Advanced Molecular Molecular Dynamics

Technical details

May 11, 2021

Today

- Discuss technical aspects required for MD simulations

How to calculate pressure?

The pressure is the derivative of the free-energy wrt the volume

Pressure definition: $$P = \frac{2}{V}(E_{\text{kin}} - W)$$

Virial definition: $$W(\mathbf{r}) = \frac{3}{2} V \frac{dU}{dV}$$ potential energy

For isotropic scaling: $$\frac{d\mathbf{r}_i}{dV} = \frac{\mathbf{r}_i}{3V}$$ (virial definitions can contain different factors)

The viral sum: $$W(\mathbf{r}) = -\frac{1}{2} \sum_i \mathbf{r}_i \cdot \mathbf{F}_i$$

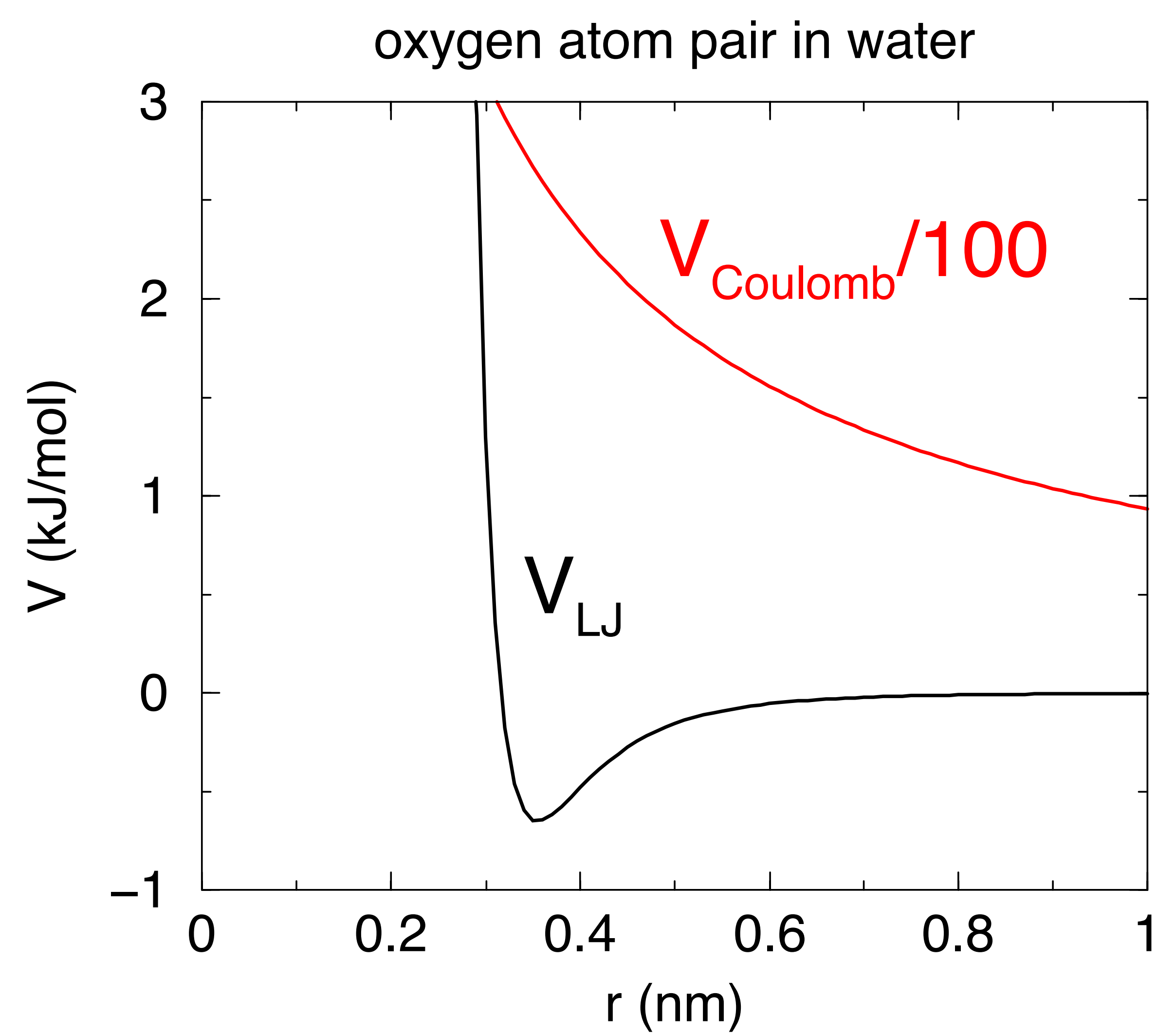
For pair interactions: $$W(\mathbf{r}) = -\frac{1}{2} \sum_{i<j} \mathbf{r}_{ij} \cdot \mathbf{F}_{ij}$$

How to calculate interactions

$$U(\mathbf{r}) = \sum_{bonds} U_{bond}(\mathbf{r}) + \sum_{angles} U_{angle}(\mathbf{r}) + \sum_{dihs} U_{dih}(\mathbf{r})$$
$$+ \sum_i \sum_{j>i} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} + \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}}$$

- We often only need: $\mathbf{F}_i = -\frac{dU}{d\mathbf{r}_i}$
- For bonded interactions: simply loop over items in the sum and calculate F (and U)
- Non-bonded for small molecules: do the double loop
- Cost $O(N^2)$: prohibitive for large systems

Non-bonded cut-off



- Cut-off interactions beyond a radius
- Fine for LJ
- Not fine for Coulomb
- Potential should be the integral of the force

$$V_{co}(r) = \begin{cases} V(r) - V(r_c) & r < r_c \\ 0 & r \geq r_c \end{cases}$$

Cut-off effect on Lennard-Jones

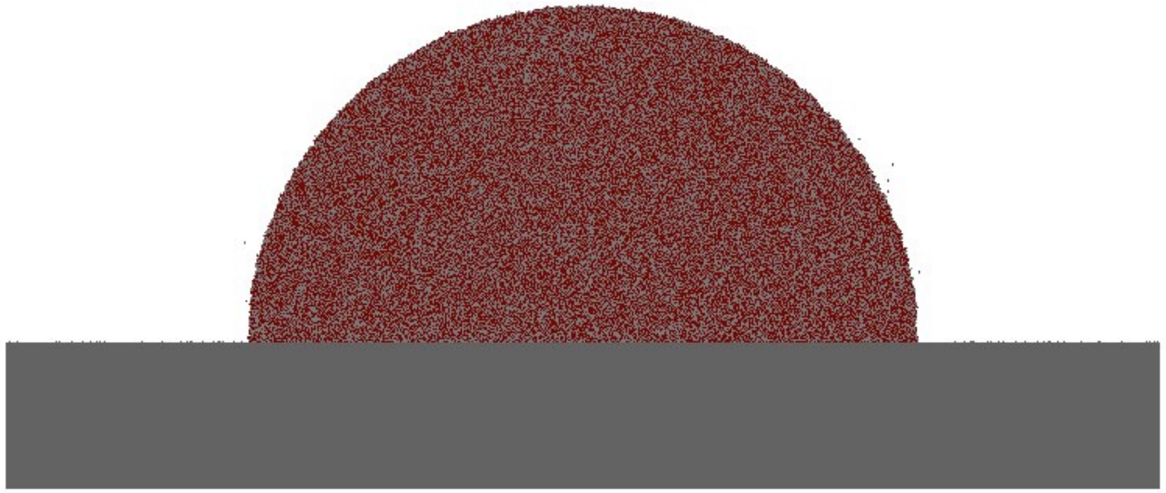
- Lennard-Jones potential decays a r^{-6}
- But one atom sees many others
- For constant density beyond the cut-off, the missing LJ energy is:

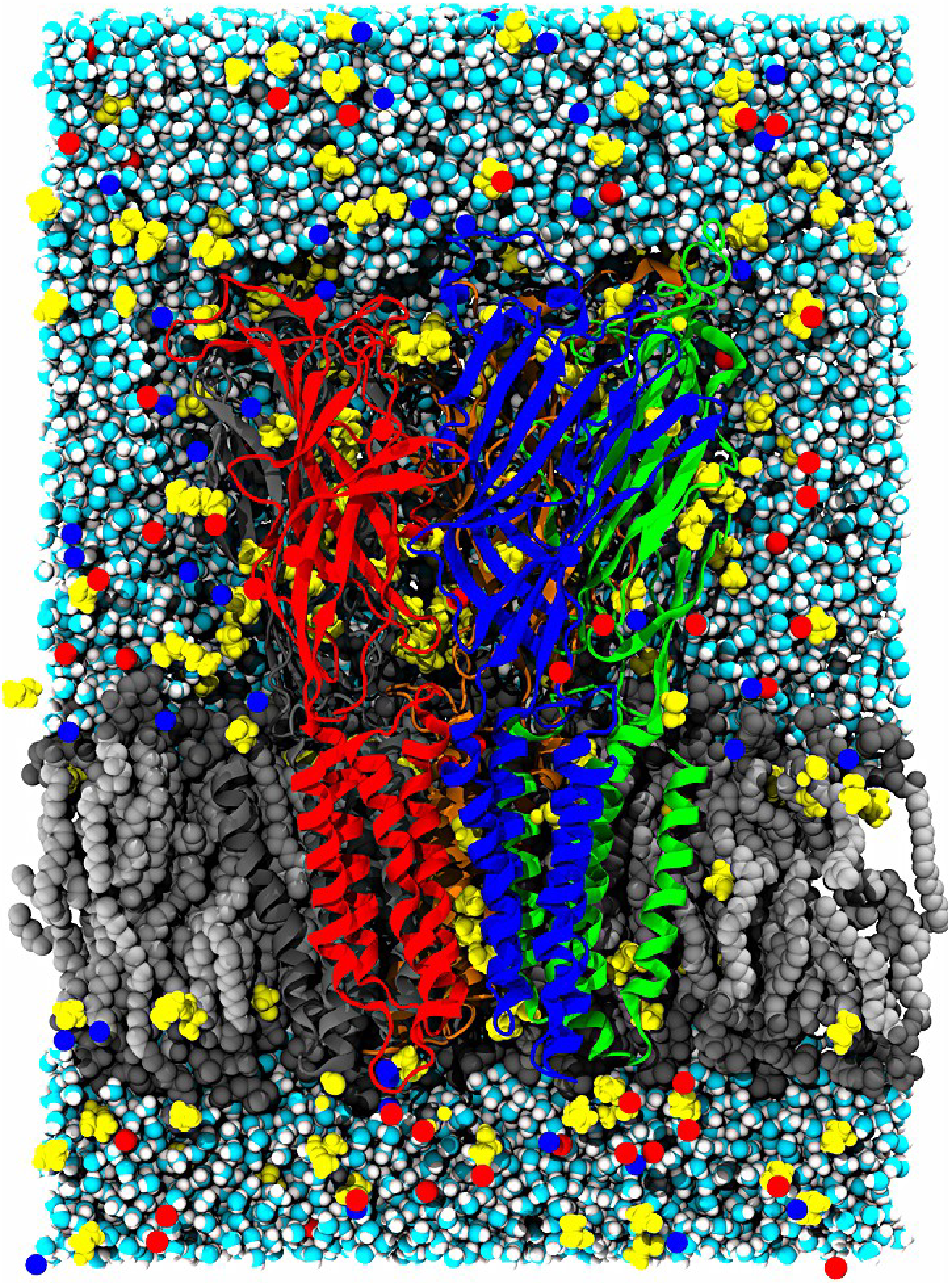
$$U_{LJ} = \int_{r_c}^{\infty} 4\pi r^2 \rho_N \overline{C_{6\,ij}} \frac{1}{r^6}\, dr = 3\rho_N \overline{C_{6\,ij}} \frac{1}{r_c^3}$$

- This missing attraction can be added:
  long-range or dispersion correction
- Virial correction is identical, but adds a factor of 6
- The pressure correction can be significant

Inhomogeneous dispersion

- Uniform correction does not work for inhomogeneous systems
- e.g. phase boundaries and lipid membranes



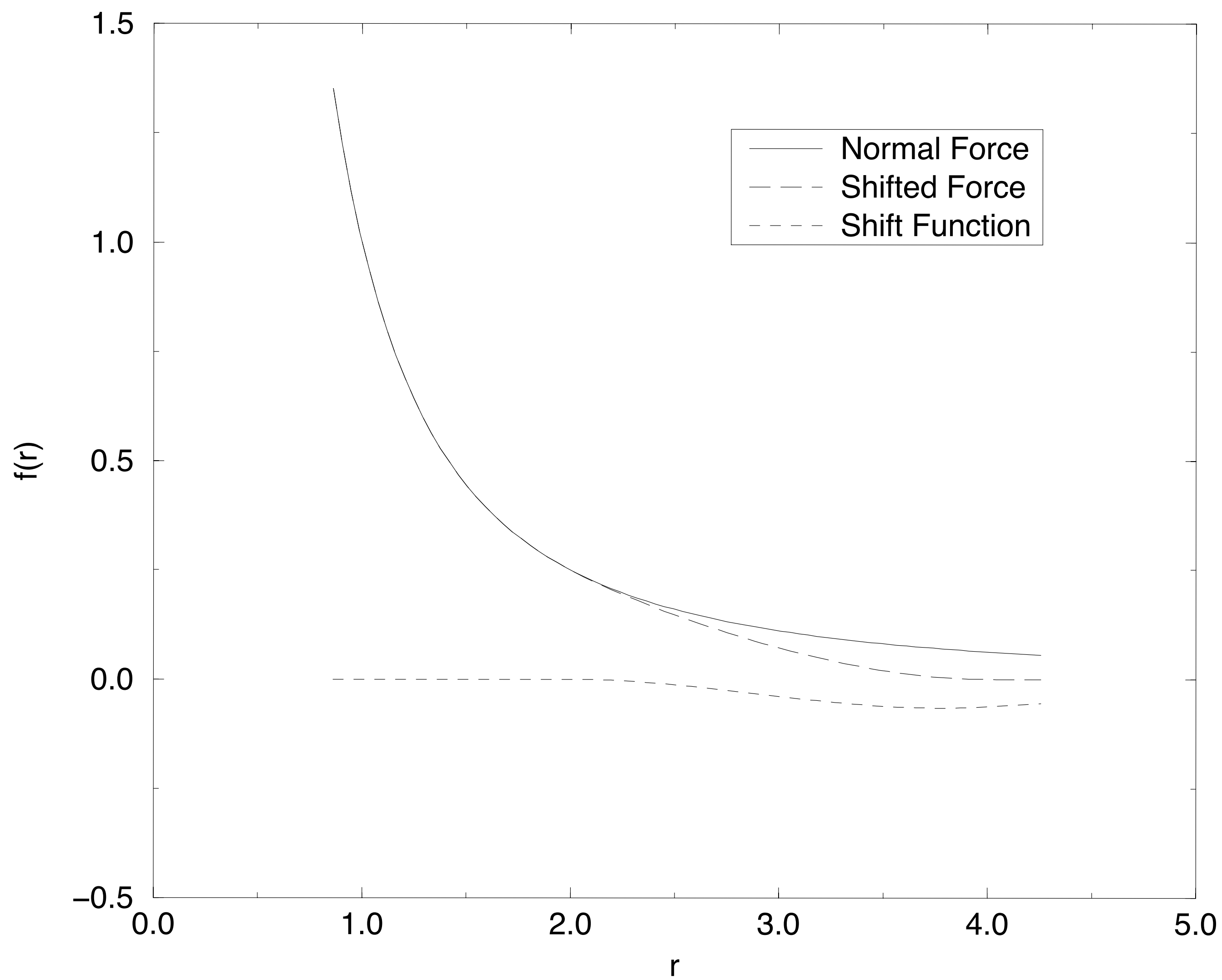


Cut-off & force fields

- The basis, including LJ parameters, of most bio-molecular force fields is decades old
- Simulation were done with cut-off's of 0.8/0.9 nm
- Force-fields were parametrized to give the correct density and ΔH_{vap} with the cut-off used
- Using a larger cut-off, or dispersion correction will thus result in a too high density
- To correct this: re-parametrize all LJ interactions
- Advice: use the right cut-off for the force field!

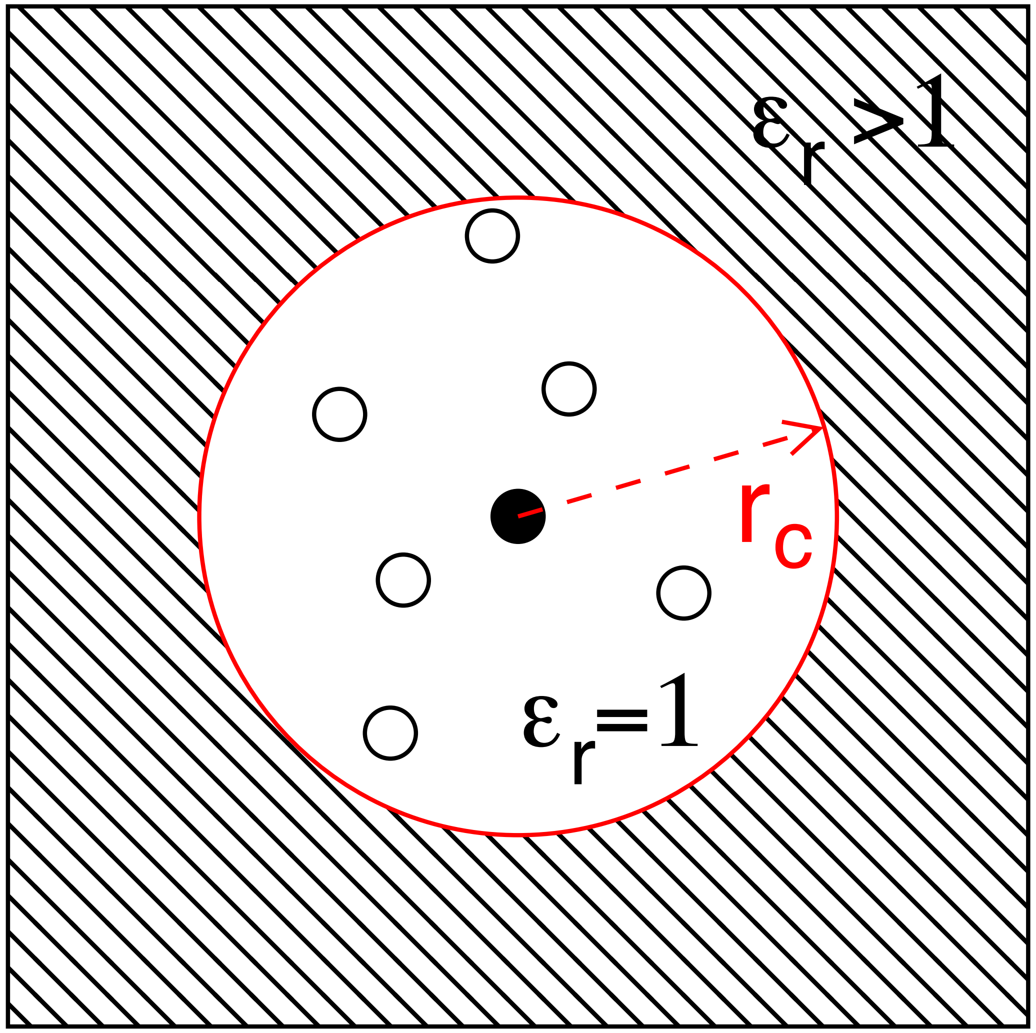
Non-zero force at cut-off

- With a "plain" cut-off: $F(r_c)!=0$
  - This could give integration errors
  - A huge problem for Coulomb
  - No real issue for LJ
    - larger issues for small r
- Solution:
  - switch F to 0
  - shift V to 0



Coulomb cut-off

- Coulomb interactions decay as 1/r
- Cut-off can't be used
- What to do?
- One option:
  - High dielectric:
    - weak electrostatics
    - use a reaction-field



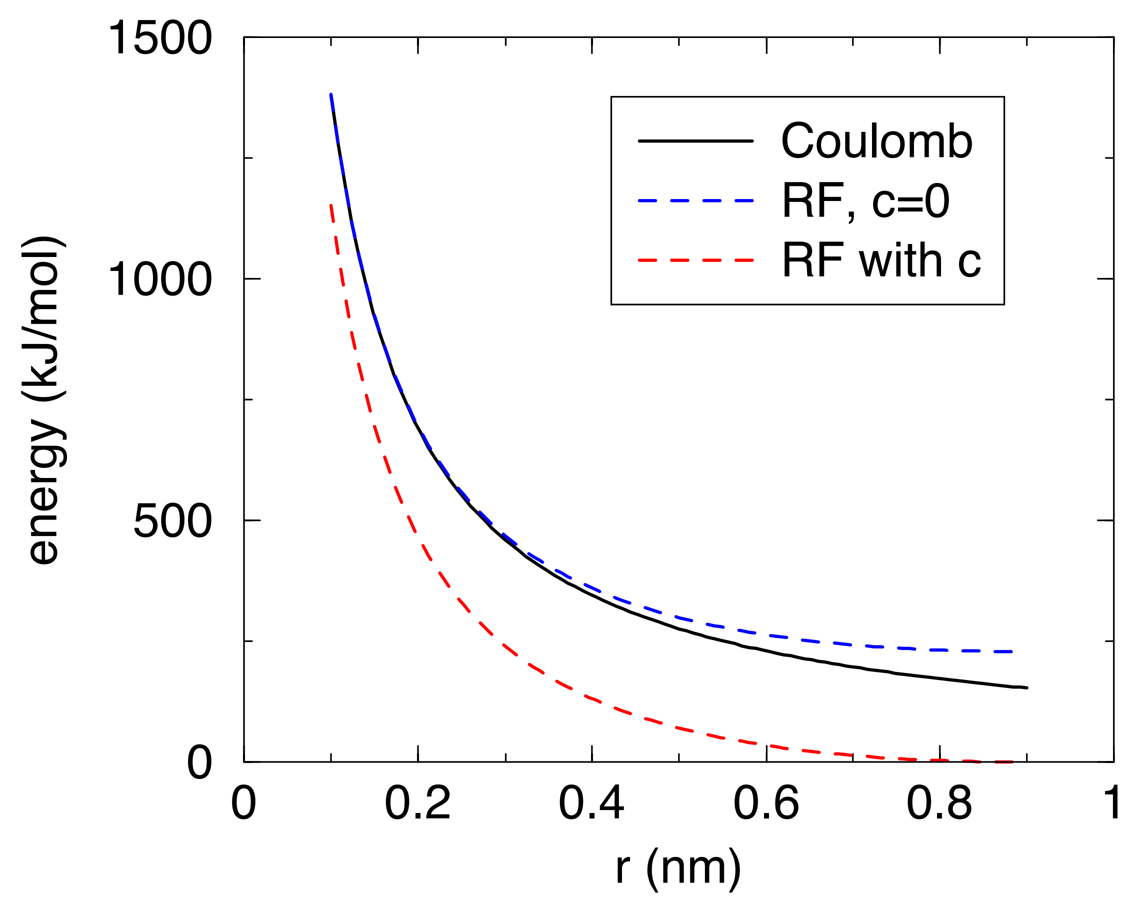
Reaction-field

$$V_{\mathrm{rf}} = \frac{q_i q_j}{4\pi\epsilon_0}\left[\frac{1}{r_{ij}} + k\, r_{ij}^2 - c\right]$$

$$k = \frac{1}{r_c^3}\frac{\epsilon_{\mathrm{rf}} - 1}{2\epsilon_{\mathrm{rf}} + 1}$$

$$c = \frac{1}{r_c}\frac{3\epsilon_{\mathrm{rf}}}{2\epsilon_{\mathrm{rf}} + 1}$$

- Developed for dipoles
- Linear dipole reaction F
- For charges:
  - Additional constant required for $V(r_c)=0$
  - Implicit assumption:
    - uniform background charge



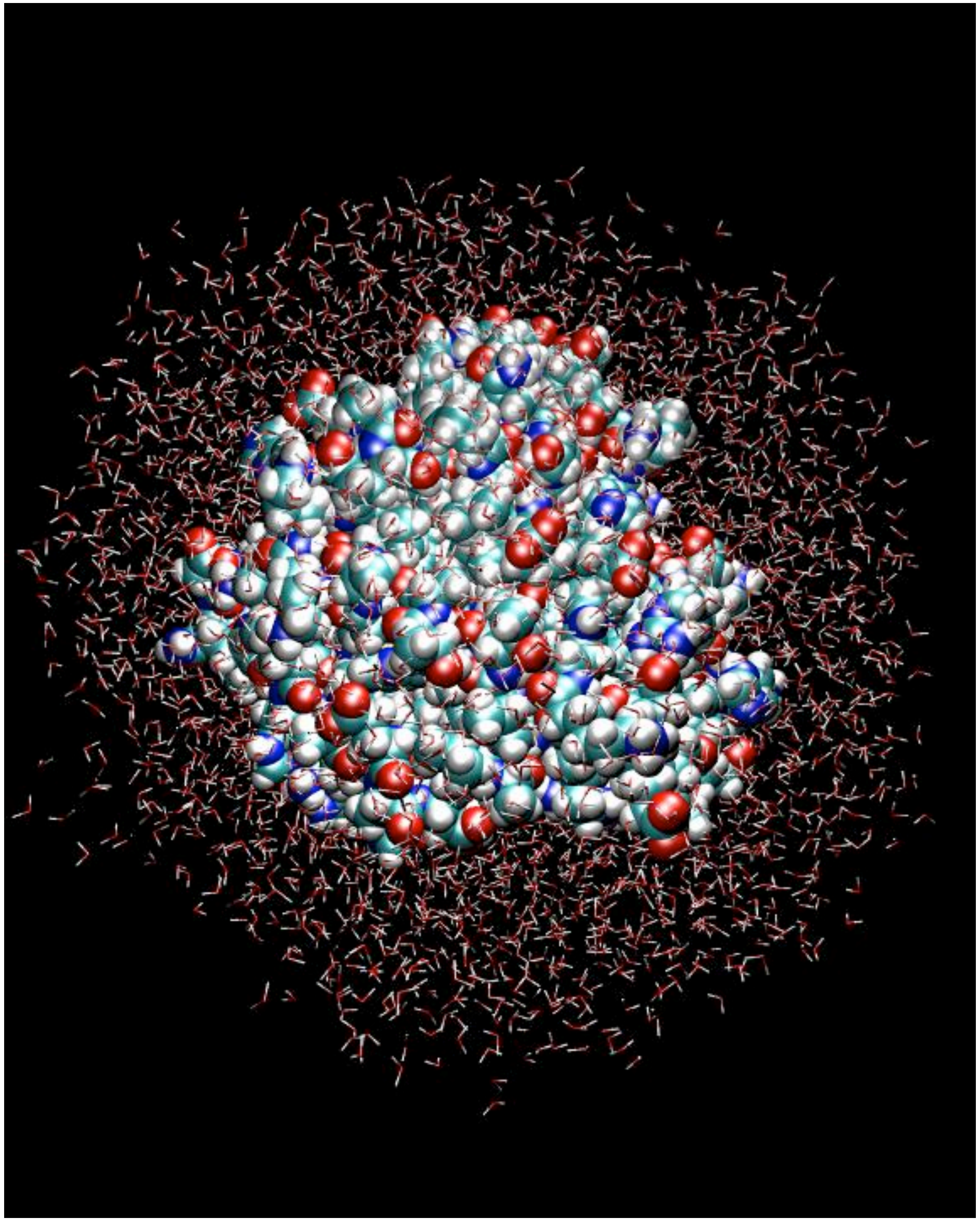
Reaction-field cntd

$$F_{\mathrm{rf}}(r_c) = \frac{q_i q_j}{4\pi\epsilon_0}\frac{1}{r_c^2}\left[1 - \frac{2\epsilon_{\mathrm{rf}} - 2}{2\epsilon_{\mathrm{rf}} + 1}\right]$$

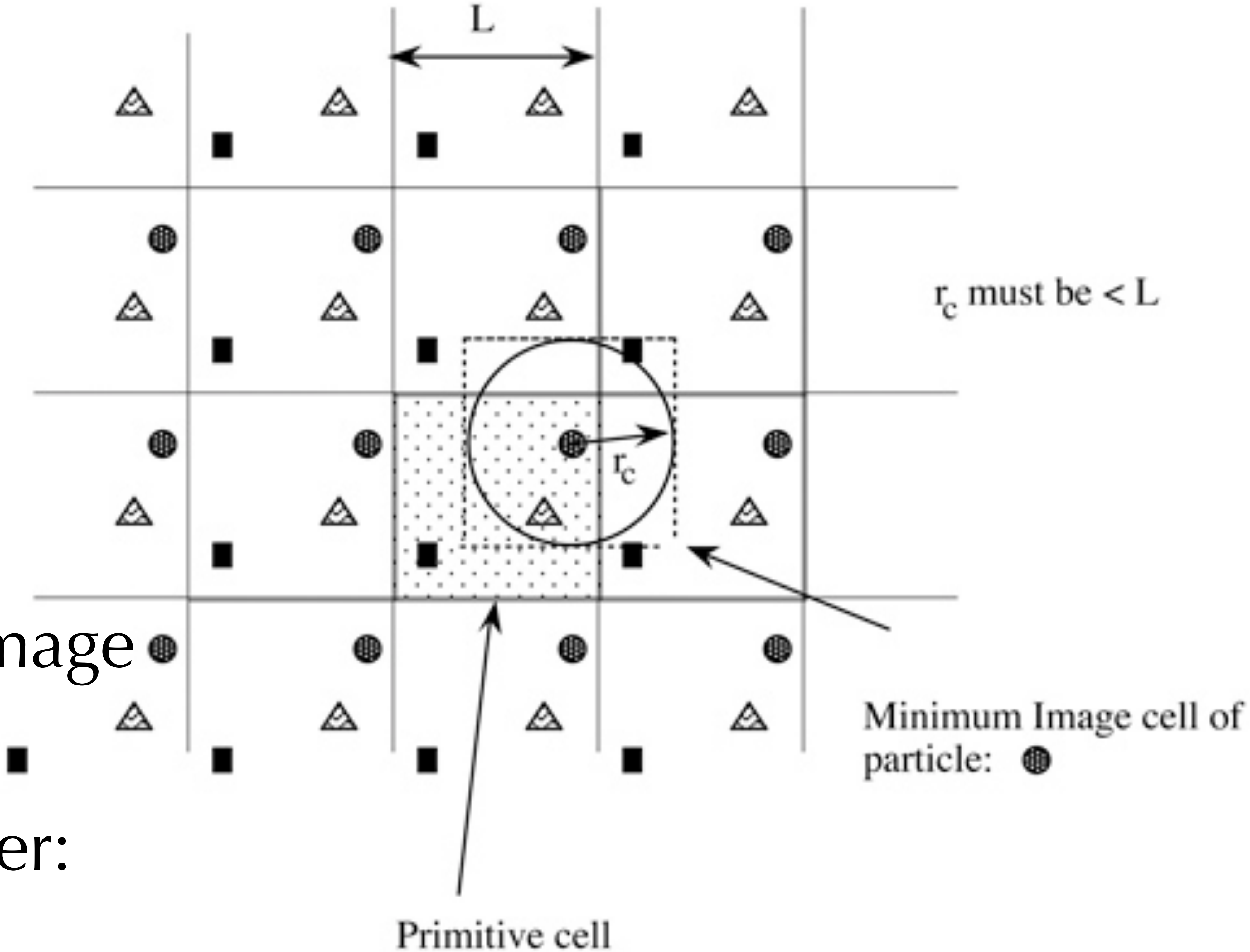
- Issue: $F(r_c)!=0$
- Solution: use $\epsilon_{\mathrm{rf}} = \infty$
  - conducting or "tin-foil" boundary condition
- But shouldn't you match the dielectric of the solvent?
  - Integration errors often worse than deviation in dielectric
  - Also, mismatch goes as: $1/\epsilon_r$

Boundary conditions

- One option:
  - end the system
- Spherical boundary
- But what happens at the boundary?
  - apolar liquids OK
  - water problematic
- What effects on the pressure?

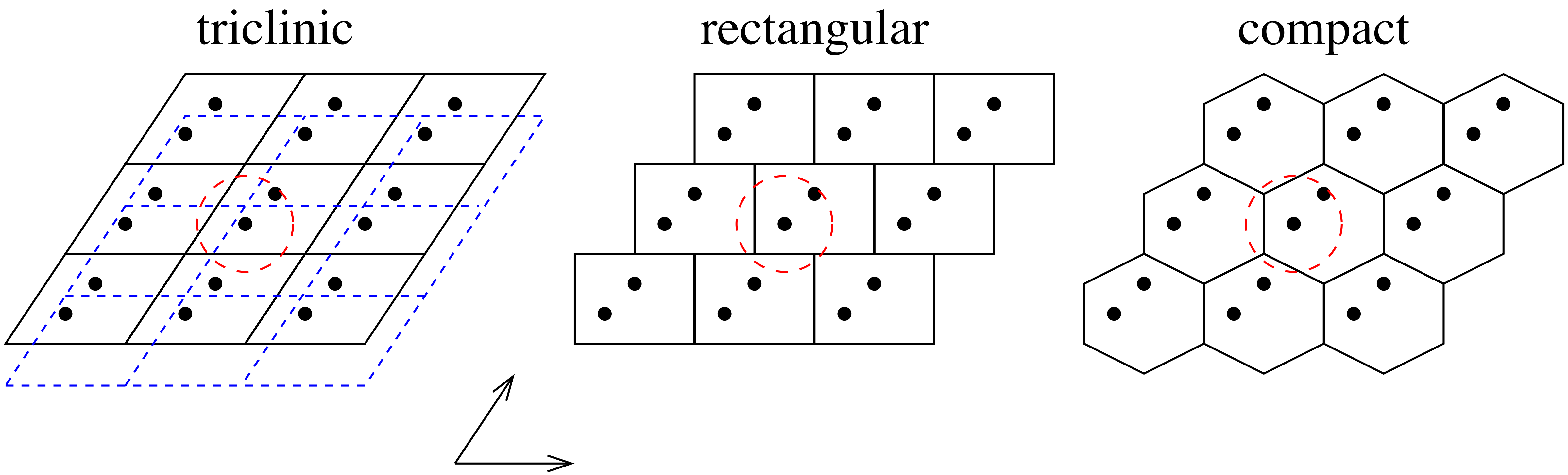


Periodic boundary conditions



- Minimum image convention
- Much simpler: use $r_c<L/2$

Periodic unit or primitive cell



- Different shapes can represent the same periodic boundary conditions
- What matters is not the shape but the periodic shift vectors
- Different shapes useful for different purposes

3D triclinic unit-cells

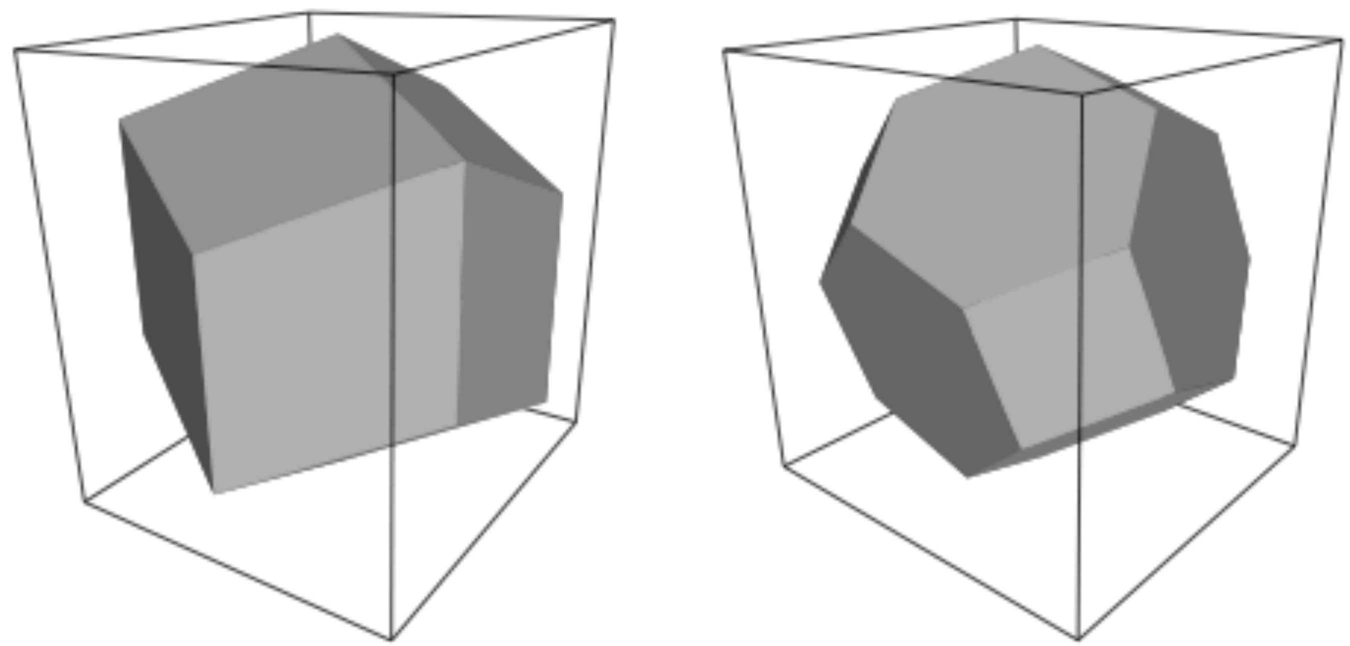
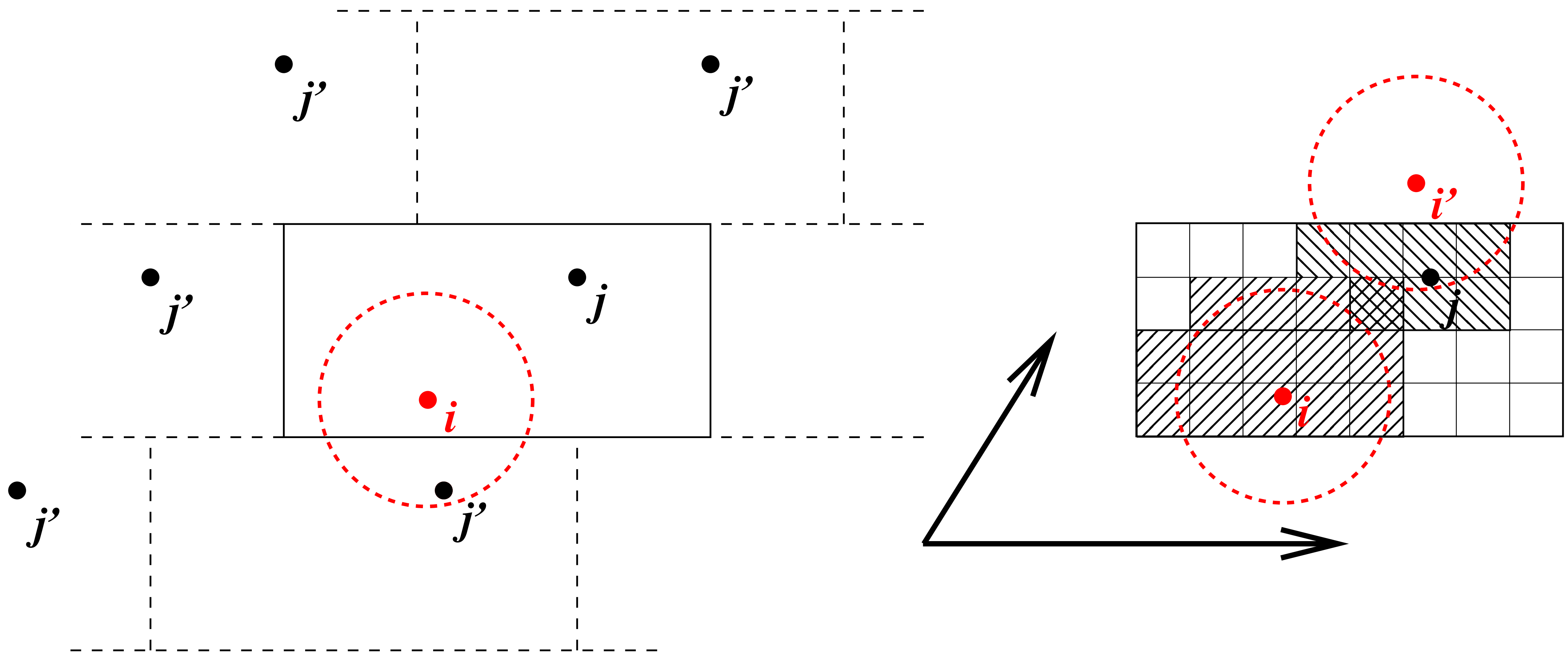


Figure 3.2: A rhombic dodecahedron and truncated octahedron (arbitrary orientations).

box type	image distance	box volume	box vectors			box vector angles		
			a	**b**	**c**	$\angle$**bc**	$\angle$**ac**	$\angle$**ab**
cubic	d	d^3	d 0 0	0 d 0	0 0 d	90°	90°	90°
rhombic dodecahedron (xy-square)	d	$\frac{1}{2}\sqrt{2}\,d^3$ $0.707\,d^3$	d 0 0	0 d 0	$\frac{1}{2}\,d$ $\frac{1}{2}\,d$ $\frac{1}{2}\sqrt{2}\,d$	60°	60°	90°
rhombic dodecahedron (xy-hexagon)	d	$\frac{1}{2}\sqrt{2}\,d^3$ $0.707\,d^3$	d 0 0	$\frac{1}{2}\,d$ $\frac{1}{2}\sqrt{3}\,d$ 0	$\frac{1}{2}\,d$ $\frac{1}{6}\sqrt{3}\,d$ $\frac{1}{3}\sqrt{6}\,d$	60°	60°	60°
truncated octahedron	d	$\frac{4}{9}\sqrt{3}\,d^3$ $0.770\,d^3$	d 0 0	$\frac{1}{3}\,d$ $\frac{2}{3}\sqrt{2}\,d$ 0	$-\frac{1}{3}\,d$ $\frac{1}{3}\sqrt{2}\,d$ $\frac{1}{3}\sqrt{6}\,d$	71.53°	109.47°	71.53°

Calculating periodic interactions



- Calculating all interactions of one i with many j:
  - you can find which j-image you need
  - easier: move i to different periodic shifts
- More on this later ...

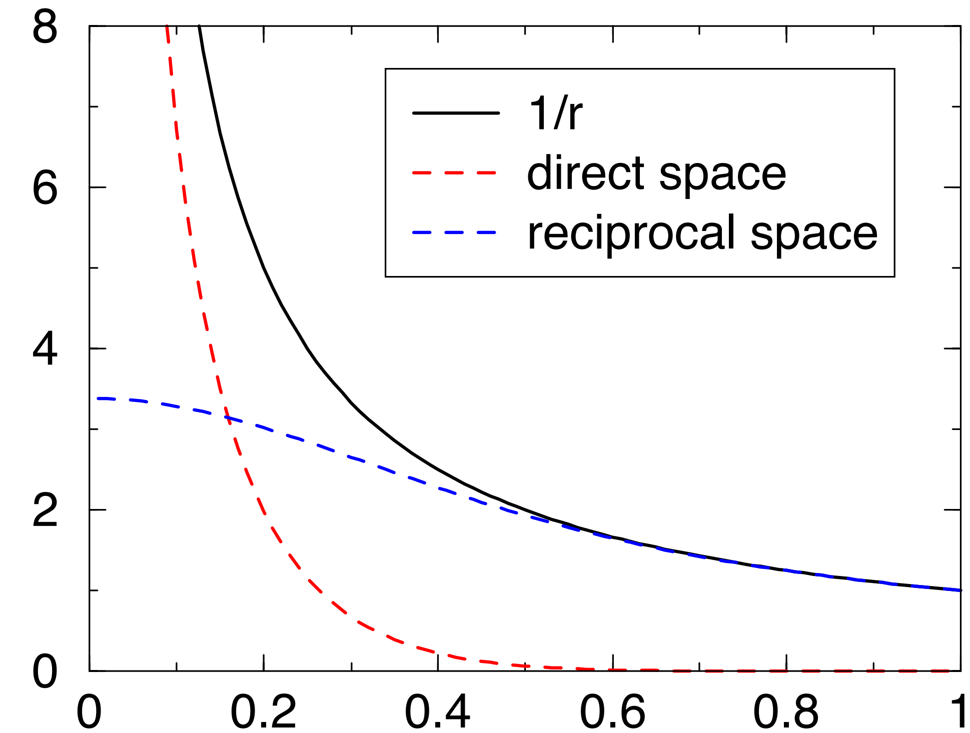
Electrostatics in periodic systems

$$V = \frac{f}{2} \sum_{n_x} \sum_{n_y} \sum_{n_z *} \sum_{i}^{N} \sum_{j}^{N} \frac{q_i q_j}{\mathbf{r}_{ij,\mathbf{n}}}$$

- We can write down the sum over all charge pairs in all periodic images
- But as Coulomb goes as 1/r, this sum is only conditionally convergent
- Direct sums have bad convergence

Ewald summation

$$V = V_{dir} + V_{rec} + V_0$$



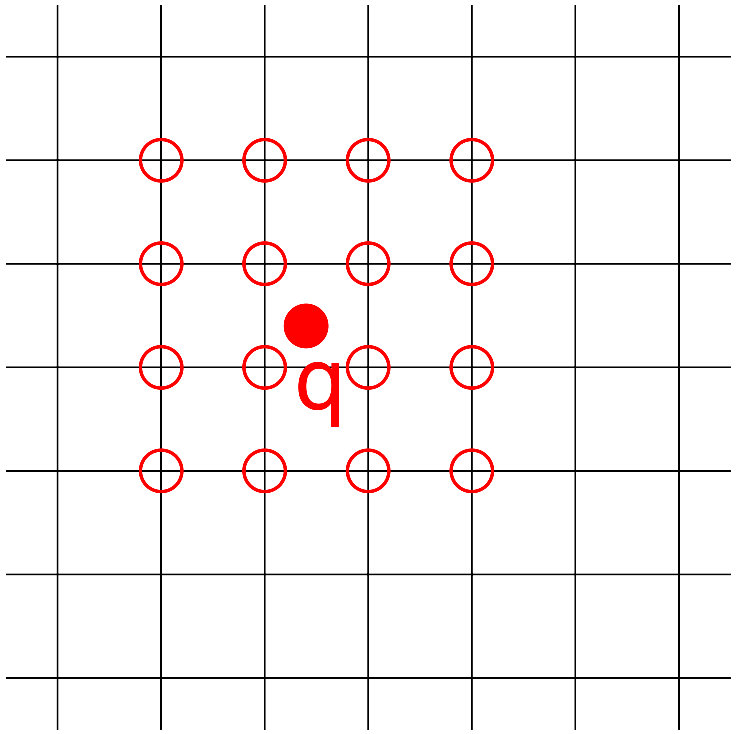
$$V_{dir} = \frac{f}{2}\sum_{i,j}^{N}\sum_{n_x}\sum_{n_y}\sum_{n_z*} q_i q_j \frac{\operatorname{erfc}(\beta r_{ij,\mathbf{n}})}{r_{ij,\mathbf{n}}}$$

$$V_{rec} = \frac{f}{2\pi V}\sum_{i,j}^{N} q_i q_j \sum_{m_x}\sum_{m_y}\sum_{m_z*} \frac{\exp\left(-(\pi\mathbf{m}/\beta)^2 + 2\pi i\mathbf{m}\cdot(\mathbf{r}_i - \mathbf{r}_j)\right)}{\mathbf{m}^2}$$

$$V_0 = -\frac{f\beta}{\sqrt{\pi}}\sum_{i}^{N} q_i^2$$

Particle mesh methods

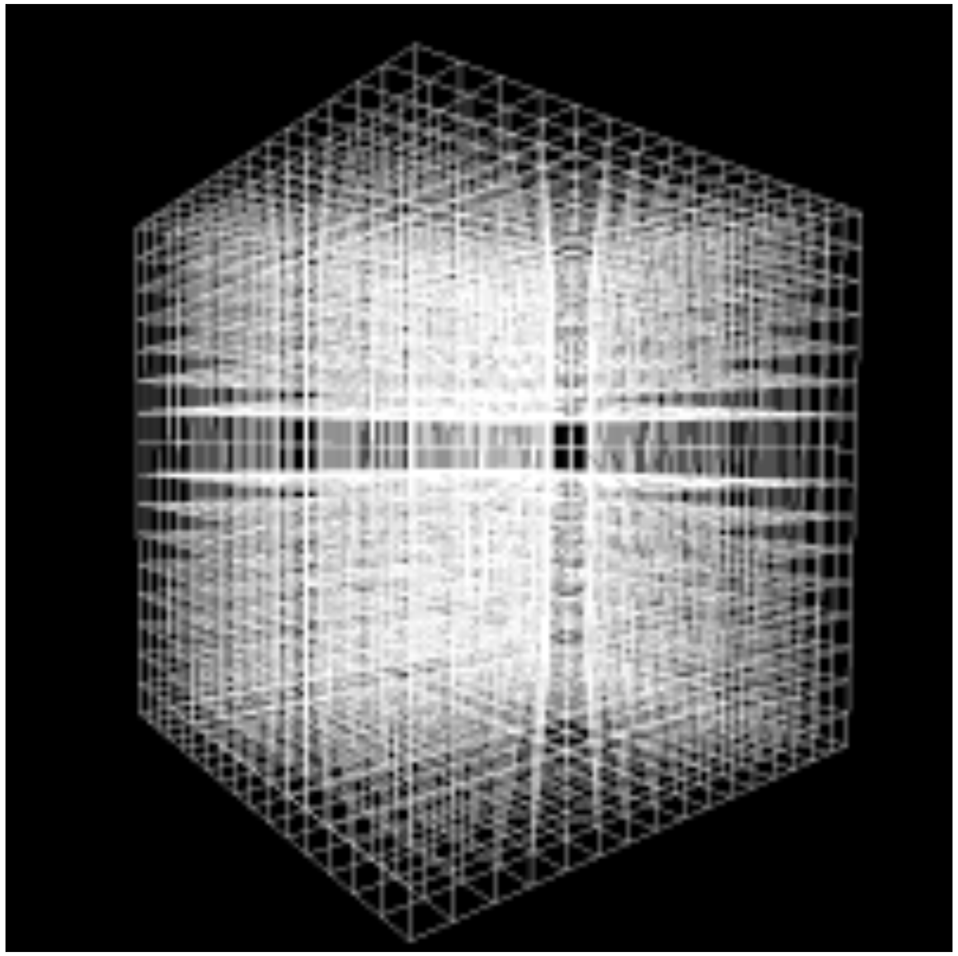
- Ewald summation is slow: $O(N^2)$
- Solution: do the reciprocal part on a mesh
  - Particle-Particle Particle-Mesh (PPPM/P3M)
  - Particle-Mesh Ewald (PME)
  - Most popular SPME (smooth) by Darden et al.
- spread charges on grid
- 3D FFT
- solve in fourier space
- 3D FFT
- gather forces from grid



PME

- PME Parameters
  - cut-off
  - smoothing parameter B
  - spreading order S
  - grid size: M_x,M_y,M_z
- Computational cost:
  - direct: $O(r_c^3)$
  - spread: O(#charges*S^3)
  - 3D FFT: O(N log(N))
- Accuracy determined by:
  - real space error beyond cut-off: $\mathrm{erfc}(B\ r_c)/r_c$
  - spreading accuracy
    - spreading order
    - smoothing parameter
    - grid spacing
- Complex, for SPME no simple analytical formula
- But, important for performance and accuracy of simulations!

PME settings in practice



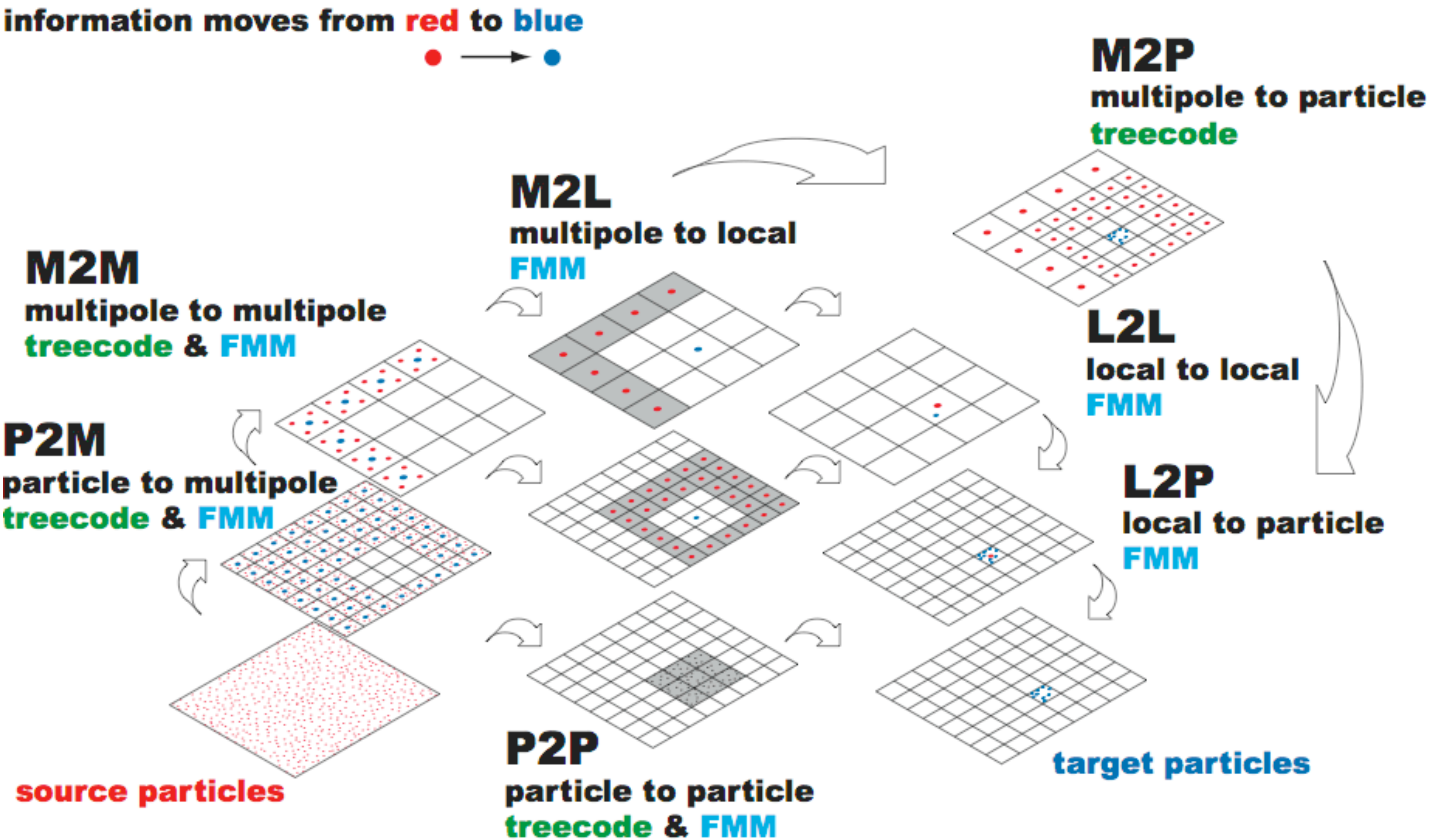
- Complex, but also costly!
- Use what others use with your system and/or software
- Typical settings:
  - order 4: spread 4^3=64 points
  - cut-off 0.9 nm
  - grid spacing 0.12 nm
    - #grid point similar to #particles
- Soon: tools to set parameters based on force accuracy
  - But what does a force accuracy of 0.1 kJ/mol/nm^2 mean?

Long-range electrostatics methods

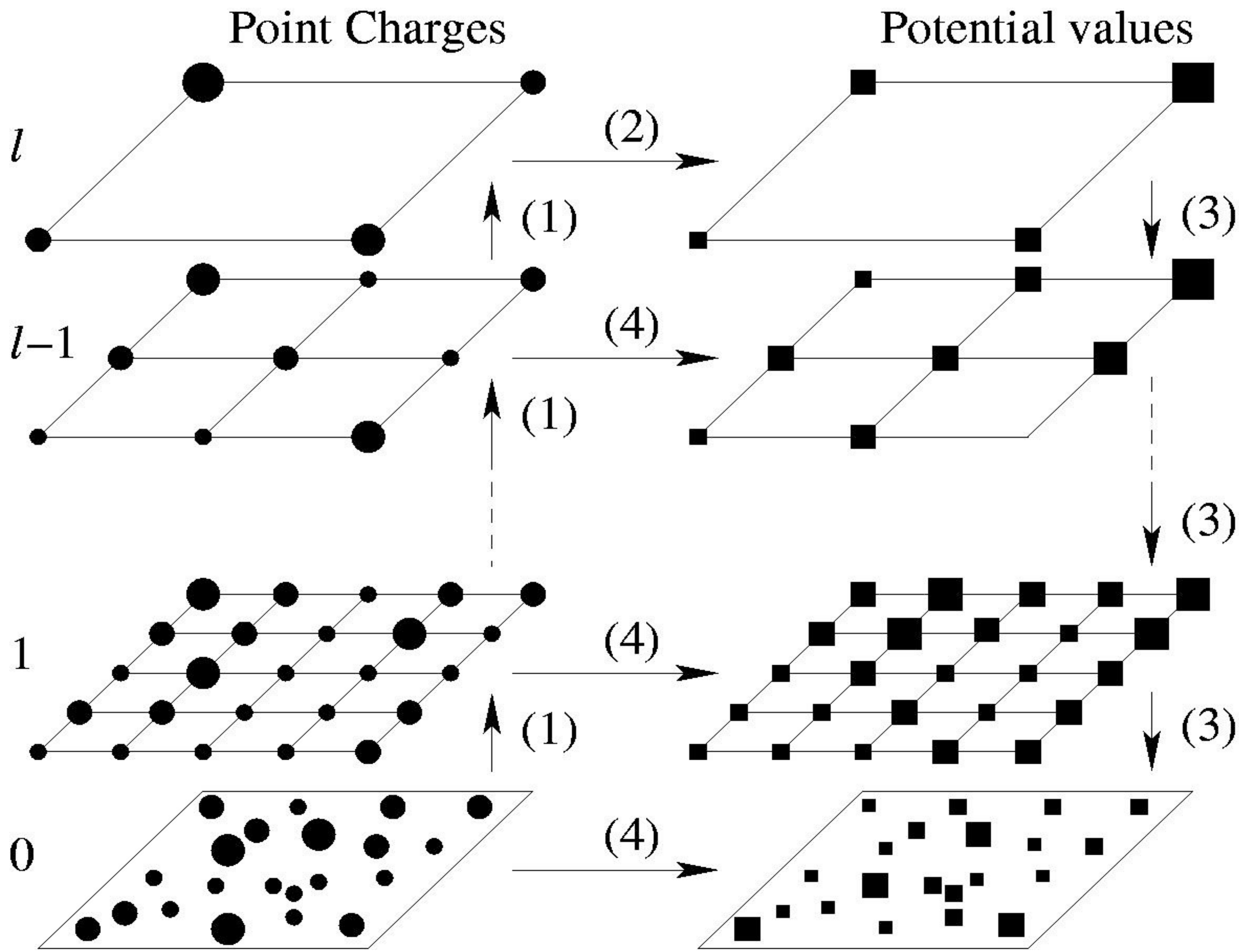
- All methods calculate the same potential and forces

	computational cost	pre-factor	communication cost
Ewald summation	$O(N^{3/2})$	small (FFT)	high
PPPM / Particle Mesh Ewald	O(N log N)	small (FFT)	high
Fast Multipole Method	O(N)	larger	low
Multigrid Electrostatics	O(N)	larger	low

Fast multipole method

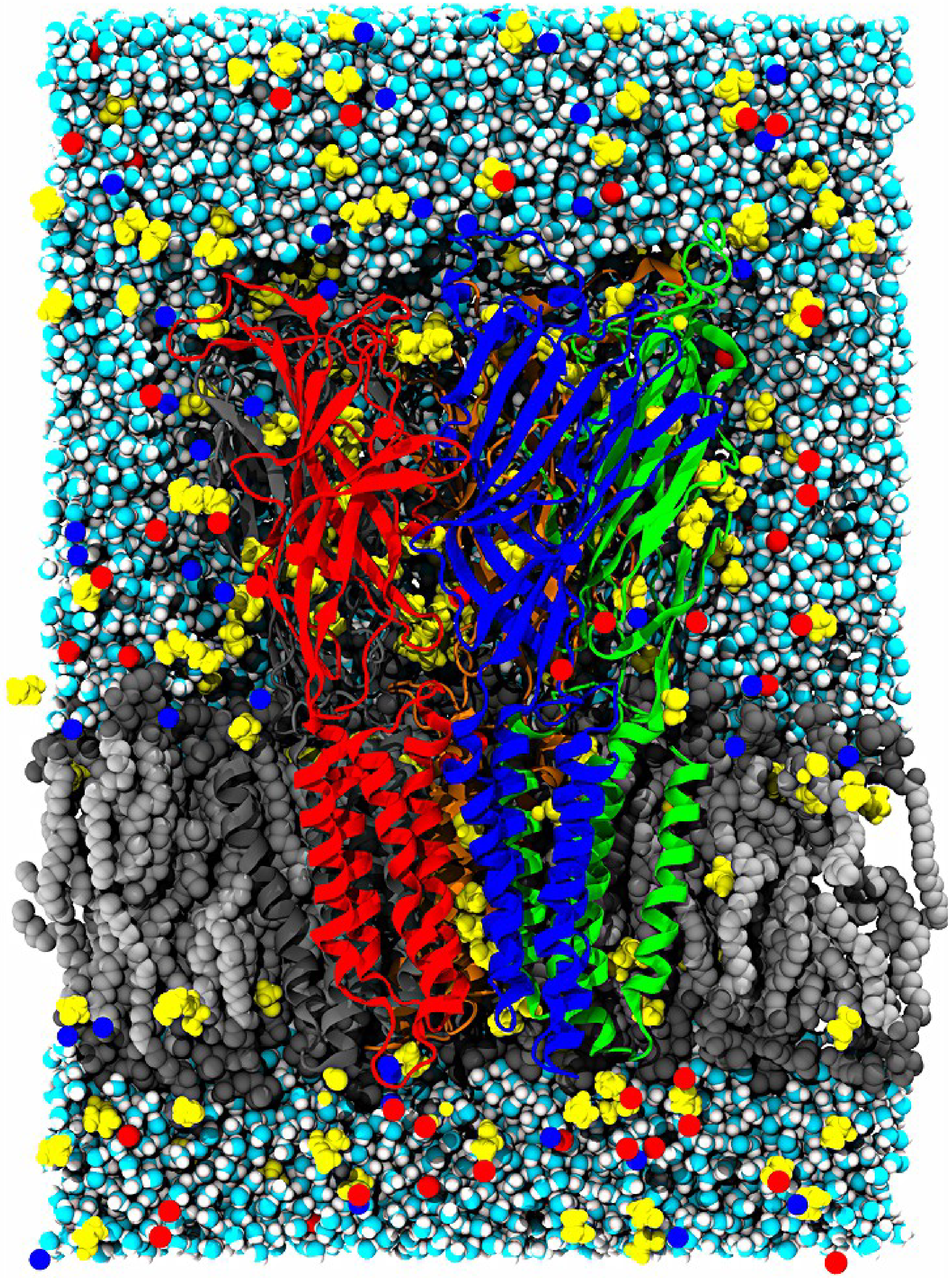


Multigrid electrostatics



PME and charged systems

- With net system charge:
  - implicit assumption:
    - uniform background charge
    - no effect on F, effect on U
- Problems with non-uniform dielectric
- Always safer to neutralize with ions!

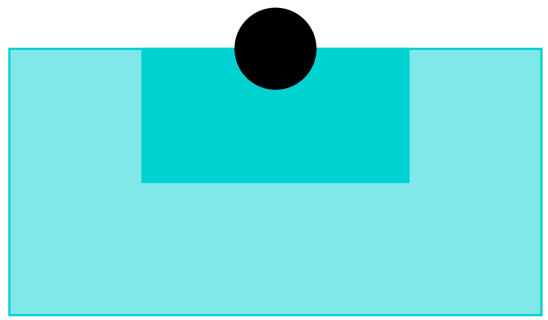


Adding (counter-)ions

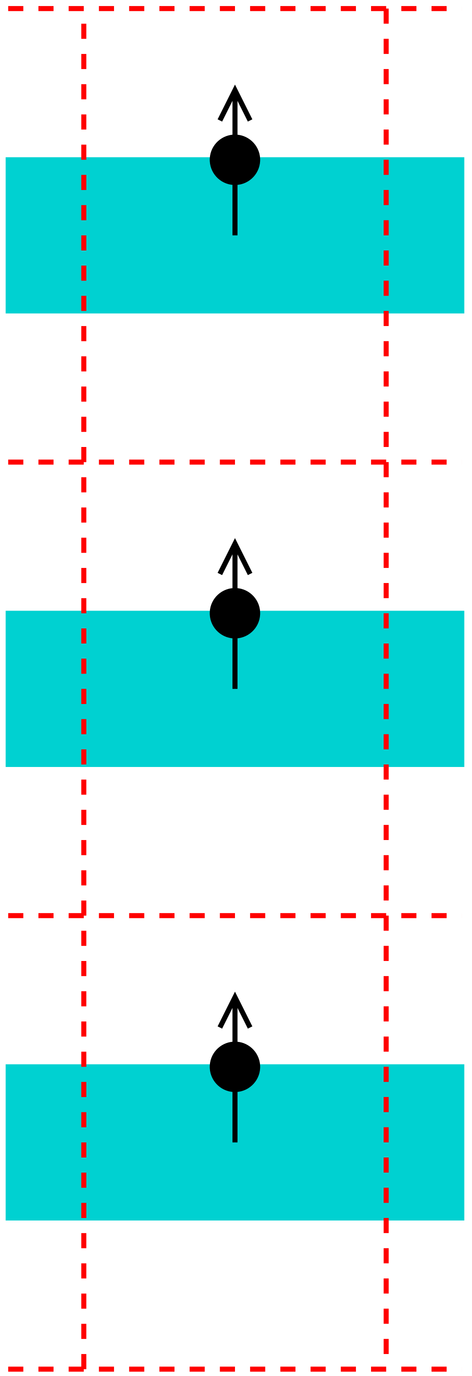
- Simply add ions Na^+ or Cl^- to neutralize
  - Adding just a few ions can lead to sampling problems
- Better: add Na^+ and Cl^- at physiological concentration, if possible

- What is physiological concentration locally?
- Some systems might need different ions
  - DNA/RNA: Na^+ or Mg^{2+}?

Electrostatics at surfaces



- One surface is impossible, at least 2
- To use PME with PBC:
  - Add 2/3 of vacuum
  - Use dipole correction Yeh&Berkowitz (JCP 111,3155)



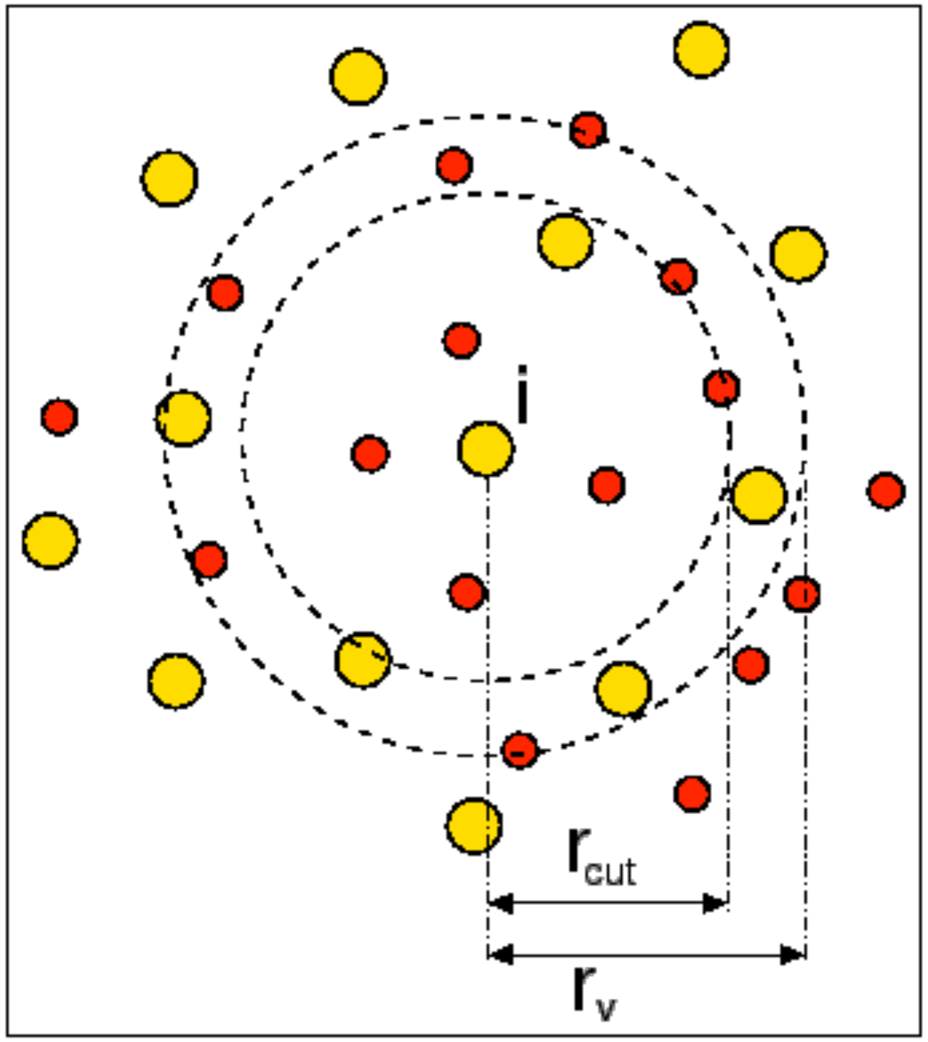
$$U_z = \frac{2\pi}{V} M_z^2 \ , \quad F_{z,i} = -\frac{4\pi q_i}{V} M_z$$

Algorithms to do efficient MD

Recap.

- What do we need to calculate?
  - bonded interactions: cheap
  - non-bonded interactions: expensive
  - maybe PME: expensive
  - integration: cheap

Calculating non-bonded interactions



- Using a Verlet list:
  - Make a "Verlet" pair list using radius: $r_{list} = r_c + r_{buf}$
  - Calculate interactions for n steps within cut-off r_c
  - When to update the list?
    - Option: when a particle moved more than $r_{buf}/2$
      - Becomes expensive for large systems

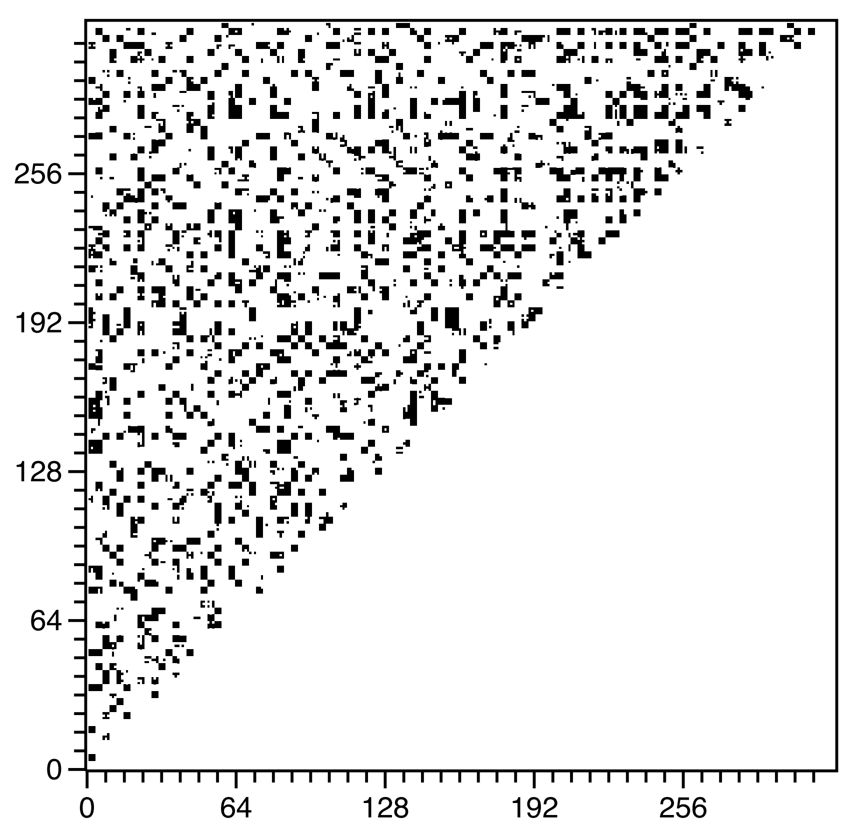
Charge groups

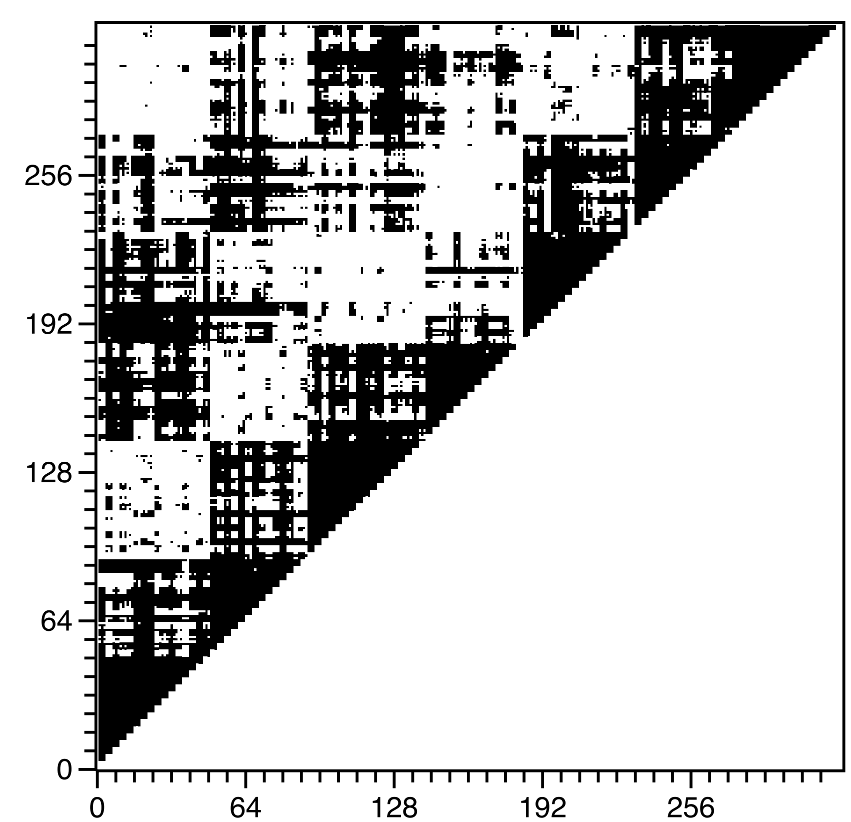
- In the early years of MD cut-offs were used: bad with cut-off electrostatics!
  - partial remedy: use neutral "charge groups"
    - e.g. group 3 atoms in water: only dipole

- Remnants of this still in the Gromacs package
  - often used without a Verlet list
    - bad for energy conservation
      - with thermostat fine for most purposes

Order of particles & interactions

- Without particle ordering interactions are randomly distributed in memory: bad performance
- Sorting the particles on a grid groups interactions
  - good for performance & parallelization

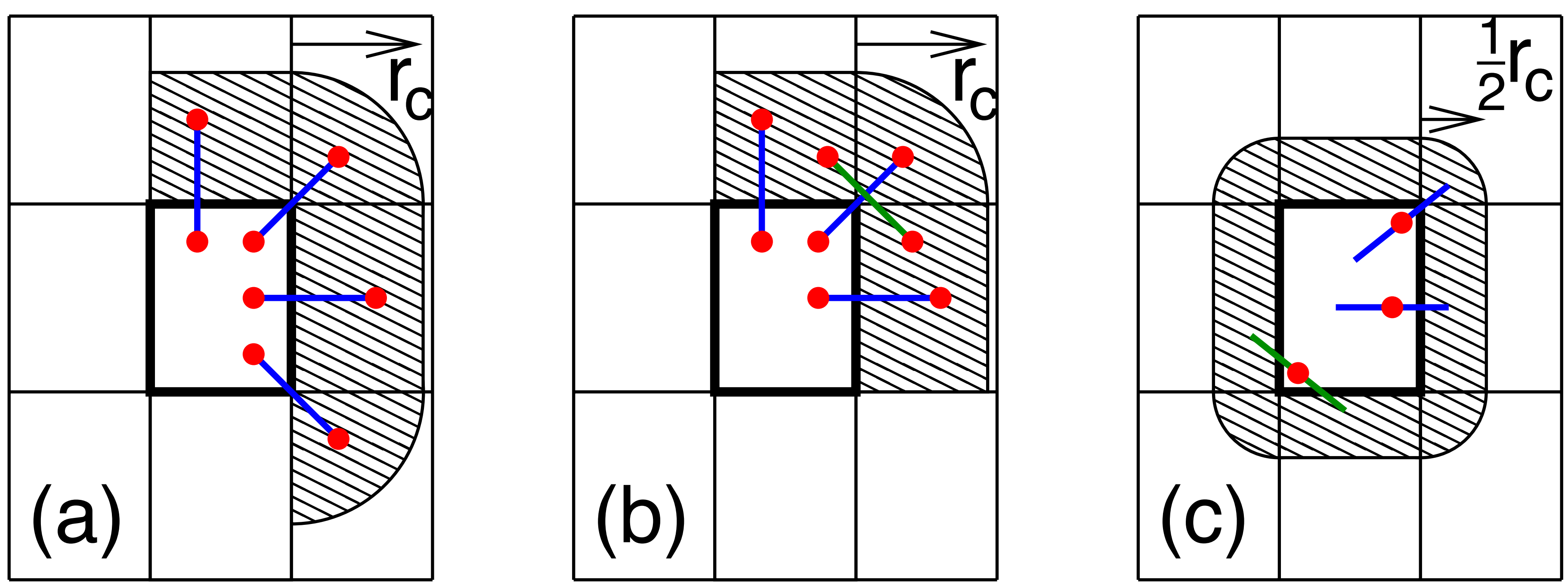




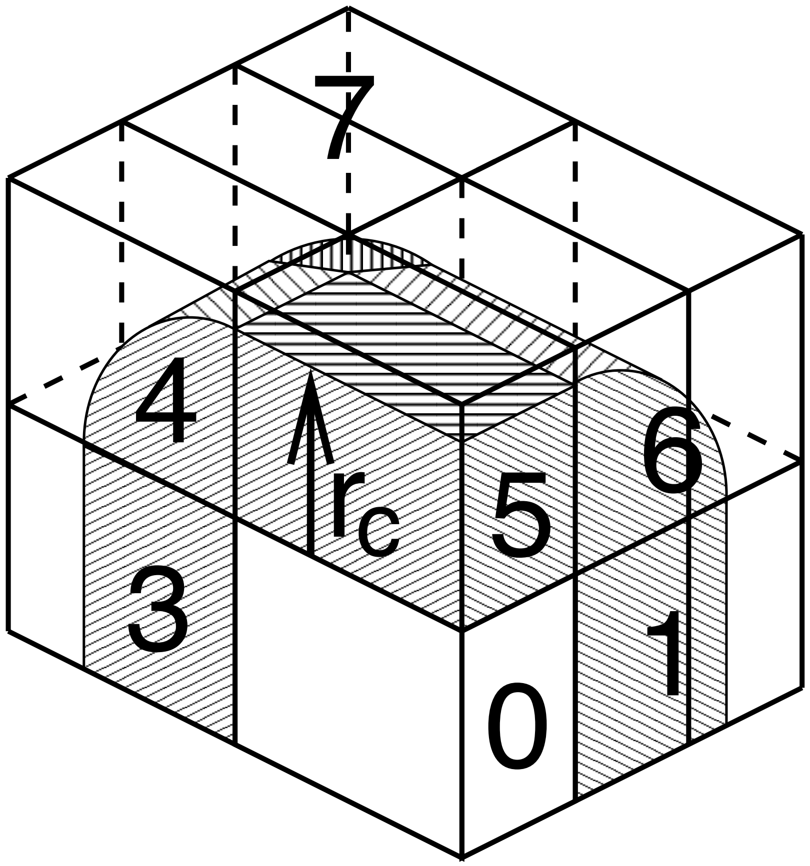
Parallel Molecular Dynamics

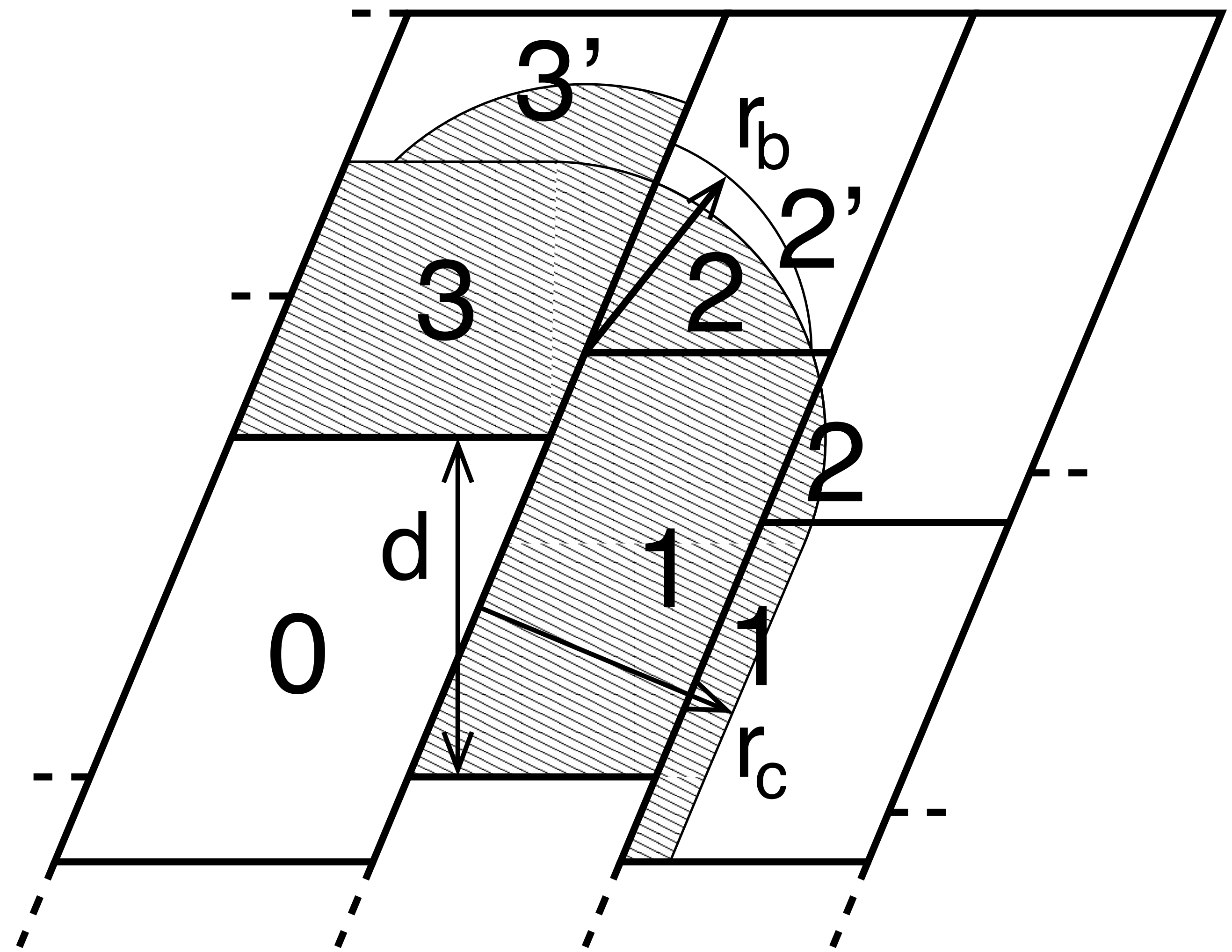
- Particle or force decomposition
  - bad memory access characteristics
- Spatial or domain decomposition
  - good memory access & communication

half shell eighth shell midpoint



Domains & load balancing





- For inhomogeneous systems load balancing is required
- Gromacs has full 3D dynamic load balancing

Reading

Read Frenkel&Smit part V.F on saving CPU time

I will put an exercise on the site this afternoon

Next lecture: November 10, 10:00 at FB55

No lecture on November 15