Improved Long-Range Solvers for Molecular Simulations

Daniel Tameling
tameling@aices.rwth-aachen.de

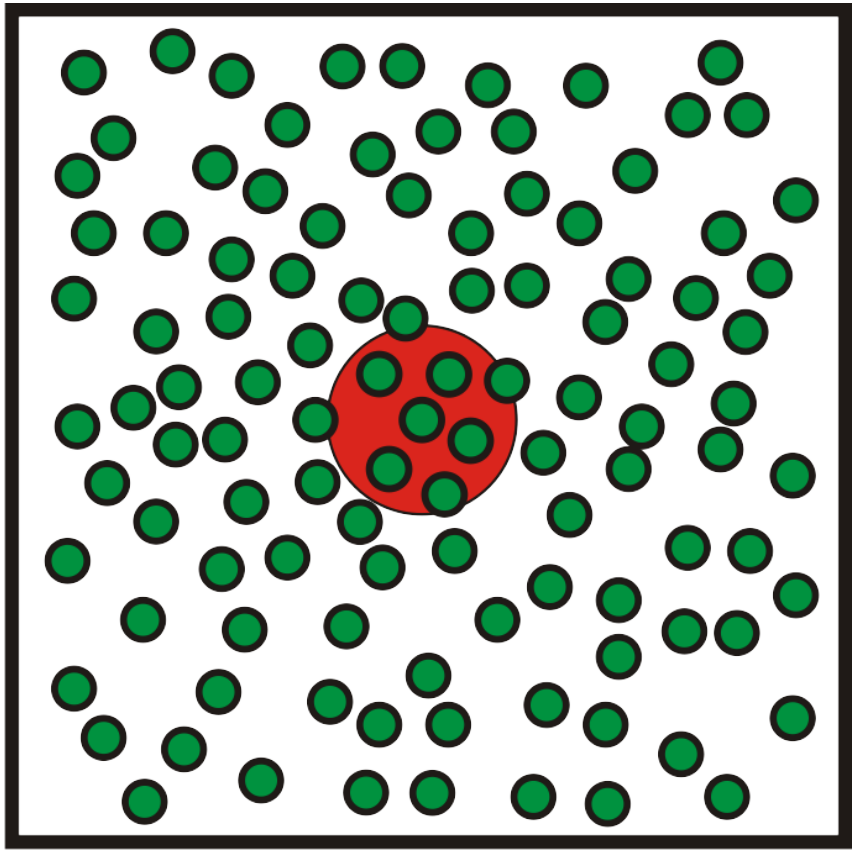
Prof. Ahmed Ismail, Ph.D., Prof. Paolo Bientinesi, Ph.D.

Annual Report

7 April 2014

Table of Contents

Dispersion Interactions

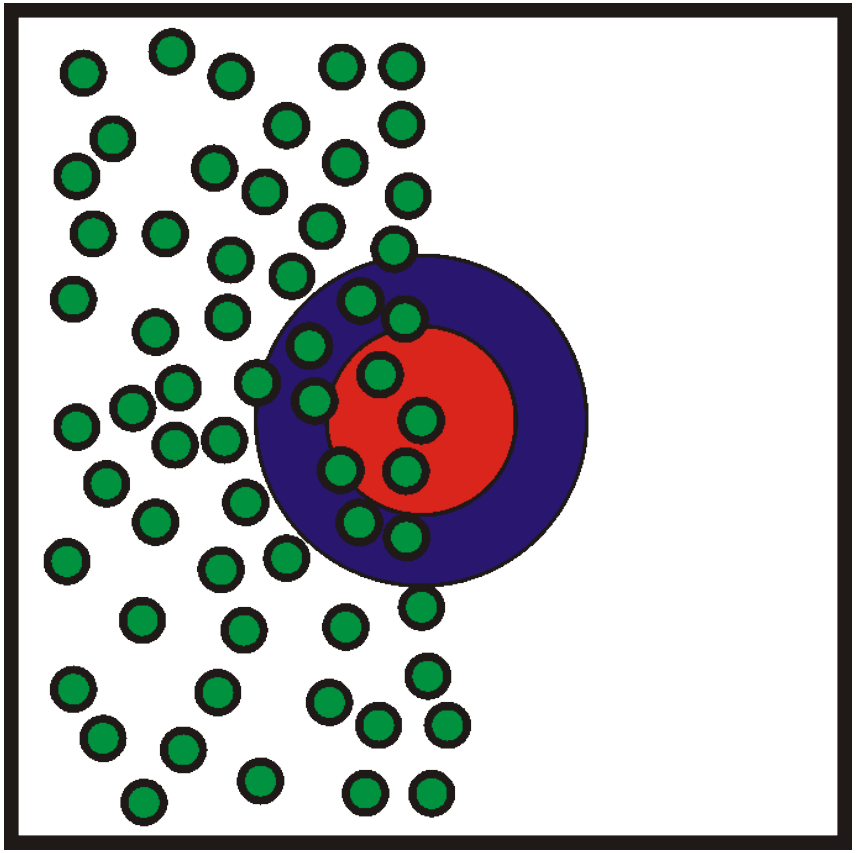


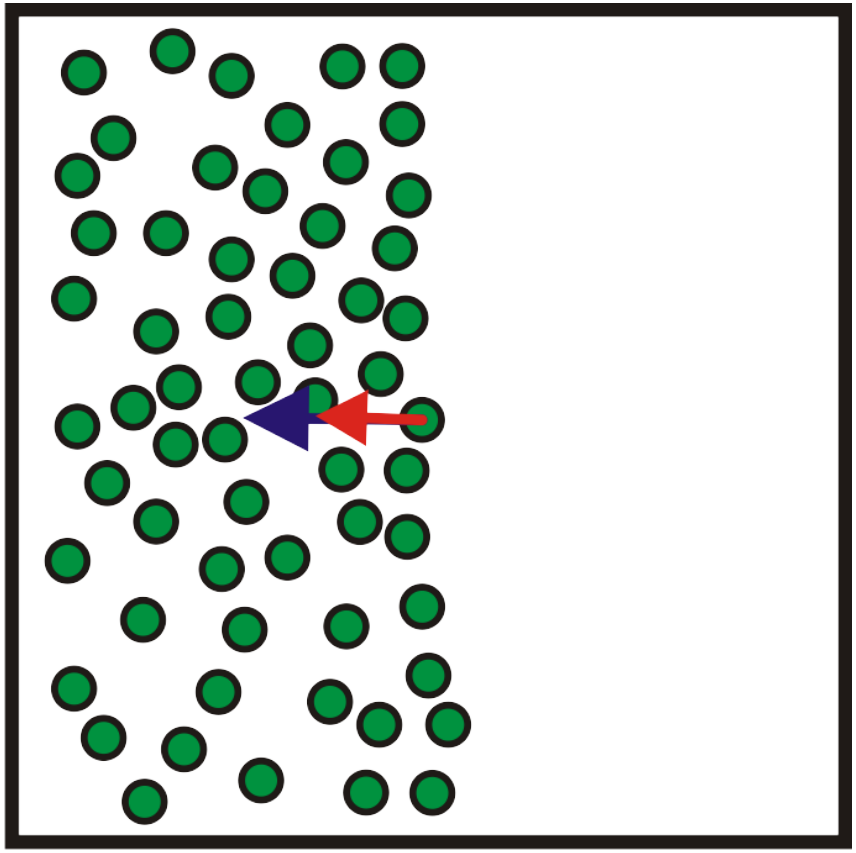
Dispersion potential:

$$V_{disp} = \frac{1}{2} \sum_{i=1}^{N} \sum_{\substack{j=1 \\ i \neq j}}^{N} \frac{C_{ij}}{r_{ij}^6}$$

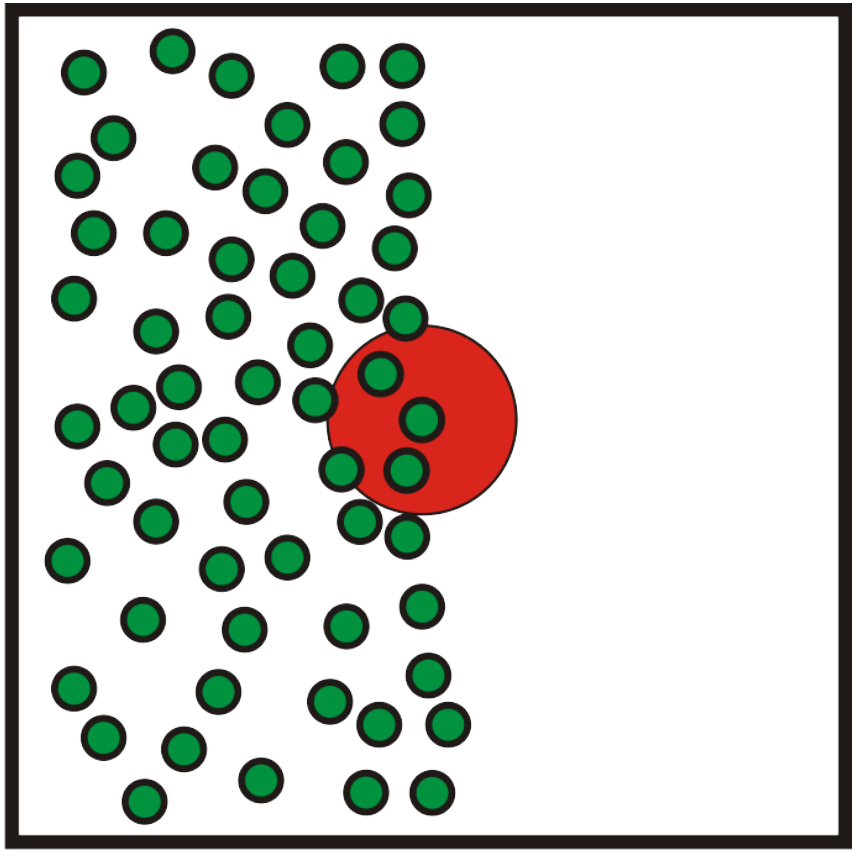
- E.g. in Lennard-Jones and Buckingham potentials
- Only attractive interaction between all pairs of atoms

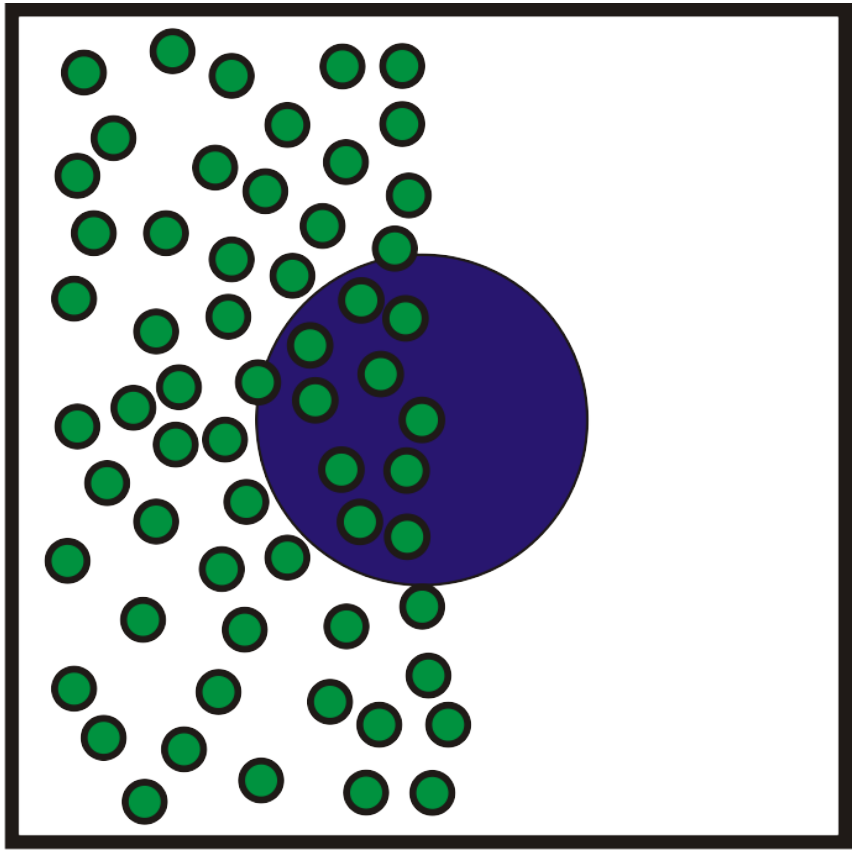
Dispersion Interactions



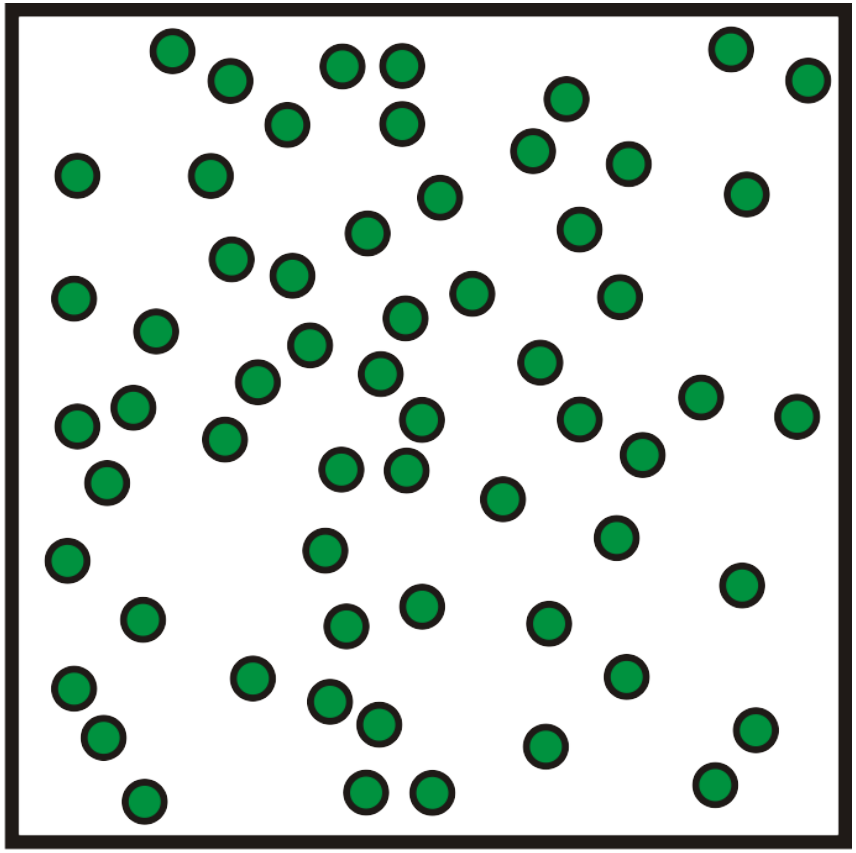


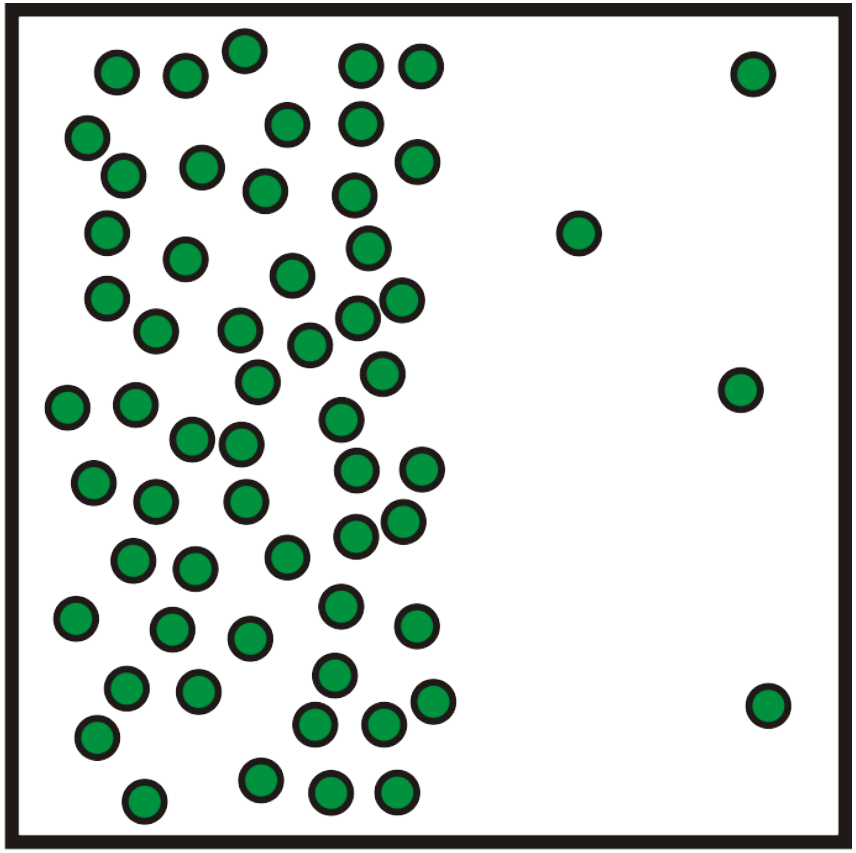
Dispersion Interactions





Dispersion Interactions





Long-Range Methods

Existing long-range methods:

- Direct evaluation
  - $\mathcal{O}\left(N^2\right)$
  - typically too expensive

Long-Range Methods

Existing long-range methods:

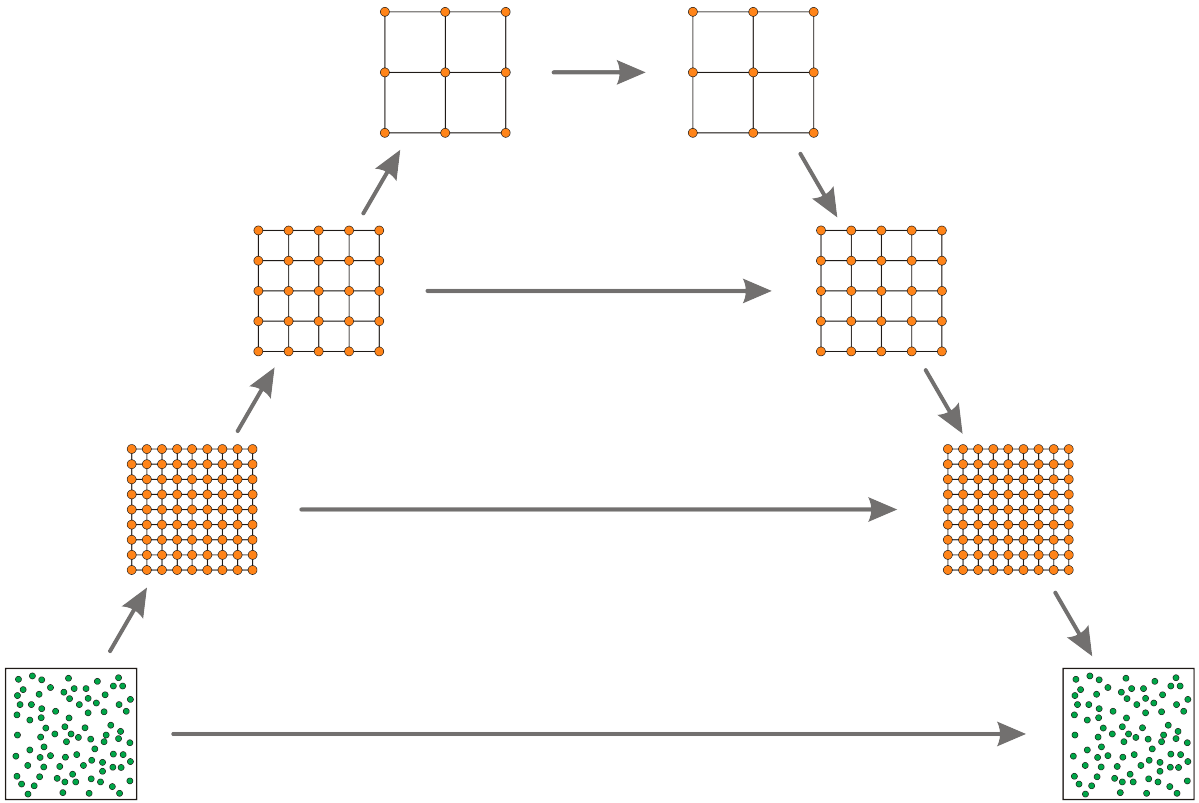
- Direct evaluation
  - $\mathcal{O}\left(N^2\right)$
  - typically too expensive
- Standard methods: Ewald Mesh methods
  - Use fast Fourier transforms
  - $\mathcal{O}\left(N \log N\right)$
  - Don't scale well

Long-Range Methods

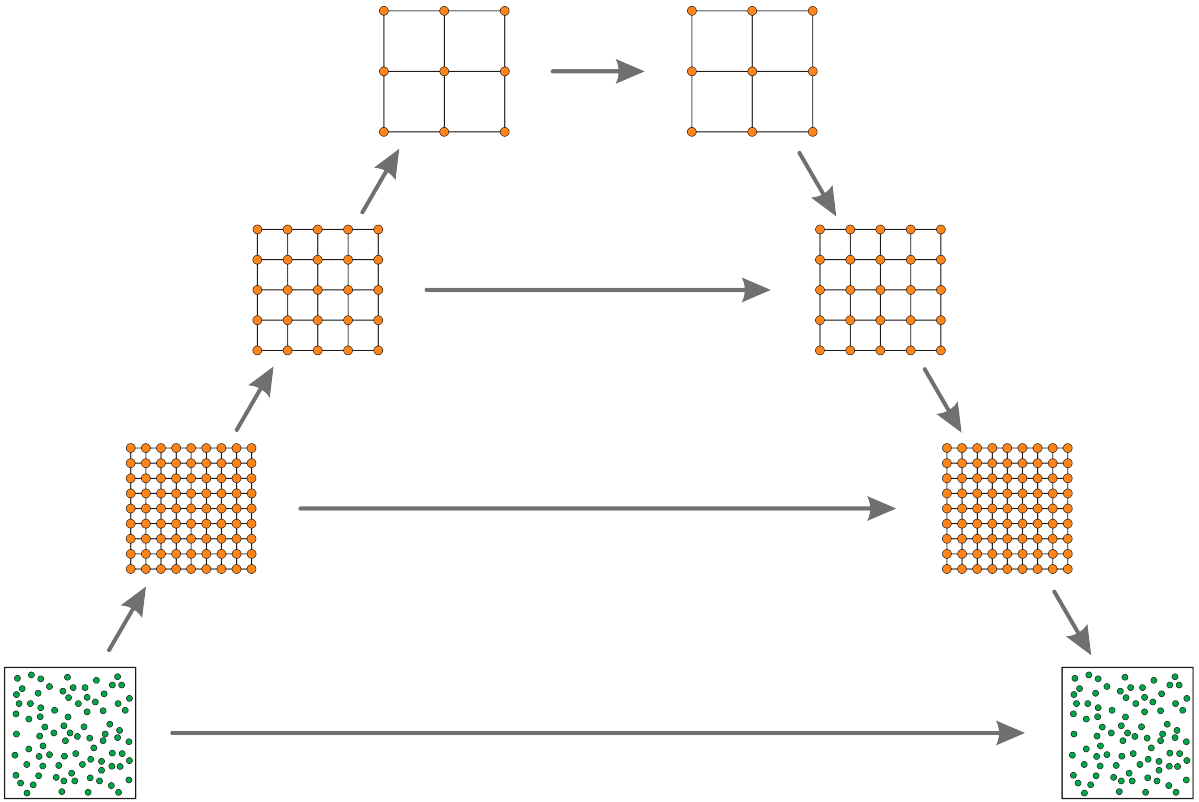
Existing long-range methods:

- Direct evaluation
  - $\mathcal{O}\left(N^2\right)$
  - typically too expensive
- Standard methods: Ewald Mesh methods
  - Use fast Fourier transforms
  - $\mathcal{O}\left(N \log N\right)$
  - Don't scale well
- Multilevel summation method
  - $\mathcal{O}\left(N\right)$

Multilevel Summation



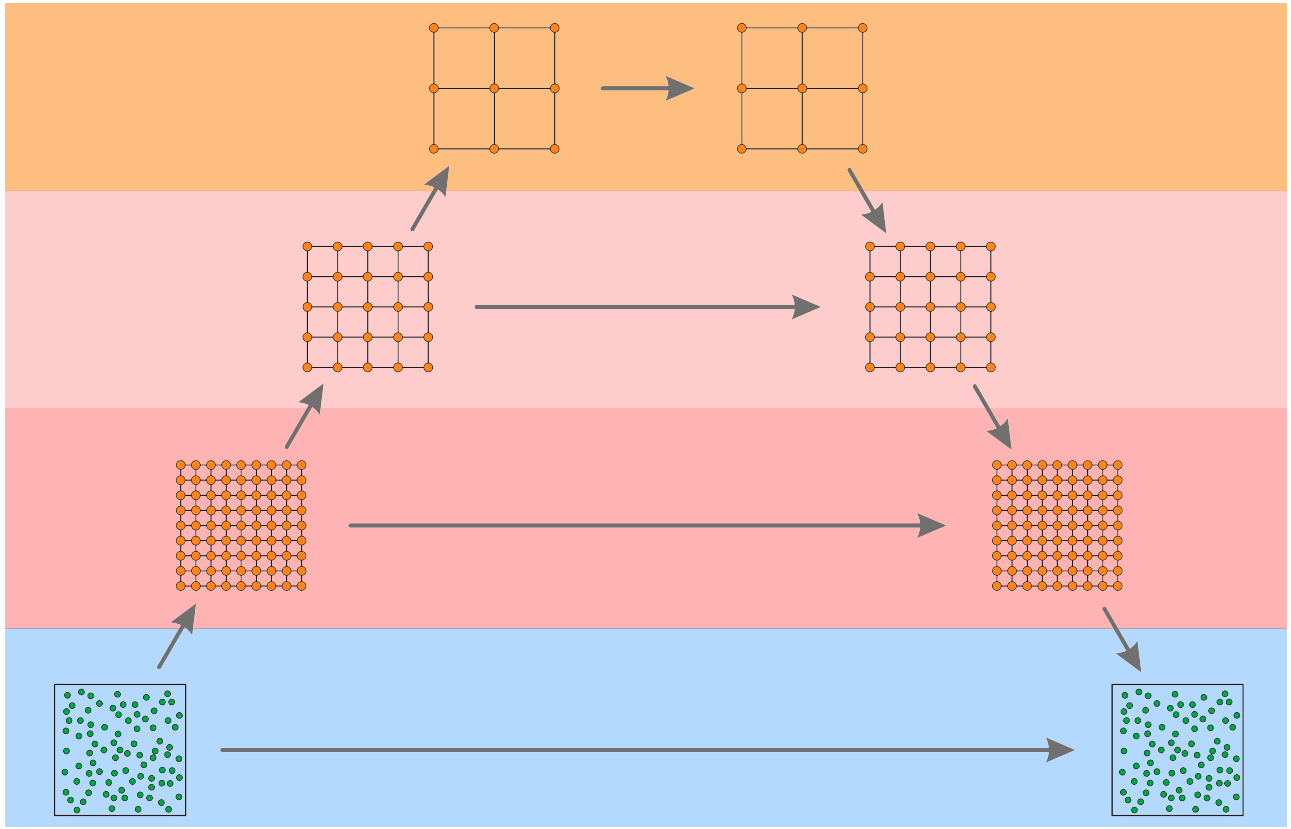
Multilevel Summation



$$\frac{1}{r^6} = \frac{1}{r^6} - \frac{1}{a^6}\gamma\left(\frac{r}{a}\right) + \frac{1}{a^6}\gamma\left(\frac{r}{a}\right) - \frac{1}{(2a)^6}\gamma\left(\frac{r}{2a}\right)$$

$$+ \frac{1}{(2a)^6}\gamma\left(\frac{r}{2a}\right) - \frac{1}{(4a)^6}\gamma\left(\frac{r}{4a}\right) + \cdots + \frac{1}{(2^{l-1}a)^6}\gamma\left(\frac{r}{2^{l-1}a}\right)$$

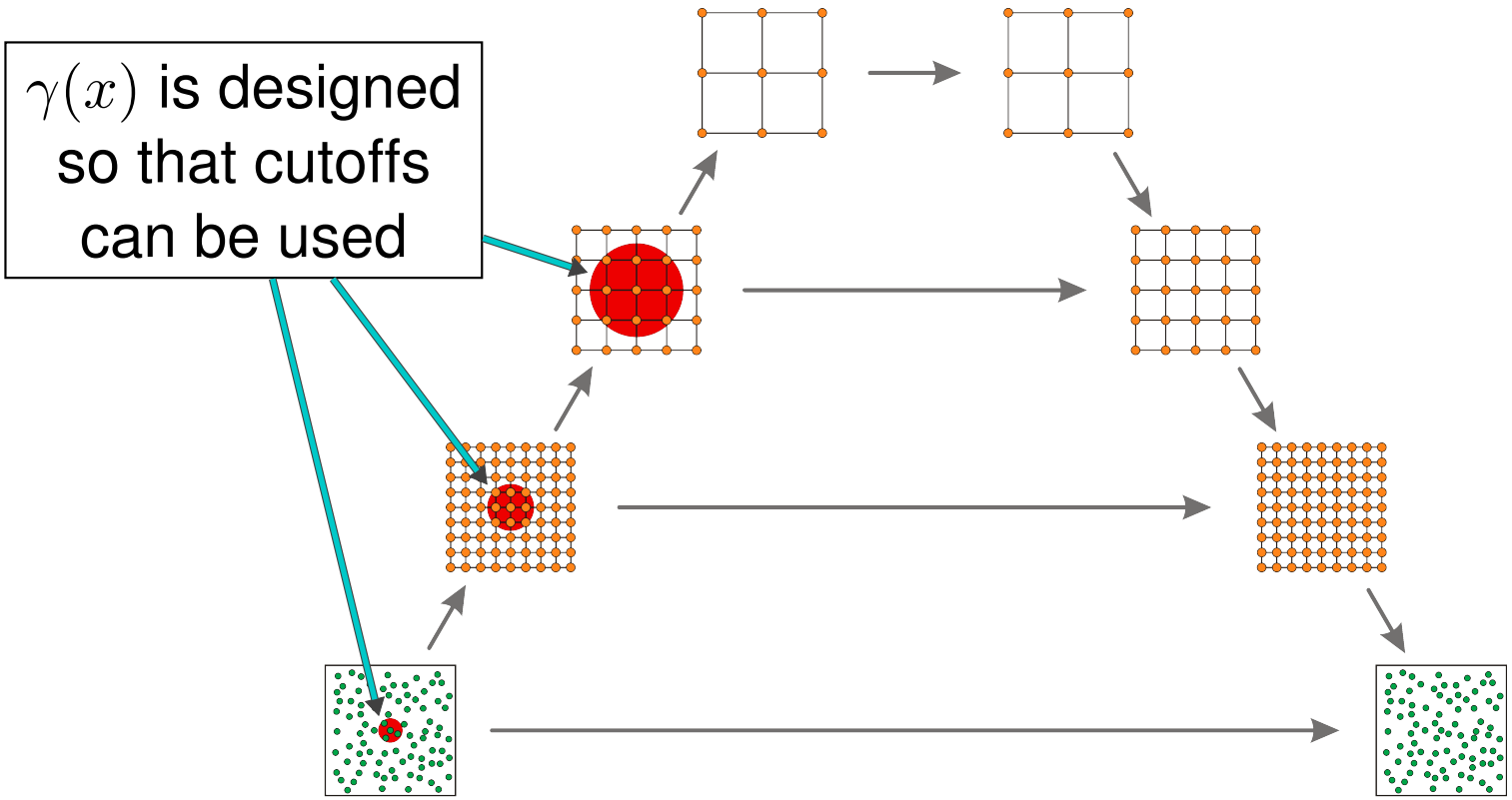
Multilevel Summation



$$\frac{1}{r^6} = \frac{1}{r^6} - \frac{1}{a^6}\gamma\left(\frac{r}{a}\right) + \frac{1}{a^6}\gamma\left(\frac{r}{a}\right) - \frac{1}{(2a)^6}\gamma\left(\frac{r}{2a}\right)$$

$$+ \frac{1}{(2a)^6}\gamma\left(\frac{r}{2a}\right) - \frac{1}{(4a)^6}\gamma\left(\frac{r}{4a}\right) + \cdots + \frac{1}{(2^{l-1}a)^6}\gamma\left(\frac{r}{2^{l-1}a}\right)$$

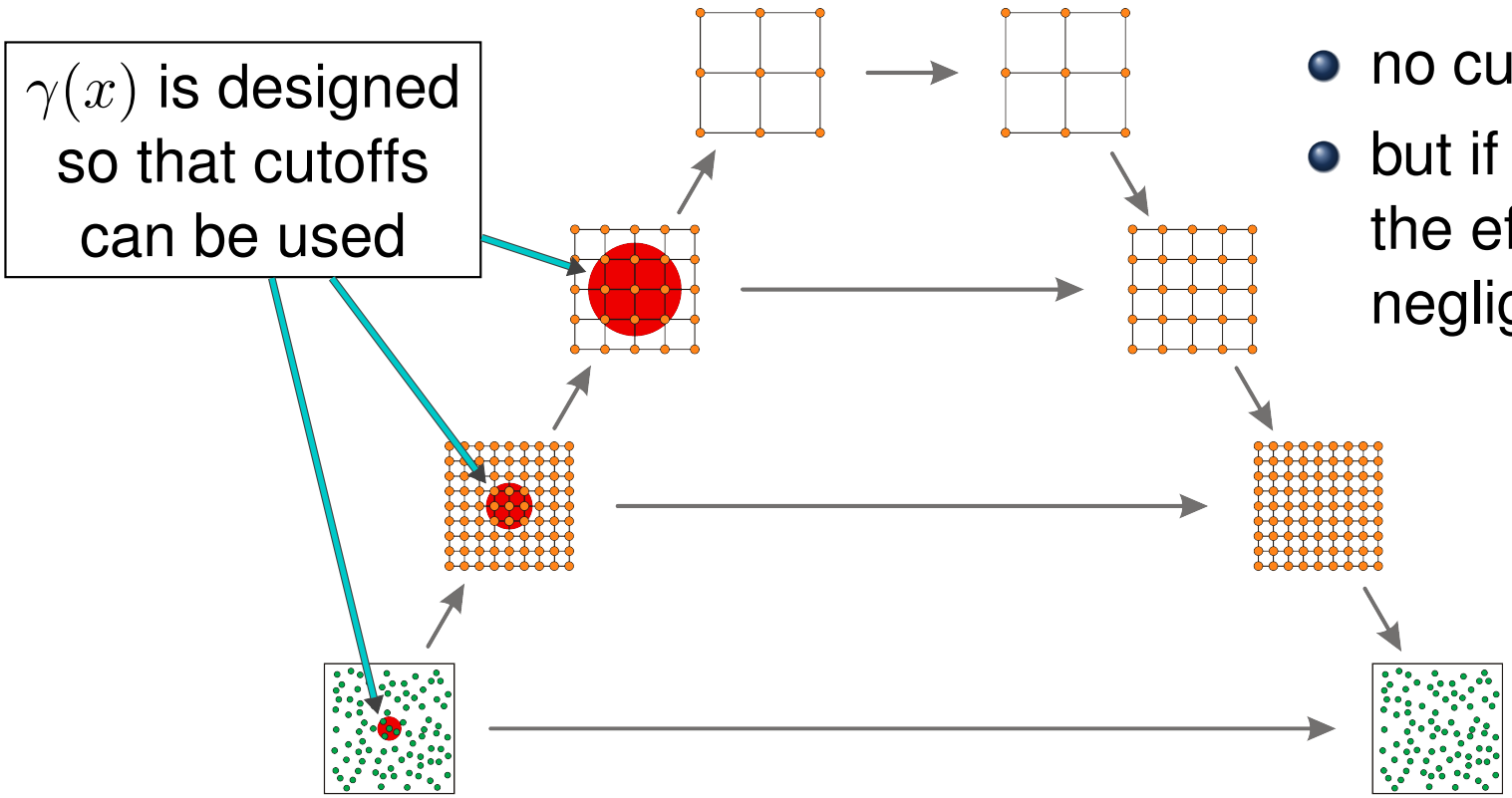
Multilevel Summation



$$\frac{1}{r^6} = \frac{1}{r^6} - \frac{1}{a^6}\gamma\left(\frac{r}{a}\right) + \frac{1}{a^6}\gamma\left(\frac{r}{a}\right) - \frac{1}{(2a)^6}\gamma\left(\frac{r}{2a}\right)$$

$$+ \frac{1}{(2a)^6}\gamma\left(\frac{r}{2a}\right) - \frac{1}{(4a)^6}\gamma\left(\frac{r}{4a}\right) + \cdots + \frac{1}{(2^{l-1}a)^6}\gamma\left(\frac{r}{2^{l-1}a}\right)$$

Multilevel Summation

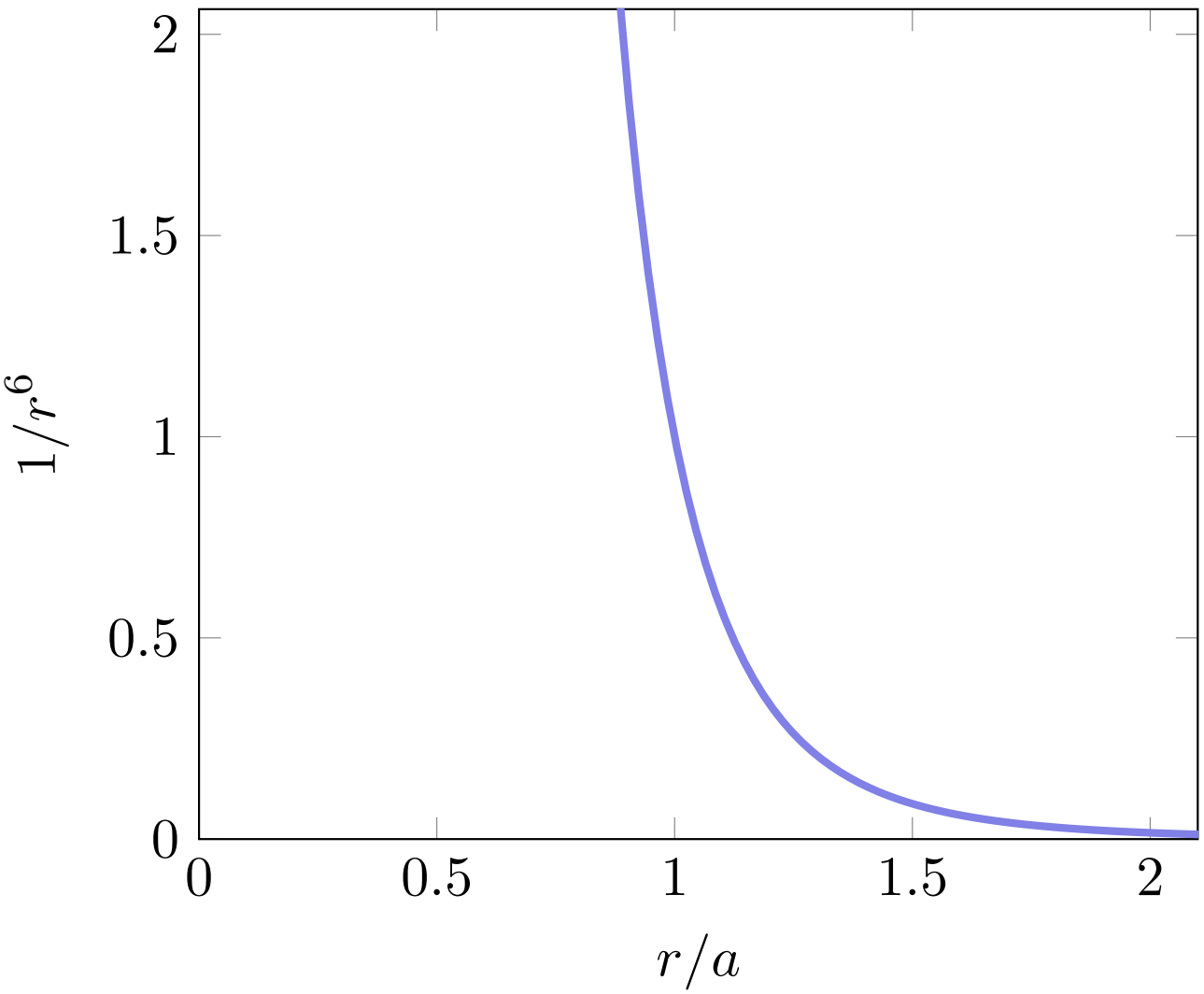


- no cutoff on last grid
- but if enough grids, the effort there is negligible

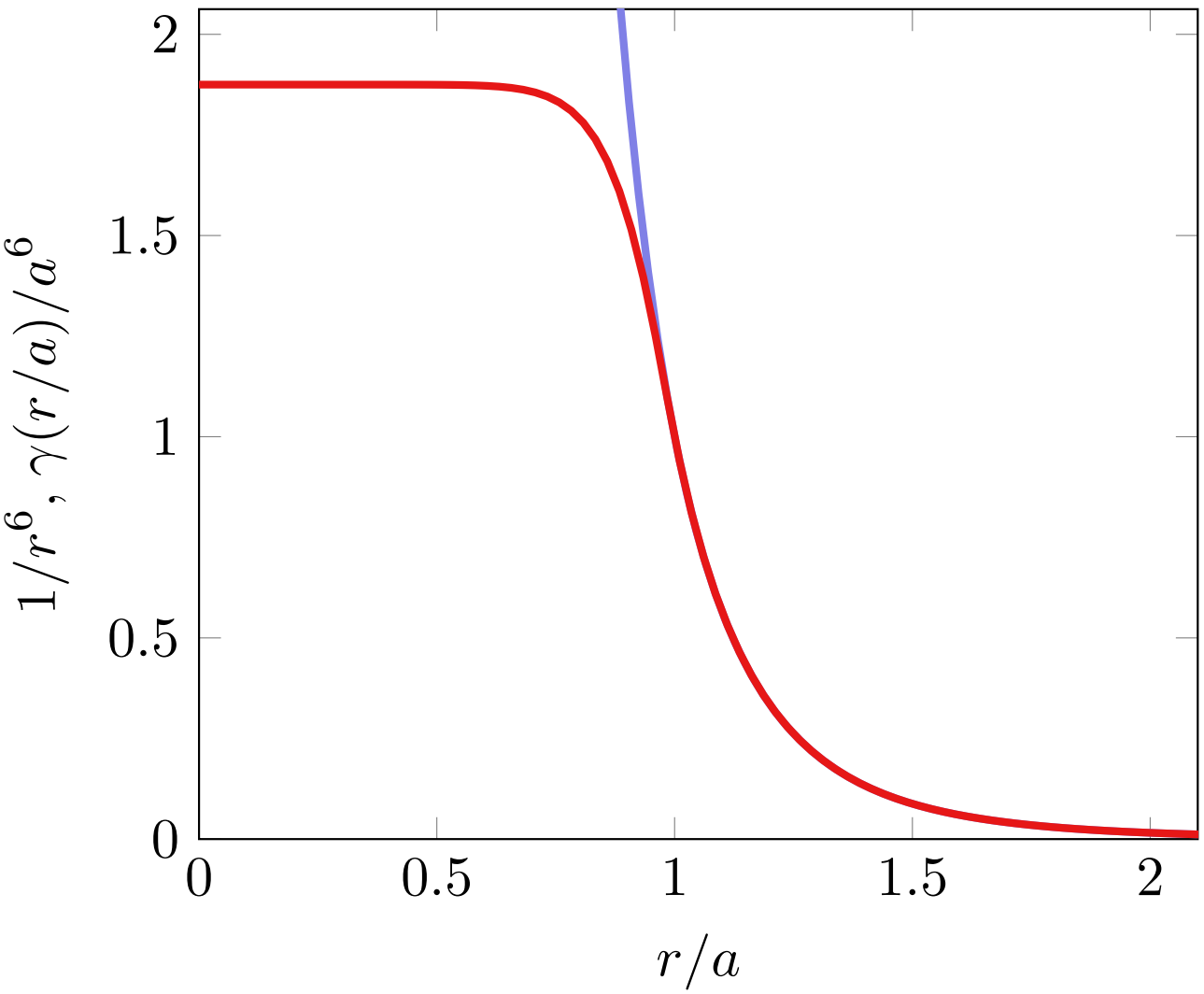
$$\frac{1}{r^6} = \frac{1}{r^6} - \frac{1}{a^6}\gamma\left(\frac{r}{a}\right) + \frac{1}{a^6}\gamma\left(\frac{r}{a}\right) - \frac{1}{(2a)^6}\gamma\left(\frac{r}{2a}\right)$$

$$+ \frac{1}{(2a)^6}\gamma\left(\frac{r}{2a}\right) - \frac{1}{(4a)^6}\gamma\left(\frac{r}{4a}\right) + \cdots + \frac{1}{(2^{l-1}a)^6}\gamma\left(\frac{r}{2^{l-1}a}\right)$$

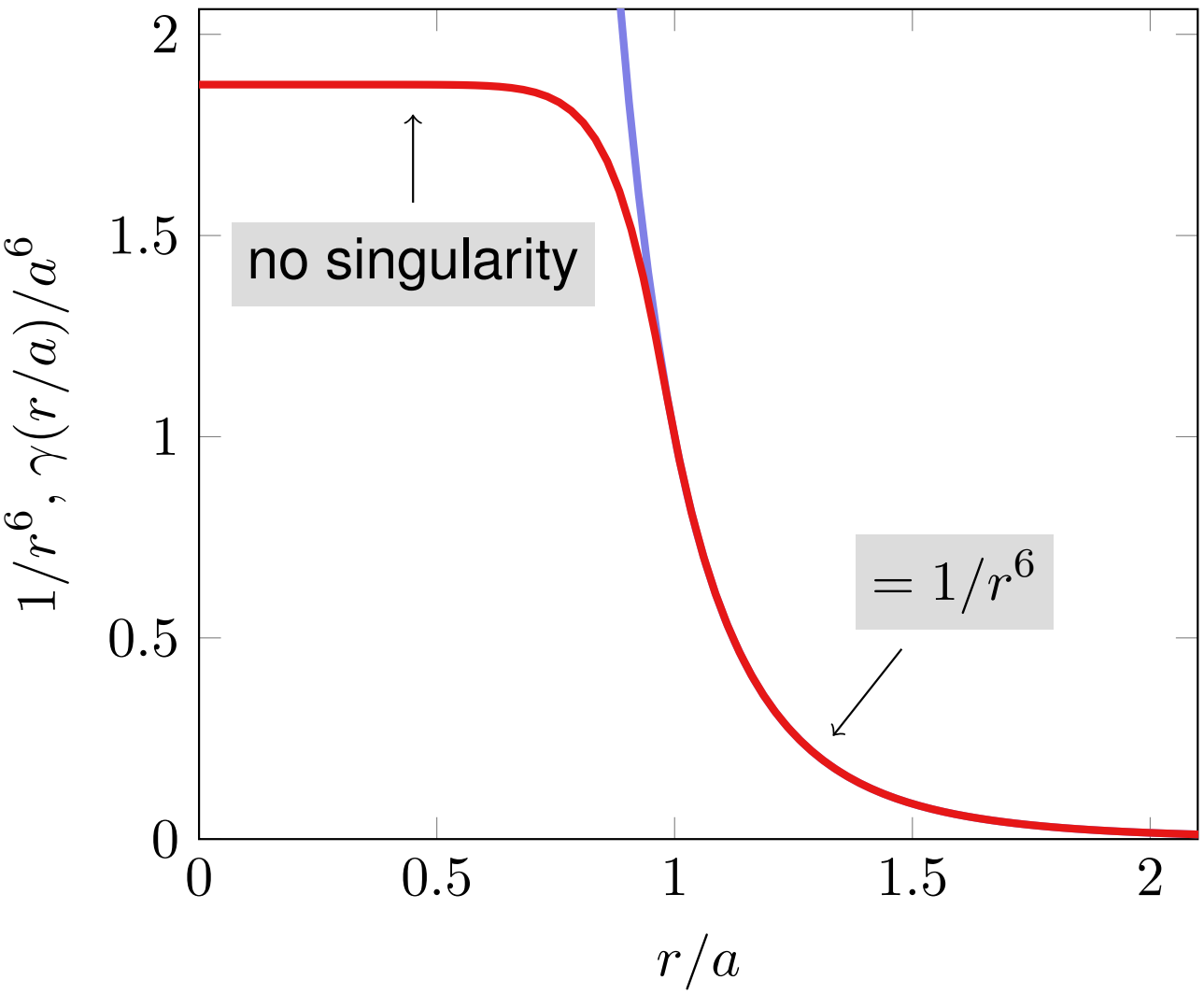
Smoothing Function γ



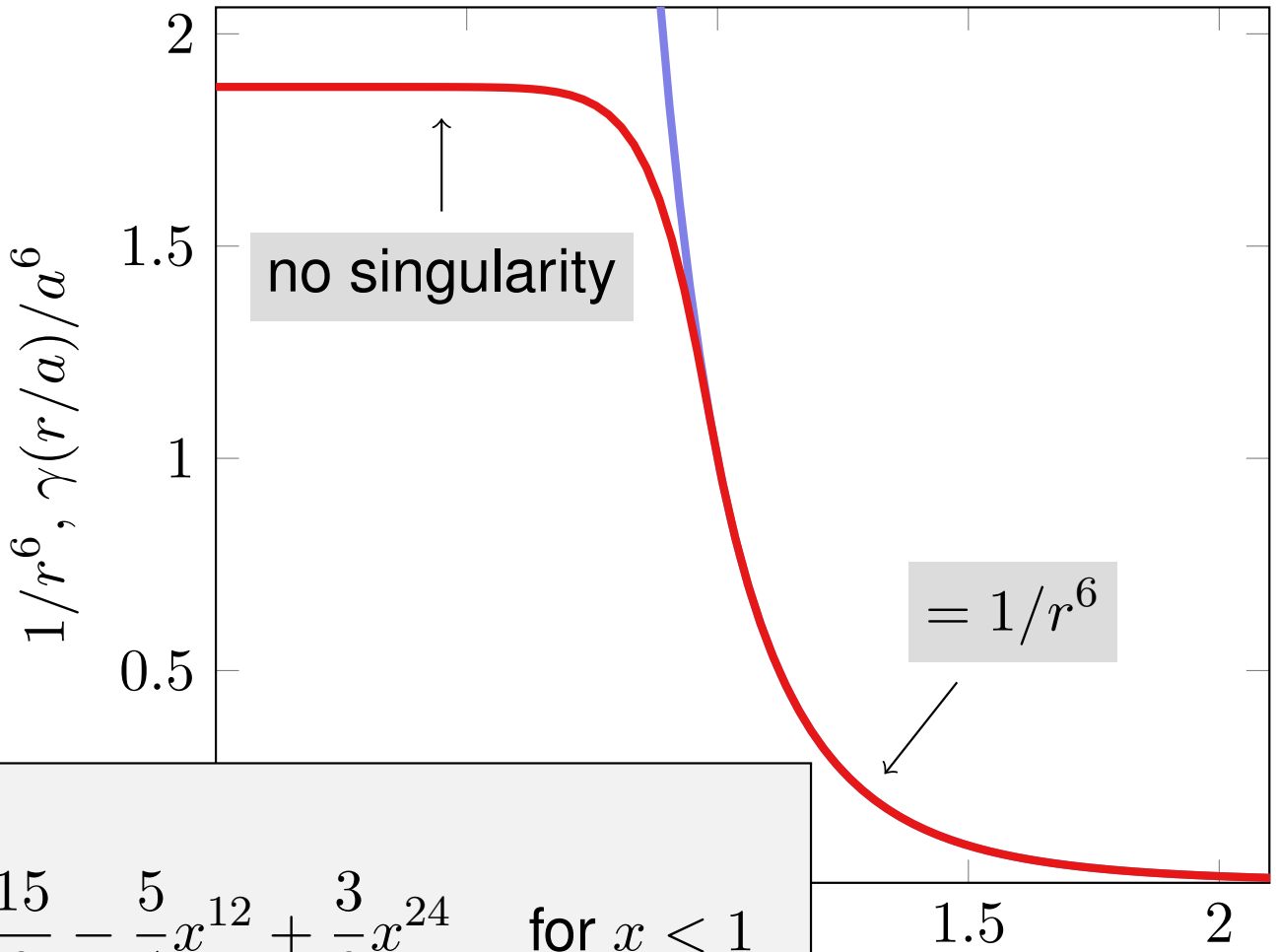
Smoothing Function γ



Smoothing Function γ



Smoothing Function γ



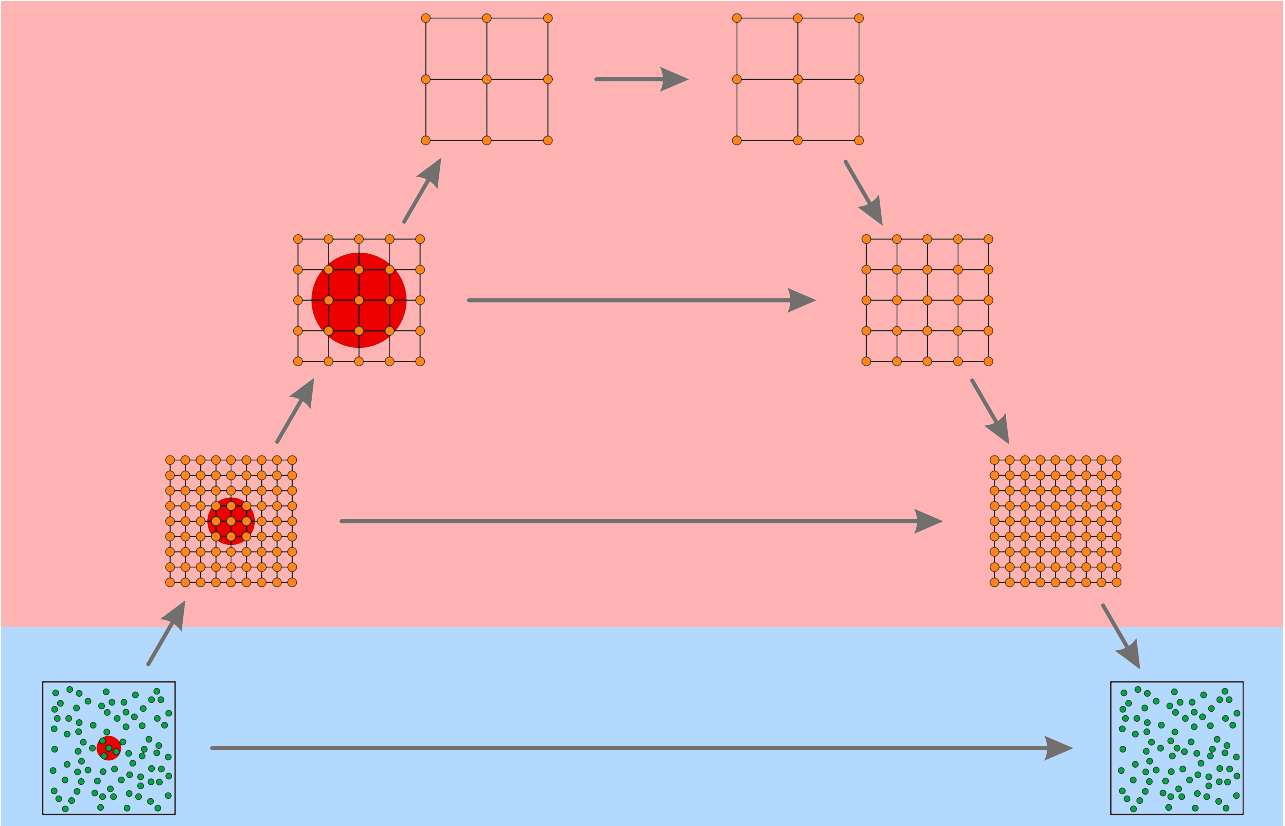
Example:

$$\gamma(x) = \begin{cases} \frac{15}{8} - \frac{5}{4}x^{12} + \frac{3}{8}x^{24} & \text{for } x < 1 \\ \frac{1}{x^6} & \text{for } x \geq 1 \end{cases}$$

Splitting

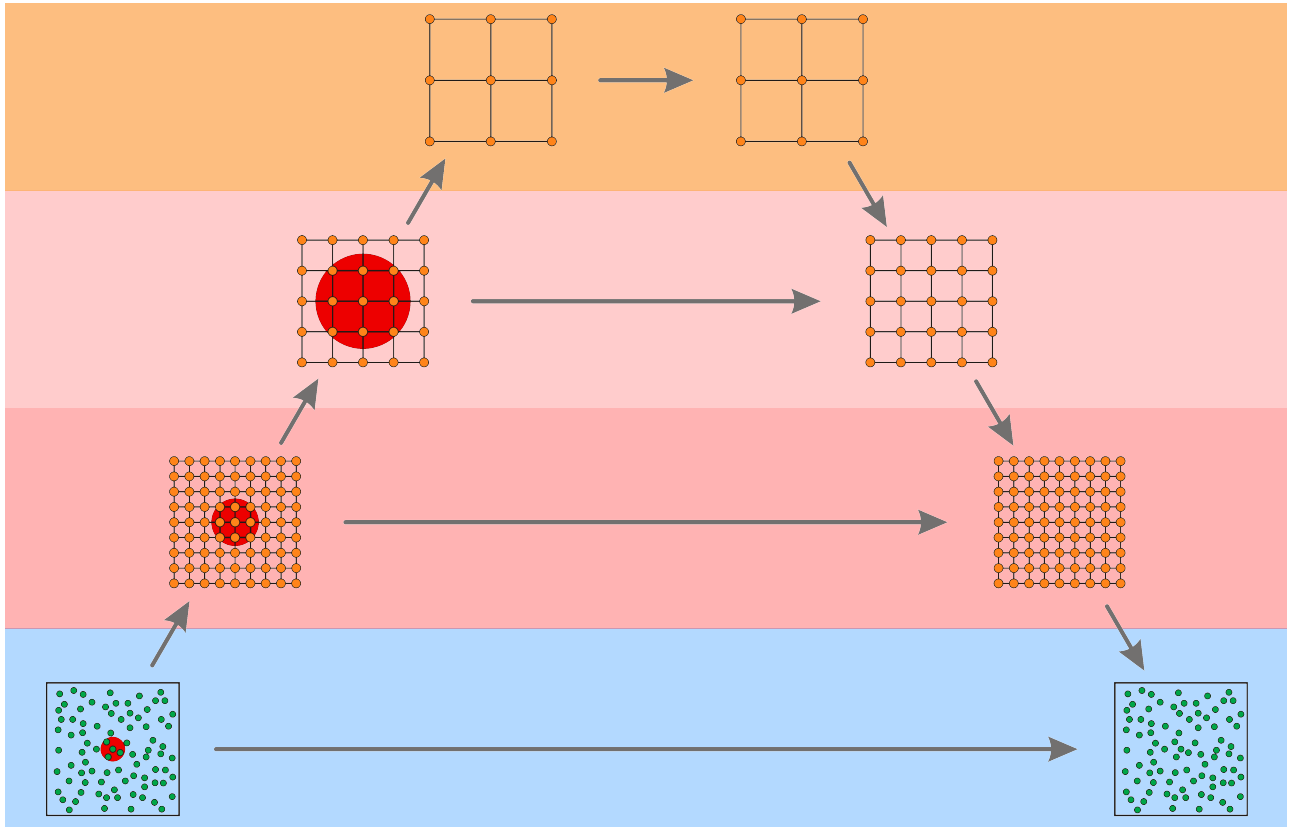
$$\frac{1}{r^6} = \frac{1}{r^6} - \frac{1}{a^6}\gamma\left(\frac{r}{a}\right) + \frac{1}{a^6}\gamma\left(\frac{r}{a}\right)$$

Splitting



$$\frac{1}{r^6} = \frac{1}{r^6} - \frac{1}{a^6}\gamma\left(\frac{r}{a}\right) + \frac{1}{a^6}\gamma\left(\frac{r}{a}\right)$$

Splitting



$$\frac{1}{r^6} = \frac{1}{r^6} - \frac{1}{a^6}\gamma\left(\frac{r}{a}\right) + \frac{1}{a^6}\gamma\left(\frac{r}{a}\right) - \frac{1}{(2a)^6}\gamma\left(\frac{r}{2a}\right)$$

$$+ \frac{1}{(2a)^6}\gamma\left(\frac{r}{2a}\right) - \frac{1}{(4a)^6}\gamma\left(\frac{r}{4a}\right) + \cdots + \frac{1}{(2^{l-1}a)^6}\gamma\left(\frac{r}{2^{l-1}a}\right)$$

Splitting

$$\frac{1}{r^6} = \frac{1}{r^6} - \frac{1}{a^6}\gamma\left(\frac{r}{a}\right) + \frac{1}{a^6}\gamma\left(\frac{r}{a}\right)$$

Splitting

$$\frac{1}{r^6} = \frac{1}{r^6} - \frac{1}{a^6}\gamma\left(\frac{r}{a}\right) + \frac{1}{a^6}\gamma\left(\frac{r}{a}\right)$$

Example:

$$\gamma(x) = \begin{cases} \frac{15}{8} - \frac{5}{4}x^{12} + \frac{3}{8}x^{24} & \text{for } x < 1 \\ \frac{1}{x^6} & \text{for } x \geq 1 \end{cases}$$

Splitting

$$\frac{1}{r^6} = \frac{1}{r^6} - \frac{1}{a^6}\gamma\left(\frac{r}{a}\right) + \frac{1}{a^6}\gamma\left(\frac{r}{a}\right) - \frac{1}{(2a)^6}\gamma\left(\frac{r}{2a}\right)$$

$$+ \frac{1}{(2a)^6}\gamma\left(\frac{r}{2a}\right)$$

Splitting

$$\frac{1}{r^6} = \frac{1}{r^6} - \frac{1}{a^6}\gamma\left(\frac{r}{a}\right) + \frac{1}{a^6}\gamma\left(\frac{r}{a}\right) - \frac{1}{(2a)^6}\gamma\left(\frac{r}{2a}\right)$$

$$+ \frac{1}{(2a)^6}\gamma\left(\frac{r}{2a}\right) - \frac{1}{(4a)^6}\gamma\left(\frac{r}{4a}\right) + \frac{1}{(4a)^6}\gamma\left(\frac{r}{4a}\right)$$

Splitting

$$\frac{1}{r^6} = \frac{1}{r^6} - \frac{1}{a^6}\gamma\left(\frac{r}{a}\right) + \frac{1}{a^6}\gamma\left(\frac{r}{a}\right) - \frac{1}{(2a)^6}\gamma\left(\frac{r}{2a}\right)$$

$$+ \frac{1}{(2a)^6}\gamma\left(\frac{r}{2a}\right) - \frac{1}{(4a)^6}\gamma\left(\frac{r}{4a}\right) + \cdots +$$

$$+ \frac{1}{(2^{l-1}a)^6}\gamma\left(\frac{r}{2^{l-1}a}\right)$$

Splitting

$$g_0(r) = \frac{1}{r^6} - \frac{1}{a^6}\gamma\left(\frac{r}{a}\right)$$

$$g_k(r) = \frac{1}{2^{6(k-1)}a^6}\gamma\left(\frac{r}{2^{(k-1)}a}\right) - \frac{1}{2^{6k}a^6}\gamma\left(\frac{r}{2^k a}\right) \text{ for } k = 1, ..., l-1$$

$$g_l(r) = \frac{1}{2^{6l-6}a^6}\gamma\left(\frac{r}{2^{l-1}a}\right)$$

Splitting

$$g_0(r) = \frac{1}{r^6} - \frac{1}{a^6}\gamma\left(\frac{r}{a}\right)$$

$$g_k(r) = \frac{1}{2^{6(k-1)}a^6}\gamma\left(\frac{r}{2^{(k-1)}a}\right) - \frac{1}{2^{6k}a^6}\gamma\left(\frac{r}{2^k a}\right) \text{ for } k = 1, ..., l-1$$

$$g_l(r) = \frac{1}{2^{6l-6}a^6}\gamma\left(\frac{r}{2^{l-1}a}\right)$$

the part of the calculation with g_0 can be done exactly

$\Rightarrow$ error arises completely from grid-based part

Approximation of g_i

$$\frac{1}{r^6} = g_0 + g_1 + g_2 + \cdots + g_{l-1} + g_l$$

Approximation of g_i

$$\frac{1}{r^6} = g_0 + g_1 + g_2 + \cdots + g_{l-1} + g_l$$

$$\frac{1}{r^6} \approx g_0 + \mathcal{I}_1 \left[g_1 + \mathcal{I}_2 \left[g_2 \ldots + \mathcal{I}_{l-1} \left[g_{l-1} + \mathcal{I}_l \left[g_l \right] \right] \ldots \right] \right]$$

Approximation of g_i

$$\frac{1}{r^6} = g_0 + g_1 + g_2 + \cdots + g_{l-1} + g_l$$

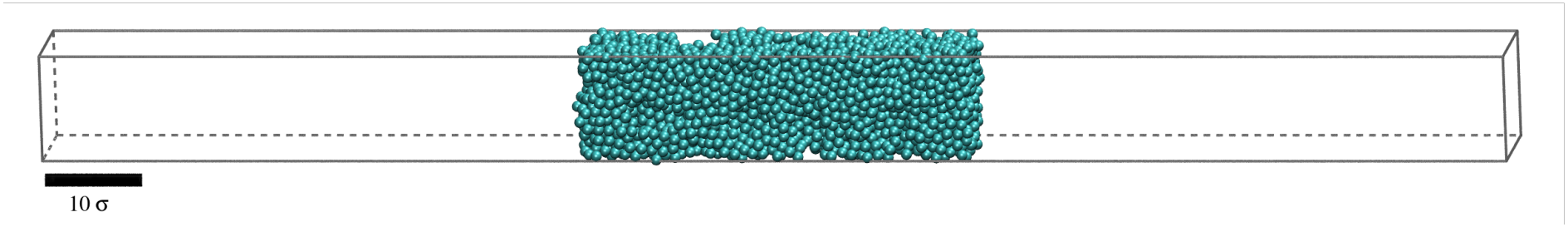
$$\frac{1}{r^6} \approx g_0 + \mathcal{I}_1\left[g_1 + \mathcal{I}_2\left[g_2 \ldots + \mathcal{I}_{l-1}\left[g_{l-1} + \mathcal{I}_l\left[g_l\right]\right] \ldots\right]\right]$$

$$\mathcal{I}_k\left[g_k\right]\left(\mathbf{x}_i^{k-1}, \mathbf{x}_j^{k-1}\right) = \sum_\mu \sum_\nu \phi_\mu^k\left(\mathbf{x}_i^{k-1}\right) g_k\left(\mathbf{x}_\mu^k, \mathbf{x}_\nu^k\right) \phi_\nu^k\left(\mathbf{x}_j^{k-1}\right)$$

$$\phi_\mu^k\left(\mathbf{x}_i^{k-1}\right) = \Phi\left(\frac{x^{k-1} - x_\mu^k}{2^k h_x}\right) \Phi\left(\frac{y^{k-1} - y_\mu^k}{2^k h_y}\right) \Phi\left(\frac{z^{k-1} - z_\mu^k}{2^k h_z}\right)$$

$$\Phi(\xi) = \begin{cases} (1 - |\xi|)\left(1 + |\xi| - \frac{3}{2}\xi^2\right), & \text{for } |\xi| \leq 1, \\ -\frac{1}{2}(|\xi| - 1)(2 - |\xi|)^2, & \text{for } 1 < |\xi| \leq 2, \\ 0 & \text{otherwise.} \end{cases}$$

Simulations



4000 particles

Lennard-Jones potential with geometric mixing

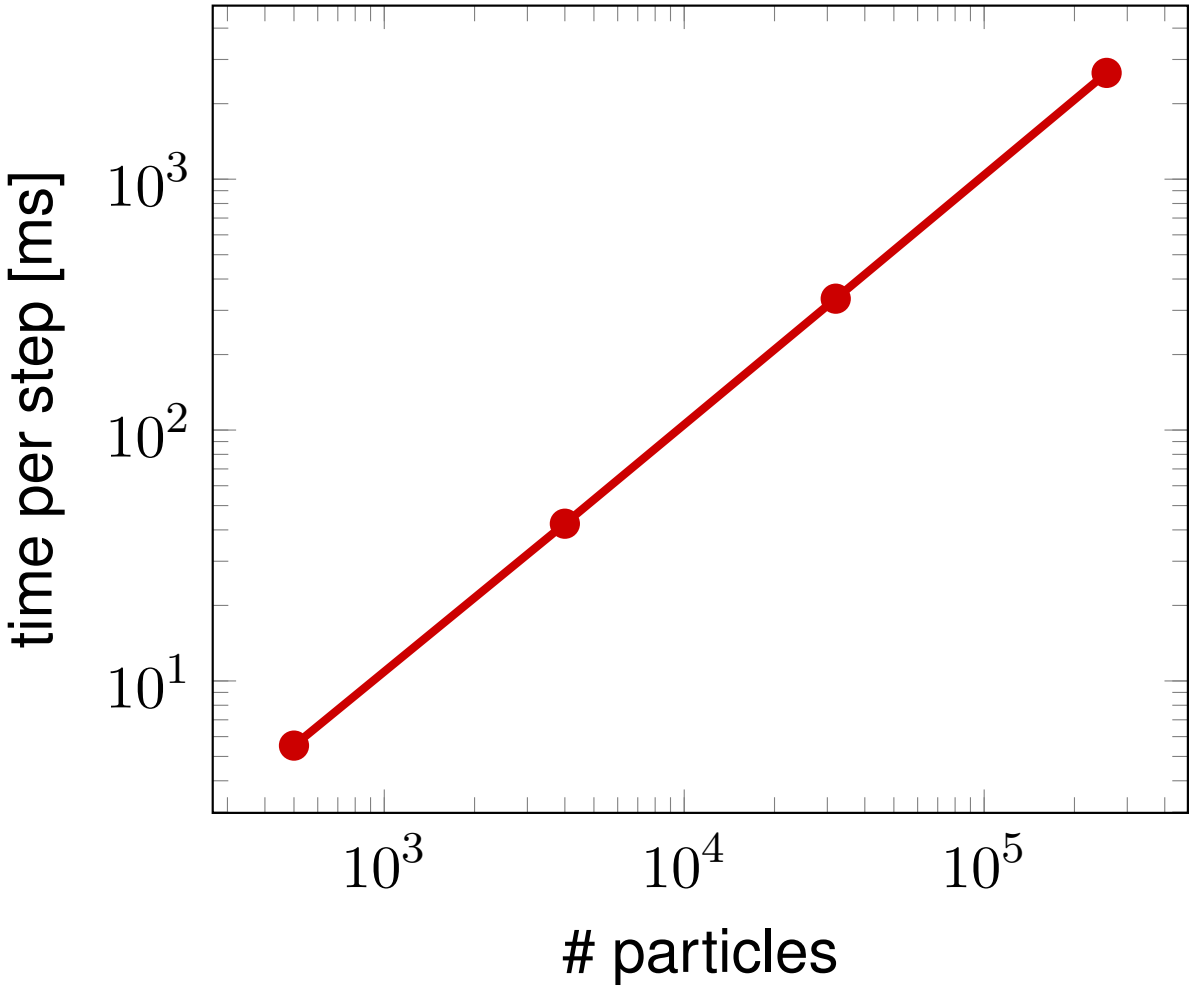
$11.01\sigma \times 11.01\sigma \times 176.16\sigma$ with periodic boundaries

In the dispersion case there are these parameters:

- cutoff a
- spacing of finest grid h
- number of grids
- function $\gamma(x)$
- function $\Phi(x)$

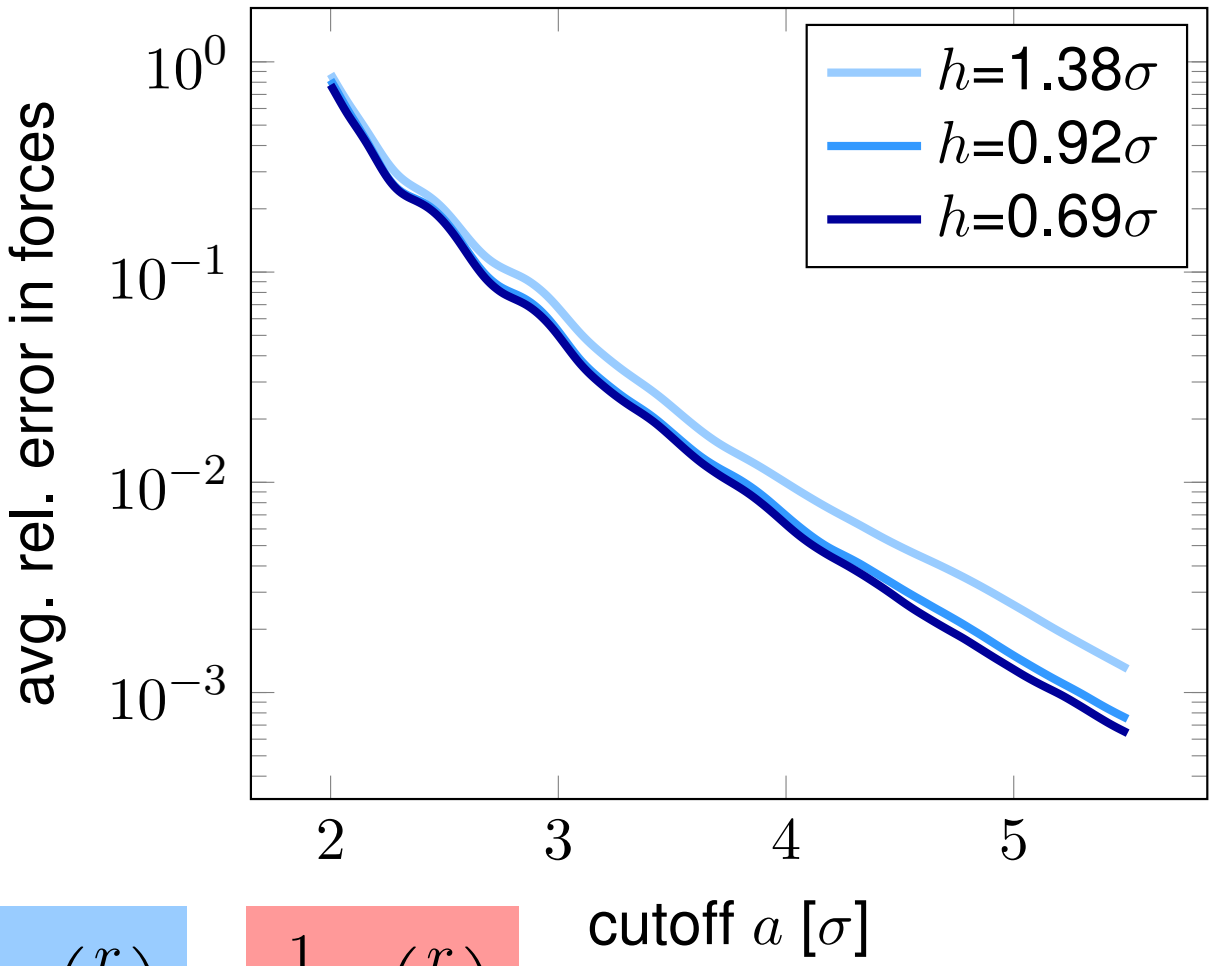
Complexity

Results achieved with serial, unoptimized prototype implementation:



Accuracy & Performance

Results achieved with serial, unoptimized prototype implementation:



$$\frac{1}{r^6} = \frac{1}{r^6} - \frac{1}{a^6}\gamma\left(\frac{r}{a}\right) + \frac{1}{a^6}\gamma\left(\frac{r}{a}\right)$$

Number of Grids

Table: Timings for the 4000-particle system with $a = 3\sigma$ and $h = 0.688\sigma$. Serial runs on Sandy-Bridge on AICES cluster.

# grids k	2	3	4	5
Kspace time per step [ms]	302.96	253.41	257.49	258.69
Pair time per step [ms]	6.71	6.73	6.73	6.94

Number of Grids

Table: Timings for the 4000-particle system with $a = 3\sigma$ and $h = 0.688\sigma$. Serial runs on Sandy-Bridge on AICES cluster.

# grids k	2	3	4	5
Kspace time per step [ms]	302.96	253.41	257.49	258.69
Pair time per step [ms]	6.71	6.73	6.73	6.94
# points on last grid	8192	1024	128	16

points within cutoff: 2801

Number of Grids

Table: Accuracy for the 4000-particle system with $a = 3\sigma$ and $h = 0.688\sigma$.

# grids k	3	4	5
$\frac{\lvert\Delta E_3 - \Delta E_k\rvert}{\Delta E_3}$	0	4.67×10^{-4}	6.05×10^{-4}
$\frac{\lVert\Delta F_3 - \Delta F_k\rVert}{\Delta F_3}$	0	2.15×10^{-6}	2.10×10^{-6}

Accuracy & Performance

General form of error bounds:

$$\epsilon = C_{error}\, d^7 \frac{h^n}{a^{n+7}}$$

n depends on: interpolation function Φ
smoothing function γ

Example:

$\epsilon = 844.9\, d^7 h^2 / a^9$

d: mean nearest-neighbor distance ($d^3 = V/N$)

Accuracy & Performance

- Performance model: expensive part are the cutoff calculations
- Modeled by:

$$c(a,h) = C_{\text{local}} \left(\frac{a}{d}\right)^3 N + C_{\text{grids}} \left(\frac{a}{h}\right)^3 \frac{V}{h^3}$$

V: volume of simulation box
d: mean nearest-neighbor distance ($d^3 = V/N$)

(roughly 90% of total time on grids is spent on finest one)

Accuracy & Performance

- Performance model: expensive part are the cutoff calculations
- Modeled by:

$$c(a,h) = C_{\text{local}} \left(\frac{a}{d}\right)^3 N + C_{\text{grids}} \left(\frac{a}{h}\right)^3 \frac{V}{h^3}$$

↑

avg. number of particles inside cutoff

V: volume of simulation box
d: mean nearest-neighbor distance ($d^3 = V/N$)

(roughly 90% of total time on grids is spent on finest one)

Accuracy & Performance

- Performance model: expensive part are the cutoff calculations
- Modeled by:

number of points inside cutoff

$$c(a,h) = C_{\text{local}} \left(\frac{a}{d}\right)^3 N + C_{\text{grids}} \left(\frac{a}{h}\right)^3 \frac{V}{h^3}$$

avg. number of particles inside cutoff

V: volume of simulation box
d: mean nearest-neighbor distance ($d^3 = V/N$)

(roughly 90% of total time on grids is spent on finest one)

Accuracy & Performance

- Performance model: expensive part are the cutoff calculations
- Modeled by:

number of points inside cutoff

$$c(a,h) = C_{\text{local}} \left(\frac{a}{d}\right)^3 N + C_{\text{grids}} \left(\frac{a}{h}\right)^3 \frac{V}{h^3}$$

avg. number of particles inside cutoff

number of points of finest grid

V: volume of simulation box
d: mean nearest-neighbor distance ($d^3 = V/N$)

(roughly 90% of total time on grids is spent on finest one)

Accuracy & Performance

- Performance model: expensive part are the cutoff calculations
- Modeled by:

$$c(a,h) = C_{\text{local}} \left(\frac{a}{d}\right)^3 N + C_{\text{grids}} \left(\frac{a}{h}\right)^3 \left(\frac{d}{h}\right)^3 N$$

V: volume of simulation box
d: mean nearest-neighbor distance ($d^3 = V/N$)

(roughly 90% of total time on grids is spent on finest one)

Accuracy & Performance

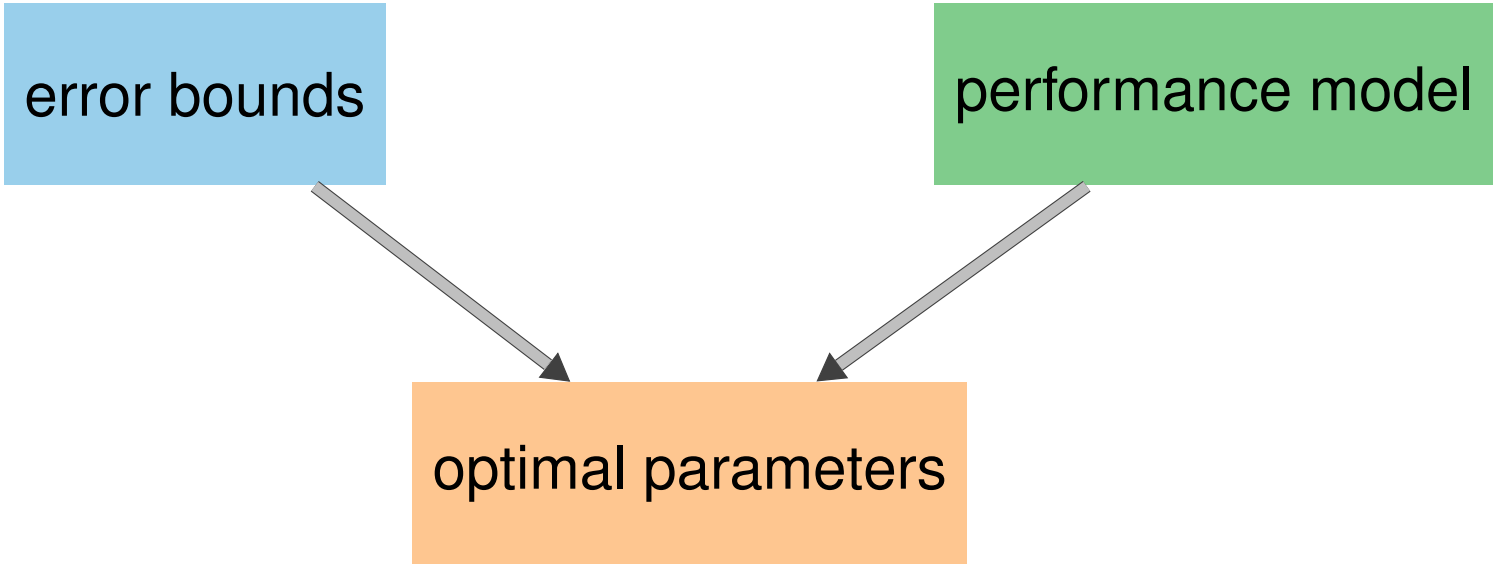
- Performance model: expensive part are the cutoff calculations
- Modeled by:

$$c(a,h) = C_{\text{local}} \left(\frac{a}{d}\right)^3 N + C_{\text{grids}} \left(\frac{a}{h}\right)^3 \left(\frac{d}{h}\right)^3 N$$

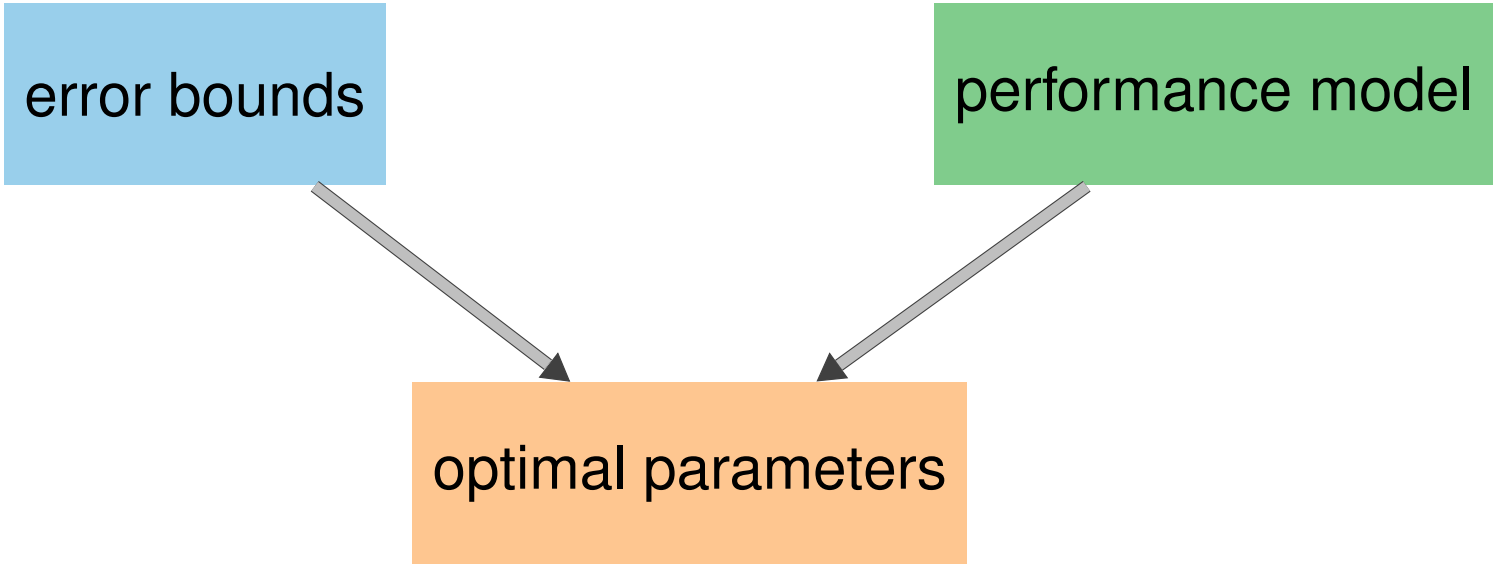
V: volume of simulation box
d: mean nearest-neighbor distance ($d^3 = V/N$)

(roughly 90% of total time on grids is spent on finest one)

Accuracy & Performance



Accuracy & Performance

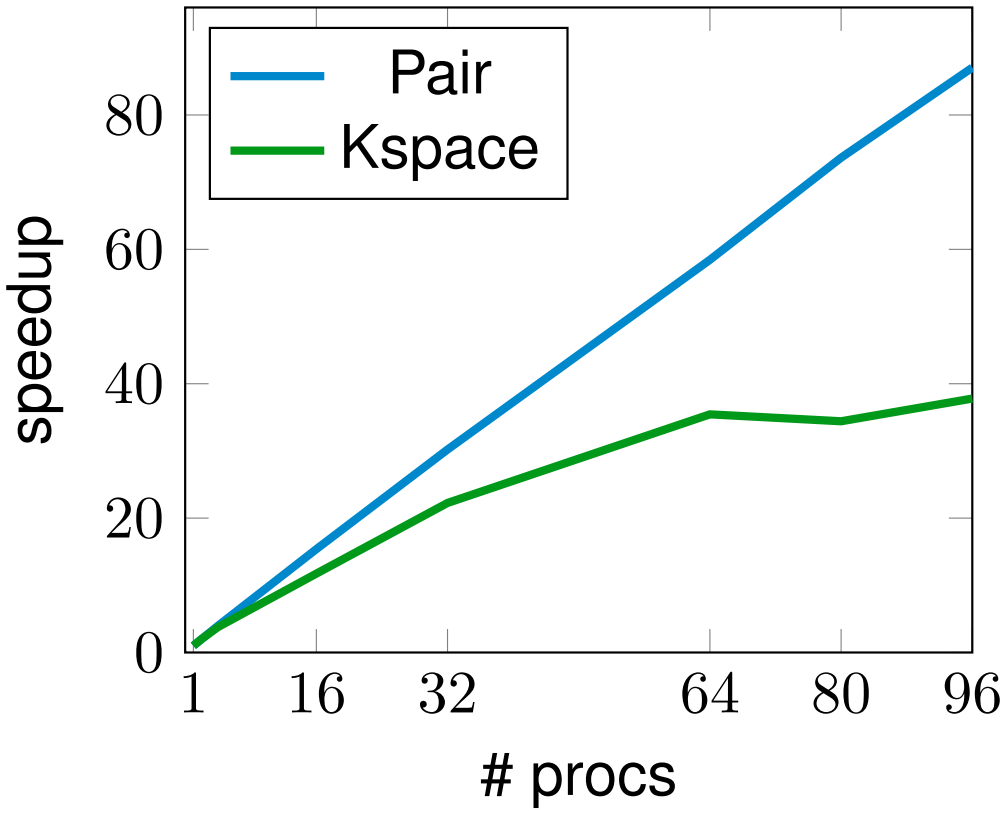


$$a = \left(\frac{C_{error}\, d^7 h^n}{\epsilon} \right)^{\frac{1}{n+7}} \qquad h = \left(\frac{C_{\text{grids}}}{C_{\text{local}}} \frac{(n+14)}{n} \right)^{1/6} d$$

Strong Scaling

Strong scaling for a 256,000-particle system with interface $h = 0.688\sigma$ ($\approx$ 4.2m points on finest grid)

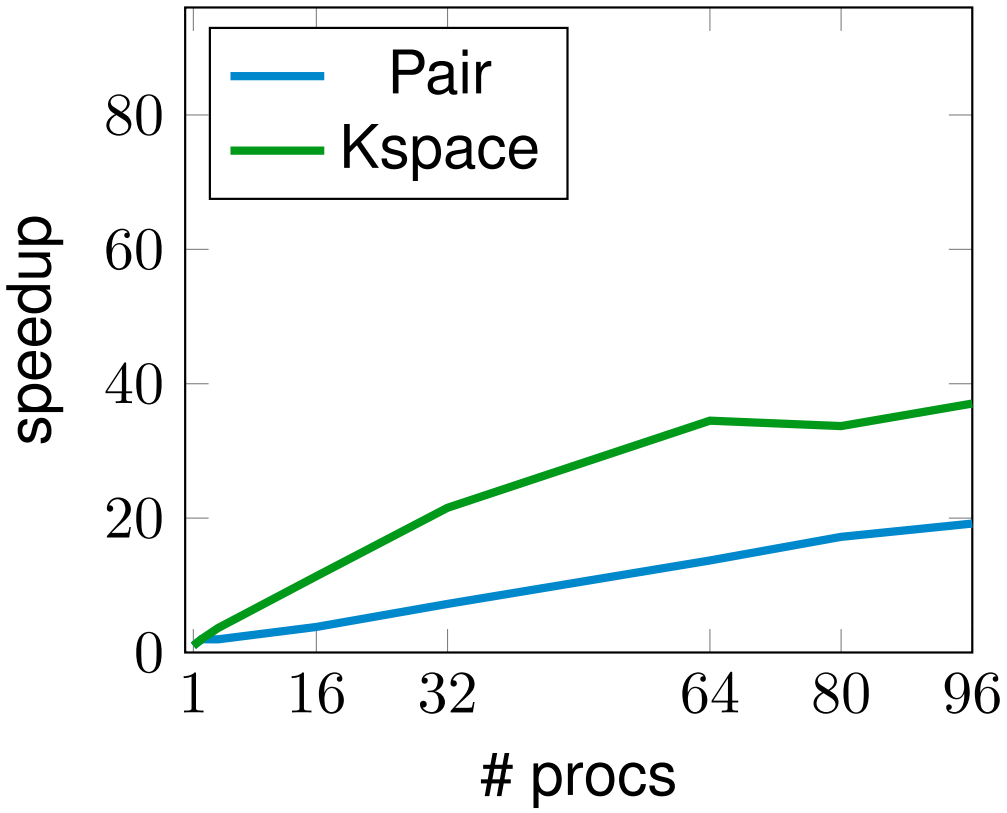
(Sandy-Bridge on AICES cluster)



Strong Scaling

Strong scaling for a 256,000-particle system with interface $h = 0.688\sigma$ ($\approx$ 4.2m points on finest grid)

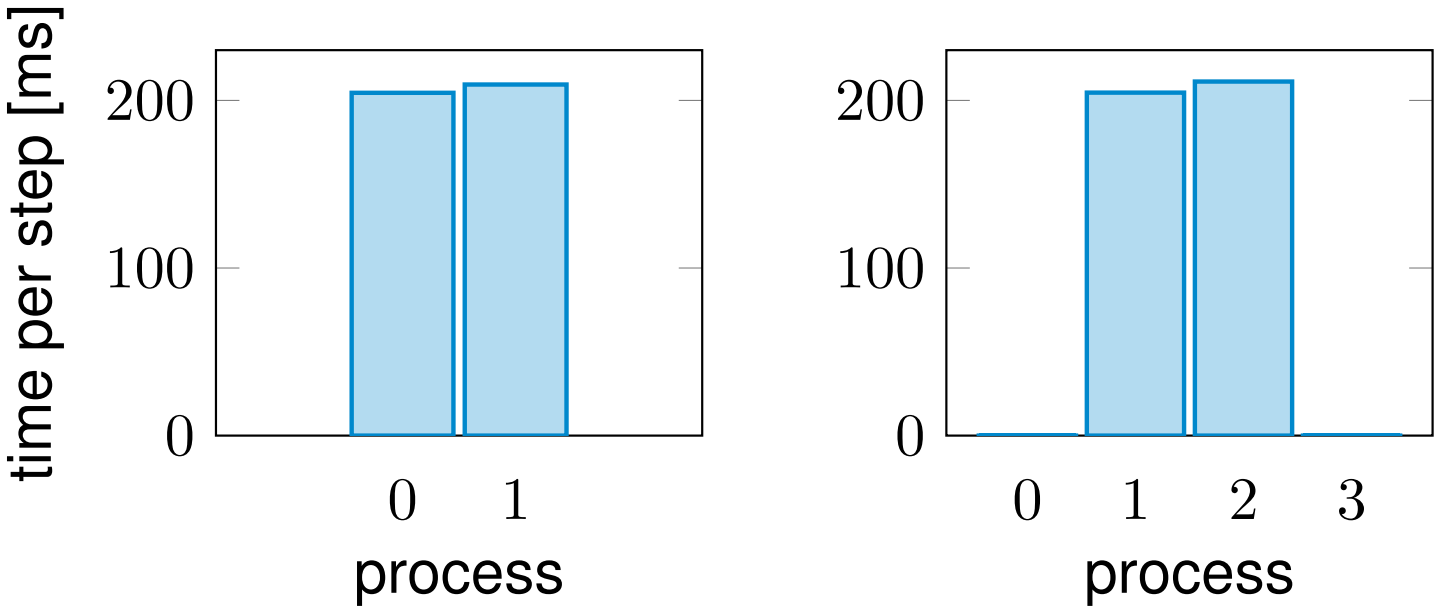
(Sandy-Bridge on AICES cluster)



Strong Scaling

Strong scaling for a 256,000-particle system with interface $h = 0.688\sigma$ ($\approx$ 4.2m points on finest grid)

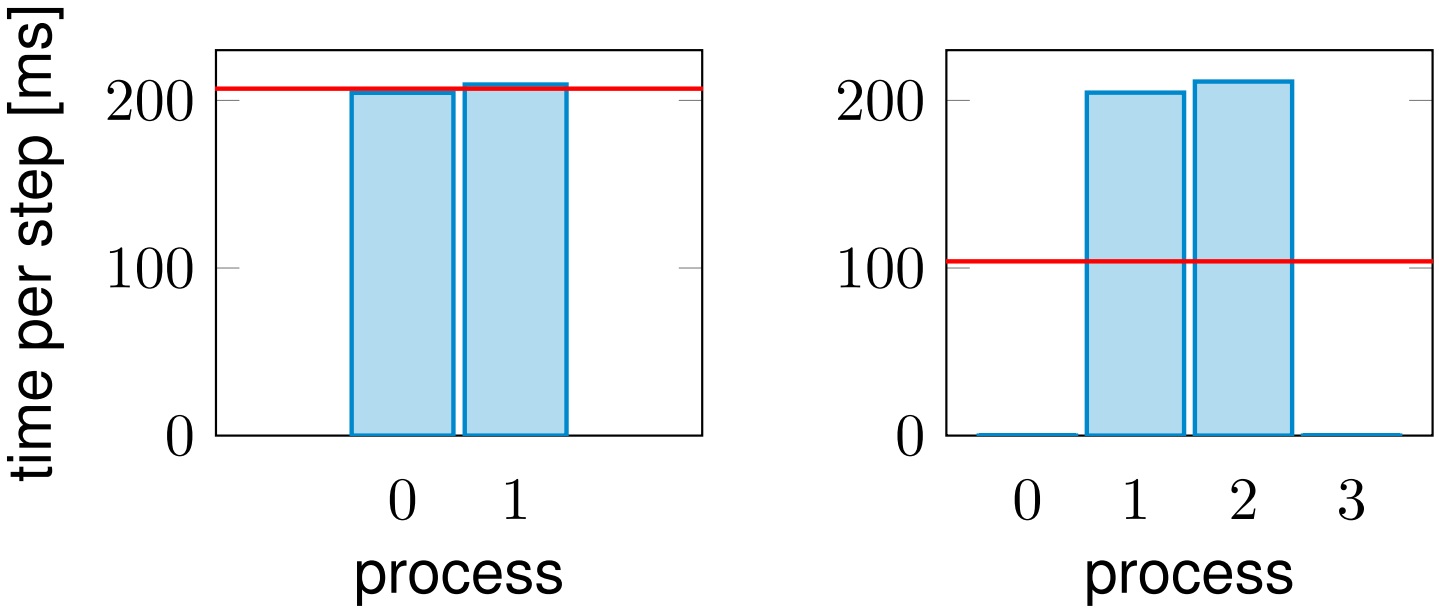
(Sandy-Bridge on AICES cluster)



Strong Scaling

Strong scaling for a 256,000-particle system with interface $h = 0.688\sigma$ ($\approx$ 4.2m points on finest grid)

(Sandy-Bridge on AICES cluster)



Strong Scaling

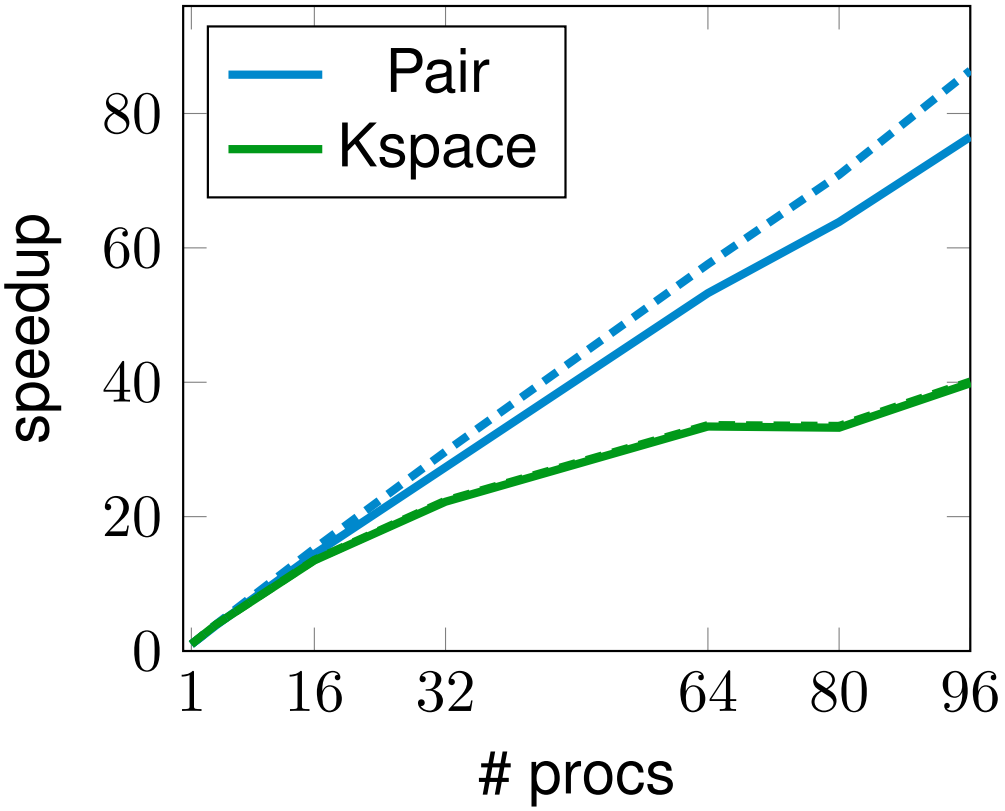
Table: Strong scaling for a 256,000-particle system with $h = 0.688\sigma$.

#p	Pair	Kspace	Neigh	Comm	Other
1	412.97	10426.	120.24	1.92	12.77
2	206.77	5225.93	60.06	1.29	6.05
4	103.86	2758.69	30.95	0.91	28.71
16	26.85	888.13	7.98	3.79	18.63
32	13.67	468.45	3.94	1.24	10.69
64	7.07	294.16	2.02	0.98	6.76
80	5.60	302.97	1.60	0.93	5.34
96	4.74	275.88	1.36	1.11	5.17

Strong Scaling

Strong scaling for a 1,024,000-particle system without interface
$h = 0.688\sigma$ ($\approx$ 4.2m points on finest grid)

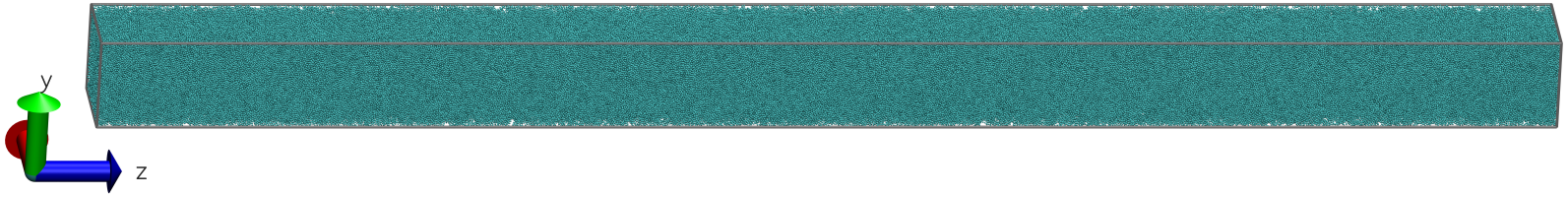
(Sandy-Bridge on AICES cluster)



Strong Scaling

Strong scaling for a 1,024,000-particle system without interface
$h = 0.688\sigma$ ($\approx$ 4.2m points on finest grid)

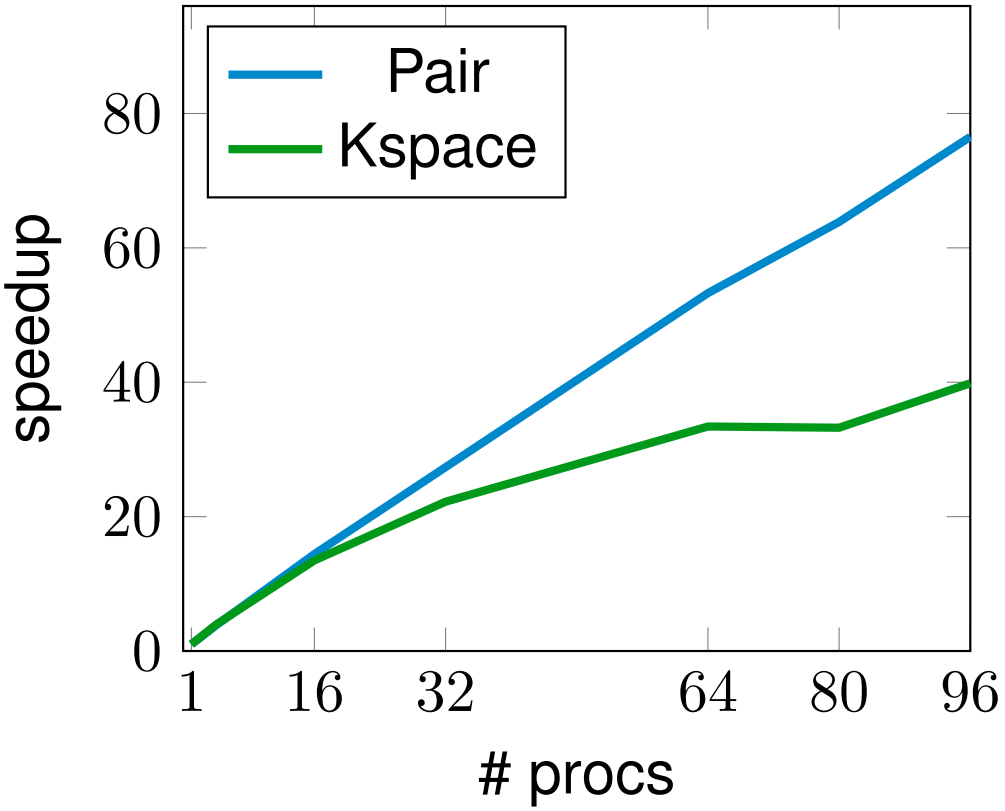
(Sandy-Bridge on AICES cluster)



Strong Scaling

Strong scaling for a 1,024,000-particle system without interface
$h = 0.688\sigma$ ($\approx$ 4.2m points on finest grid)

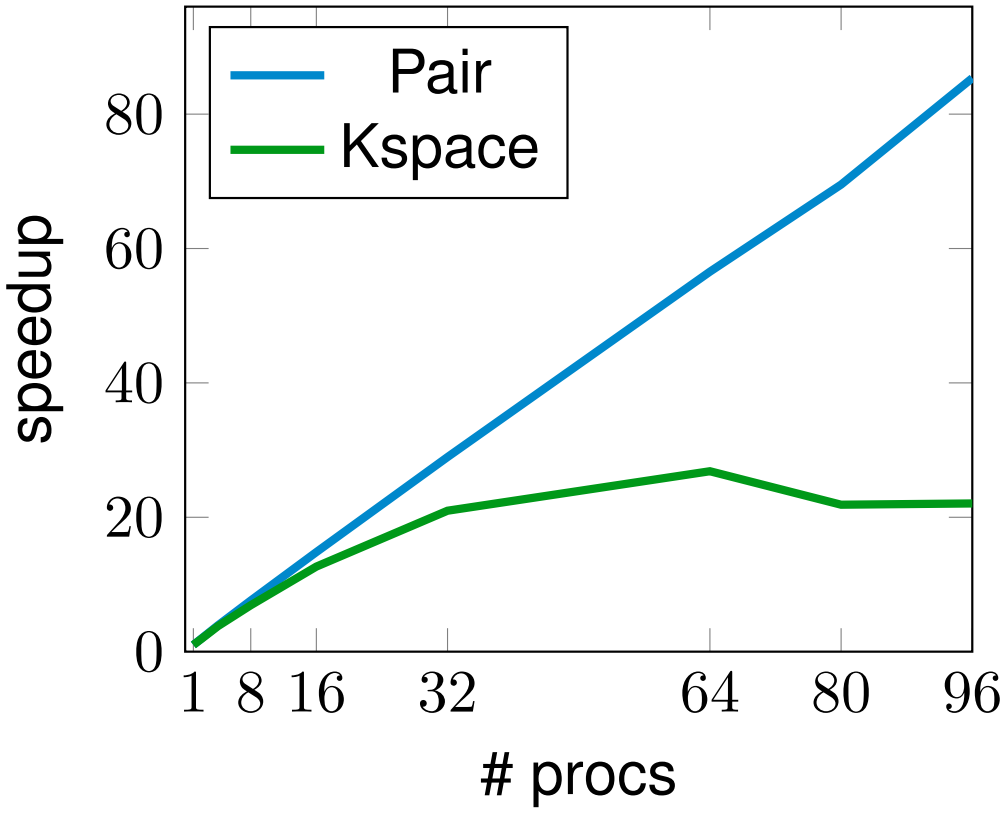
(Sandy-Bridge on AICES cluster)



Strong Scaling

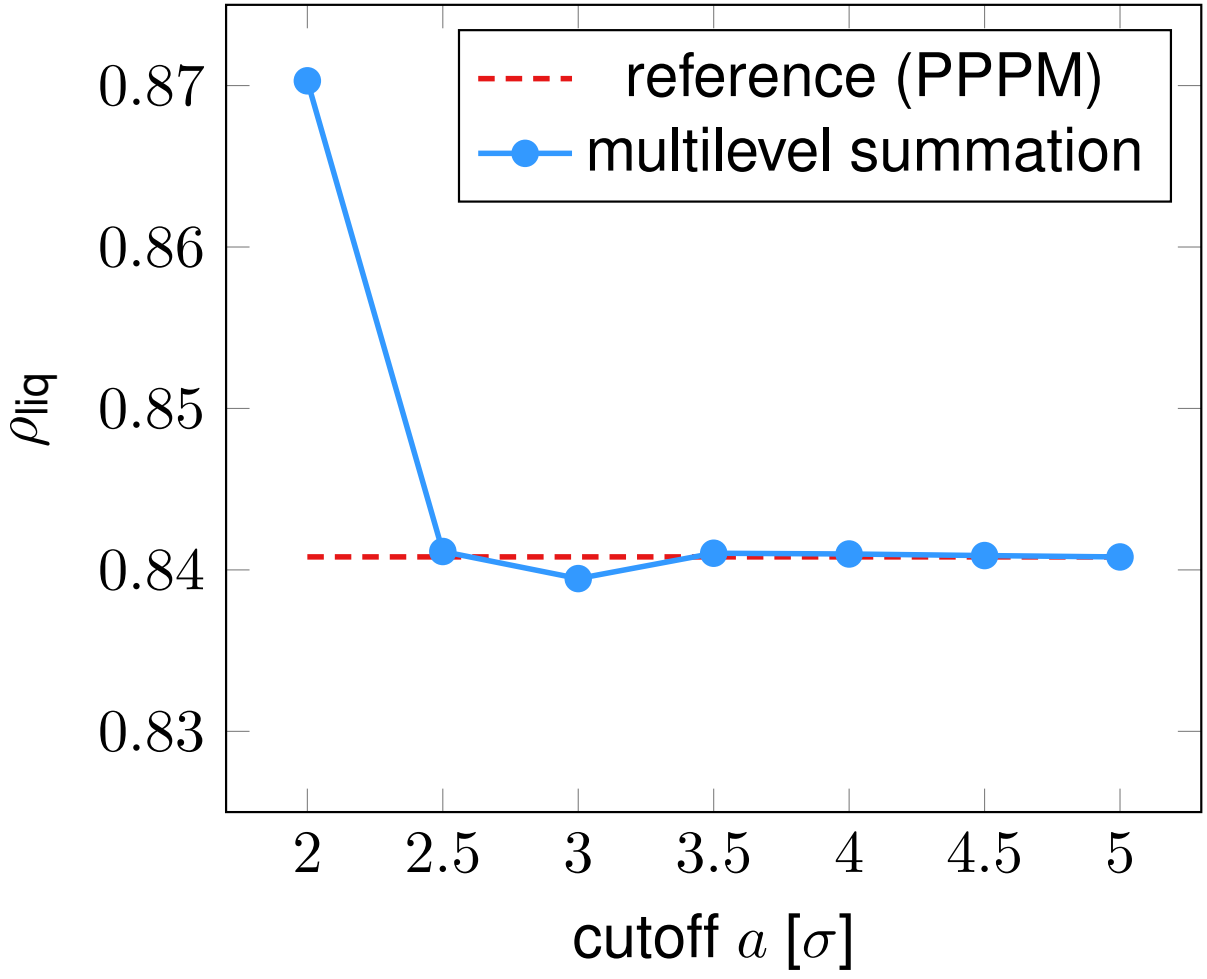
Strong scaling for a 1,024,000-particle system without interface
$h = 1.376\sigma$ ($\approx$ 0.5m points on finest grid)

(Harpertown on AICES cluster)



Liquid Density

Liquid density for 4,000 particles at $T^* = 0.7$
for various values of the cutoff a with $h = 1.38\sigma$:



Conclusions

- I implemented a MPI parallelized version of the multilevel summation for dispersion interactions
- Appealing properties of Multilevel Summation:
  - Linear scaling in number of particles
  - Mainly local communications
- Error bounds and performance model deliver recommendations for parameter selection

Future Work

- Error estimator requires further attention
- Better model for performance?
- What factors have an influence on C_{local} and C_{grids}?
- Improve the scaling
- Load-balancing for inhomogeneous systems
- How to choose $\gamma(x)$ and $\Phi(x)$?
- Different ways of construction $\gamma(x)$?

Acknowledgement

Financial support from the
Deutsche Forschungsgemeinschaft
(German Research Foundation)
through grant GSC 111
is gratefully acknowledged.

Deutsche Forschungsgemeinschaft
DFG