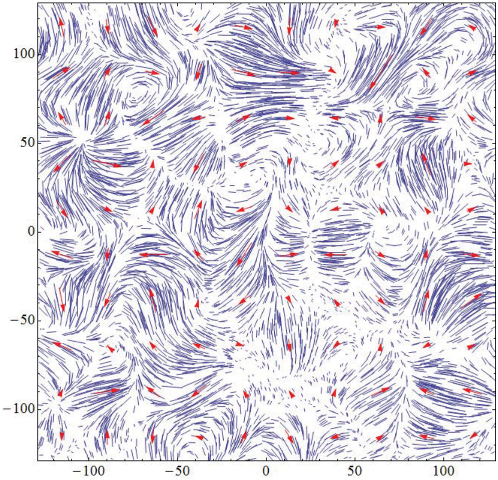
$$\frac{du_{jp}}{dt} = (1-s)F_{jp}/\rho_p + sF_j/\rho_p/N(t)$$

$$+ \frac{\partial}{\partial x_i}\left(s(1-s)\alpha \sum_{p=1,N(t)} u_{jp}/N(t)\frac{\partial}{\partial x_i}\left(\tilde{\rho} - \sum_{p=1,N(t)} \rho_p\right)\right)\frac{1}{\rho_p N(t)}$$

$$- \frac{\partial}{\partial x_i}\left(s(1-s)\beta \frac{\partial}{\partial x_i}\left(\tilde{u}_j\tilde{\rho} - \sum_{p=1,N(t)} u_{jp}\rho_p\right)\right)\frac{1}{\rho_p N(t)},$$

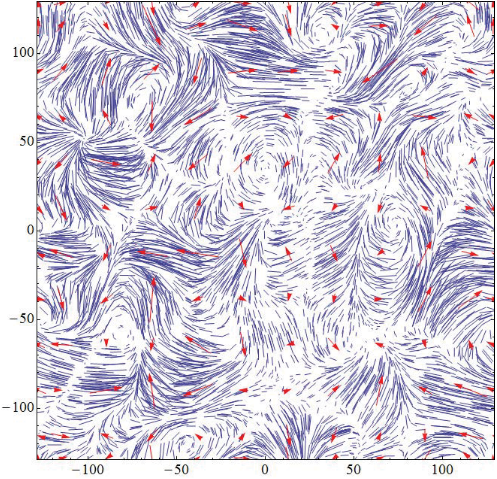
Results: 2D Lennard-Jones liquid

strong coupling: $s = 0.8$

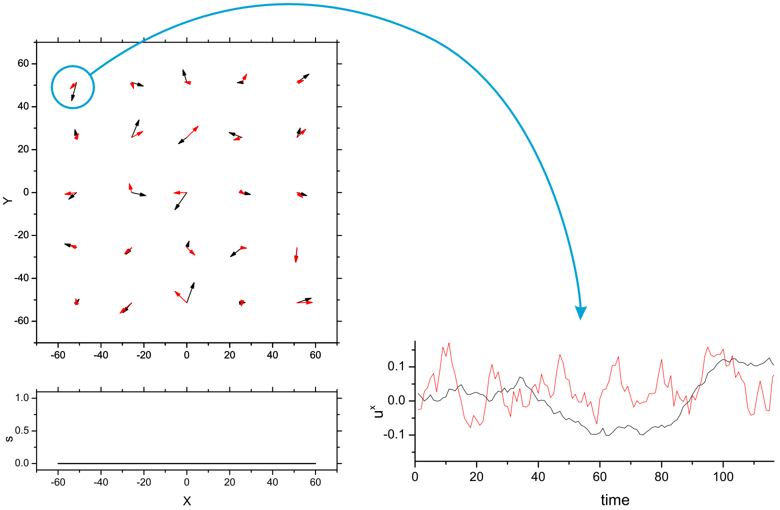


Results: 2D Lennard-Jones liquid

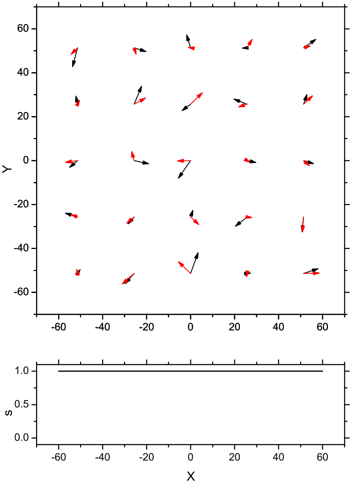
weaker coupling: $s = 0.6$

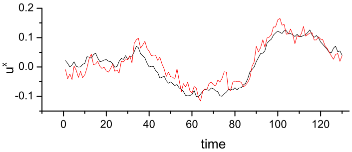


Velocities at various s

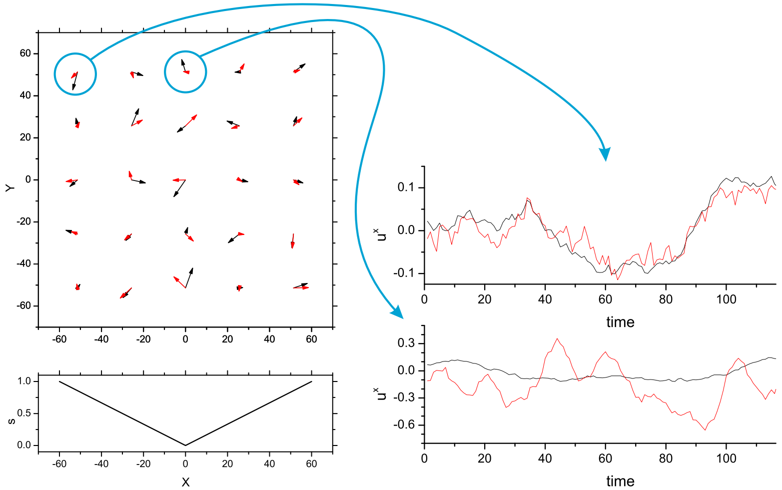


Velocities at various s



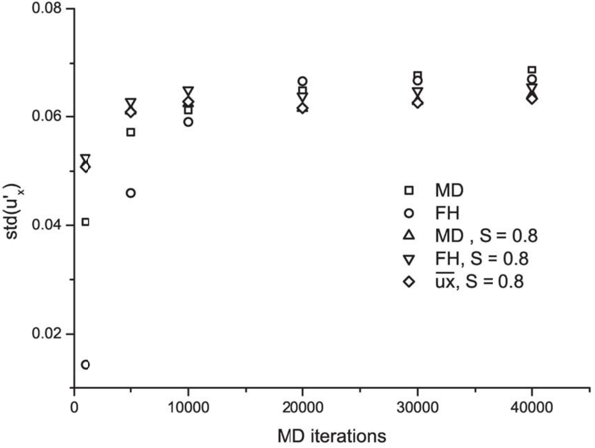


Velocities at various s



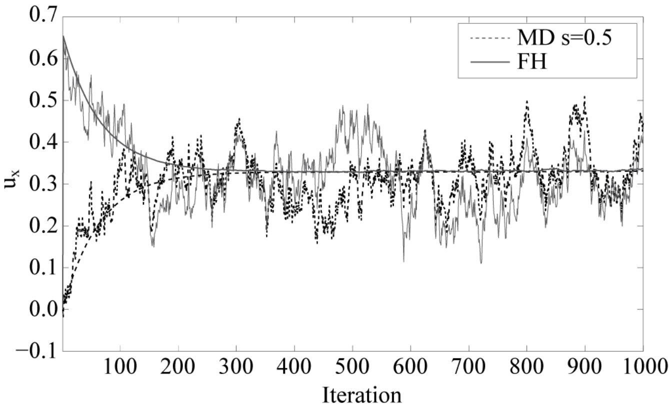
Results: 3D Lennard-Jones liquid

Convergence of fluctuations towards the same limit



Results: 3D Lennard-Jones liquid

Convergence of fluctuations towards the same limit



Conclusions

- Atomistic and continuum representations of liquid can be connected without artificial barriers or ad hoc correction forces in space and time.
- The domains of each representation can be defined arbitrarily in space and time.
- Challenges: *multiphysics* (non-stationary MD + hydrodynamics), multiscale *computing* (efficient multi space-time algorithms in parallel environment).

Thoughts for the future

The need to organize activities in multiscale modelling:

- a database of existing well documented algorithms and codes in multiscale modelling;
- optimising existing algorithms for computational efficiency (GPU, parallel scalability, distributed computing, etc);
- a database of benchmark cases (from the simplest molecular models to realistic biomolecular systems);
- training: hands-on workshops, schools on existing software.