TUM seminar Theory and Computation of Materials and Molecules

Electronic coarse-graining of semiconductors: methodology and applications

Andriy Zhugayevych

Max Planck Institute for Polymer Research, Polymer Theory Department

June 11, 2024

Acknowledgments

MPIP and MPG for emergency fellowship for Ukrainian scientists

Denis Andrienko's group

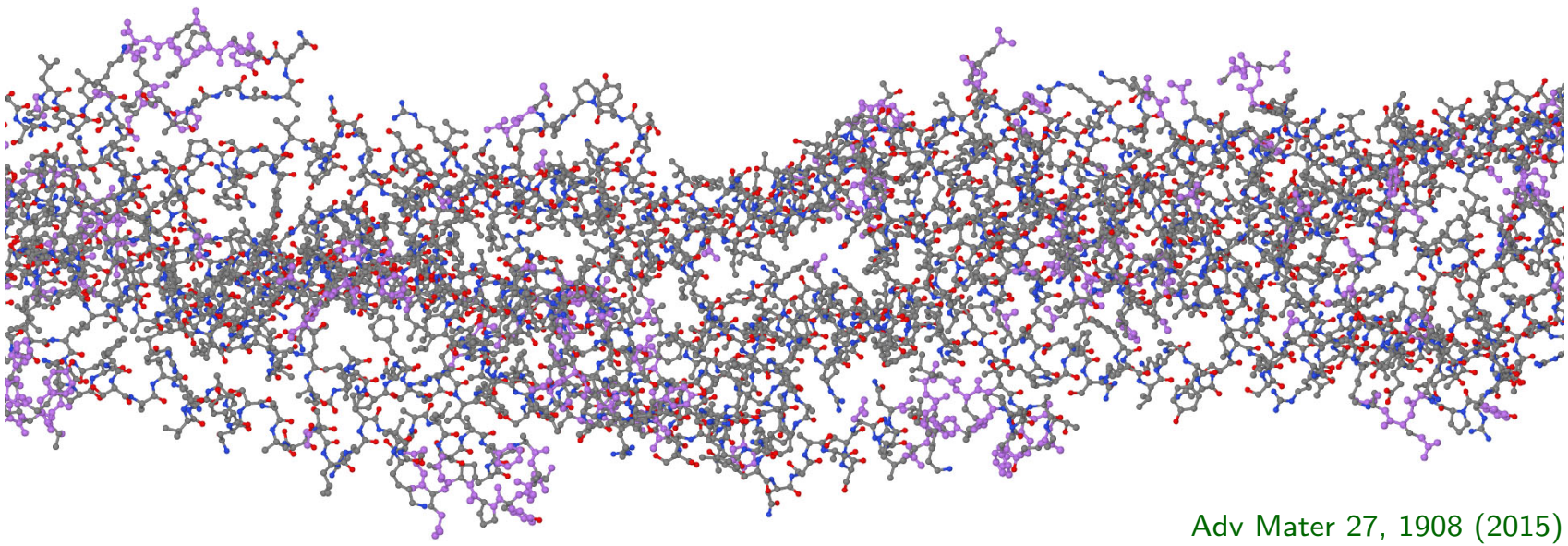
Sergei Tretiak (LANL/CINT user project)

Outline

- Introduction and illustrative examples
- Methodology
- Applications

Electronic Coarse-Graining (ECG) in examples

- When do we need ECG? – When system is too complex for use of conventional electronic structure methods (e.g. polymers)
- What is the "price" for ECG? – Only frontier orbitals are calculated (enough for majority of electronic phenomena)
- ECG example 1/4: biopolymers (localized states $\Rightarrow$ trivial ECG)



Adv Mater 27, 1908 (2015)

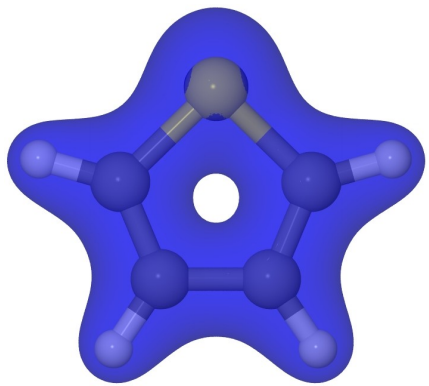
10k non-H atoms $\Longrightarrow$ 50k valence molecular orbitals (MOs) or 1M AOs
but frontier MOs are localized on a small fraction of AOs (violet color)

ECG: more precise formulation

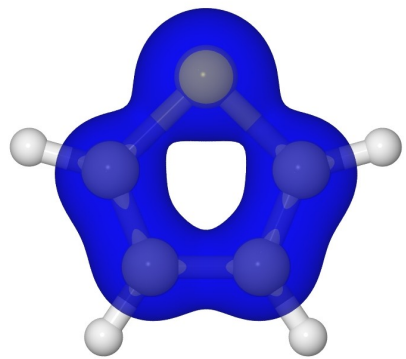
ECG idea/goal: Get minimal basis providing accurate description of a particular electronic property under molecular fluctuations

(large-scale/low-energy electronic phenomena: UV-Vis spectra, transport)

Output: electronic basis + electronic Hamiltonian



All-electron density

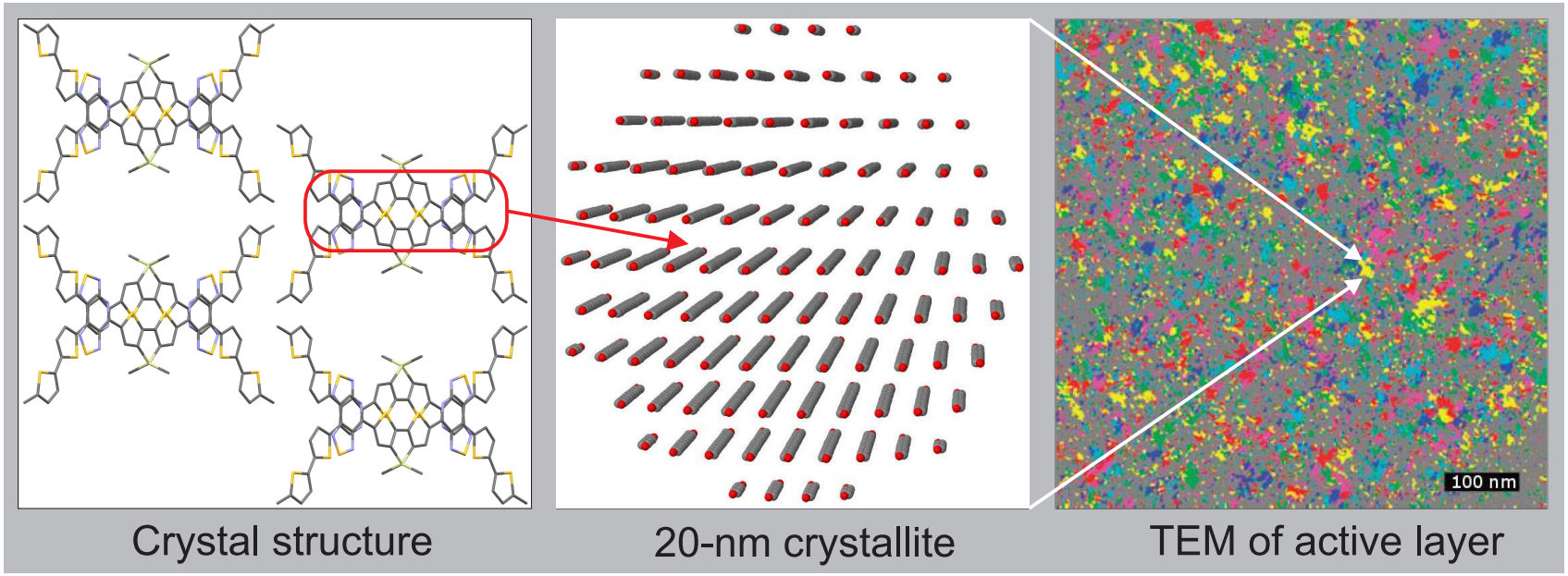


CG density (2 MOs)

Requirements:

- Robustness of CG basis wrt molecular fluctuations
- Robustness and scalability of CG algorithm
- Quality control of CG basis and matrix elements
- Accurate extrapolation to infinite system (if needed)

ECG example 2/4: electronic transport in a molecular solid



- Coarse grain electrons to one site per molecule
- Simplify molecular motions to harmonic vibrations
- Linearize coupling between electrons and molecular motions

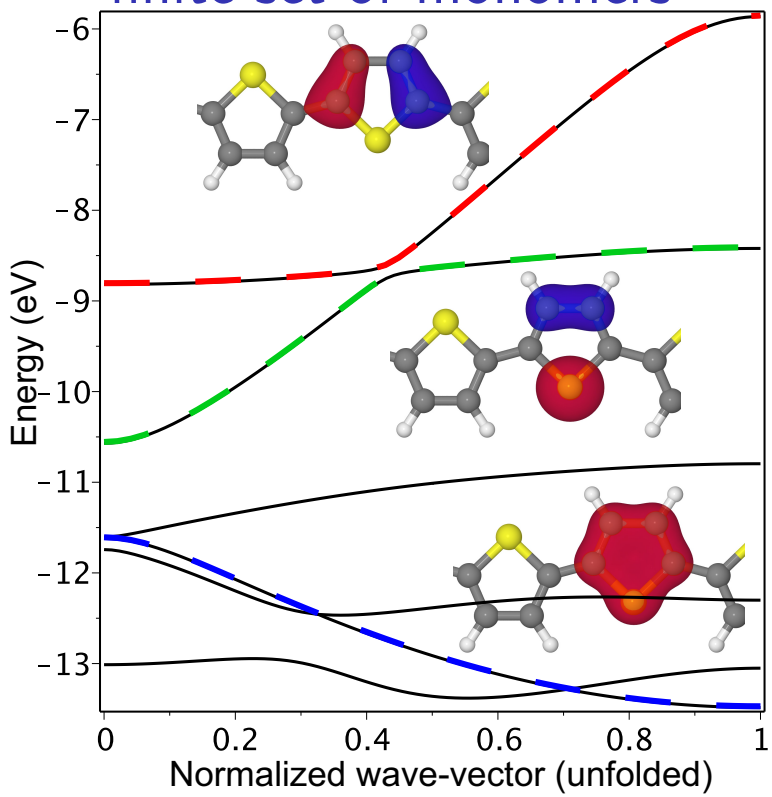
$$\Longrightarrow \sum_{ij} H^{1p}_{ij} c^\dagger_i c_j + \sum_\alpha \hbar\omega_\alpha \left(b^\dagger_\alpha b_\alpha + \frac{1}{2} \right) + \sum_{ij\alpha} \hbar\omega_\alpha g_{ij\alpha} \left(b^\dagger_\alpha + b_\alpha \right) c^\dagger_i c_j$$

Then solve this Hamiltonian (e.g. in small polaron hopping approximation)

Annu Rev Phys Chem 66, 305 (2015) first review – Chem Rev 104, 4971 (2004)

ECG in examples 3/4: less trivial coarse-graining

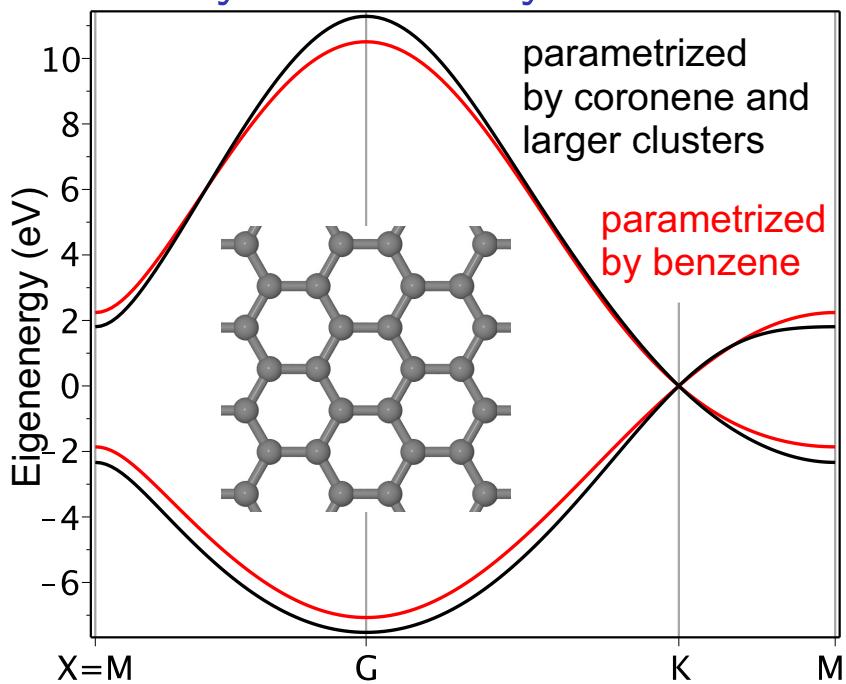
Combinatorial screening of polymers and oligomers from a finite set of monomers



Single polythiophene chain (π-system)

rational design – Chem Sci 8, 1146 (2017), Solar Energy 198, 605 (2020)

Electronic structure of infinite systems with methods available only for finite systems*



Graphene (π-system)

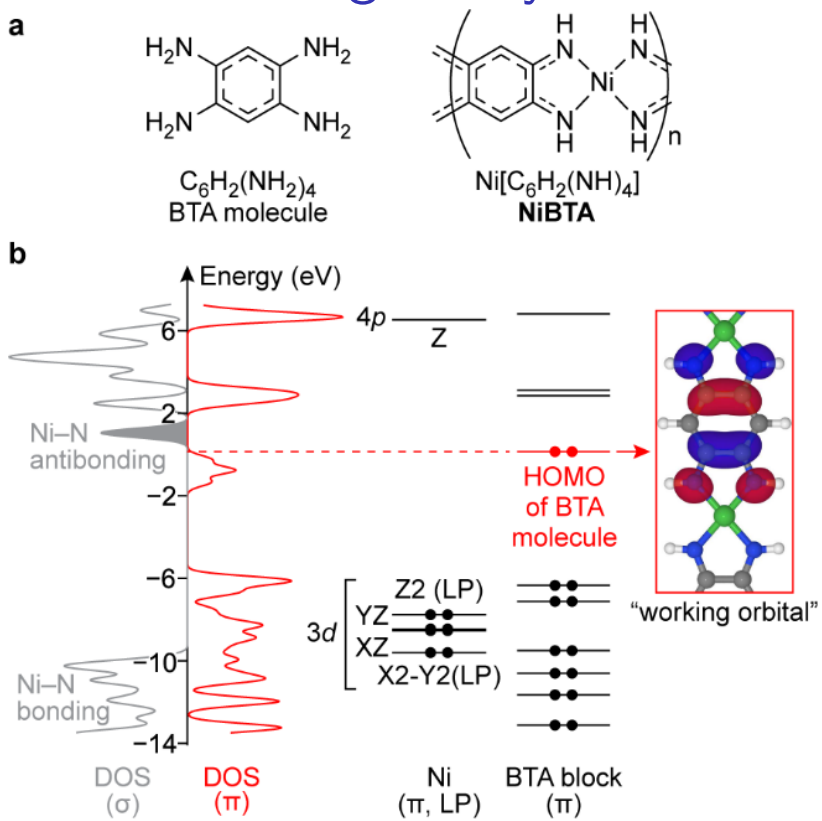
polarons in phosphorene – JPCL 12, 4674 (2021)

*e.g. PBE-D3 is good for crystal structure but bad for e-properties JCTC 19, 8481 (2023)

ECG in examples 4/4: nontrivial coarse-graining

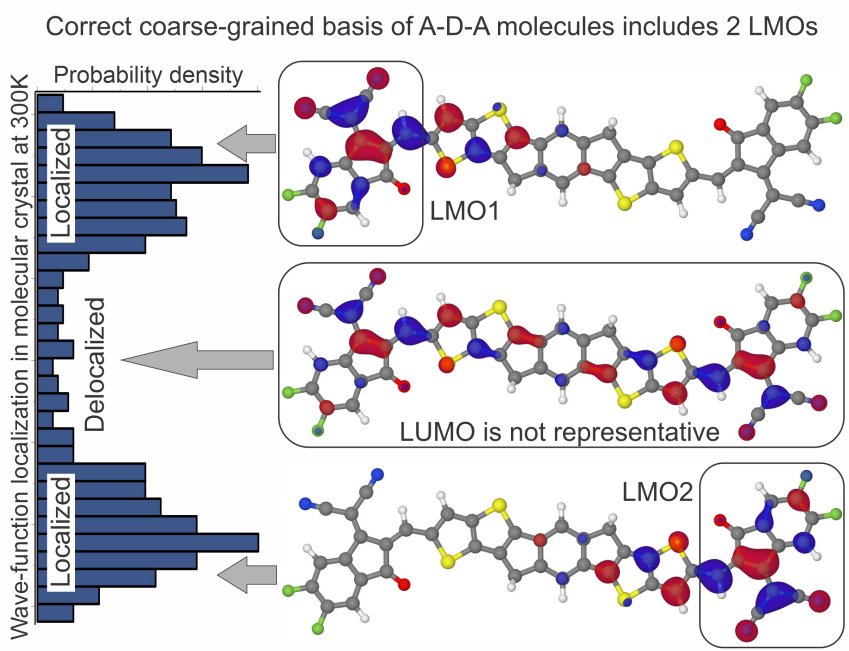
(and applications)

Metal-organic systems



energy storage – Chem Sci 13, 8161 (2022)

Extended molecules



solar cells – JCP 159, 024107 (2023)

... and other structurally complex systems

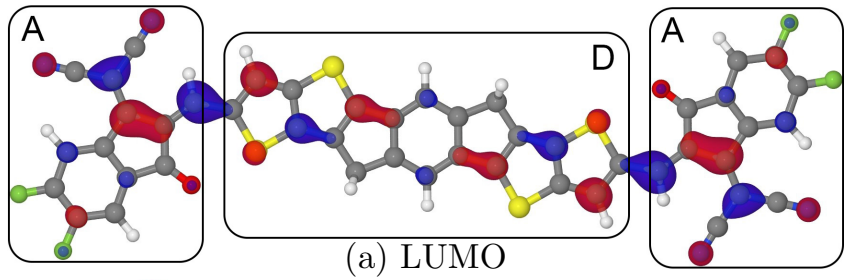
Methodology of ECG

Requirements:

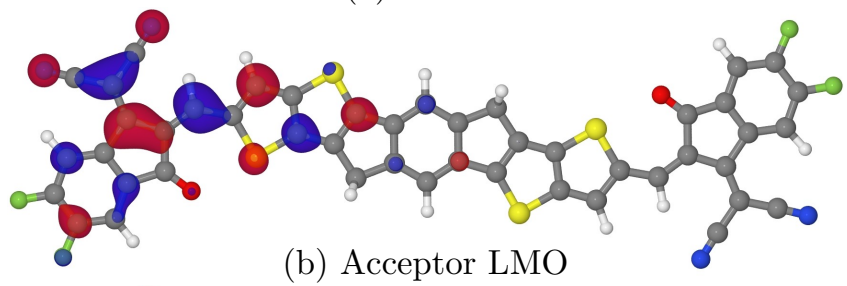
- Robustness of CG basis wrt molecular fluctuations
  — Case study of A-D-A molecules (e.g. NFA for solar cells)
- Robustness and scalability of CG algorithm
  — Implementation
- Quality control of CG basis and matrix elements
- Accurate extrapolation to infinite system (if needed)
  — Efficient parametrization of tight-binding models

Robustness of CG basis wrt molecular fluctuations:

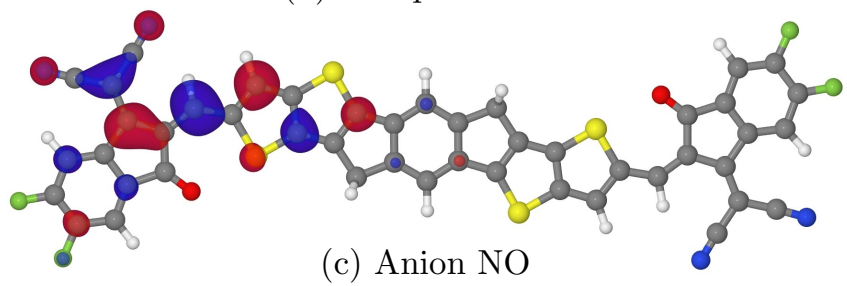
Case study of A-D-A molecules (non-fullerene acceptors in solar cells) JCP 159, 024107 (2023)



(a) LUMO



(b) Acceptor LMO



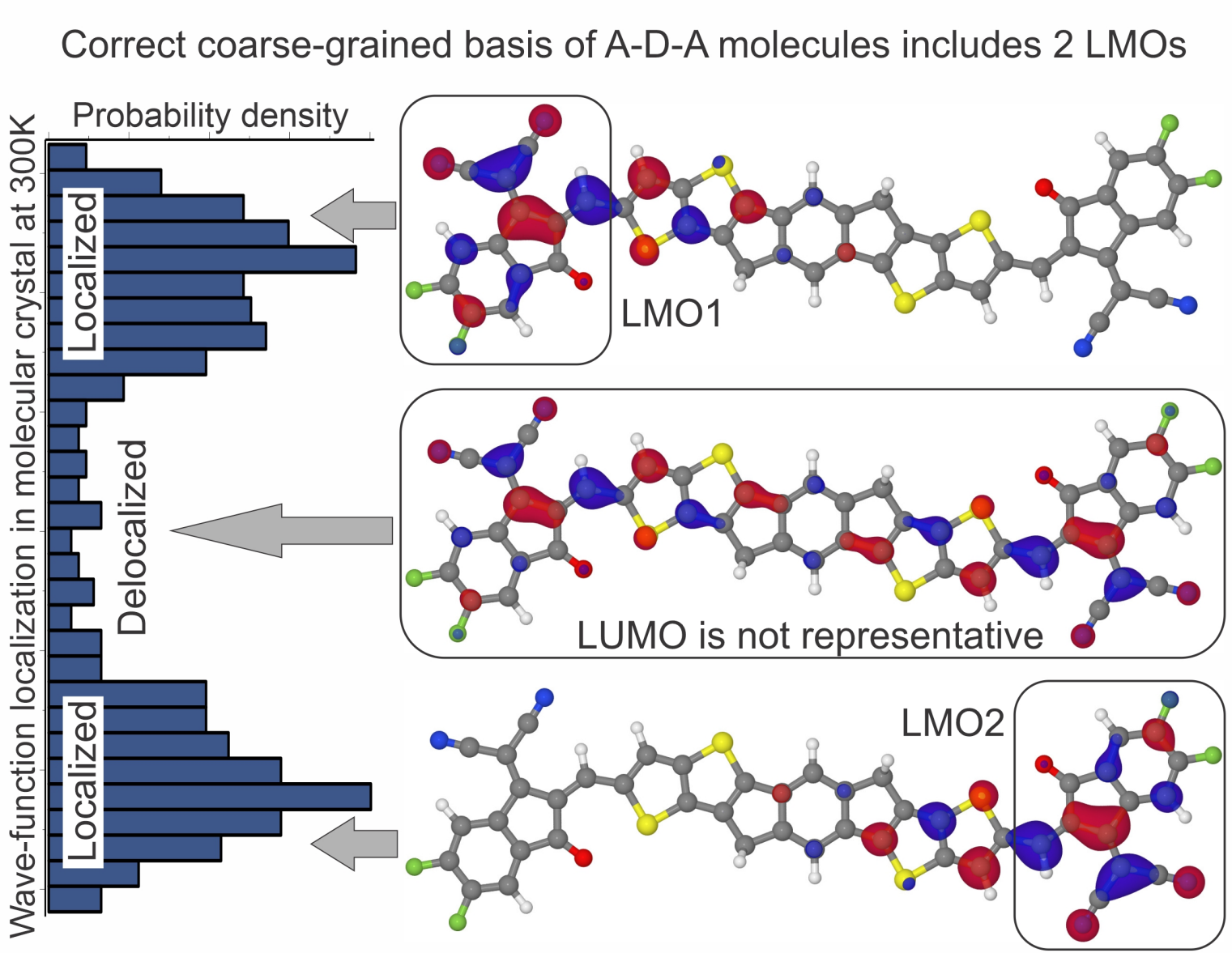
(c) Anion NO

Determine qualitative and quantitative errors of inaccurate ECG:

- Is electron localized on acceptor or delocalized over molecule?
- How inaccurate ECG influences charge transport parameters?
- To be compared: 1-site vs 2-site, 2-LUMO vs 2-LMO models

MO=Molecular Orbital, LUMO=Lowest Unoccupied MO, LMO=Localized MO

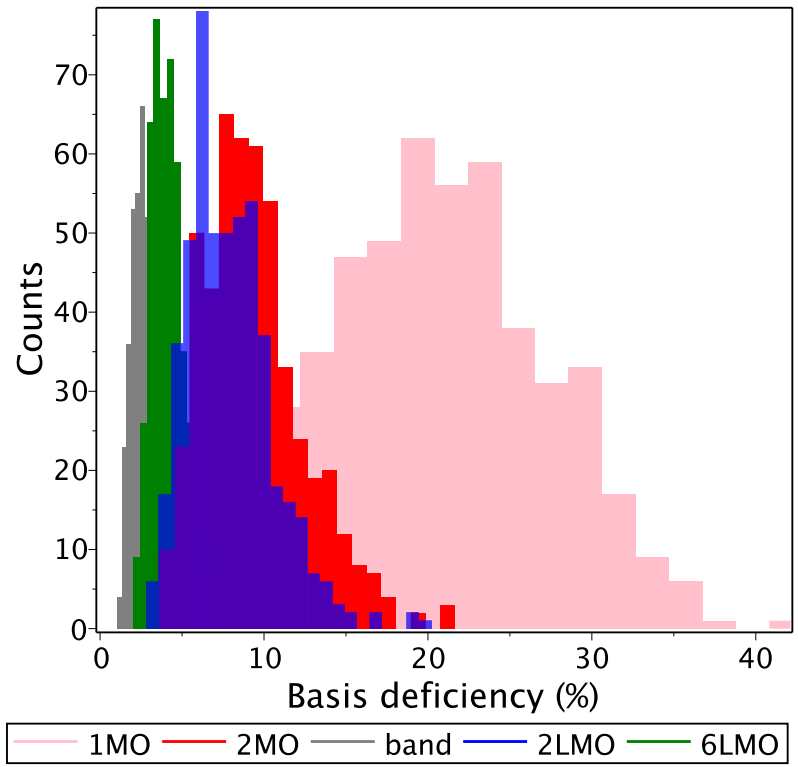
Main result



For molecular solids, the geometry is sampled by classical MD.
For molecules, quantum effects are considered by sampling thermal vibrations.

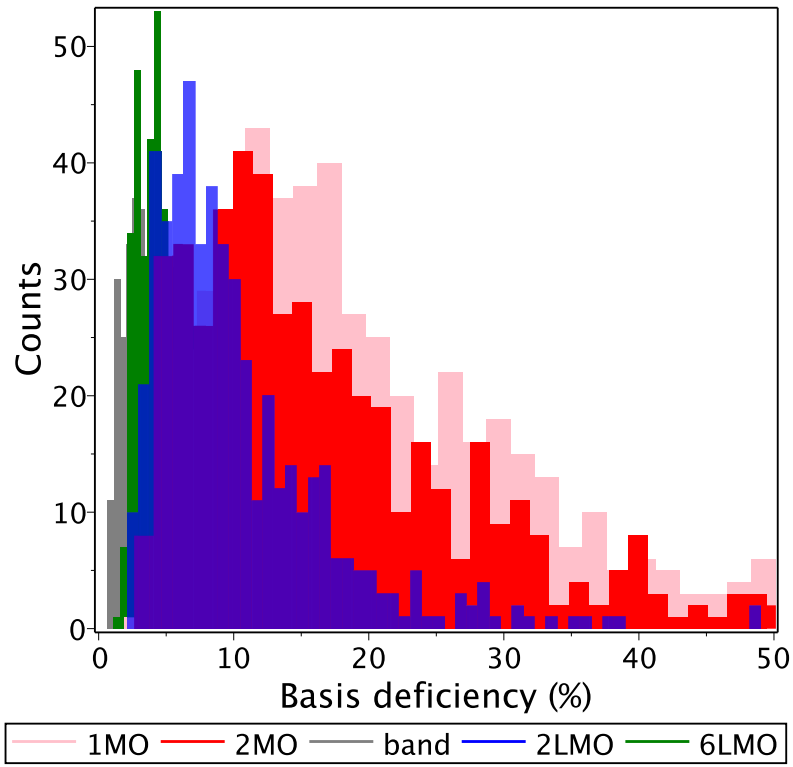
More rigorously: 2-LMO model is the most appropriate

Statistical analysis of the basis deficiency in a coarse-grained description of electron MO and NO (i.e. LUMO and anion NO):



Electron NO in crystal

$\Longrightarrow$ 1-site model fails



Electron MO in amorphous solid

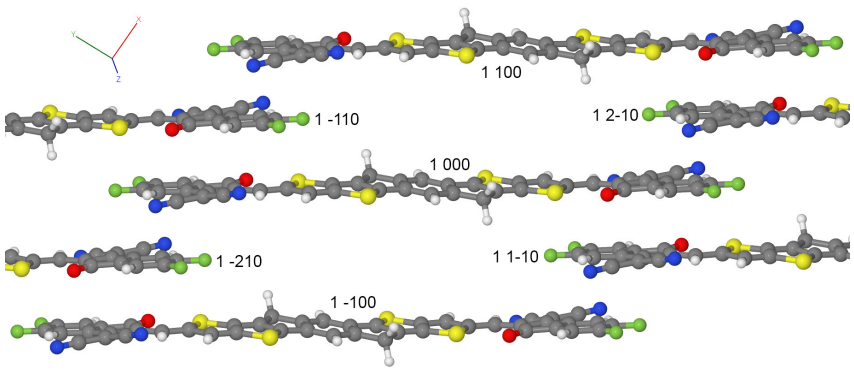
$\Longrightarrow$ 2-LUMO model fails

How inaccurate ECG influences calculated mobility?

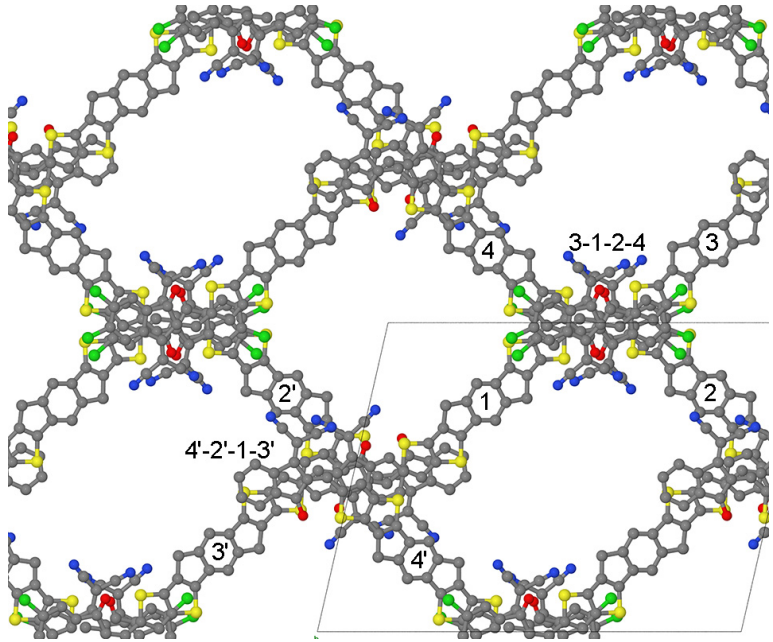
In a solid A-D-A molecules form 2 intermolecular contacts per A

$\Longrightarrow$ coordination number of electronic connectivity graphs is 3

$\Longrightarrow$ either honeycomb (2D) or K4/Laves (3D) lattices



Brickwork packing
(ITIC-4F)



Wire-mesh packing
(ITIC-2Cl/asym, o-IDTBR, Y6)

Analysis of effective masses and hopping amplitudes (∼mobility) shows that the main effect is that 1-site model underestimates intermolecular couplings by factor of 2

Methodology of ECG

Requirements:

- Robustness of CG basis wrt molecular fluctuations
- Robustness and scalability of CG algorithm
  — Implementation
- Quality control of CG basis and matrix elements
- Accurate extrapolation to infinite system (if needed)

Robustness and scalability of ECG algorithm

Algorithm

1. Determination of coarse-grained (CG) basis
2. Parametrization of tight-binding Hamiltonian
3. Quality control (for new systems)

ECG might often look trivial but here "the devil is in the details", so there are many challenges in practical implementation:

- Large systems (large supercells)
- Entangled bands (overlapping in energy or hybridized)
- Cluster approximation to infinite systems
- Complex cases (e.g. excited states, no fixed SCF)
- Parametrization of complex models (e.g. electron-phonon)
- SCF and total energy in CG-basis

Determination of coarse-grained (CG) basis

In what follows we consider coarse-graining of molecular orbitals (MOs)

Algorithm

1. MOs of interest must be representable in CG basis
   $\Longrightarrow$ CG MOs are usually localized MOs (LMOs)
2. Tight binding parameters (ε_i and t_{ij}) are obtained by projection of Fock matrix of model fragments onto LMOs
3. Effective Hamiltonian of a large system is obtained by its fragmentation into the model fragments

Why LMO-based approach is working:

- Slater determinant is invariant under rotations of 1e orbitals
- Locality of phenomena: locality of 1e-density matrix and Hamiltonian, often only local SCF is important

  [V Heine, Solid State Phys 35, 1 (1980)]

Localization of molecular orbitals for coarse-graining

- Localization procedure is not unique Acc Chem Res 47, 2758 (2014) (trade off between spatial and energy localization)
- Projection approaches are scalable and robust
- Localization to unit cell is performed by Wannier functions
- The largest problem is bands entanglement (e.g. π and σ)
- LMOs can be used to get SCF (MOZYME code in MOPAC, fragmentation methods Chem Rev 112, 632 (2012))

Algorithm

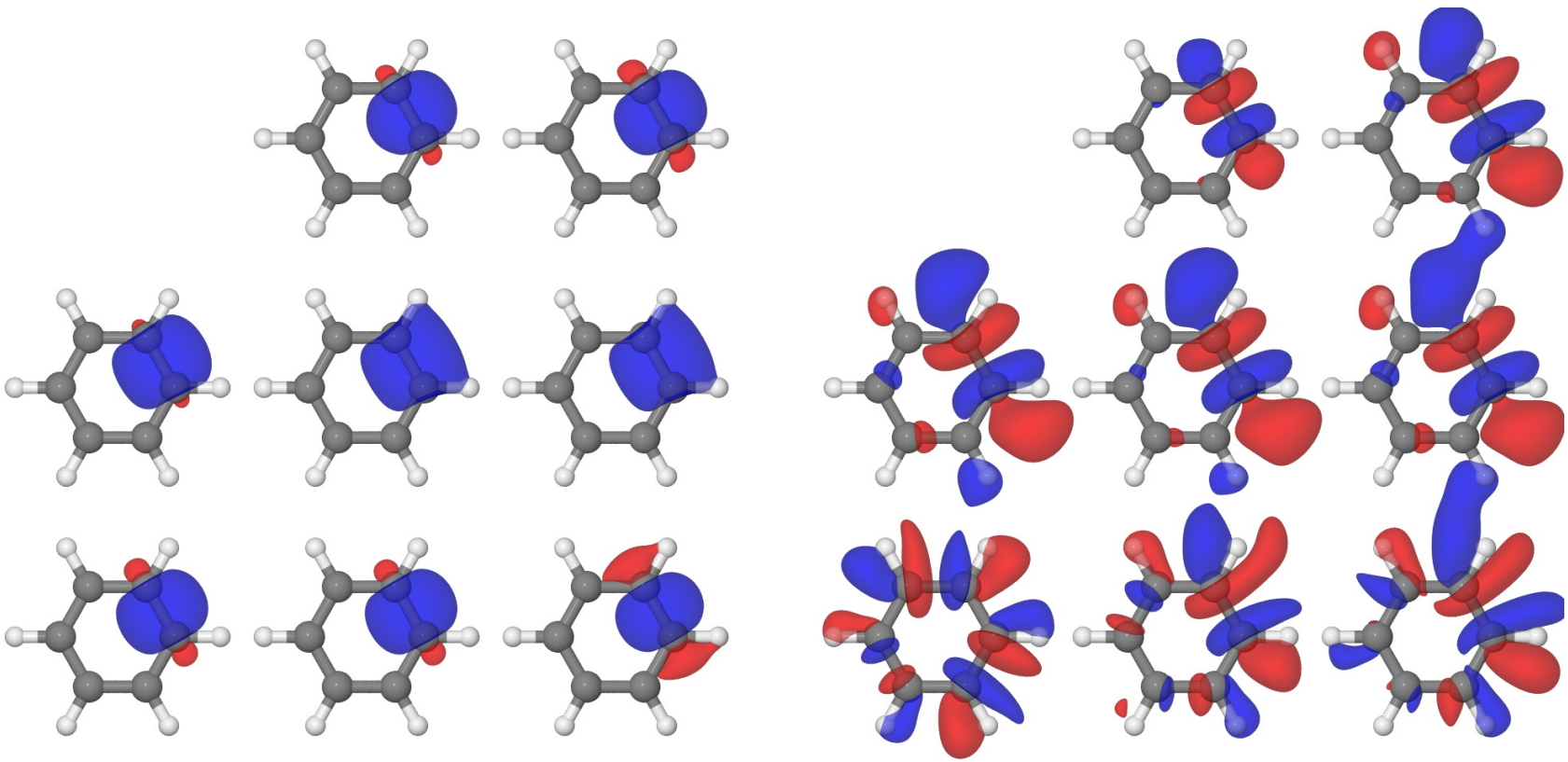
1. Reduce to unit cell (Wannier, or cluster in PBC or OBC)
2. Define projector (quality of the projector is critical)
3. Select initial MO subspace, project, and refine MOs
4. Convert LMOs to CG-basis (e.g. by block diagonalization)

Projecting MO to LMO

Open-source code MolMod/LocalizeMO

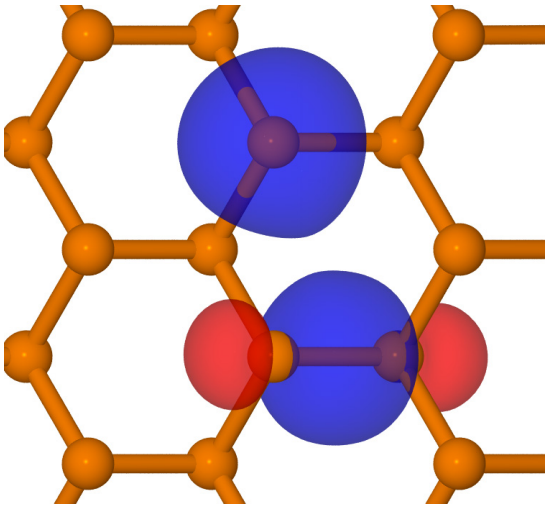
Project on/in/out by SVD of M^+SL or $M^+P'SPM$,
where S – overlap, M – MOs, L – pre-LMOs, P – projector

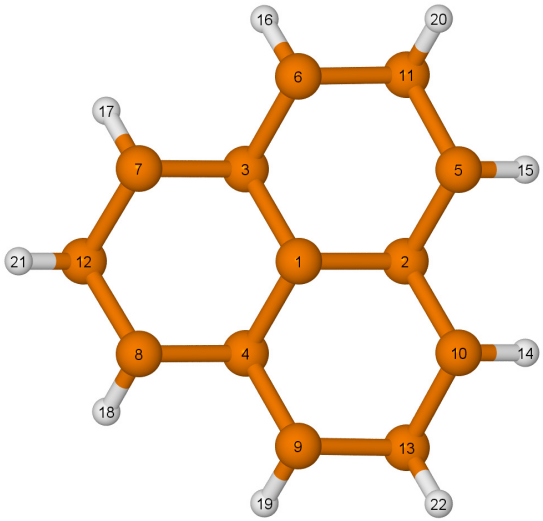
Possible projector: subset of AOs (geometrical fragmentation)
Possible pre-LMOs: LMOs by other method ('ad hoc' approach)
Example: benzene σ and σ^* LMOs (NBO/on/in/out/...)

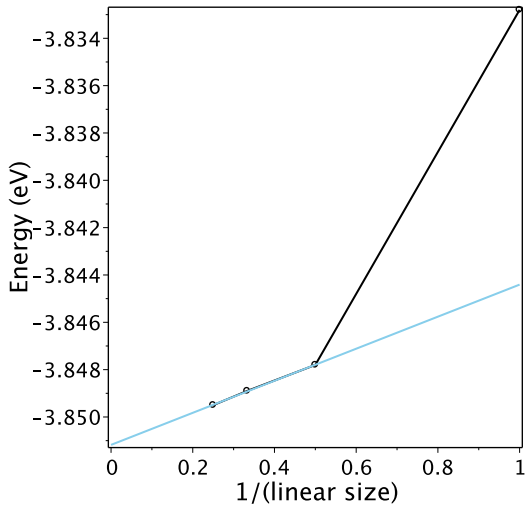


Locality and fast convergence of tight-binding parameters...

Illustrated by NBO analysis of the valence band of blue (A7) phosphorene





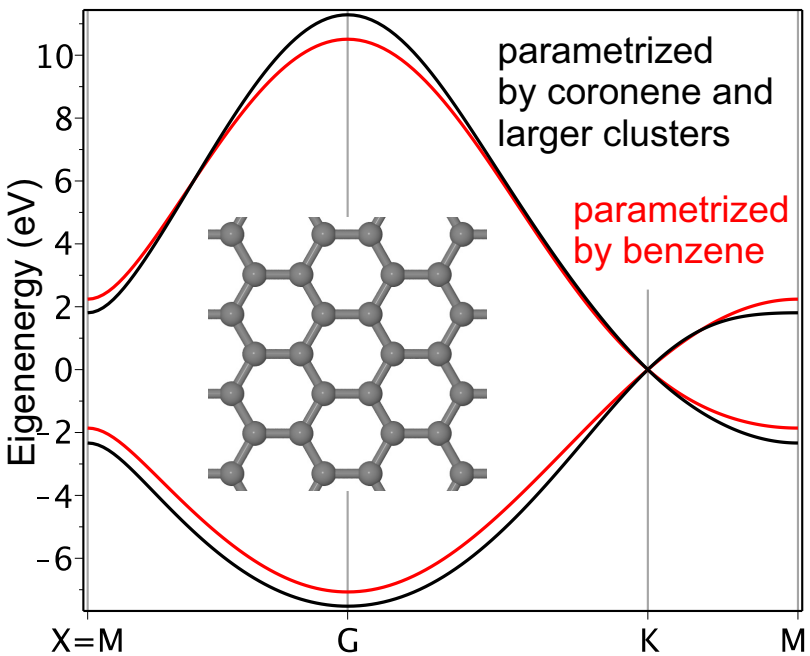


- There are two symmetry unique NBOs
- Error in estimation energy difference between them is 270 meV for the smallest cluster and 42 meV for the next one in series
- Error in estimation coupling between them is 18 meV for the smallest cluster and 3 meV for the next one in series

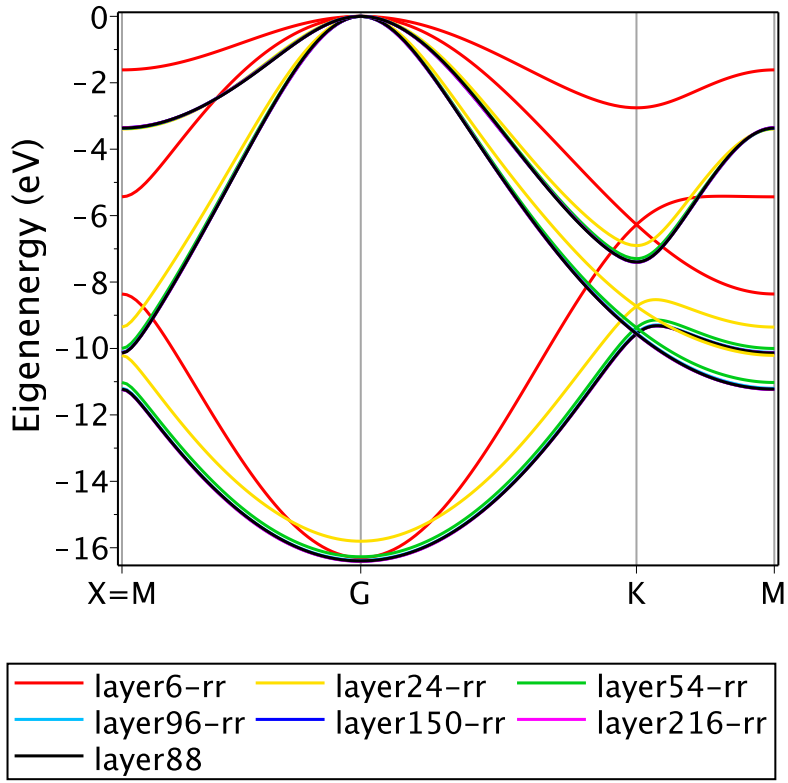
Already the smallest fragment can be used for transfer integrals

... small clusters are enough for parametrization

Example of hydrogen-passivated graphene clusters – graphene nanoflakes



Graphene π-band
(24-atom cluster is enough)



Graphene valence σ-band
(54-atom cluster is enough)

Methodology of ECG

Requirements:

- Robustness of CG basis wrt molecular fluctuations
- Robustness and scalability of CG algorithm
- Quality control of CG basis and matrix elements
- Accurate extrapolation to infinite system (if needed)
  — Efficient parametrization of tight-binding models

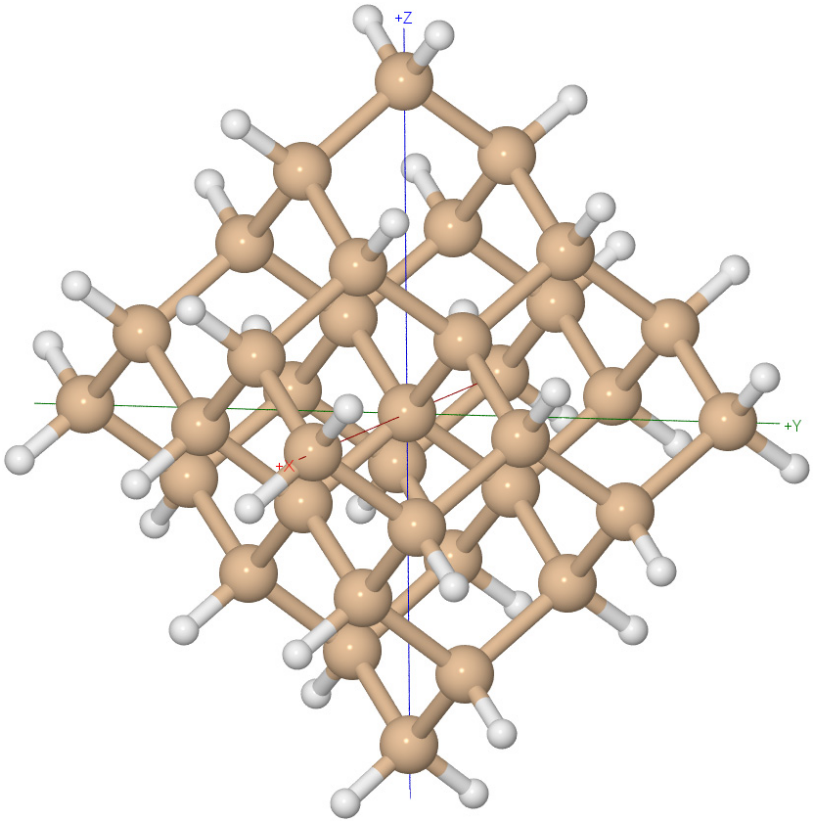
Efficient parametrization of tight-binding (TB) models

Approaches to first-principles parametrization:

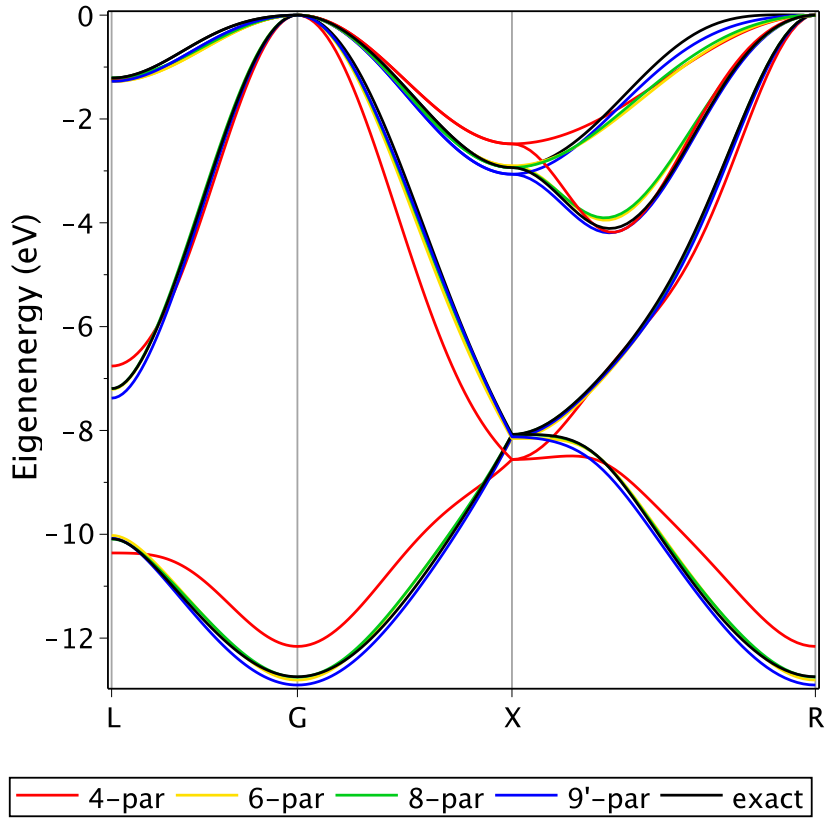
- Fit by one-electron energies (mainstream)
  - \+ Only energies are needed
  - – Only energies are accurate for sure
  - – Systematic accuracy improvement requires ML-fitting
  - ▶ Best for simple and minimal TB models
- Derive from Fock matrix (natural to ECG)
  - \+ Numerically exact in principle (no fitting)
  - \+ Transferable TB elements (e.g. cluster-to-solid)
  - – More complicated in implementation
  - – Wave-function is needed
  - ▶ Best for complex and accurate TB models

Parametrization of TB models from cluster calculations

https://cmsos.github.io/tbm



Input: Cluster calculations (DFT)



Output: Symmetry-unique TB elements

Technical challenges and solutions

https://cmsos.github.io/tbm

Use symmetry for TB model formulation, solution, and parametrization

- 10 bonds in diamond $\Longrightarrow$ 100 symmetry-unique dimers $\Longrightarrow$ 5000 dimers per primitive cell $\Longrightarrow$ 10^7 comparisons
- FiniteGroups package identifies 1500 symmetry-unique dimers in diamond per minute
- Wave-function symmetry is currently supported only for 1 orbital per site models – enough for valence *sp*-orbitals

Use library of electronic structure prototypes

- Structurally different materials might have the same TB model
  - Triangular lattice – herringbone packing
  - Honeycomb lattice – brickwork packing of A-D-A molecules
  - K4 lattice – wiremesh packing of A-D-A molecules
- We can compare materials by their effective Hamiltonians rather than computed properties

Applications of ECG

Calculate electronic structure of infinite systems:

- Solids of large flexible molecules such as A-D-A molecules

  Chem Mater 33, 966 (2021); JCP 159, 024107 (2023)

- Polymers: electronic structure of a single polymer chain

  Chem Sci 8, 1146 (2017), Solar Energy 198, 605 (2020)

- Polymers: electronic structure of crystalline polymers

  work in progress

Analyze localized states in infinite systems:

- Metal-organic polymers and frameworks

  Chem Sci 13, 8161 (2022)

- Polarons in non-polar semiconductors

  JPCL 12, 4674 (2021)

- Traps in organic semiconductors

  work in progress

Summary

- Electronic Coarse-Graining is a powerful tool (sometimes the only one) for calculation and analysis of electronic structure of complex systems:
  - Anything that can be monomerized (polymers, COFs, MOFs)
  - Solids and clusters of large molecules
  - Defects in solids
  - Infinite systems from finite-size calculations

Project web-page:

- Electronic Coarse-Graining
  https://zhugayevych.me/research/ECG.html
- Computational Materials Science of Organic Semiconductors
  https://cmsos.github.io