Development of efficient multiscale methods and extrapolation techniques for multiphysics molecular chemistry

Étienne Polack Doctoral research project defence, January 2022 Chemistry science of matter and its elements;

Chemistry science of matter and its elements; understanding it to

Chemistry

Chemistry science of matter and its elements; understanding it to

• Synthesise materials (e.g., plastics, 2D materials)

Chemistry

Chemistry science of matter and its elements; understanding it to

- Synthesise materials (e.g., plastics, 2D materials)
- Optimise industrial processes (e.g., sewage plants)

Chemistry science of matter and its elements; understanding it to

- Synthesise materials (e.g., plastics, 2D materials)
- Optimise industrial processes (e.g., sewage plants)
- Discover new drugs (e.g., opioid antagonist naloxone)

Chemistry science of matter and its elements; understanding it to

- Synthesise materials (e.g., plastics, 2D materials)
- Optimise industrial processes (e.g., sewage plants)
- Discover new drugs (e.g., opioid antagonist naloxone)

Interest for drug design applications

Chemistry science of matter and its elements; understanding it to

- Synthesise materials (e.g., plastics, 2D materials)
- Optimise industrial processes (e.g., sewage plants)
- Discover new drugs (e.g., opioid antagonist naloxone)



Molecular docking

Interest for drug design applications

Need for predictive models

https://commons.wikimedia.org/wiki/File:Docking_representation_2.png

Chemistry science of matter and its elements; understanding it to

- Synthesise materials (e.g., plastics, 2D materials)
- Optimise industrial processes (e.g., sewage plants)
- Discover new drugs (e.g., opioid antagonist naloxone)



Molecular docking

Interest for drug design applications

- Need for predictive models
- Wide size range of interesting compounds

https://commons.wikimedia.org/wiki/File:Docking_representation_2.png

Chemistry science of matter and its elements; understanding it to

- Synthesise materials (e.g., plastics, 2D materials)
- Optimise industrial processes (e.g., sewage plants)
- Discover new drugs (e.g., opioid antagonist naloxone)



Molecular docking

Interest for drug design applications

- Need for predictive models
- Wide size range of interesting compounds
- \rightarrow Quick increase in complexity

https://commons.wikimedia.org/wiki/File:Docking_representation_2.png

Time

System size



https://commons.wikimedia.org/wiki/File:Aspirin-B-3D-balls.png



NEWTON equation: parameterised through force fields, medium accuracy, only nuclei

• AMBER • CHARMM • AMOEBA

Millions of atoms

https://commons.wikimedia.org/wiki/File:Aspirin-B-3D-balls.png

https://commons.wikimedia.org/wiki/File:Coronavirus._SARS-CoV-2.png



Interaction with an electric field to model solvent

COSMO

https://commons.wikimedia.org/wiki/File:Aspirin-B-3D-balls.png

https://commons.wikimedia.org/wiki/File:Coronavirus._SARS-CoV-2.png

https://commons.wikimedia.org/wiki/File:Water_drop_001.jpg



Focus

Context QM/MM(AMOEBA), multiphysics quantum mechanics/molecular mechanics coupling

https://commons.wikimedia.org/wiki/File:Aspirin-B-3D-balls.png

https://commons.wikimedia.org/wiki/File:Coronavirus._SARS-CoV-2.png

https://commons.wikimedia.org/wiki/File:Water_drop_001.jpg



Focus

Context QM/MM(AMOEBA), multiphysics quantum mechanics/molecular mechanics coupling

I Part of system described with quantum mechanics (e.g., active site)

https://commons.wikimedia.org/wiki/File:Aspirin-B-3D-balls.png

https://commons.wikimedia.org/wiki/File:Coronavirus._SARS-CoV-2.png

https://commons.wikimedia.org/wiki/File:Water_drop_001.jpg



Focus

Context QM/MM(AMOEBA), multiphysics quantum mechanics/molecular mechanics coupling

- | Part of system described with quantum mechanics (e.g., active site)
- II Part of system described with classical mechanics (e.g., solvent)

https://commons.wikimedia.org/wiki/File:Aspirin-B-3D-balls.png

https://commons.wikimedia.org/wiki/File:Coronavirus._SARS-CoV-2.png

https://commons.wikimedia.org/wiki/File:Water_drop_001.jpg

Outline



Outline



Goal Speeding-up QM/MM(AMOEBA) multiscale ab-initio molecular dynamics



Goal Speeding-up QM/MM(AMOEBA) multiscale ab-initio molecular dynamics

Outline

- 0 General introduction to quantum mechanics and classical mechanics
- I Speeding-up quantum mechanics simulations
- II Speeding-up classical mechanics simulations

Not discussed: Polarisable classical molecular dynamics, Multiscale ab-initio molecular dynamics and Extrapolation using the magic points method (chapters II, III and V)

Outline



Goal Speeding-up QM/MM(AMOEBA) multiscale ab-initio molecular dynamics

Outline

- 0 General introduction to quantum mechanics and classical mechanics
- I Speeding-up quantum mechanics simulations
- II Speeding-up classical mechanics simulations

Not discussed: Polarisable classical molecular dynamics, Multiscale ab-initio molecular dynamics and Extrapolation using the magic points method (chapters II, III and V)



Goal Speeding-up QM/MM(AMOEBA) multiscale ab-initio molecular dynamics

Outline

- 0 General introduction to quantum mechanics and classical mechanics
- I Speeding-up quantum mechanics simulations
- II Speeding-up classical mechanics simulations

Not discussed: Polarisable classical molecular dynamics, Multiscale ab-initio molecular dynamics and Extrapolation using the magic points method (chapters II, III and V)

- Molecular system¹ with
 - *M* nuclei of charge *Z_k* at positions **R**_{*k*}
 - N electrons of charge -1 au

¹No particles' spins, BORN-OPPENHEIMER approximation.

- Molecular system¹ with
 - *M* nuclei of charge *Z_k* at positions **R**_{*k*}
 - N electrons of charge -1 au
- Electrons described by antisymmetric² wave functions $\psi \in L^2(\mathbb{R}^{3N}, \mathbb{C})$

 $^{^1 \}rm No$ particles' spins, BORN–OPPENHEIMER approximation.

²PAULI's exclusion principle

- Molecular system¹ with
 - *M* nuclei of charge *Z_k* at positions **R**_{*k*}
 - + N electrons of charge -1 au
- Electrons described by antisymmetric² wave functions $\psi \in L^2(\mathbb{R}^{3N}, \mathbb{C})$
- Ground state of SCHRÖDINGER equation is of interest (system of lowest energy)

¹No particles' spins, BORN-OPPENHEIMER approximation.

²PAULI's exclusion principle

- Molecular system¹ with
 - *M* nuclei of charge *Z_k* at positions **R**_{*k*}
 - N electrons of charge -1 au
- Electrons described by antisymmetric² wave functions $\psi \in L^2(\mathbb{R}^{3N}, \mathbb{C})$
- Ground state of SCHRÖDINGER equation is of interest (system of lowest energy)

SCHRÖDINGER equation (ground state)

 $\inf\{\left<\psi,\mathcal{H}\psi\right> \mid \psi\in\mathsf{L}^2_a(\mathbb{R}^{3N},\mathbb{C}),\ |\psi|=1\}$

¹No particles' spins, BORN-OPPENHEIMER approximation.

²PAULI's exclusion principle

- Molecular system¹ with
 - *M* nuclei of charge *Z_k* at positions **R**_{*k*}
 - N electrons of charge -1 au
- Electrons described by antisymmetric² wave functions $\psi \in L^2(\mathbb{R}^{3N},\mathbb{C})$
- Ground state of SCHRÖDINGER equation is of interest (system of lowest energy)

SCHRÖDINGER equation (ground state)

$$\inf\{\langle \psi, \mathcal{H}\psi \rangle \mid \psi \in \mathsf{L}^2_a(\mathbb{R}^{3N}, \mathbb{C}), \ |\psi| = 1\}$$

$$\mathcal{H} := \underbrace{-\frac{1}{2} \sum_{i=1}^{N} \Delta_{\mathbf{r}_{i}}}_{\text{kinetic energy}} + \underbrace{\sum_{i=1}^{N} \sum_{k=1}^{M} \frac{-Z_{k}}{|\mathbf{r}_{i} - \mathbf{R}_{k}|}}_{\text{interaction n-e}} + \underbrace{\sum_{1 \leq i < j \leq N} \frac{1}{|\mathbf{r}_{i} - \mathbf{r}_{j}|}}_{\text{interaction e-e}}$$

¹No particles' spins, BORN-OPPENHEIMER approximation.

²PAULI's exclusion principle

- Molecular system¹ with
 - *M* nuclei of charge *Z_k* at positions **R**_{*k*}
 - N electrons of charge -1 au
- Electrons described by antisymmetric² wave functions $\psi \in L^2(\mathbb{R}^{3N}, \mathbb{C})$
- Ground state of SCHRÖDINGER equation is of interest (system of lowest energy)

SCHRÖDINGER equation (ground state)

$$\inf\{\langle \psi, \mathcal{H}\psi \rangle \mid \psi \in \mathsf{L}^2_a(\mathbb{R}^{3N}, \mathbb{C}), \ |\psi| = 1\}$$

$$\mathcal{H} := \underbrace{-\frac{1}{2} \sum_{i=1}^{N} \Delta_{\mathbf{r}_{i}}}_{\text{kinetic energy}} + \underbrace{\sum_{i=1}^{N} \sum_{k=1}^{M} \frac{-Z_{k}}{|\mathbf{r}_{i} - \mathbf{R}_{k}|}}_{\text{interaction n-e}} + \underbrace{\sum_{\substack{1 \le i < j \le N}} \frac{1}{|\mathbf{r}_{i} - \mathbf{r}_{j}|}}_{\text{interaction e-e}}$$
$$\mathsf{L}_{a}^{2}(\mathbb{R}^{3N}, \mathbb{C}) := \bigwedge_{1}^{N} \mathsf{L}^{2}(\mathbb{R}^{3}, \mathbb{C})$$

¹No particles' spins, BORN-OPPENHEIMER approximation.

²PAULI's exclusion principle

SCHRÖDINGER equation is great

SCHRÖDINGER equation is great, but ψ defined on $L^2_a(\mathbb{R}^{3N},\mathbb{C})$

Schrödinger equation is great, but ψ defined on $L^2_a(\mathbb{R}^{3N},\mathbb{C})$

 \Rightarrow curse of dimensionality!

SCHRÖDINGER equation is great, but ψ defined on $L^2_a(\mathbb{R}^{3N},\mathbb{C})$

 \Rightarrow curse of dimensionality!

Acetylsalicylic acid C₉H₈O₄ (aspirin)

Contains 21 nuclei and 94 electrons

SCHRÖDINGER equation is great, but ψ defined on $L^2_a(\mathbb{R}^{3N}, \mathbb{C})$

 \Rightarrow curse of dimensionality!

```
Acetylsalicylic acid C<sub>9</sub>H<sub>8</sub>O<sub>4</sub> (aspirin)
```

Contains 21 nuclei and 94 electrons SCHRÖDINGER problem (BORN–OPPENHEIMER) of dimension greater than $2^{3\times94}\approx 10^{85}$

SCHRÖDINGER equation is great, but ψ defined on L²_a(\mathbb{R}^{3N} , \mathbb{C})

 \Rightarrow curse of dimensionality!

```
Acetylsalicylic acid C_9H_8O_4 (aspirin)
```

Contains 21 nuclei and 94 electrons SCHRÖDINGER problem (BORN-OPPENHEIMER) of dimension greater than $2^{3\times94}\approx10^{85}$ Larger than 10^{80} , the EDDINGTON number (protons in observable universe)

```
Schrödinger equation is great, but \psi defined on L^2_a(\mathbb{R}^{3N},\mathbb{C})
```

```
\Rightarrow curse of dimensionality!
```

```
Acetylsalicylic acid C<sub>9</sub>H<sub>8</sub>O<sub>4</sub> (aspirin)
Contains 21 nuclei and 94 electrons
SCHRÖDINGER problem (BORN–OPPENHEIMER) of dimension greater than 2^{3\times94} \approx 10^{85}
Larger than 10^{80}, the EDDINGTON number (protons in observable universe)
```

Need for approximations

Density functional theory Considers electronic density ρ defined on \mathbb{R}^3
Density functional theory Considers electronic density ρ defined on \mathbb{R}^3

Electronic density

$$\rho_{\psi}(\mathbf{r}) := N \int_{\mathbb{R}^{3(N-1)}} \left| \psi(\mathbf{r}, \mathbf{r}_{2}, \dots, \mathbf{r}_{N}) \right|^{2} \mathrm{d}\mathbf{r}_{2} \dots \mathrm{d}\mathbf{r}_{N}$$
with $\psi \in \mathsf{L}^{2}_{*}(\mathbb{R}^{3N}, \mathbb{C})$

Density functional theory

Density functional theory Considers electronic density ρ defined on \mathbb{R}^3

Electronic density

$$\begin{split} \rho_\psi(\mathbf{r}) &\coloneqq N \, \int_{\mathbb{R}^{3(N-1)}} \Bigl| \psi(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N) \Bigr|^2 \, \mathrm{d} \mathbf{r}_2 \dots \, \mathrm{d} \mathbf{r}_N \\ \text{with } \psi \in \mathsf{L}^2_a(\mathbb{R}^{3N}, \mathbb{C}) \end{split}$$



https://commons.wikimedia.org/wiki/File:ElectronDensityAniline.PNG

Equation with density unknown

$$\inf_{\rho} \left\{ \mathcal{E}_{KS}(\rho) := \underbrace{T_{KS}(\rho) + J(\rho) + E_{xc}(\rho)}_{\mathcal{F}(\rho)} + \int \rho V \left| \rho \ge 0, \ \rho^{1/2} \in \mathsf{H}^1(\mathbb{R}^3), \ \int \rho = N \right\}$$

Кони-Sнам functional

$$\inf_{\rho} \left\{ \mathcal{E}_{KS}(\rho) := \underbrace{T_{KS}(\rho) + J(\rho) + E_{xc}(\rho)}_{\mathcal{F}(\rho)} + \int \rho V \middle| \rho \ge 0, \ \rho^{1/2} \in \mathrm{H}^{1}(\mathbb{R}^{3}), \ \int \rho = N \right\}$$

• External potential (interaction with nuclei)

$$\inf_{\rho} \left\{ \mathcal{E}_{KS}(\rho) := \underbrace{\mathcal{T}_{KS}(\rho) + J(\rho) + E_{xc}(\rho)}_{\widehat{\mathcal{F}}(\rho)} + \int \rho V \middle| \rho \ge 0, \ \rho^{1/2} \in \mathsf{H}^1(\mathbb{R}^3), \ \int \rho = N \right\}$$

- External potential (interaction with nuclei)
- Approximation of the kinetics energy

$$\inf_{\rho} \left\{ \mathcal{E}_{KS}(\rho) := \underbrace{T_{KS}(\rho) + J(\rho) + E_{xc}(\rho)}_{\mathcal{F}(\rho)} + \int \rho V \left| \rho \ge 0, \ \rho^{1/2} \in \mathsf{H}^1(\mathbb{R}^3), \ \int \rho = N \right\}$$

- External potential (interaction with nuclei)
- · Approximation of the kinetics energy
- Approximations of the repulsive electronic energy

$$\inf_{\rho} \left\{ \mathcal{E}_{KS}(\rho) := \underbrace{T_{KS}(\rho) + J(\rho) + \underline{E_{xc}(\rho)}}_{\widehat{\mathcal{F}(\rho)}} + \int \rho V \left| \rho \ge 0, \ \rho^{1/2} \in \mathsf{H}^{1}(\mathbb{R}^{3}), \ \int \rho = N \right\}$$

- External potential (interaction with nuclei)
- Approximation of the kinetics energy
- · Approximations of the repulsive electronic energy
- Exchange-correlation energy functional (approximations)

Кони-Sнам functional

$$\inf_{\rho} \left\{ \mathcal{E}_{KS}(\rho) := \underbrace{T_{KS}(\rho) + J(\rho) + E_{xc}(\rho)}_{\mathcal{F}(\rho)} + \int \rho V \left| \rho \ge 0, \ \rho^{1/2} \in \mathsf{H}^1(\mathbb{R}^3), \ \int \rho = N \right\}$$

Rewritten nonlinear eigenproblem¹

Find N lowest eigenvalues and molecular orbitals solving the nonlinear equations

$$\underbrace{\left(-\frac{1}{2}\Delta + v_{KS}[\rho](\mathbf{r})\right)}_{\mathcal{H}_{KS}[\rho]}\phi_{i}(\mathbf{r}) = \varepsilon_{i}\phi_{i}(\mathbf{r}), \quad \text{for } i \in \llbracket 1 \dots N \rrbracket$$

 $\mathrm{in}\;\mathcal{W}_{\!N} \coloneqq \left\{ \Phi \mathrel{\mathop:}= (\phi_1, \ldots, \phi_N) \in \mathsf{L}^2(\mathbb{R}^3, \mathbb{C})^N \, \big| \, \langle \phi_i, \phi_j \rangle = \delta_{ij}, 1 \le i, j \le N \right\}$

¹EULER-LAGRANGE on energy functional

Кони-Sнам functional

$$\inf_{\rho} \left\{ \mathcal{E}_{KS}(\rho) := \underbrace{T_{KS}(\rho) + J(\rho) + E_{xc}(\rho)}_{\mathcal{F}(\rho)} + \int \rho V \left| \rho \ge 0, \ \rho^{1/2} \in \mathsf{H}^1(\mathbb{R}^3), \ \int \rho = N \right\}$$

Rewritten nonlinear eigenproblem¹

Find N lowest eigenvalues and molecular orbitals solving the nonlinear equations

$$\underbrace{\left(-\frac{1}{2}\Delta + v_{KS}[\rho](\mathbf{r})\right)}_{\mathcal{H}_{KS}[\rho]} \phi_i(\mathbf{r}) = \varepsilon_i \phi_i(\mathbf{r}), \quad \text{for } i \in \llbracket 1 \dots N \rrbracket$$

 $\text{ in }\mathcal{W}_{N} := \left\{ \Phi := (\phi_{1}, \dots, \phi_{N}) \in \mathsf{L}^{2}(\mathbb{R}^{3}, \mathbb{C})^{N} \, \big| \, \langle \phi_{i}, \phi_{j} \rangle = \delta_{ij}, 1 \leq i, j \leq N \right\}$

N problems on $\mathsf{L}^2(\mathbb{R}^3,\mathbb{C})$ instead of one problem on $\mathsf{L}^2(\mathbb{R}^{3N},\mathbb{C})$

¹EULER-LAGRANGE on energy functional

Кони-Sнам functional

$$\inf_{\rho} \left\{ \mathcal{E}_{KS}(\rho) := \underbrace{T_{KS}(\rho) + J(\rho) + E_{xc}(\rho)}_{\mathcal{F}(\rho)} + \int \rho V \left| \rho \ge 0, \ \rho^{1/2} \in \mathsf{H}^1(\mathbb{R}^3), \ \int \rho = N \right\}$$

Rewritten nonlinear eigenproblem¹

Find N lowest eigenvalues and molecular orbitals solving the nonlinear equations

$$\underbrace{\left(-\frac{1}{2}\Delta + v_{KS}[\rho](\mathbf{r})\right)}_{\mathcal{H}_{KS}[\rho]}\phi_{i}(\mathbf{r}) = \varepsilon_{i}\phi_{i}(\mathbf{r}), \quad \text{for } i \in \llbracket 1 \dots N \rrbracket$$

 $\mathrm{in}\;\mathcal{W}_{\!N} \coloneqq \left\{ \Phi := (\phi_1, \dots, \phi_N) \in \mathsf{L}^2(\mathbb{R}^3, \mathbb{C})^N \, \big| \, \langle \phi_i, \phi_j \rangle = \delta_{ij}, 1 \le i, j \le N \right\}$

N problems on $\mathsf{L}^2(\mathbb{R}^3,\mathbb{C})$ instead of one problem on $\mathsf{L}^2(\mathbb{R}^{3N},\mathbb{C})$

Electronic density $\rho(\mathbf{r}) = 2 \sum_{i=1}^{N} |\phi_i(\mathbf{r})|^2$

¹EULER-LAGRANGE on energy functional

Solutions as linear combination of atomic orbitals (LCAO)

$$\Phi \in \mathcal{V}^N \Rightarrow \phi_i(\mathbf{r}) := \sum_{\mu=1}^{\mathcal{N}} C_{\mu i} \chi_{\mu}(\mathbf{r}), \quad i \in \llbracket 1 \dots N \rrbracket$$

Solutions as linear combination of atomic orbitals (LCAO)

$$\Phi \in \mathcal{V}^N \Rightarrow \phi_i(\mathbf{r}) \coloneqq \sum_{\mu=1}^N C_{\mu i} \chi_\mu(\mathbf{r}), \quad i \in \llbracket 1 \dots N \rrbracket$$



https://commons.wikimedia.org/wiki/File:Dihydrogen-phase-3D-balls.png

Solutions as linear combination of atomic orbitals (LCAO)

$$\Phi \in \mathcal{V}^N \Rightarrow \phi_i(\mathbf{r}) \coloneqq \sum_{\mu=1}^{\mathcal{N}} C_{\mu i} \chi_{\mu}(\mathbf{r}), \quad i \in \llbracket 1 \dots N \rrbracket$$





https://commons.wikimedia.org/wiki/File:Dihydrogen-phase-3D-balls.png https://commons.wikimedia.org/wiki/File:Nicotine_(HOMO-LUMO).png

Classical molecular dynamics

Human immunodeficiency virus 1 (≈ 100 nm) on T cell ($\approx 5 \,\mu$ m)



https://commons.wikimedia.org/wiki/File:HIV_H9_T-cell.jpg

Human immunodeficiency virus 1 (≈ 100 nm) on T cell ($\approx 5 \mu$ m), capsid (≈ 4 mil. atoms)



https://commons.wikimedia.org/wiki/File:HIV_H9_T-cell.jpg

https://commons.wikimedia.org/wiki/File:Protein_Imager_high_quality_illustration_example_1_HIV_capsid.png

Cannot account for electrons



Benzene C₆H₆

https://commons.wikimedia.org/wiki/File:Elektronendichtedarstellungen_von_Benzol.png

Cannot account for electrons; replaced by notion of (parameterised) chemical bounds





Benzene C₆H₆

https://commons.wikimedia.org/wiki/File:Elektronendichtedarstellungen_von_Benzol.png https://commons.wikimedia.org/wiki/File:Benzene-aromatic-3D-balls.png



https://commons.wikimedia.org/wiki/File:Bond_stretching_energy.svg



https://commons.wikimedia.org/wiki/File:Bond_stretching_energy.svg https://commons.wikimedia.org/wiki/File:Torsionsschwingung_ethylen.png

Molecular orbitals of water (quantum mechanics)



https://commons.wikimedia.org/wiki/File:Molecular_Orbitals_for_Water.png

Molecular orbitals of water (quantum mechanics)



Force fields Choices of representation (intra- and intermolecular terms) along with parameterisations

First order: point charges

https://commons.wikimedia.org/wiki/File:Molecular_Orbitals_for_Water.png

Molecular orbitals of water (quantum mechanics)



Force fields Choices of representation (intra- and intermolecular terms) along with parameterisations

First order: point charges

For anisotropy: static multipoles

https://commons.wikimedia.org/wiki/File:Molecular_Orbitals_for_Water.png

Molecular orbitals of water (quantum mechanics)



Force fields Choices of representation (intra- and intermolecular terms) along with parameterisations

First order: point charges

For anisotropy: static multipoles

Second order: polarisation (changes of properties with the geometry)

https://commons.wikimedia.org/wiki/File:Molecular_Orbitals_for_Water.png

Force fields Choices of representation (intra- and intermolecular terms) along with parameterisations

First order: point charges

For anisotropy: static multipoles

Second order: polarisation (changes of properties with the geometry)

AMOEBA model Polarisable force field with static multipoles up to quadrupoles, implemented in TINKER

Speeding-up quantum mechanics simulations



Context

KOHN-SHAM equations

$$\underbrace{\left(-\frac{1}{2}\Delta + v_{KS}[\rho](\mathbf{r})\right)}_{\mathcal{H}_{KS}[\rho]}\phi_{i}(\mathbf{r}) = \varepsilon_{i}\phi_{i}(\mathbf{r}), \quad \text{for } i \in \llbracket 1 \dots N \rrbracket.$$

Context

KOHN-SHAM equations

$$\underbrace{\left(-\frac{1}{2}\Delta + v_{KS}[\rho](\mathbf{r})\right)}_{\mathcal{H}_{KS}[\rho]}\phi_{i}(\mathbf{r}) = \varepsilon_{i}\phi_{i}(\mathbf{r}), \quad \text{for } i \in \llbracket 1 \dots N \rrbracket.$$

After discretisation, molecular system with

- N electrons
- \mathcal{N} basis elements (atomic orbitals)

Context

KOHN-SHAM equations

$$\underbrace{\left(-\frac{1}{2}\Delta + v_{KS}[\rho](\mathbf{r})\right)}_{\mathcal{H}_{KS}[\rho]}\phi_{i}(\mathbf{r}) = \varepsilon_{i}\phi_{i}(\mathbf{r}), \quad \text{for } i \in \llbracket 1 \dots N \rrbracket.$$

After discretisation, molecular system with

- N electrons
- *N* basis elements (atomic orbitals)

Self-consistent field equations

Find $C \in \mathcal{M}$ and diagonal matrix $E \in \mathbb{R}^{N \times N}$ such that

$$\begin{cases} H_{KS}(D)C = SCE \\ C^{\mathsf{T}}SC = \mathrm{Id}_{N} \\ D = CC^{\mathsf{T}} \end{cases}$$

with $\mathcal{M} := \left\{ C \in \mathbb{R}^{\mathcal{N} \times N} \, \big| \, C^{\mathsf{T}} S C = \mathrm{Id}_N \right\}^a$

 $^{a}S_{\mu\nu}\coloneqq\int_{\mathbb{R}^{3}}\chi_{\mu}\chi_{\nu}$

Context

KOHN-SHAM equations

$$\underbrace{\left(-\frac{1}{2}\Delta + v_{KS}[\rho](\mathbf{r})\right)}_{\mathcal{H}_{KS}[\rho]} \phi_{i}(\mathbf{r}) = \varepsilon_{i}\phi_{i}(\mathbf{r}), \quad \text{for } i \in \llbracket 1 \dots N \rrbracket.$$

After discretisation, molecular system with

- N electrons
- *N* basis elements (atomic orbitals)

Self-consistent field equations

Find $\mathbf{C} \in \mathcal{M}$ and diagonal matrix $E \in \mathbb{R}^{N \times N}$ such that

$$\begin{cases} H_{KS}(D)C = SCE \\ C^{\mathsf{T}}SC = \mathrm{Id}_{N} \\ D = CC^{\mathsf{T}} \end{cases}$$

with $\mathcal{M} := \left\{ C \in \mathbb{R}^{\mathcal{N} \times N} \mid C^{\mathsf{T}}SC = \mathrm{Id}_N \right\}^a$

 $^{a}S_{\mu\nu}\coloneqq\int_{\mathbb{R}^{3}}\chi_{\mu}\chi_{\nu}$

Context

KOHN-SHAM equations

$$\underbrace{\left(-\frac{1}{2}\Delta + v_{KS}[\rho](\mathbf{r})\right)}_{\mathcal{H}_{KS}[\rho]} \phi_{i}(\mathbf{r}) = \frac{\varepsilon_{i}}{\varepsilon_{i}}\phi_{i}(\mathbf{r}), \quad \text{for } i \in \llbracket 1 \dots N \rrbracket.$$

After discretisation, molecular system with

- N electrons
- *N* basis elements (atomic orbitals)

Self-consistent field equations

Find $C \in \mathcal{M}$ and diagonal matrix $\underline{E} \in \mathbb{R}^{N \times N}$ such that

$$\begin{cases} H_{KS}(D)C = SCE \\ C^{\mathsf{T}}SC = \mathrm{Id}_{N} \\ D = CC^{\mathsf{T}} \end{cases}$$

with $\mathcal{M} := \left\{ C \in \mathbb{R}^{\mathcal{N} \times N} \, \big| \, C^{\mathsf{T}} S C = \mathrm{Id}_N \right\}^a$

 $^{a}S_{\mu\nu}\coloneqq\int_{\mathbb{R}^{3}}\chi_{\mu}\chi_{\nu}$

Fixed-point algorithm for the self-consistent field equations

Context

Initial value C_0 , with $C_0^{\mathsf{T}}SC_0 = \mathrm{Id}_N$ and construct a sequence $(C_n)_{n \in \mathbb{N}}$ solving

$$\begin{aligned} &(H_{KS}(D_n)C_{n+1} = SC_{n+1}\mathsf{E}_{n+1} \\ &C_{n+1}^{\mathsf{T}}SC_{n+1} = \mathrm{Id}_N \\ &D_{n+1} = C_{n+1}C_{n+1}^{\mathsf{T}} \end{aligned}$$

Fixed-point algorithm for the self-consistent field equations

Context

Initial value C_0 , with $C_0^{\mathsf{T}}SC_0 = \mathrm{Id}_N$ and construct a sequence $(C_n)_{n \in \mathbb{N}}$ solving

$$\begin{cases} H_{KS}(D_n)C_{n+1} = SC_{n+1}E_{n+1} \\ C_{n+1}^TSC_{n+1} = \mathrm{Id}_N \\ D_{n+1} = C_{n+1}C_{n+1}^T \\ \hline \\ Q_{n+1} = C_{n+1}C_{n+1}^T \\ \hline \\ \\ \\ \\ \hline \\ \hline \\ \\ \hline \\ \\ \hline \\ \hline \\ \hline \\ \hline \\ \\ \hline \\ \hline \\ \\ \hline \hline \\ \hline \\ \hline \\ \hline \\ \hline \\ \hline \hline \\ \hline \hline \\ \hline \hline \\ \hline \\ \hline \\ \hline \\ \hline \hline \\ \hline \\ \hline \\ \hline \\ \hline \hline \\ \hline$$
Fixed-point algorithm for the self-consistent field equations

Context

Initial value C_0 , with $C_0^{\mathsf{T}}SC_0 = \mathrm{Id}_N$ and construct a sequence $(C_n)_{n \in \mathbb{N}}$ solving

$$\begin{cases} H_{KS}(D_n)C_{n+1} = SC_{n+1}E_{n+1} \\ C_{n+1}^TSC_{n+1} = \mathrm{Id}_N \\ D_{n+1} = C_{n+1}C_{n+1}^T \\ \hline D_n = C_nC_n^T \\ \downarrow \\ \hline \\ Assemble H_{KS}(D_n) \\ \downarrow \\ \hline \\ Find \ C_{n+1} \ solving \ eigenproblem \ for \ H_{KS}(D_n) \\ \downarrow \\ \hline \\ \hline \\ If \ converged \ stop, \ otherwise \\ same \ process \ with \ D_{n+1} \\ \end{cases}$$

Algorithm stops when residue $|D_{n+1} - D_n|$ smaller than desired threshold

Fixed-point algorithm for the self-consistent field equations

Context

Initial value C_0 , with $C_0^{\mathsf{T}}SC_0 = \mathrm{Id}_N$ and construct a sequence $(C_n)_{n \in \mathbb{N}}$ solving



Algorithm stops when residue $|D_{n+1} - D_n|$ smaller than desired threshold

To improve (or even *insure*) convergence, more subtle ways to update the density matrices D_n (damping, DIIS, etc.)...

Fixed-point algorithm for the self-consistent field equations

Context

Initial value C_0 , with $C_0^{\mathsf{T}}SC_0 = \mathrm{Id}_N$ and construct a sequence $(C_n)_{n \in \mathbb{N}}$ solving



Algorithm stops when residue $|D_{n+1} - D_n|$ smaller than desired threshold

To improve (or even *insure*) convergence, more subtle ways to update the density matrices D_n (damping, DIIS, etc.)...

... or provide algorithm with a good initial guess C_0 .

• Self-consistent field iterations costly and bottlenecks in density functional theory

Context

- Self-consistent field iterations costly and bottlenecks in density functional theory
- Computations can take months, no small gain

- Self-consistent field iterations costly and bottlenecks in density functional theory
- Computations can take months, no small gain
- With molecular dynamics we expect density matrices to be *close* and want to reuse them from previous steps

Problem statement

- Self-consistent field iterations costly and bottlenecks in density functional theory
- Computations can take months, no small gain
- With molecular dynamics we expect density matrices to be *close* and want to reuse them from previous steps



Simple scheme thanks to LCAO: at time-step t^{k+1} , use previous converged density D_n^k

Context

Problem statement

- Self-consistent field iterations costly and bottlenecks in density functional theory
- Computations can take months, no small gain
- With molecular dynamics we expect density matrices to be *close* and want to reuse them from previous steps



Simple scheme thanks to LCAO: at time-step t^{k+1} , use previous converged density D_n^k

Aim

Fewer self-consistent field iterations on same molecular system using appropriate initial guess

Context

Up to normalisation, density matrices are points on a Grassmannian manifold

$$S^{1/2}DS^{1/2} \in Gr := \left\{ D \in \mathbb{R}^{\mathcal{N} \times N} \mid D = D^{\mathsf{T}}, D^2 = D, \operatorname{Tr}(D) = N \right\}.$$



Context

Up to normalisation, density matrices are points on a Grassmannian manifold

$$S^{1/2}DS^{1/2} \in Gr := \{ D \in \mathbb{R}^{\mathcal{N} \times N} \mid D = D^{\mathsf{T}}, D^2 = D, \operatorname{Tr}(D) = N \}$$



Classical linear combinations extrapolation scheme: objects not on manifold!

$$\sum c_i D_i \notin Gr$$

Context

Up to normalisation, density matrices are points on a Grassmannian manifold

$$S^{1/2}DS^{1/2} \in Gr := \{ D \in \mathbb{R}^{\mathcal{N} \times N} \mid D = D^{\mathsf{T}}, D^2 = D, \operatorname{Tr}(D) = N \}$$



Classical linear combinations extrapolation scheme: objects not on manifold!

$$\sum c_i D_i \notin Gr$$

Let's use the tangent space!

Exponential and logarithmic maps



Local diffeomorphism between manifold $\mathcal{G}r$ and affine space, the tangent space, cheap to compute



Local diffeomorphism between manifold $\mathcal{G}r$ and affine space, the tangent space, cheap to compute

• Exponential function

$$\begin{aligned} & \operatorname{Exp}_{\mathcal{G}r,0} \colon \mathsf{T}_{D_0}\mathcal{G}r \to \mathcal{G}r \\ & \Gamma \mapsto CC^{\mathsf{T}'} \end{aligned}$$

where $C = [C_0 V \cos(\Sigma) + U \sin(\Sigma)]V^T$, with $\Gamma = U \Sigma V^T$ the singular value decomposition of Γ



Local diffeomorphism between manifold $\mathcal{G}r$ and affine space, the tangent space, cheap to compute

• Exponential function

$$\operatorname{Exp}_{Gr,0} \colon \mathsf{T}_{D_0} \mathcal{G}r \to \mathcal{G}r$$
$$\Gamma \mapsto CC^{\mathsf{T}'}$$

where $C = [C_0 V \cos(\Sigma) + U \sin(\Sigma)] V^T$, with $\Gamma = U \Sigma V^T$ the singular value decomposition of Γ

Logarithm function

$$\begin{aligned} \operatorname{Log}_{\mathcal{G}r,0} \colon \mathcal{G}r \to \mathsf{T}_{D_0}\mathcal{G}r \\ D \mapsto \operatorname{Log}_{\mathcal{G}r,0}(D) \end{aligned}$$

defined similarly

• Providing a good initial guess

- Providing a good initial guess
- Useful for large multiphysics molecular systems

- Providing a good initial guess
- Useful for large multiphysics molecular systems
- Negligible overhead

- Providing a good initial guess
- Useful for large multiphysics molecular systems
- Negligible overhead
- Flexible and as simple as possible for end users

Each time-step, nuclei's positions for *free*; we want the corresponding density matrices $D_{\mathbf{R}}$

Molecular descriptors	
$\mathbb{R}^{3M} \to Gr(N, \mathcal{N})$ $\mathbf{R} \mapsto D_{\mathbf{R}}$	

Each time-step, nuclei's positions for *free*; we want the corresponding density matrices $D_{\mathbf{R}}$ We want linear combinations on tangent space, and density matrices as initial guesses

Molecular descriptors

$$\begin{split} &\mathbb{R}^{3M} \to \mathsf{T}_{D_0} \mathcal{G} r \to \mathcal{G} r(N, \mathcal{N}) \\ &\mathbb{R} \mapsto \ \Gamma_{\mathbb{R}} \ \mapsto D_{\mathbb{R}} = \mathrm{Exp}_{\mathcal{G} r, 0}(\Gamma_{\mathbb{R}}) \end{split}$$

$$\mathbf{R} \mapsto \Gamma_{\mathrm{app}}(\mathbf{R}) = \sum_{i=1}^{N_t} c_{\mathbf{R},i} \, \Gamma_{\mathbf{R}(t_i)} \in \mathsf{T}_{D_0} \mathcal{G} r$$

Each time-step, nuclei's positions for *free*; we want the corresponding density matrices $D_{\mathbf{R}}$ We want linear combinations on tangent space, and density matrices as initial guesses Map should take advantages of properties of the density matrices (*e.g.*, translation invariance)

Molecular descriptors $d_{\mathbf{R}}$

$$\begin{split} \mathbb{R}^{3M} &\to \mathcal{M} \to \mathsf{T}_{D_0} \mathcal{G} r \to \mathcal{G} r(N, \mathcal{N}) \\ \mathbb{R} &\mapsto d_{\mathbb{R}} \mapsto \Gamma_{\mathbb{R}} \quad \mapsto D_{\mathbb{R}} = \mathrm{Exp}_{\mathcal{G} r, 0}(\Gamma_{\mathbb{R}}) \end{split}$$

Notion from machine learning

Each time-step, nuclei's positions for *free*; we want the corresponding density matrices $D_{\mathbf{R}}$ We want linear combinations on tangent space, and density matrices as initial guesses Map should take advantages of properties of the density matrices (*e.g.*, translation invariance)

Molecular descriptors $d_{\mathbf{R}}$

$$\begin{split} \mathbb{R}^{3M} &\to \mathcal{M} \to \mathsf{T}_{D_0} \mathcal{G} r \to \mathcal{G} r(N, \mathcal{N}) \\ \mathbb{R} &\mapsto d_{\mathbb{R}} \mapsto \Gamma_{\mathbb{R}} \quad \mapsto D_{\mathbb{R}} = \mathrm{Exp}_{\mathcal{G} r, 0}(\Gamma_{\mathbb{R}}) \end{split}$$

Coulomb matrix at time-step t_k

$$(d_{\mathbf{R}})_{ij} = \begin{cases} 0.5Z_i^{2.4} & \text{if } i = j \\ \\ Z_iZ_j \\ \hline |\mathbf{R}^i(t_k) - \mathbf{R}^j(t_k)| & \text{otherwise} \end{cases}$$

Each time-step, nuclei's positions for *free*; we want the corresponding density matrices $D_{\mathbf{R}}$ We want linear combinations on tangent space, and density matrices as initial guesses Map should take advantages of properties of the density matrices (*e.g.*, translation invariance)

Molecular descriptors $d_{\mathbf{R}}$

$$\begin{split} \mathbb{R}^{3M} &\to \mathcal{M} \to \mathsf{T}_{D_0} \mathcal{G} r \to \mathcal{G} r(N, \mathcal{N}) \\ \mathbb{R} &\mapsto d_{\mathbb{R}} \mapsto \Gamma_{\mathbb{R}} \quad \mapsto D_{\mathbb{R}} = \mathrm{Exp}_{\mathcal{G} r, 0}(\Gamma_{\mathbb{R}}) \end{split}$$

Coulomb matrix at time-step t_k

$$(d_{\mathbf{R}})_{ij} = \begin{cases} 0.5Z_i^{2.4} & \text{if } i = j \\ \\ \frac{Z_iZ_j}{|\mathbf{R}^i(t_k) - \mathbf{R}^j(t_k)|} & \text{otherwise} \end{cases}$$

G-Ext method we developed

- Use of molecular descriptors
- · Linear approximation on tangent space
- Retraction on manifold Gr

We want coefficients $c_{\mathbf{R},i}$ that can approximate the density matrices on the tangent space

$$\mathbf{R} \mapsto \Gamma_{\mathrm{app}}(\mathbf{R}) = \sum_{i=1}^{N_t} c_{\mathbf{R},i} \Gamma_{\mathbf{R}(t_i)} \in \mathsf{T}_{D_0} \mathcal{G} r$$

We want coefficients $c_{\mathbf{R},i}$ that can approximate the density matrices on the tangent space

$$\mathbf{R} \mapsto \Gamma_{\mathrm{app}}(\mathbf{R}) = \sum_{i=1}^{N_t} c_{\mathbf{R},i} \Gamma_{\mathbf{R}(t_i)} \in \mathsf{T}_{D_0} \mathcal{G} r$$

Least-squares method to solve

$$\min_{c_{\mathbf{R}} \in \mathbb{R}^{N_t}} \left| d_{\mathbf{R}} - \sum_{i=1}^{N_t} c_{\mathbf{R},i} d_{\mathbf{R}(t_i)} \right|^2$$

We want coefficients $c_{\mathbf{R},i}$ that can approximate the density matrices on the tangent space

$$\mathbf{R} \mapsto \Gamma_{\mathrm{app}}(\mathbf{R}) = \sum_{i=1}^{N_t} c_{\mathbf{R},i} \Gamma_{\mathbf{R}(t_i)} \in \mathsf{T}_{D_0} \mathcal{G} r$$

Least-squares method to solve

$$\min_{C_{\mathbf{R}}\in\mathbb{R}^{N_t}} \left| d_{\mathbf{R}} - \sum_{i=1}^{N_t} c_{\mathbf{R},i} d_{\mathbf{R}(t_i)} \right|^2$$

We use the same coefficients for density matrices on the tangent space and use the retracted density matrix

$$D_{\mathrm{app}}(\mathbf{R}) = \mathrm{Exp}_{\mathcal{G}r,0} \left(\sum_{i=1}^{N_t} \boldsymbol{c}_{\mathbf{R},i} \, \boldsymbol{\Gamma}_{\mathbf{R}(t_i)} \right)$$

as an initial guess to the self-consistent field algorithm

Molecular system	Quantum atoms	Classical atoms	\mathcal{N}
OCP	129	4915	1038
APPA	31	16 449	309
DMABN	21	6843	185
3HF	28	15018	290



OCP: orange carotenoid protein

Molecular system	Quantum atoms	Classical atoms	\mathcal{N}
OCP	129	4915	1038
APPA	31	16 449	309
DMABN	21	6843	185
3HF	28	15018	290



OCP: orange carotenoid protein

• 10⁻⁵ root-mean-square • 1 ps dynamics • 0.5 fs time-step • ref. method

	OCP		DMABN		APPA		3HF	
Method	Average	σ	Average	σ	Average	σ	Average	σ
XLBO	3.82	0.66	3.98	0.16	3.00	0.03	4.00	0.14
XLBO/MW	2.95	0.31	3.76	0.56	3.00	0.34	3.96	0.31
G-Ext(3)	2.57	0.84	3.54	0.78	2.95	0.50	3.09	0.41
G-Ext(4)	2.48	0.88	3.14	0.62	2.51	0.50	3.25	0.68
G-Ext(5)	2.25	0.96	3.23	0.75	2.51	0.50	3.30	0.72
G-Ext(6)	2.20	0.96	2.99	0.02	2.51	0.50	3.14	0.56

Molecular system	Quantum atoms	Classical atoms	\mathcal{N}
OCP	129	4915	1038
APPA	31	16 449	309
DMABN	21	6843	185
3HF	28	15018	290



OCP: orange carotenoid protein

• 10⁻⁵ root-mean-square • 1 ps dynamics • 0.5 fs time-step

• ref. method

	OCP		DMABN		APPA		3HF	
Method	Average	σ	Average	σ	Average	σ	Average	σ
XLBO	3.82	0.66	3.98	0.16	3.00	0.03	4.00	0.14
XLBO/MW	2.95	0.31	3.76	0.56	3.00	0.34	3.96	0.31
G-Ext(3)	2.57	0.84	3.54	0.78	2.95	0.50	3.09	0.41
G-Ext(4)	2.48	0.88	3.14	0.62	2.51	0.50	3.25	0.68
G-Ext(5)	2.25	0.96	3.23	0.75	2.51	0.50	3.30	0.72
G-Ext(6)	2.20	0.96	2.99	0.02	2.51	0.50	3.14	0.56

Decrease in self-consistent field iterations

Molecular system	Quantum atoms	Classical atoms	\mathcal{N}
OCP	129	4915	1038
APPA	31	16 449	309
DMABN	21	6843	185
3HF	28	15 018	290



OCP: orange carotenoid protein

• 10⁻⁵ root-mean-square • 1 ps dynamics

• 0.5 fs time-step

• ref. method

	OCF	OCP DMABN		N	APP	4	3HF	
Method	Average	σ	Average	σ	Average	σ	Average	σ
XLBO	3.82	0.66	3.98	0.16	3.00	0.03	4.00	0.14
XLBO/MW	2.95	0.31	3.76	0.56	3.00	0.34	3.96	0.31
G-Ext(3)	2.57	0.84	3.54	0.78	2.95	0.50	3.09	0.41
G-Ext(4)	2.48	0.88	3.14	0.62	2.51	0.50	3.25	0.68
G-Ext(5)	2.25	0.96	3.23	0.75	2.51	0.50	3.30	0.72
G-Ext(6)	2.20	0.96	2.99	0.02	2.51	0.50	3.14	0.56

Decrease in self-consistent field iterations; adds up with long molecular dynamics!



Lower number of self-consistent field iterations

Lower number of self-consistent field iterations

Future work

• Time reversibility

Lower number of self-consistent field iterations

Future work

- Time reversibility
- Geometry optimisation

Lower number of self-consistent field iterations

Future work

- Time reversibility
- Geometry optimisation
- Bypass self-consistent field algorithm (*e.g.*, empirical interpolation method for Fock matrices, minimisation on tangent space)

Speeding-up classical mechanics simulations






Molecular system with N nuclei of point charge \mathbf{M}_{0i} at positions \mathbf{r}_i

N-body problem of computing the electrostatic energy

$$\mathcal{E}_{\text{elec}} \coloneqq \frac{1}{2} \sum_{1 \leqslant i \neq j \leqslant N} \frac{\mathbf{M}_{0i} \mathbf{M}_{0j}}{|\mathbf{r}_j - \mathbf{r}_i|}.$$

Bottleneck, scales as $\mathscr{O}(N^2)$

Molecular system with N nuclei of point charge M_{0i} at positions \mathbf{r}_i

N-body problem of computing the electrostatic energy

$$\mathcal{E}_{\text{elec}} \coloneqq \frac{1}{2} \sum_{1 \leqslant i \neq j \leqslant N} \frac{\mathbf{M}_{0i} \mathbf{M}_{0j}}{|\mathbf{r}_j - \mathbf{r}_i|}.$$

Bottleneck, scales as $\mathscr{O}(N^2)$

For AMOEBA force field, generalised to (static and non-static) multipoles \mathbf{M}_n with derivation operators \mathbf{D}_i^n

$$\mathbf{\hat{L}}_i := \mathbf{M}_{0i} + \mathbf{M}_{1i} \cdot \mathbf{D}_i + \mathbf{M}_{2i} \cdot \mathbf{D}_i^2$$

Molecular system with N nuclei of point charge M_{0i} at positions \mathbf{r}_i

N-body problem of computing the electrostatic energy

$$\mathcal{E}_{\text{elec}} \coloneqq \frac{1}{2} \sum_{1 \leqslant i \neq j \leqslant N} \frac{\mathbf{M}_{0i} \mathbf{M}_{0j}}{|\mathbf{r}_j - \mathbf{r}_i|}.$$

Bottleneck, scales as $\mathcal{O}(N^2)$

For AMOEBA force field, generalised to (static and non-static) multipoles \mathbf{M}_n with derivation operators \mathbf{D}_i^n

$$\mathbf{\hat{L}}_i := \mathbf{M}_{0i} + \mathbf{M}_{1i} \cdot \mathbf{D}_i + \mathbf{M}_{2i} \cdot \mathbf{D}_i^2$$

So

$$\mathcal{E}_{\text{elec}} \coloneqq \frac{1}{2} \sum_{1 \leq i \neq j \leq N} \hat{\mathbf{L}}_i \hat{\mathbf{L}}_j \left(\frac{1}{|\mathbf{r}_j - \mathbf{r}_i|} \right).$$

Molecular system with N nuclei of point charge M_{0i} at positions \mathbf{r}_i

N-body problem of computing the electrostatic energy

$$\mathcal{E}_{\text{elec}} \coloneqq \frac{1}{2} \sum_{1 \leqslant i \neq j \leqslant N} \frac{\mathbf{M}_{0i} \mathbf{M}_{0j}}{|\mathbf{r}_j - \mathbf{r}_i|}.$$

Bottleneck, scales as $\mathcal{O}(N^2)$

For AMOEBA force field, generalised to (static and non-static) multipoles \mathbf{M}_n with derivation operators \mathbf{D}_i^n

$$\mathbf{\hat{L}}_i := \mathbf{M}_{0i} + \mathbf{M}_{1i} \cdot \mathbf{D}_i + \mathbf{M}_{2i} \cdot \mathbf{D}_i^2$$

So

$$\mathcal{E}_{\text{elec}} \coloneqq \frac{1}{2} \sum_{1 \leq i \neq j \leq N} \hat{\mathbf{L}}_i \hat{\mathbf{L}}_j \left(\frac{1}{|\mathbf{r}_j - \mathbf{r}_i|} \right).$$

Cuttoff methods not accurate for molecular dynamics because *slow* decrease of the inverse function

Smooth particle mesh EWALD

- Reference method
 Periodic problem
 Smooth function splitting the potential

Smooth particle mesh EWALD

Reference method
 Periodic problem
 Smooth function splitting the potential

With fast FOURIER transforms: $\mathscr{O}(N \log N)$ instead of $\mathscr{O}(N^2)$

Smooth particle mesh EWALD

Reference method
 Periodic problem
 Smooth function splitting the potential

With fast FOURIER transforms: $\mathscr{O}(N \log N)$ instead of $\mathscr{O}(N^2)$



> 64 million atoms, months of computations on thousands of GPUS

Juan R. Perilla and Klaus Schulten. "Physical Properties of the HIV-1 Capsid from All-Atom Molecular Dynamics Simulations". In: Nature Communications 8 (July 19, 2017), p. 15959

· Linear computational complexity hierarchical method

- Linear computational complexity hierarchical method
- Common in physics

- Linear computational complexity hierarchical method
- Common in physics; less so in chemistry

- · Linear computational complexity hierarchical method
- Common in physics; less so in chemistry
- E.g., ScalFMM solves N-body problem for point charges

Modification of ScalFMM¹ and TINKER for AMOEBA

¹P2M, L2P and P2P operators

Modification of ScalFMM¹ and TINKER for AMOEBA

Terms	Potential	Energy	Forces
Charges	•	٠	٠
Dipoles	1	1	1
Quadrupoles	1	1	, /
General multipoles	1	×	×

present ✓ implemented ✗ not implemented

¹P2M, L2P and P2P operators

Results — Linear scaling

- Single time-step 10⁻⁵ for polarisation 216 cores
 - Total execution time of TINKER(AMOEBA)-ScalFMM
 - Conservative parameters
 • Relaxed parameters



Results — Linear scaling

- Single time-step 10⁻⁵ for polarisation 216 cores
 - Total execution time of TINKER(AMOEBA)-ScalFMM
 - Conservative parameters
 • Relaxed parameters



æ

Results — Linear scaling

- Single time-step 10⁻⁵ for polarisation 216 cores
 - Total execution time of TINKER(AMOEBA)-ScalFMM
 - Conservative parameters
 • Relaxed parameters



- No polarisation Triple system sizes along cores Relaxed parameters
 - Total execution time of TINKER(AMOEBA)-ScalFMM



- No polarisation Triple system sizes along cores Relaxed parameters
 - Total execution time of TINKER(AMOEBA)-ScalFMM



Q.

- No polarisation Triple system sizes along cores Relaxed parameters
 - Total execution time of TINKER(AMOEBA)-ScalFMM



Results — Strong scalability

Single time-step
 Relaxed set of parameters

Left 96 000 atoms, double loop with TINKER

Right Scalability starts at four cores



Results — Strong scalability

Single time-step
 Relaxed set of parameters

Left 96 000 atoms, double loop with TINKER

Right Scalability starts at four cores



Results — Strong scalability

Single time-step
 Relaxed set of parameters

Left 96 000 atoms, double loop with TINKER

Right Scalability starts at four cores



(but not a deal-breaker)

Results — Energy conservation of non-solvated DHFR

- DHFR system (2489 atoms)
- 0.10 fs time-steps
- Conservative parameters
- Relaxed parameters

Point charges

Static multipoles up to quadrupoles



Results — Energy conservation of non-solvated DHFR

- DHFR system (2489 atoms)
- 0.10 fs time-steps
- Conservative parameters
- Relaxed parameters

Point charges

Static multipoles up to quadrupoles



😛 Energy conserved

Results — Energy conservation of solvated DHFR

- Solvated DHFR system (23 558 atoms)
 0.25 fs time-steps
 - 10^{-8} for polarisation 10^{-5} for polarisation

Conservative parameters

Relaxed parameters



Results — Energy conservation of solvated DHFR

- Solvated DHFR system (23 558 atoms)
 0.25 fs time-steps
 - 10^{-8} for polarisation 10^{-5} for polarisation

Conservative parameters

Relaxed parameters



Results — Energy conservation of solvated DHFR

- Solvated DHFR system (23 558 atoms)
 0.25 fs time-steps
 - 10^{-8} for polarisation 10^{-5} for polarisation

Conservative parameters

Relaxed parameters



🔁 Energy drift (needs investigation)

• Linear scaling for non-periodic systems

- Linear scaling for non-periodic systems
- Straightforward and high-performance implementation

- Linear scaling for non-periodic systems
- Straightforward and high-performance implementation

Future work

• Adaptive grid

- Linear scaling for non-periodic systems
- Straightforward and high-performance implementation

Future work

- Adaptive grid
- Periodic case for systematic comparisons to smooth particle mesh EWALD

- Linear scaling for non-periodic systems
- Straightforward and high-performance implementation

Future work

- Adaptive grid
- Periodic case for systematic comparisons to smooth particle mesh Ewald
- Real-case tests





Major results

Decrease in time for densities of quantum atoms



Major results

Decrease in time for densities of quantum atoms and model classical atoms, solvent in particular, with linear complexity on supercomputers



Major results

Decrease in time for densities of quantum atoms and model classical atoms, solvent in particular, with linear complexity on supercomputers

 \Rightarrow Allows for longer multiphysics QM/MM(AMOEBA) simulations, bigger systems or more solvent for more accurate physics
Thank you for your attention!

Publications

- Lagardère, Louis et al. "Scalable Evaluation of Polarization Energy and Associated Forces in Polarizable Molecular Dynamics: II. Toward Massively Parallel Computations Using Smooth Particle Mesh Ewald". In: *Journal of Chemical Theory and Computation* 11.6 (June 9, 2015), pp. 2589–2599.
- Lindgren, Eric B. et al. "An Integral Equation Approach to Calculate Electrostatic Interactions in Many-Body Dielectric Systems". In: Journal of Computational Physics 371 (Oct. 15, 2018), pp. 712–731.
- Loco, Daniele et al. "A QM/MM Approach Using the AMOEBA Polarizable Embedding: From Ground State Energies to Electronic Excitations". In: *Journal of Chemical Theory and Computation* 12.8 (Aug. 9, 2016), pp. 3654–3661.
- Narth, Christophe et al. "Scalable Improvement of SPME Multipolar Electrostatics in Anisotropic Polarizable Molecular Mechanics Using a General Short-Range Penetration Correction up to Quadrupoles". In: *Journal of Computational Chemistry* 37.5 (2016), pp. 494–506.
- Polack, Étienne, Yvon Maday, and Andreas Savin. FLEIM: A Stable, Accurate and Robust Extrapolation Method at Infinity for Computing the Ground State of Electronic Hamiltonians. Dec. 24, 2021. URL: http://arxiv.org/abs/2112.13139.
- Polack, Étienne et al. "An Approximation Strategy to Compute Accurate Initial Density Matrices for Repeated Self-Consistent Field Calculations at Different Geometries". In: *Molecular Physics* 118.19-20 (Oct. 17, 2020), e1779834.
- Polack, Étienne et al. "Grassmann Extrapolation of Density Matrices for Born–Oppenheimer Molecular Dynamics". In: *Journal of Chemical Theory and Computation* 17.11 (Nov. 9, 2021), pp. 6965–6973.