Nonadiabatic molecular dynamics simulations with Libra

Mohammad Shakiba

SUNY Buffalo

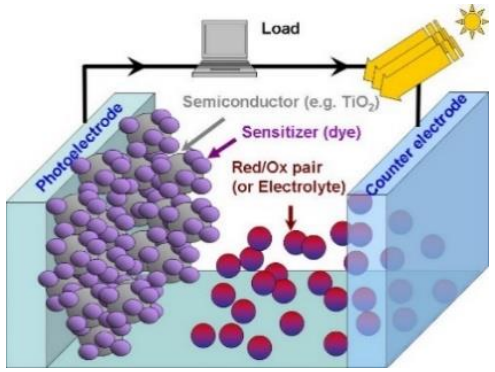
Akimov Research Group

Libra summer workshop 2024

University at Buffalo The State University of New York

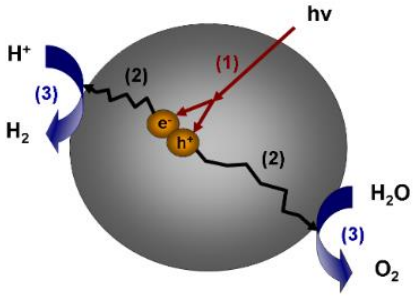
Motivation

- **Solar energy materials**

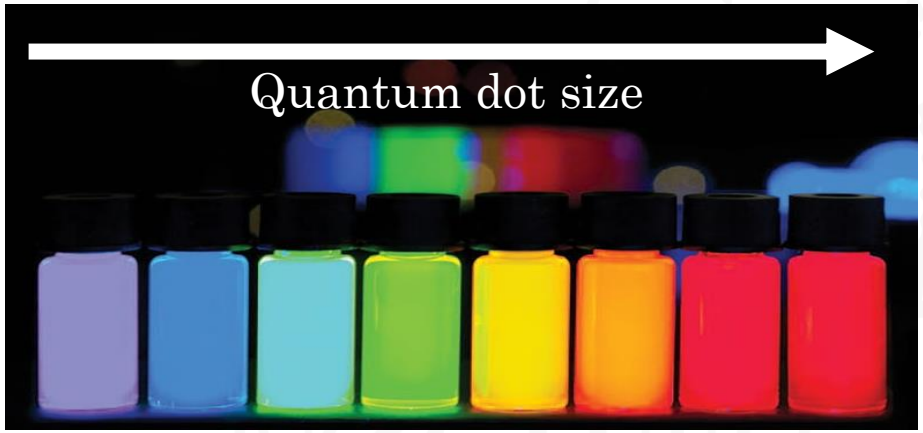


- **Photocatalysis**

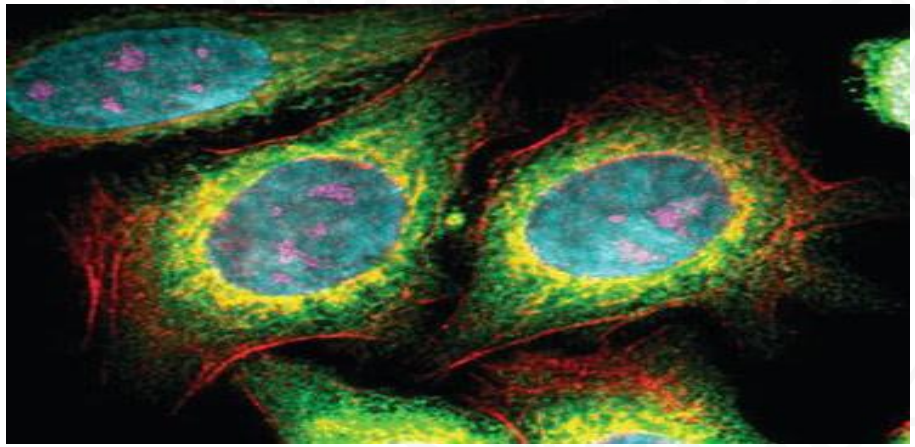
Artificial leaf



- **Light-emitting diodes**



- **Bio-imaging**



Nonadiabatic dynamics

- The total wavefunction:

$$\Psi(\mathrm{r},\mathrm{R(t)}) = \sum_i c_i(t)\psi_i(r,R(t))$$

$$H_{el}(r,R)\psi_i(r,R) = E_i\psi_i$$

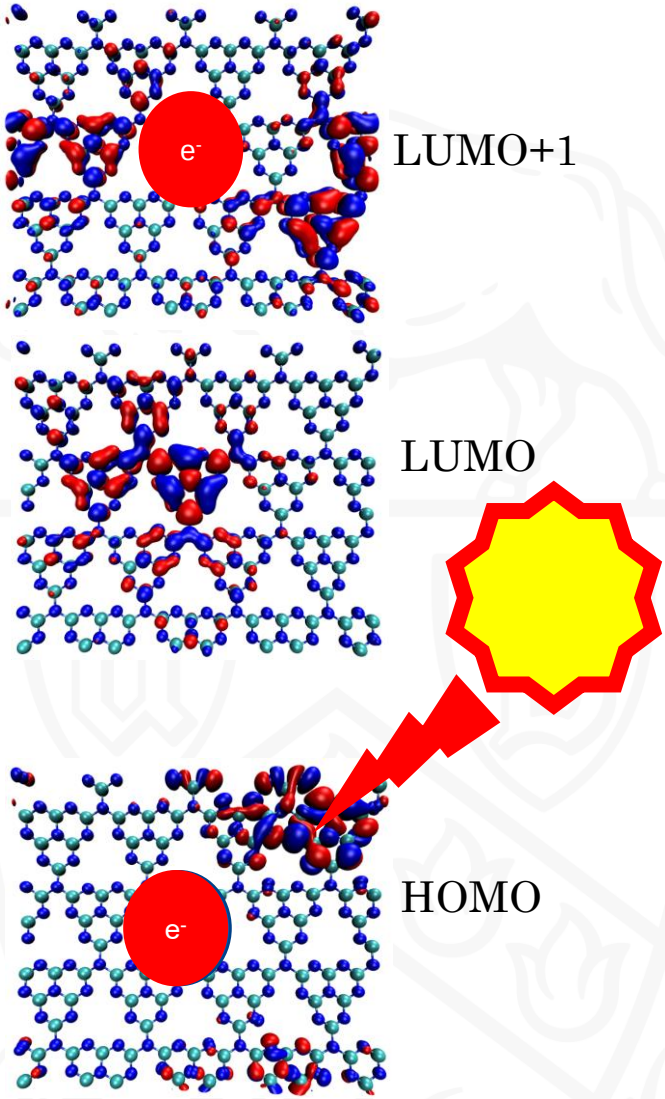
- The evolution of the electronic amplitudes:

$$i\hbar\frac{\partial c_i(t)}{\partial t} = \sum_j \left(E_i(t)\delta_{i,j} - i\hbar d_{ij}(t)\right)c_j(t)$$

$$d_{ij}(t) = \langle\psi_i\left|\frac{\partial}{\partial t}\psi_j\right\rangle$$

- Hammes-Schiffer and Tully method:

$$d_{ij}\left(t+\frac{\Delta t}{2}\right) \approx \frac{\langle\psi_i(t)|\psi_j(t+\Delta t)\rangle - \langle\psi_i(t+\Delta t)|\psi_j(t)\rangle}{2\Delta t}$$



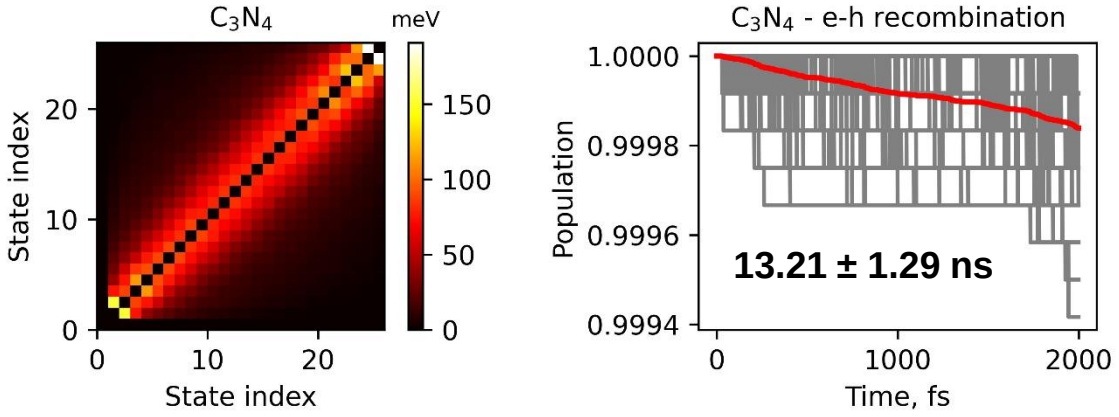
Tully, J. Chem. Phys. 1990, 93, 1061
Hammes-Schiffer, Tully, J. Chem. Phys. 1994, 101, 4657–4667

Trajectory surface hopping

- A stochastic method used for evolution of the simplified time-dependent Schrodinger equation
- A trajectory can stochastically hop from state i to state j with the hopping probability:

$$P_{i\to j}(t, t+\Delta t) = \max\left(0, \frac{\Delta t}{c_i c_i^*} \mathrm{Im}\left(c_i c_j^* d_{ji} - c_j c_i^* d_{ij}\right)\right)$$

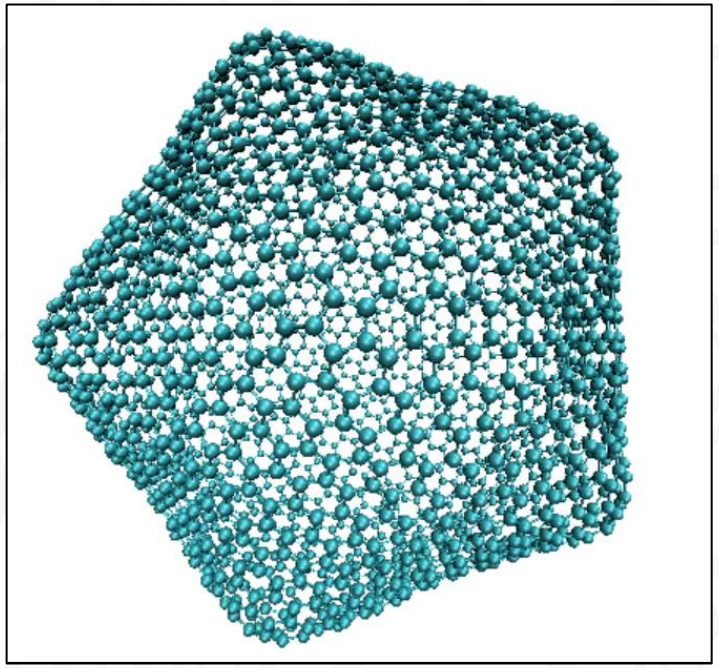
- Nonradiative electron-hole recombination, hot-electron, and hot-hole relaxation.



Tully, J. Chem. Phys. 1990, 93, 1061–1071.
Shakiba, Stippell, Li, Akimov, J. Chem. Theory Compu. 2022, 18, 5157-5180

Challenges

- NA-MD simulations are limited to small-medium sized structures
- High computational cost of the simulation of excited states dynamics in nanoscale systems
- Different approximations are made
  - Pure functionals, tight-binding family of approaches, etc
  - Single-particle description of excited states
  - Neglect of back reaction approximation
  - etc



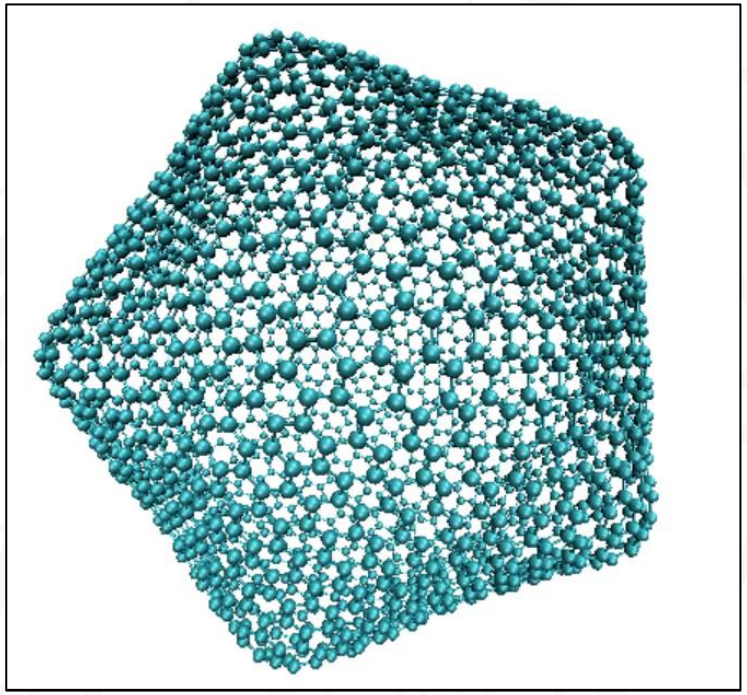
Goal

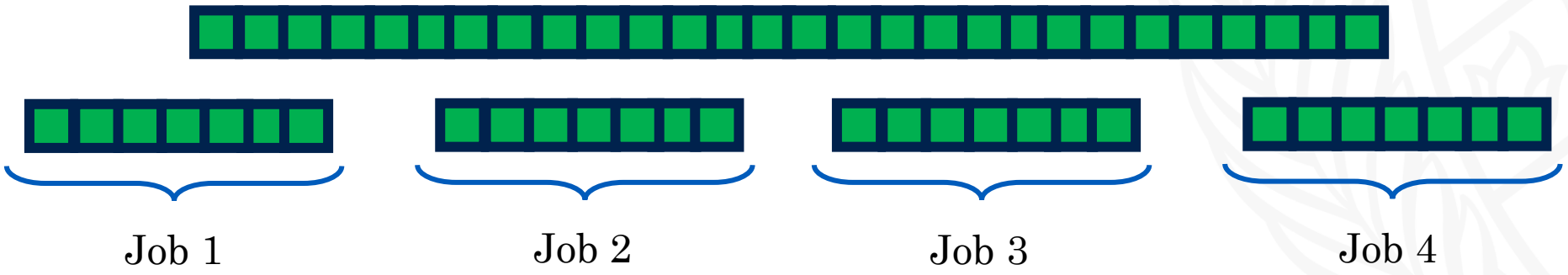
- To perform NA-MD calculations for large nanomaterials and periodic solids
- Can we speed up the calculations using machine-learning techniques?

Part I: Nonadiabatic molecular dynamics in large scale nanomaterials and periodic solids

Nonadiabatic dynamics in nanoscale systems

- Step 1:
  - Generate Pre-computed molecular dynamics trajectory: DFT, DFTB, xTB, or force fields for very large systems e.g. proteins
- Step 2:
  - Generate Kohn-Sham molecular orbitals, their overlap, time-overlap, nonadiabatic couplings
  - TD-DFT states and their configuration
- Step 3:
  - Build and computing the excited states properties
- Step 4:
  - Doing the dynamics with trajectory surface hopping

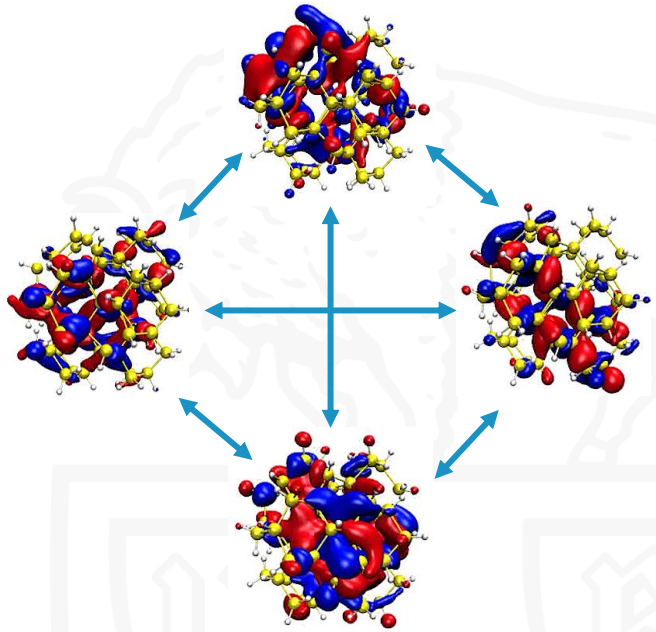




Time-overlap integration

- Grid-based approach using *.cube* files
  - Easy to implement
  - Most codes can output these file
  - Not suitable for large structures with large number of states
- Double-molecule approach
  - Easy to use and can be used in different codes
  - Very time-consuming for large structures
  - Not suitable for periodic structures
- Analytical approach
  - Suitable for large systems and large number of states
  - Recurrence relations for computing integrals: Libint2 package

$$S_{ij} = \int_{-\infty}^{\infty} \psi_i^* \psi_j dv$$

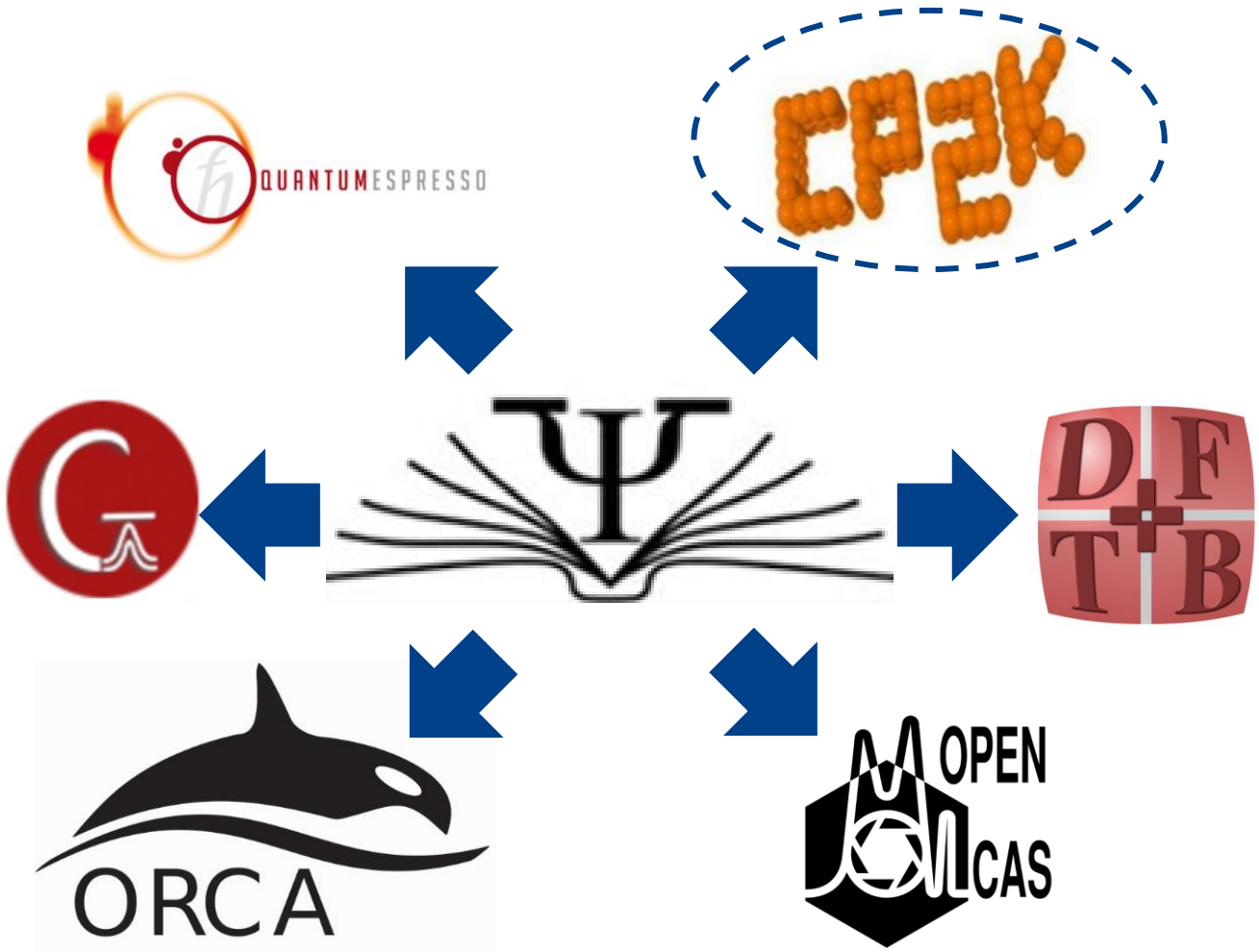


$$\boldsymbol{S^{MO}} = \boldsymbol{c^T S^{AO} c}$$

$$\boldsymbol{S^{AO}} = \begin{bmatrix} \langle\psi_1|\psi_1\rangle & \cdots & \langle\psi_1|\psi_n\rangle \\ \vdots & \ddots & \vdots \\ \langle\psi_n|\psi_1\rangle & \cdots & \langle\psi_n|\psi_n\rangle \end{bmatrix}$$

Smith, Shakiba, Akimov, J. Chem. Theory Comp. 2021, 17, 678–693
Libint, Version 2.6.0 Edward F. Valeev, http://libint.valeyev.net.
Shakiba, Stippell, Li, Akimov, J. Chem. Theory Comp. 2022, 18, 5157-5180

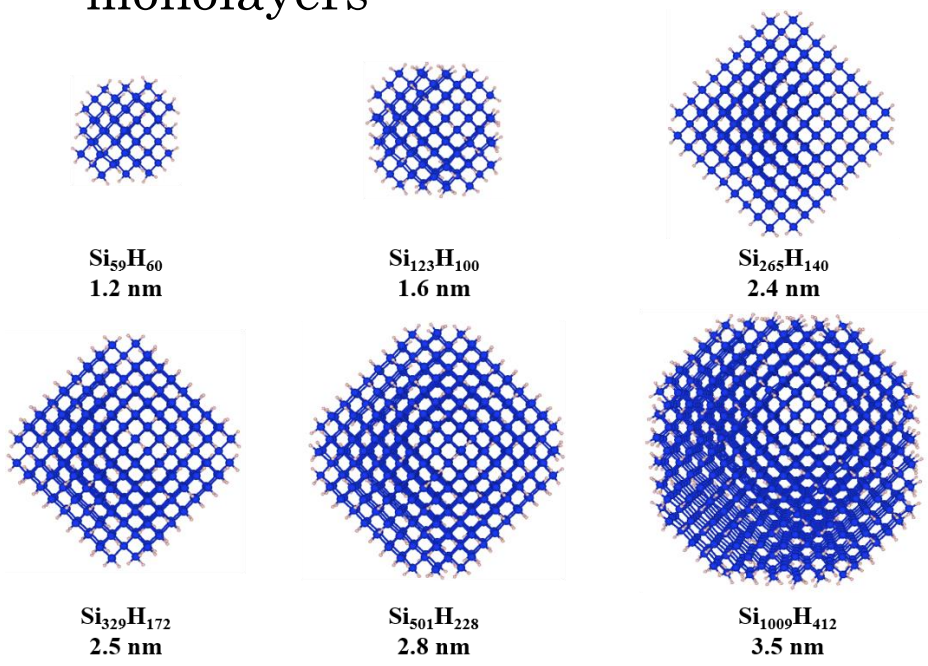
Implementation

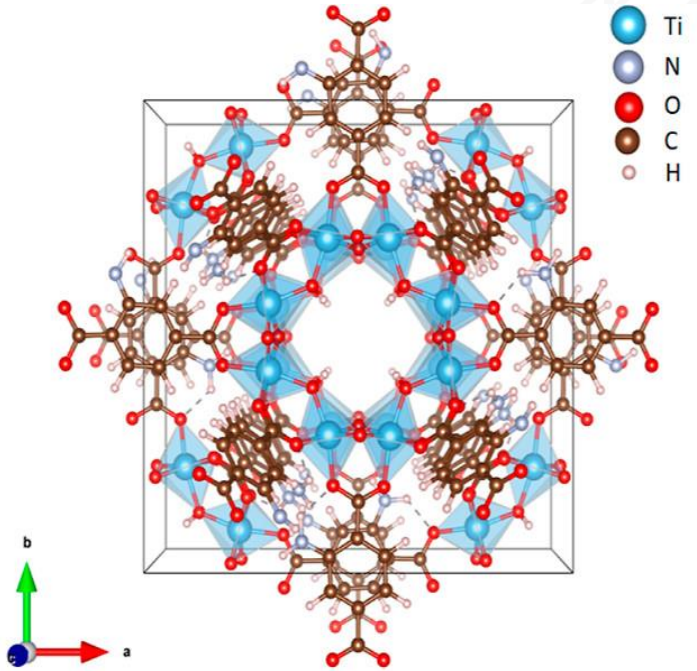


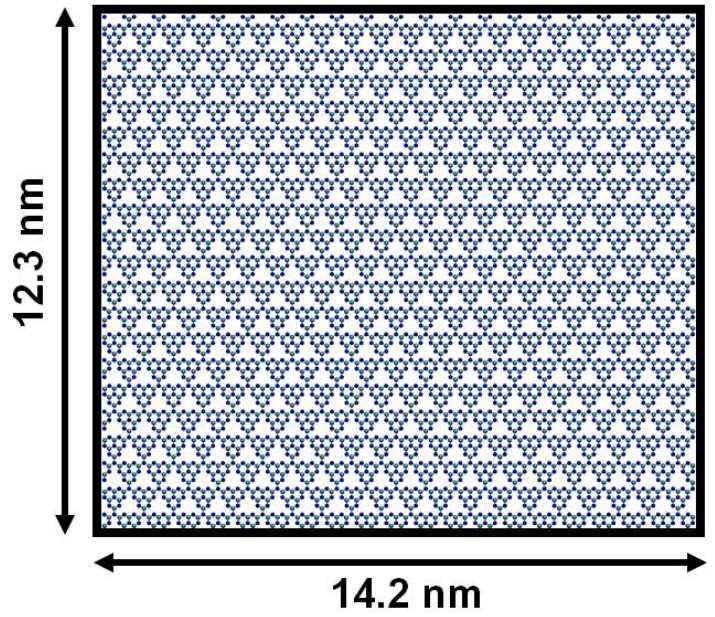

```python
import os
from libra_py import CP2K_methods
from libra_py.workflows.nbra import step2

# Setup the parameters
params = {'istep': 1, 'fstep': 2000, 'nprocs': 9, 'mpi_executable': 'srun',
          'lowest_orbital': 512-20, 'highest_orbital': 512+21,
          'res_dir': os.getcwd() + '/results',  'isUKS': False,
          'is_periodic': True, 'periodicity_type': 'XY'
          'is_spherical': True, 'isxTB': True, 'remove_molden': True,
          'cp2k_exe': 'cp2k.psmp', 'cp2k_ot_input_template': 'es_ot.inp',
          'cp2k_diag_input_template': 'es_diag.inp',
          'trajectory_xyz_filename': 'C3N4-2x2-pos.xyz',
          'cube_visualization': True, 'vmd_input_template': 'vmd.tcl',
          'vmd_exe': 'vmd', 'states_to_plot': [512,513],
          'plot_phase_corrected': True, 'remove_cube': True}

if params['is_periodic']:
    params['A_cell_vector'] = [28.483, 0.000, 0.000]
    params['B_cell_vector'] = [0.000, 24.669, 0.000]
    params['C_cell_vector'] = [0.000, 0.000, 15.000]
    # Set the origin and generate the translational vectors
    origin = [0,0,0]
    # Number of periodic images per each X, -X, Y, -Y, Z, and -Z directions
    num_periodic_images = [1,1,1]
    params['translational_vectors'] =
    CP2K_methods.generate_translational_vectors(origin, num_periodic_images,
                                                params['periodicity_type'])

# Run the calculations
step2.run_cp2k_libint_step2(params)
```

Applications

- Hot-electron cooling dynamics in silicon nanocrystals (Si NCs)
- Electron-hole recombination dynamics in metal organic frameworks and carbon nitride monolayers







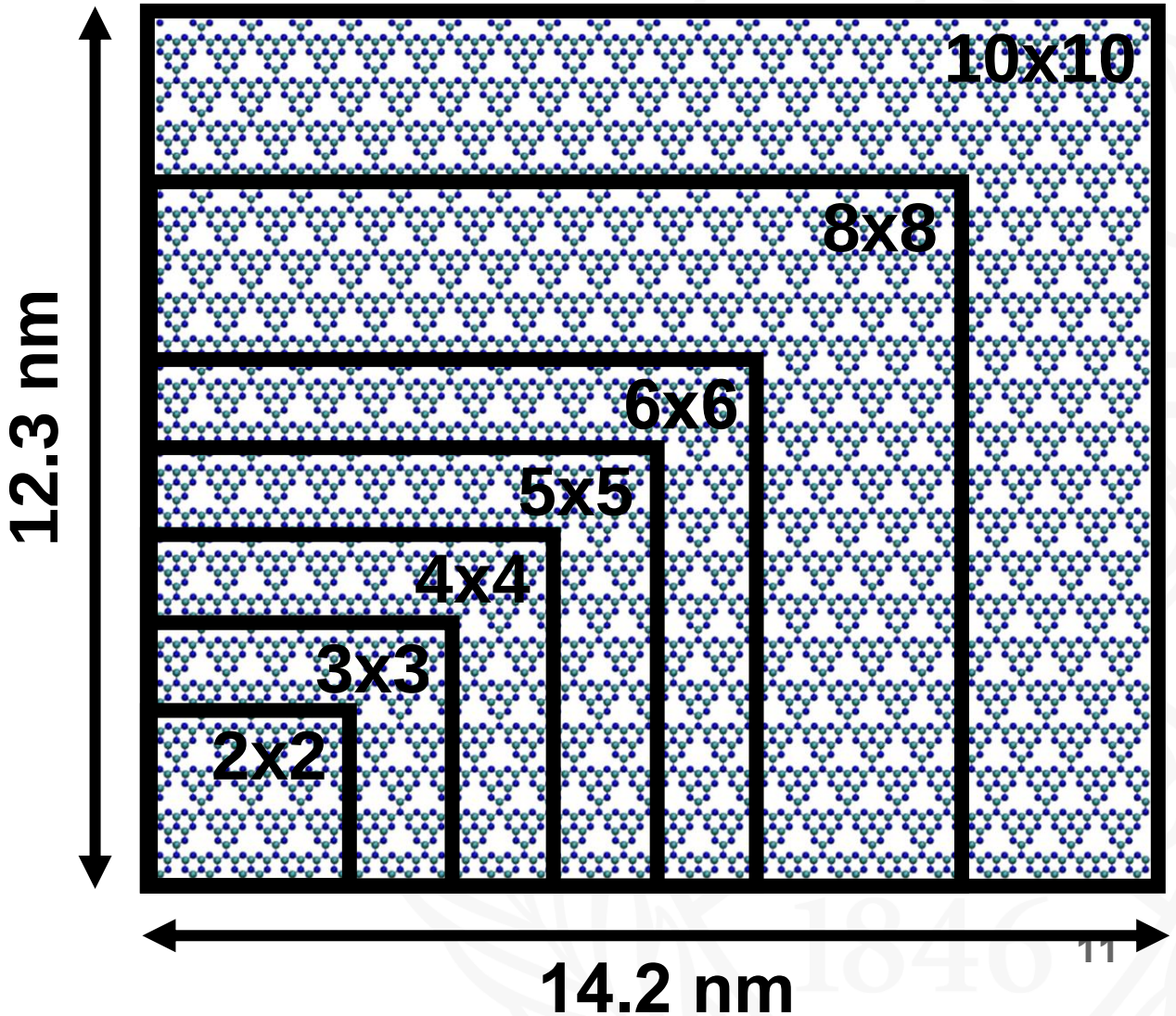
Bannwarth, Ehlert, Grimme, J. Chem. Theory Comput. 2019, 15, 1652–1671
Shakiba, Stippell, Li, Akimov, J. Chem. Theory Compu. 2022, 18, 5157-5180

Charge carrier concentration in C_3N_4 monolayers

- Intrinsic charge carrier concentration in monolayers are in the range of ~10^{10}-10^{12} cm^{-2} in experimental studies
- Theoretical studies overestimate this value due to simulation of excited states in small cell size

224 # Atoms 5600

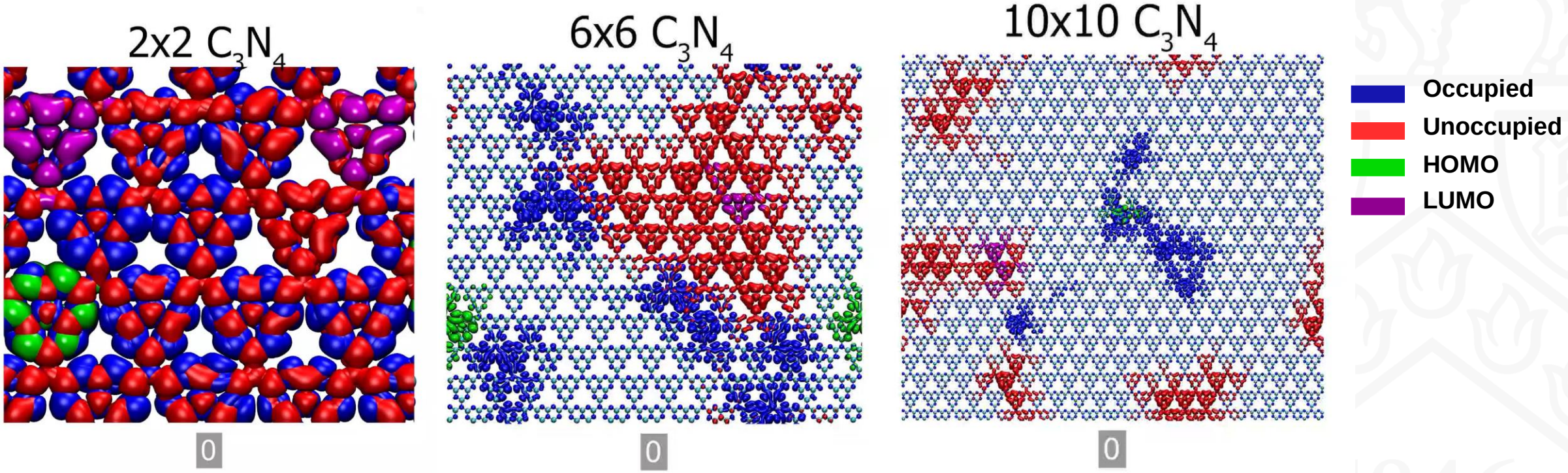
512 # Orbitals 12800



Shakiba, Akimov, JPCC 2023, 127, 9083-9096.

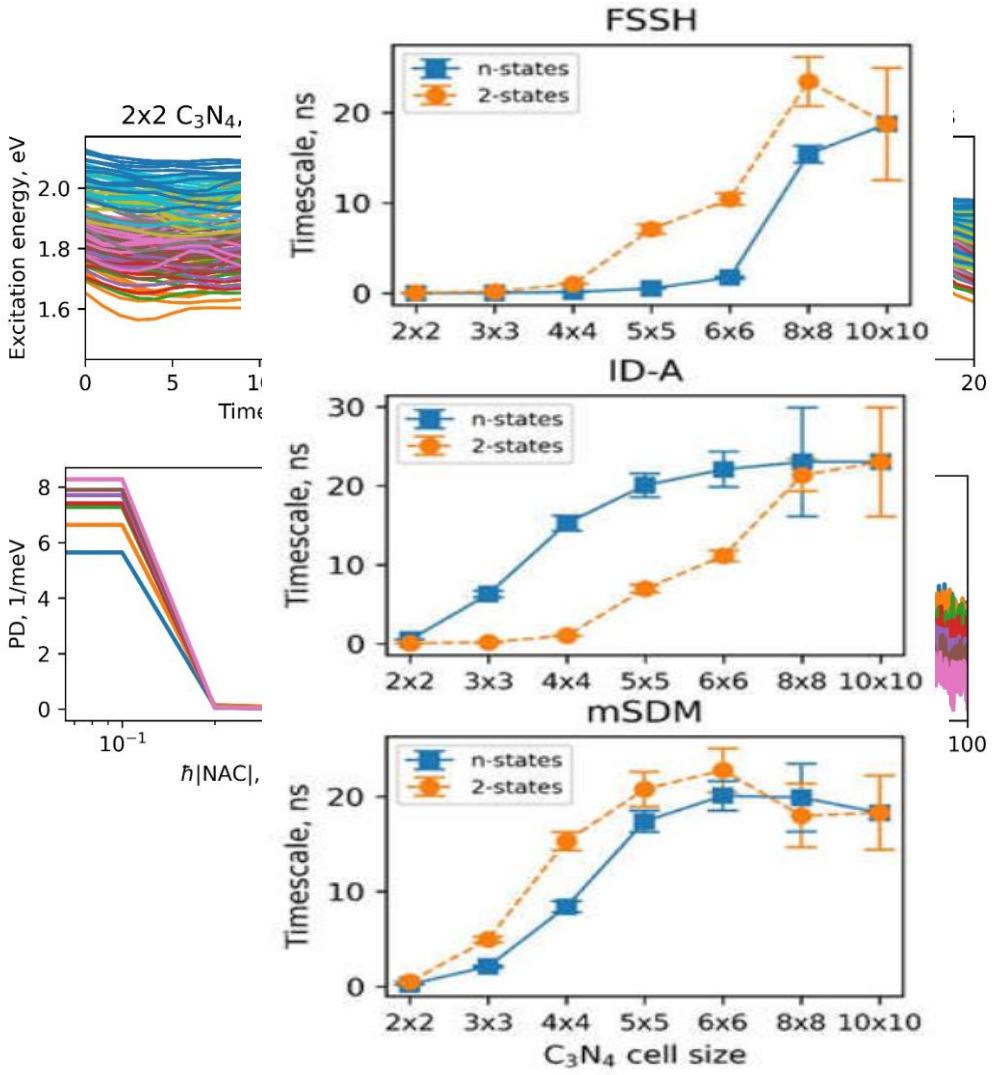
Molecular orbitals in C_3N_4 monolayers

- Occupied orbitals are more localized on several melem (triangular repeated motif of 3 fused rings) units
- Unoccupied orbitals are delocalized over multiple connected melem units which is better observed in larger supercells

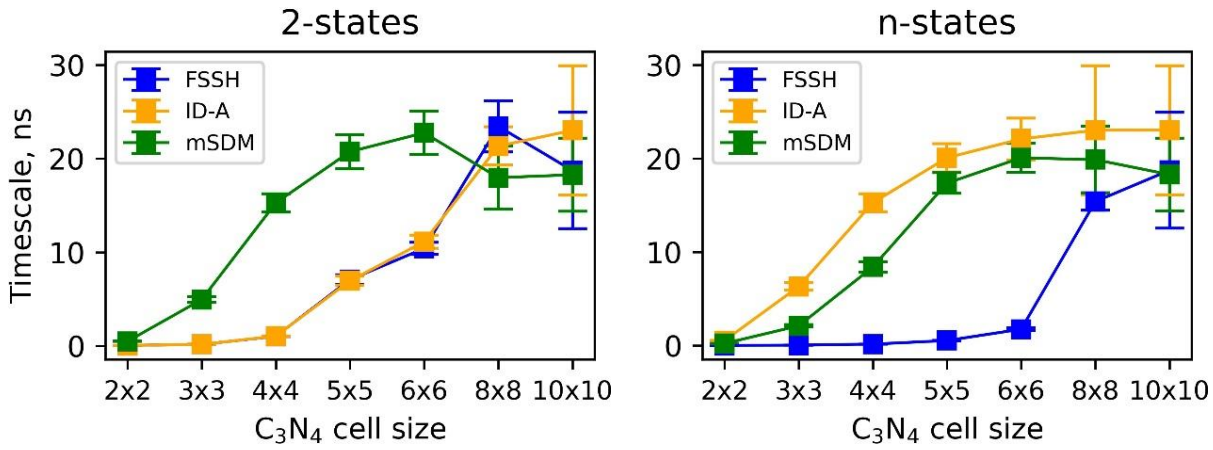


Shakiba, Akimov, JPCC 2023, 127, 9083-9096.

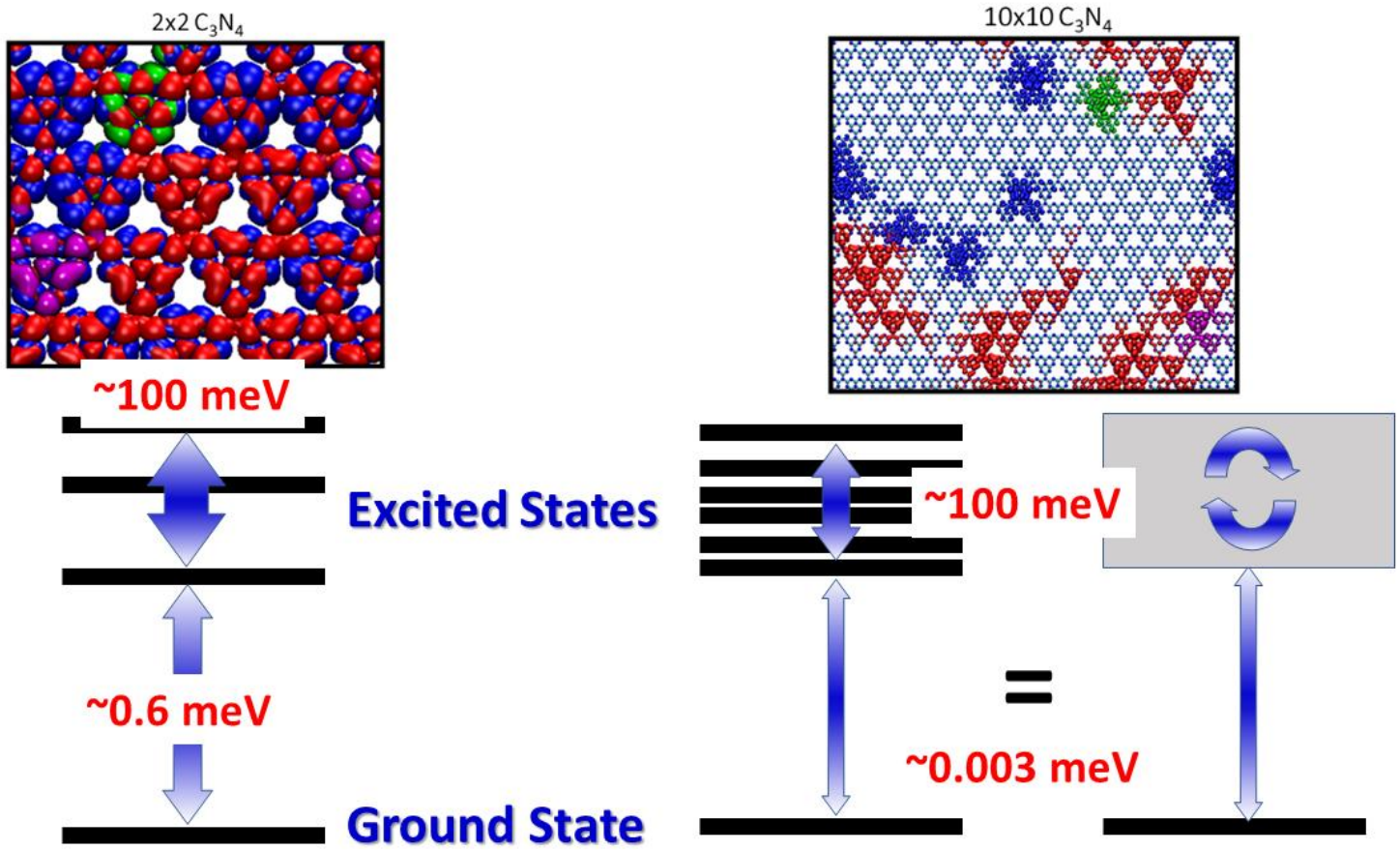
e-h recombination dynamics in C_3N_4 monolayers

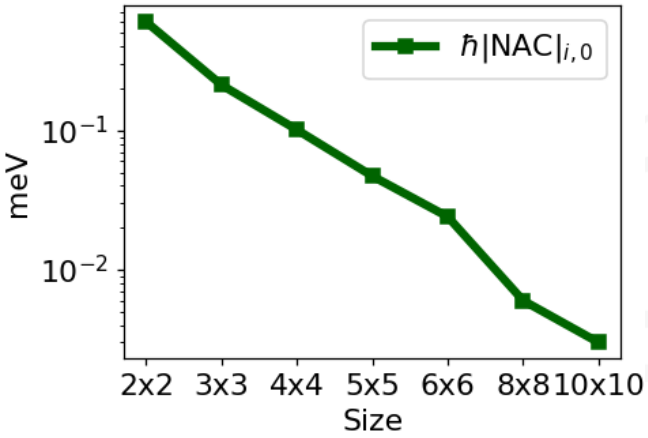


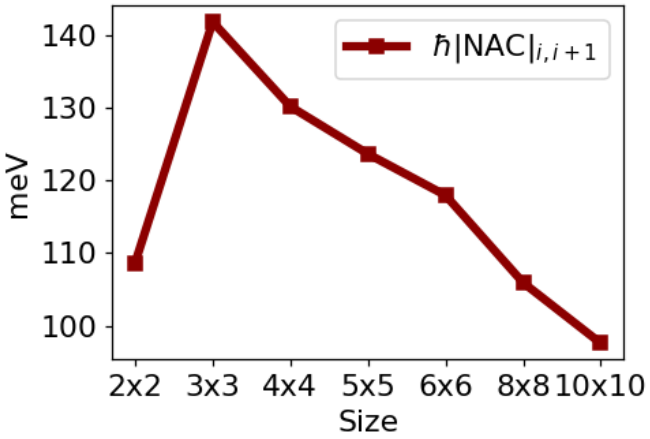
- 100 single-particle excitation states are built including all excitations from the first 10 occupied to the first 10 unoccupied molecular orbitals
- NAC values decrease by increasing supercell size
  - Small time-overlaps between molecular orbitals in large supercells
- Recombination dynamics becomes size-independent for all methods for very large supercell sizes



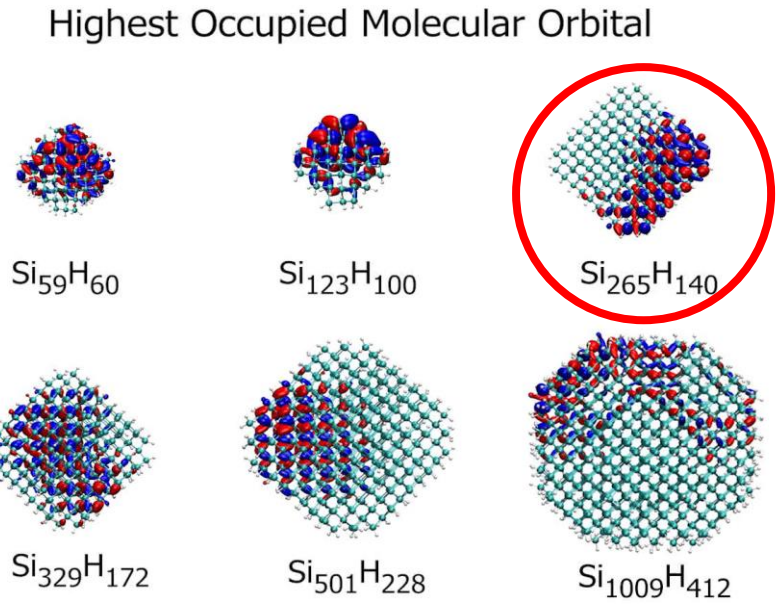
e-h recombination dynamics in C_3N_4 monolayers

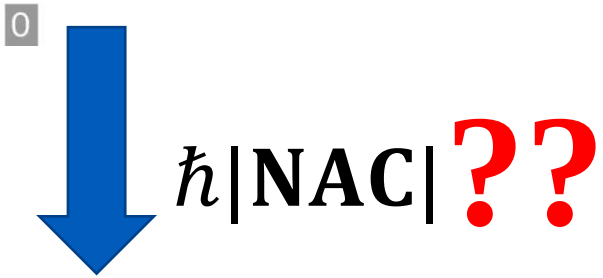


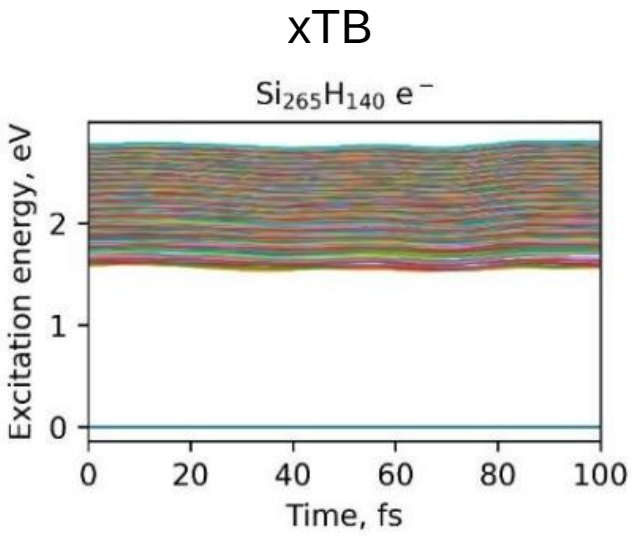


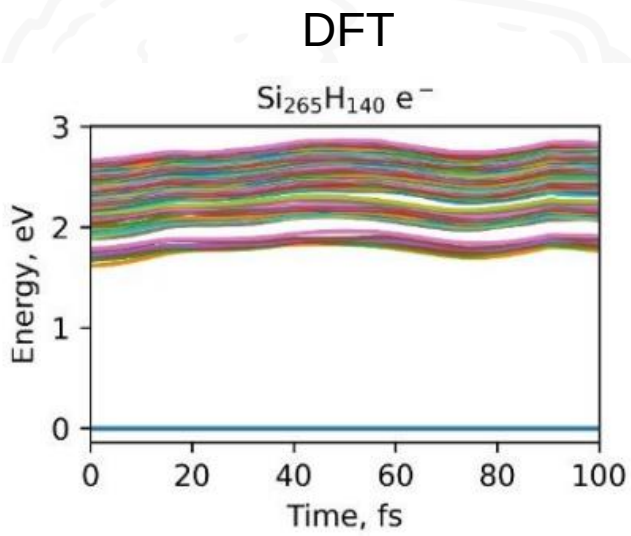


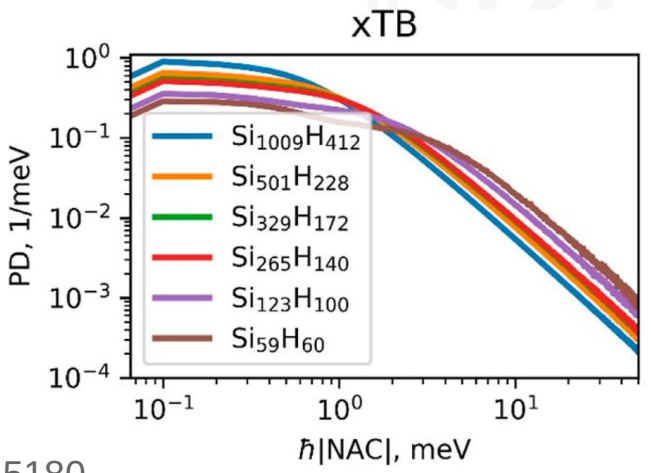
Hot-electron cooling dynamics in Si NCs

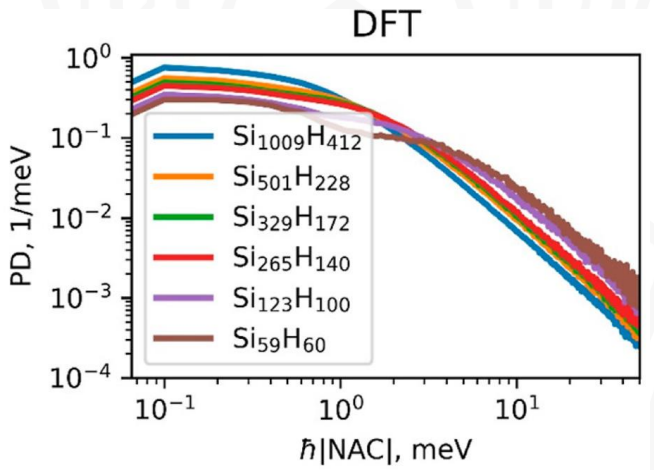












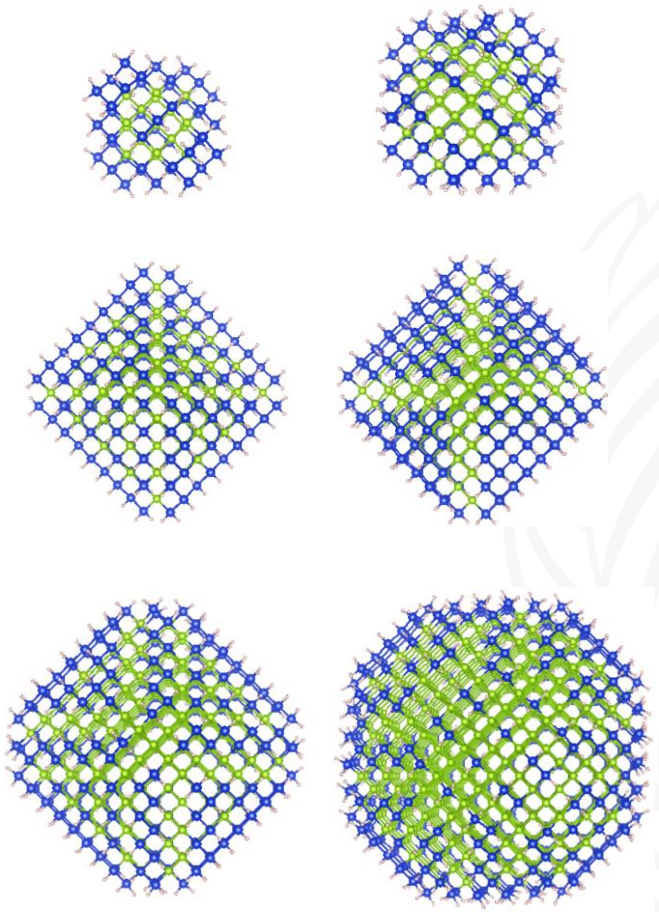
Shakiba, Stippell, Li, Akimov, J. Chem. Theory Compu. 2022, 18, 5157-5180

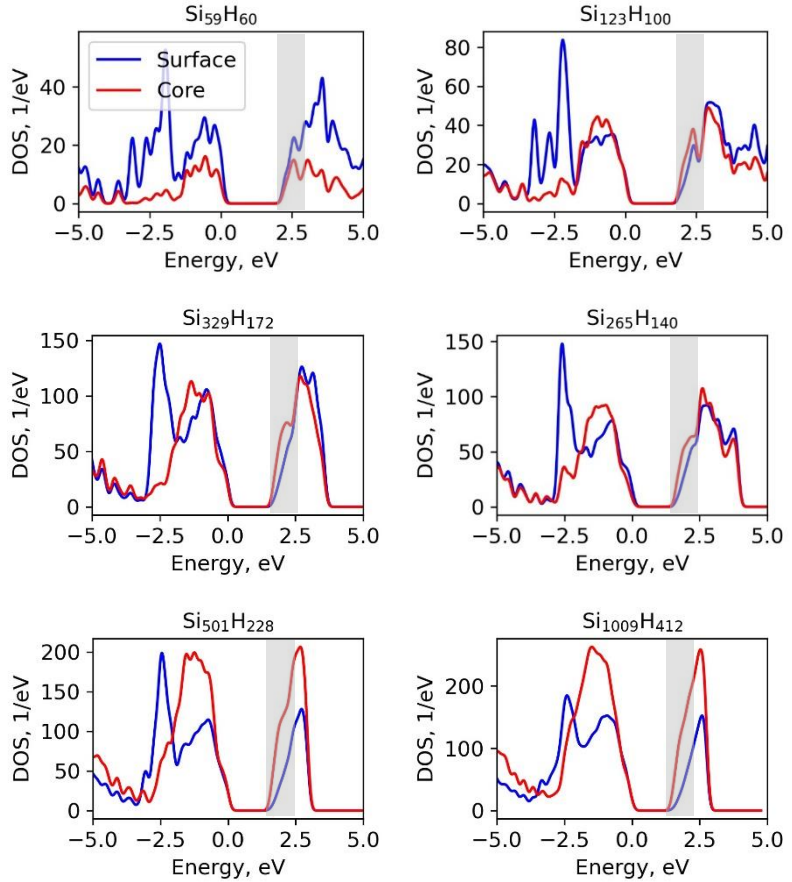
Core vs surface atoms

- By increasing size the surface/core DOS decreases
- Core atoms movements are slower than surface atoms

$$\sigma_i = \sqrt{\langle(\vec{r_i} - \langle\vec{r_i}\rangle)\rangle}$$

System	Si-surface	Si-core
$Si_{59}H_{60}$	0.273	0.172
$Si_{123}H_{100}$	0.152	0.113
$Si_{265}H_{140}$	0.144	0.114
$Si_{329}H_{172}$	0.186	0.141
$Si_{501}H_{228}$	0.262	0.190
$Si_{1009}H_{412}$	0.169	0.127

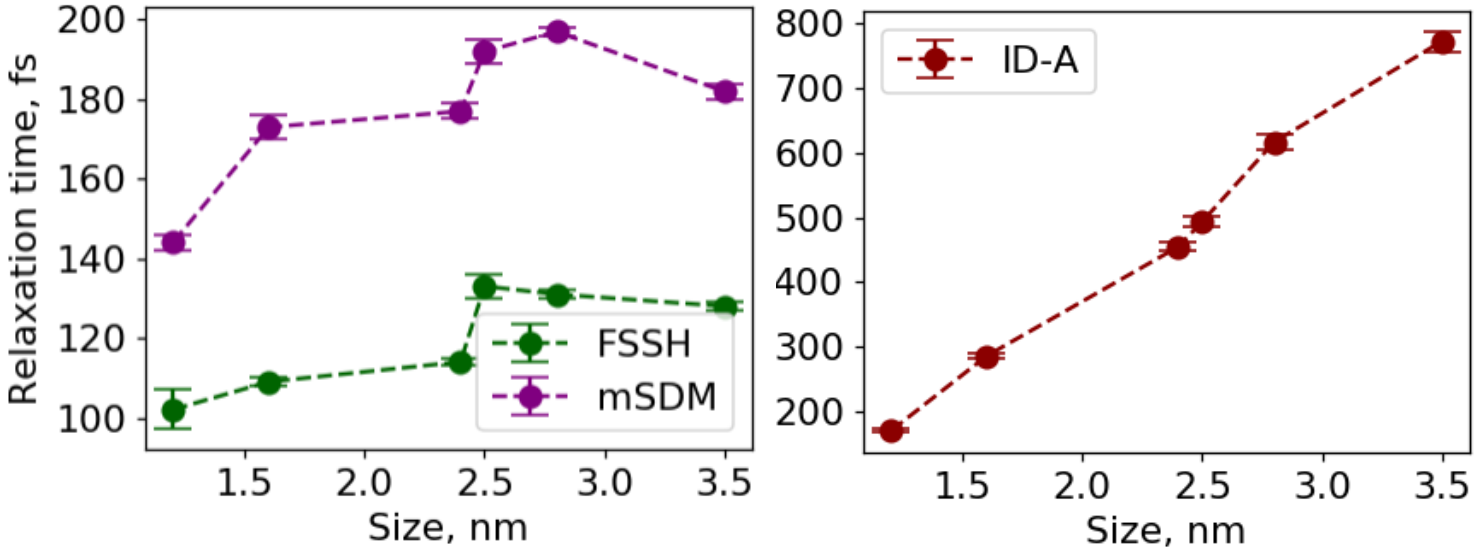




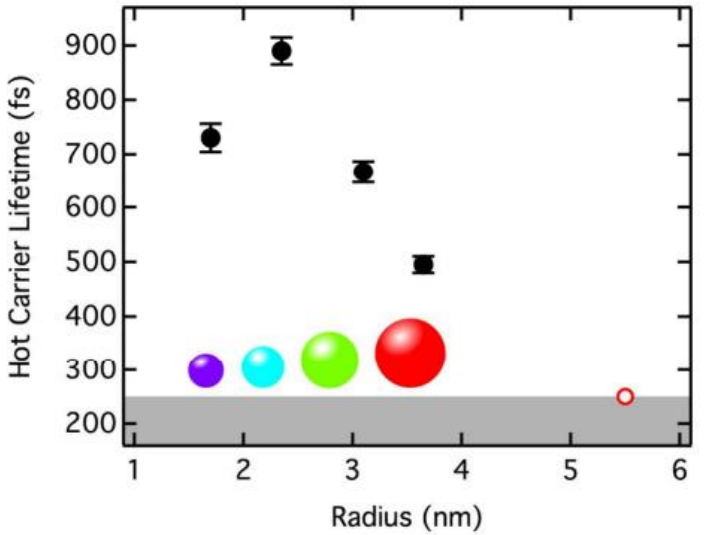
NA-MD results vs experiment

- Experimental data show a non-monotonic dependence of the relaxation time with respect to NC size
- FSSH and mSDM: Non-monotonic dependence of relaxation time on NC size
- ID-A: Monotonic dependence of relaxation time with increasing NC size

Theory



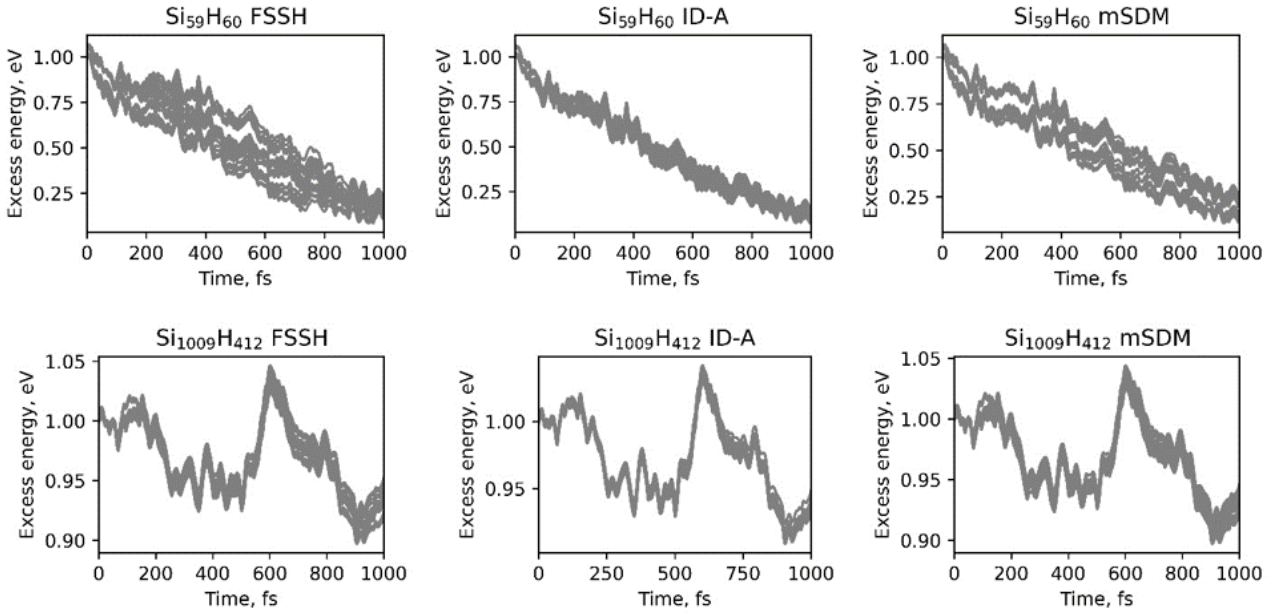
Experiment

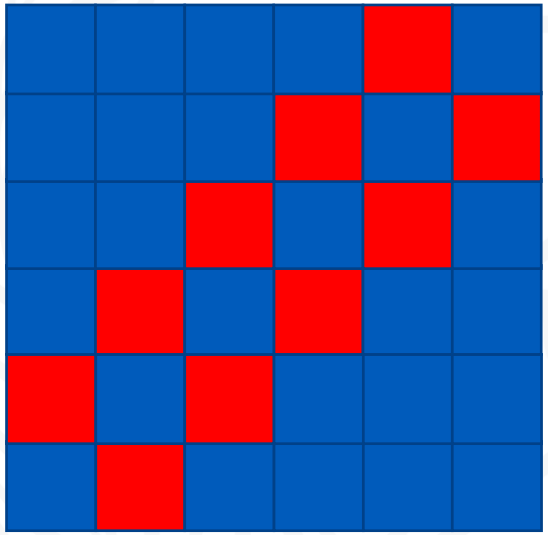


Bergren, Palomaki, Neale, Furtak, Beard, ACS Nano 2016, 10, 2316–2323

NA-MD results vs experiment

- By going to larger NCs, the possibility of relaxation happening through non-adjacent states increase
- NA-MD done with only adjacent state transition allowed does not decay for larger structures

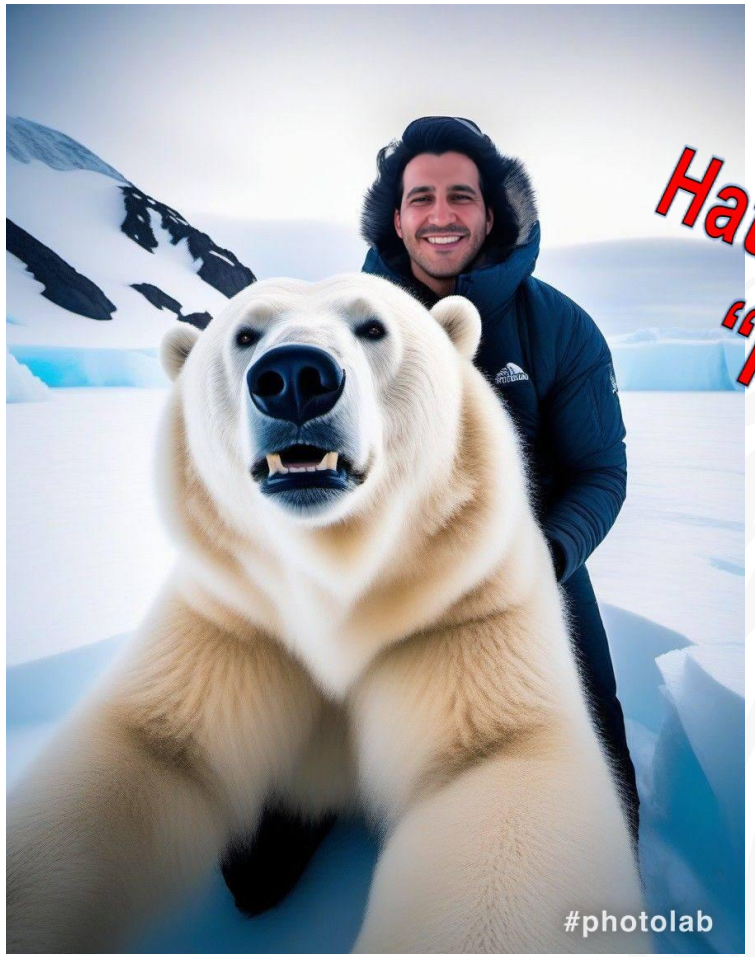




Part I Summary

- A new methodology for NA-MD simulations in large nanostructures and periodic solids is implemented in the open-source Libra code
- The NAC values decrease by increasing Si NC size and C_3N_4 monolayer supercell size
- A non-monotonic dependence of hot-electron relaxation dynamics on Si NC size was observed in FSSH and mSDM in agreement with experiment
- Non-adjacent transitions play an important role in the relaxation dynamics in Si NCs
- By varying charge carrier concentration using different supercell sizes, a saturation of the recombination timescale was observed showing the size-independence of recombination dynamics in very large structures (n-states → effective 2-states)

Part II: Nonadiabatic molecular dynamics with machine-learned Kohn-Sham Hamiltonian mapping



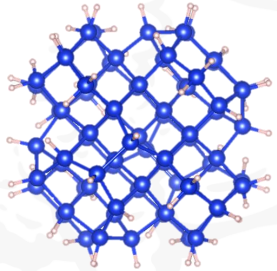
Haters gonna say:
"IT'S FAKE!"

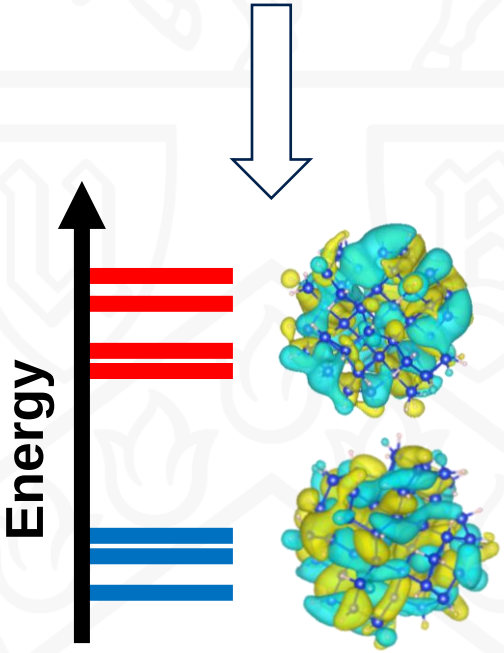
Me and my lovely pet ❤️🥰

Machine-learning in electronic structure calculations

- Single or multiple targets prediction: Band gaps, nonadiabatic couplings, dipole moment etc
- Machine-learning potentials and forcefields, one-electron reduced density matrix prediction
  - SchNorb, Schpacknet, Mlatom, QMLearn, etc
- What they have in common?
  - Trying to "surrogate" quantum chemistry software with ML models
  - Poorly transferrable
  - Not as general as they should be
  - Not applicable to all types of systems and only applicable to few atomic types
  - For large systems, they require the preparation of many data set which is time-consuming
  - Use Neural networks
    - Deep layers with large number of neurons, long training time, large % of training data (up to 90%), cumbersome preprocessing of the data, overfitting problem etc.

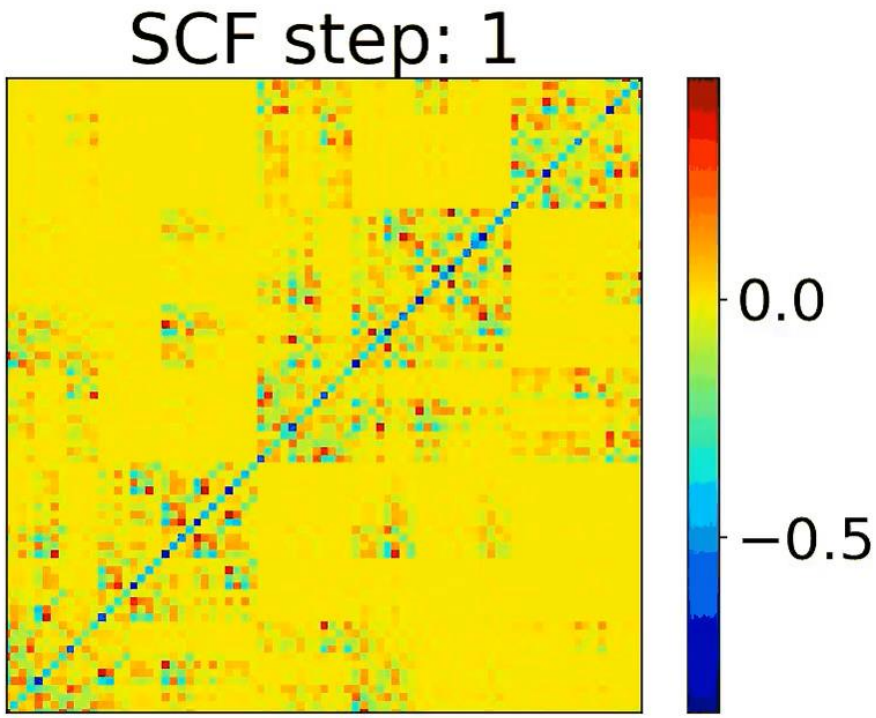
- Atomic coordinates and weights
- Potential and basis set descriptors
- etc



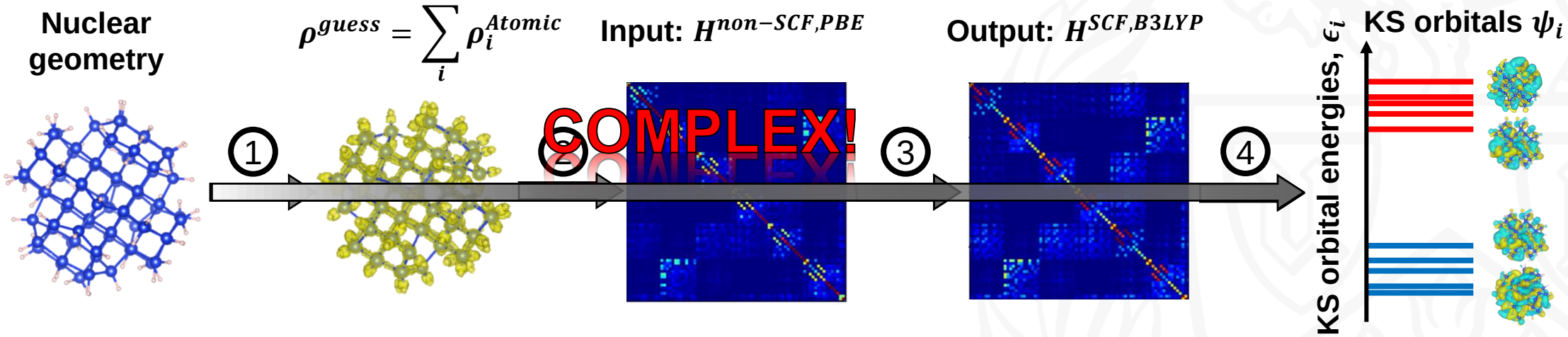


SCF cycle

```
SCF WAVEFUNCTION OPTIMIZATION - B3LYP

Step          Time (sec)      Convergence
--------------------------------------------
1 Guess/Diag.   5.7          1.59674484     -30
2 Broy./Diag.   127.5        2.77154376     -30
3 Broy./Diag.   127.6        0.72899999      31
4 Broy./Diag.   128.0        0.12822688      32
.....................................
.....................................
.....................................
39 Broy./Diag.  129.4        0.00000115      33
40 Broy./Diag.  129.4        0.00000105     -33
41 Broy./Diag.  128.7        0.00000057     -33
```



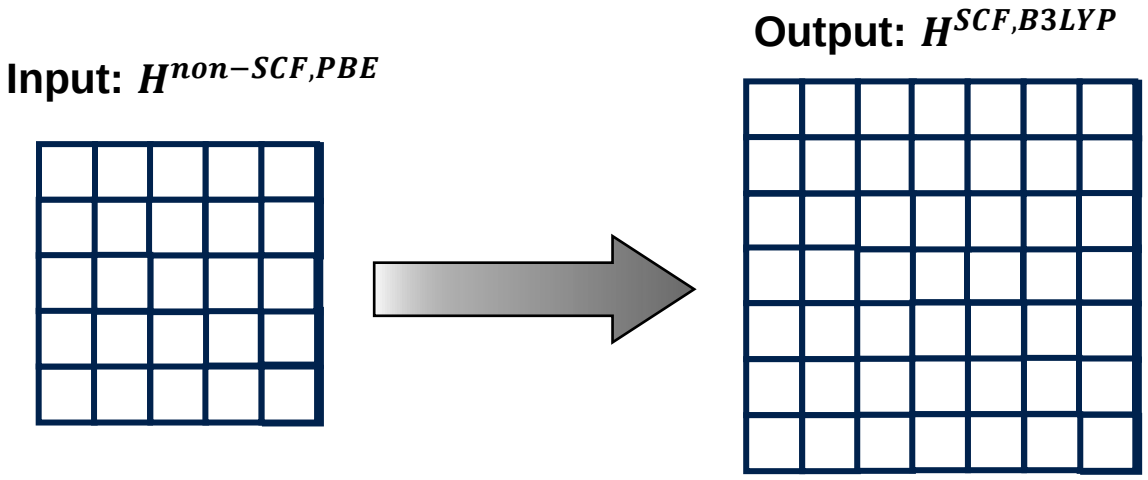
What are the inputs?

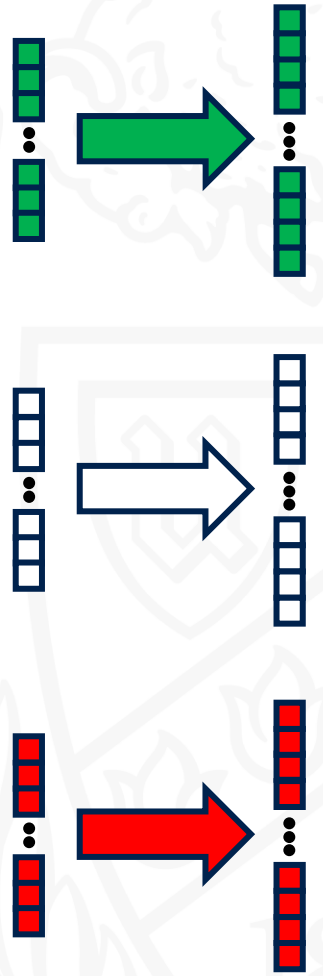


Shakiba, Akimov, J. Chem. Theory Compu., 2024, 8, 2992-3007

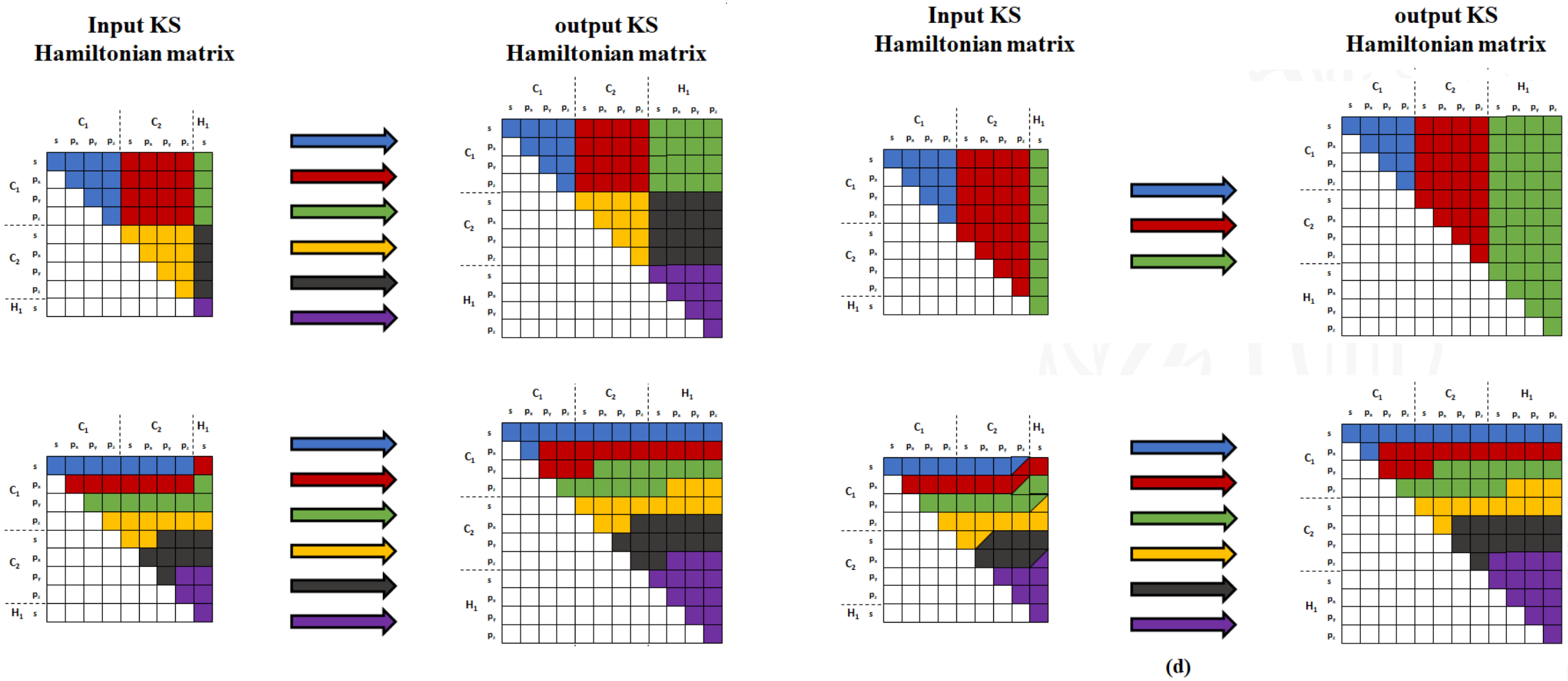
Is this efficient?

- Up to millions of elements for very large systems
- Pick the upper triangular part of the KS Hamiltonian matrix due to symmetry
- Split them into multiple partitions
- Train a separate model for each partition → Each model can be separately trained in parallel
- Rebuild the matrix and diagonalize it





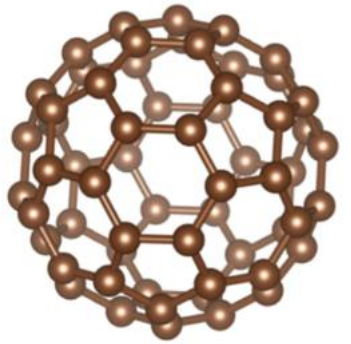
Different partitioning methods

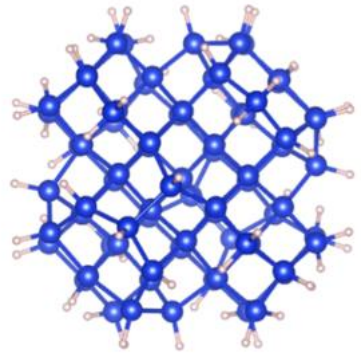


(d)

Models and the systems

- Kernel ridge regressor with a linear kernel
- C_{60} fullerene with a basis set size of 240
- $Si_{75}H_{64}$ with a basis set size of 1039
- Step 1:
  - Generate a precomputed nuclear trajectory with 2000 geometries with PBE functional (similar to what is done in typical NA-MD simulations in nanoscale systems)
- Step 2:
  - Equal partitioning of the input and target Hamiltonian matrices
    - 30 partitions just for test!
- Step 3:
  - Train the models for 50 (2.5%), 100 (5%), 250 (12.5%), 500 (25%), 750 (37.5%), 1000 (50%) randomly selected geometries
- Step 4:
  - Use the model to generate the Hamiltonian matrices and molecular orbitals which can be done using high-throughput computation





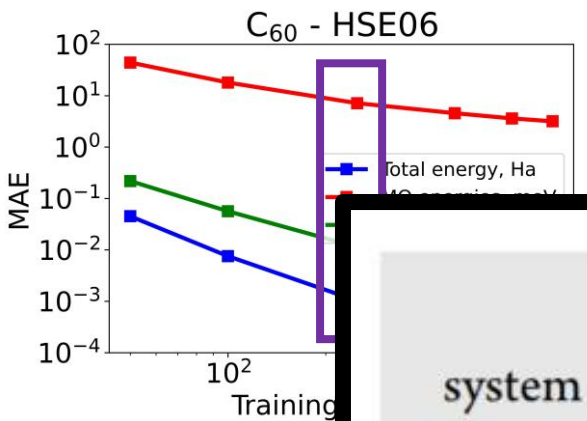
Error measures

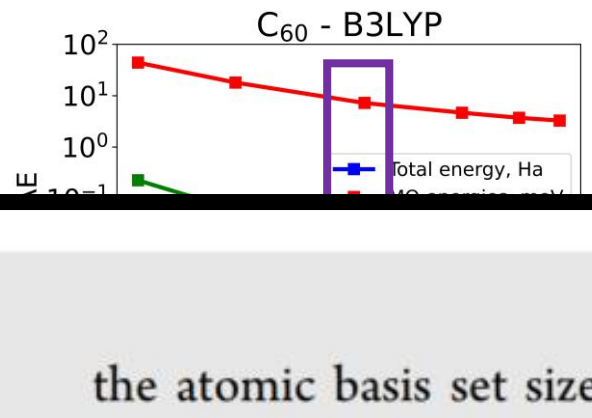
- Mean absolute error of the following:
  - 1- Total energy
    - Feed back the ML molecular orbitals to the quantum chemistry software
  - 2- Molecular orbitals energies
  - 3- Overlap of the ML and reference molecular orbitals

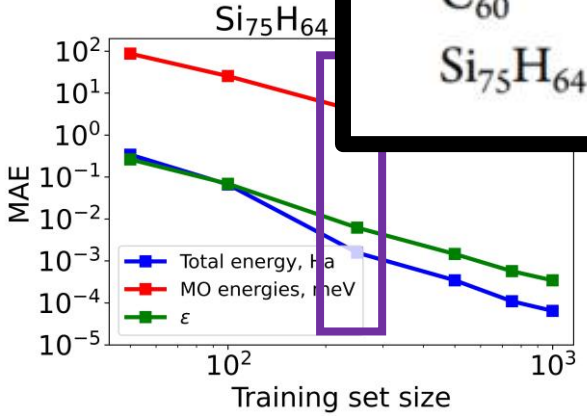
$$\epsilon_{i,overlap} = 1.0 - |\langle \psi_{i,ML} | \psi_{i,ref} \rangle|$$

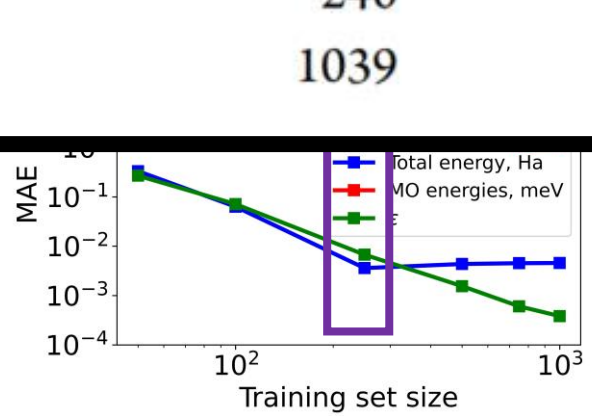
$$\epsilon = \frac{1}{N_{MO}} \sum_{i}^{N_{MO}} \epsilon_i$$

ML results

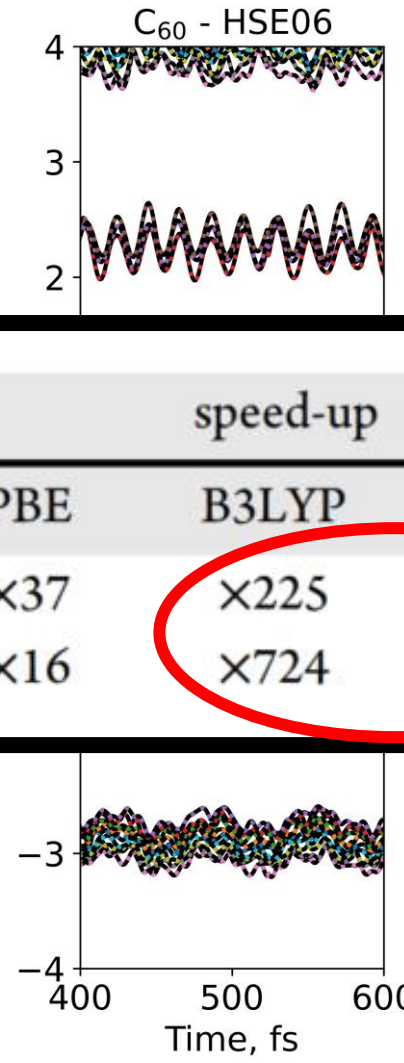


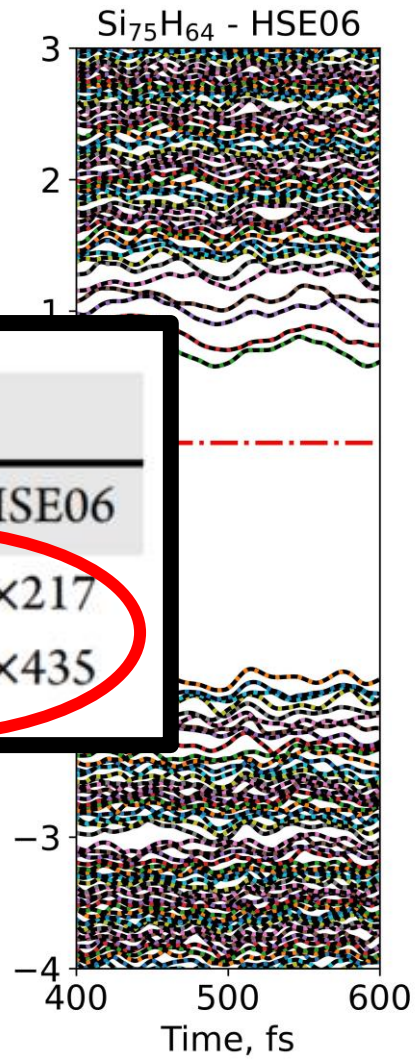






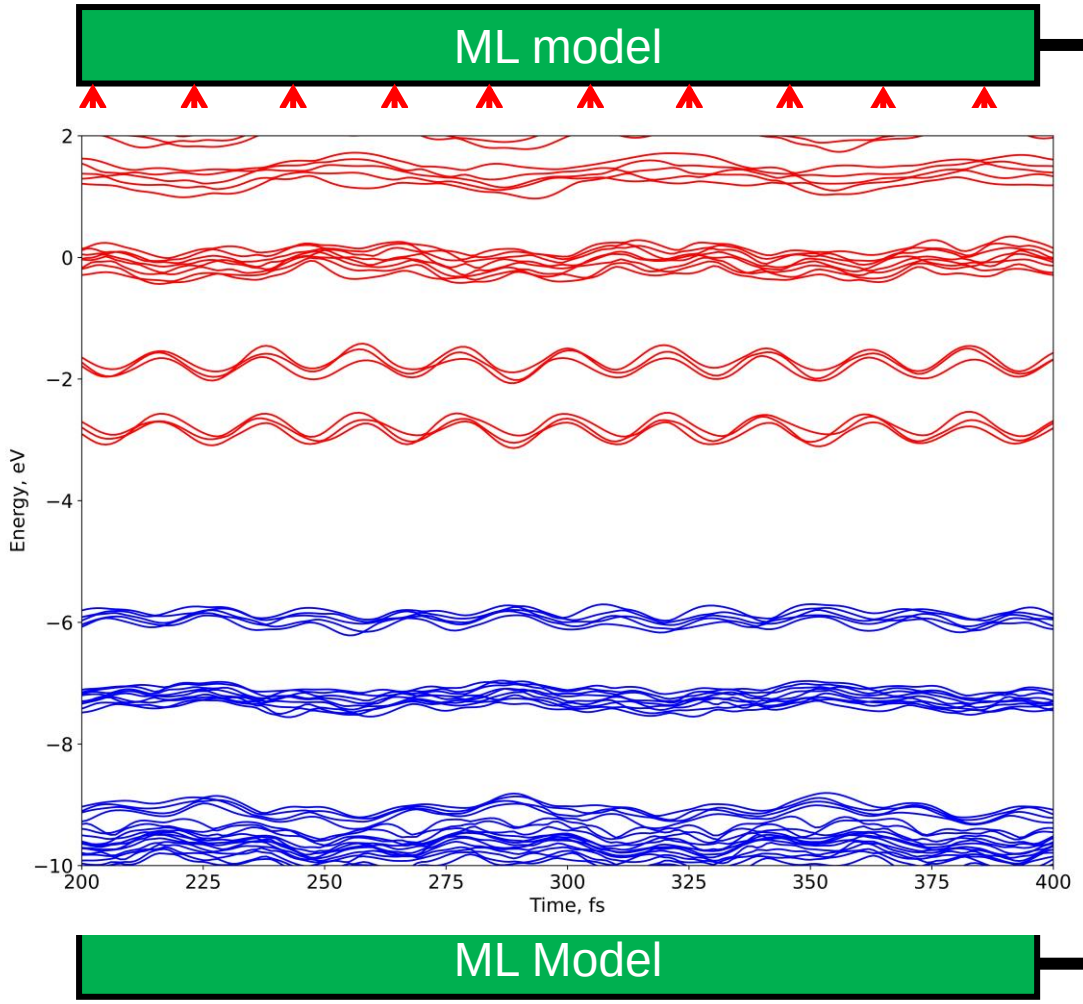
system	the atomic basis set size	speed-up PBE	B3LYP	HSE06
C_{60}	240	×37	×225	×217
$Si_{75}H_{64}$	1039	×16	×724	×435

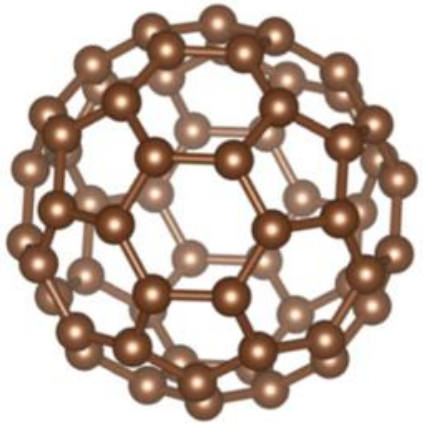




Shakiba, Akimov, J. Chem. Theory Compu., 2024, 8, 2992-3007

Let's review the idea



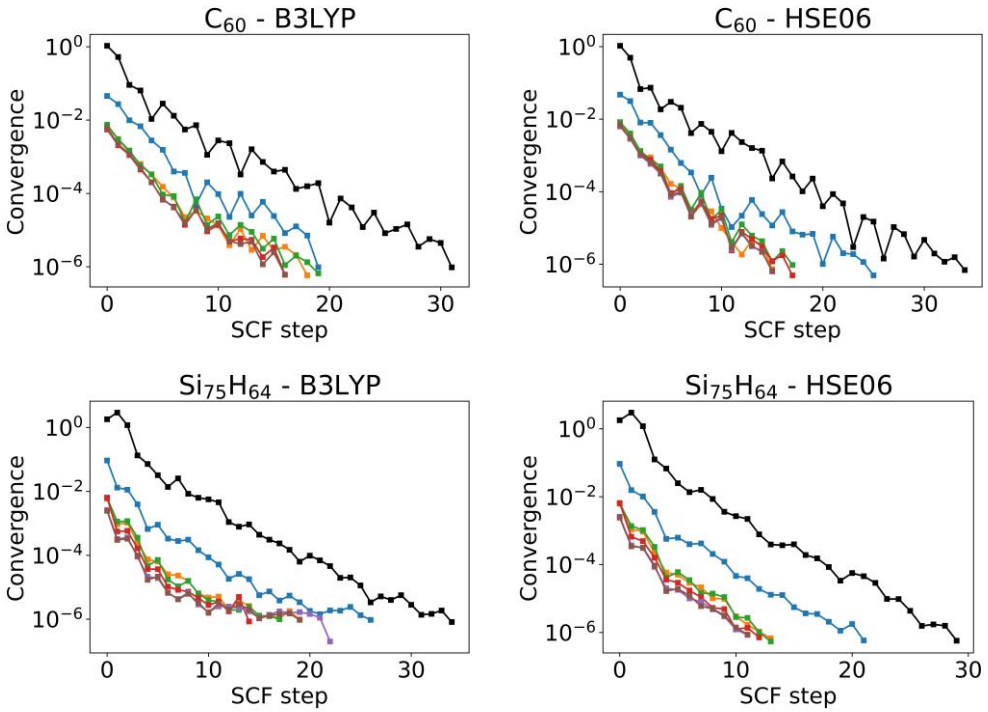


- CP2K: PBE atomic guess, 2.17 sec
- CP2K: Converged B3LYP, 529 sec
- ML mapping: 0.08 sec
- Diagonalization: 0.1 sec
- ML training: 12 sec

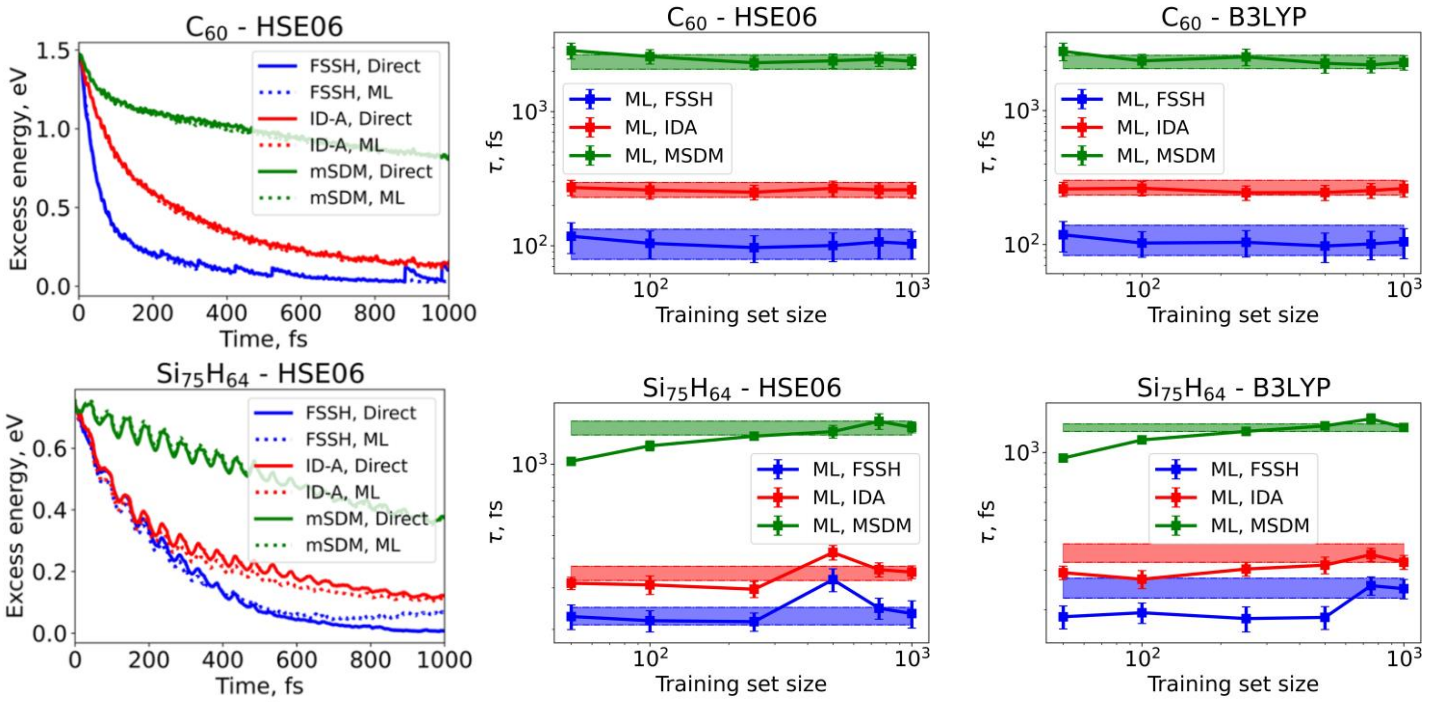
Shakiba, Akimov, J. Chem. Theory Compu., 2024, 8, 2992-3007

ML results

- SCF convergence with ML MOs
  - 1.5-2 times faster convergence
  - ML MOs are good guesses!



- Excess energy relaxation dynamics
  - Time scales from dynamics in ML MOs are in within the error margin of the reference time scales



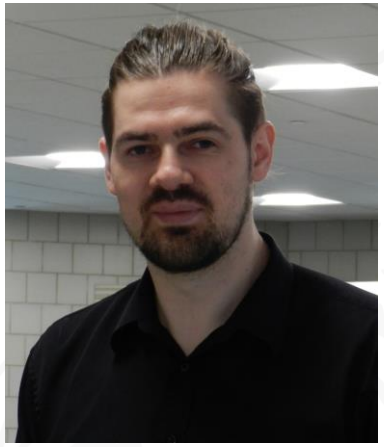
Shakiba, Akimov, J. Chem. Theory Compu., 2024, 8, 2992-3007

Part II Summary

- A simple, efficient, and scalable ML approach for mapping non-self-consistent Kohn-Sham Hamiltonians constructed with one kind of density functional to the nearly self consistent Hamiltonians constructed with another kind of density functional.
  - Speeds up the calculations by several orders of magnitude
  - Is conceptually simpler than alternative ML approaches
  - Is applicable to different systems and sizes and can be used for mapping Hamiltonians constructed with arbitrary density functionals
  - Requires a modest training data, learns fast, and generates molecular orbitals and their energies with the accuracy nearly matching that of conventional calculations
  - When applied to nonadiabatic dynamics simulation of excitation energy relaxation in large systems yields the corresponding time scales within the margin of error of the conventional calculations

Acknowledgement

- Prof. Alexey V. Akimov
- Current group members:
  - Mohammad Shakiba
  - Qingxin Zhang
  - Daeho Han
  - Kosar Yasin
  - Layla Heidarizadeh
- Former group members:
  - Dr. Brenden Smith
  - Dr. Wei Li
  - Elizabeth Stippell



University at Buffalo
The State University of New York

NSF

UB Center for Computational Research
University at Buffalo The State University of New York

Thank You!

Questions?