

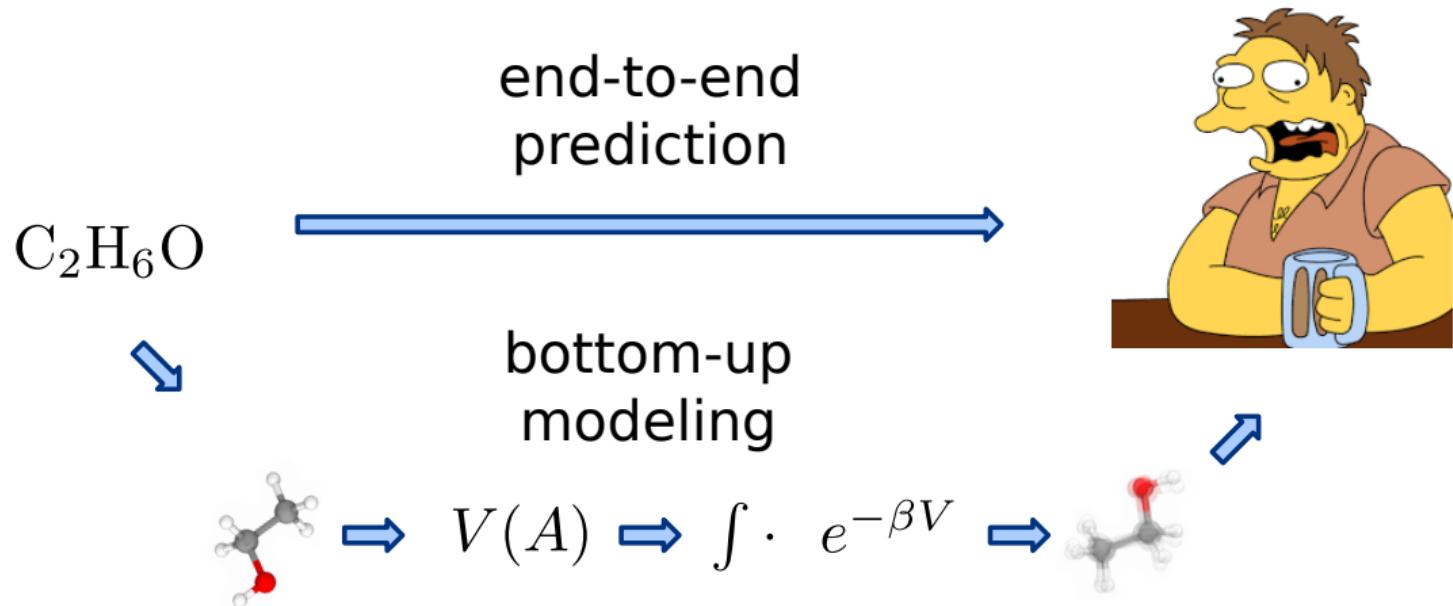
Introduction to atomic-scale machine learning

Michele Ceriotti
cosmo.epfl.ch

- Supervised and unsupervised learning in chemistry: from properties to insights
 - Representing atomic structures
 - Linear methods: PCA, ridge regression, PCovR
 - The kernel trick
- Atom-centered representations for atomistic machine learning
 - Mathematical requirements for a well-behaved featurization
 - Construction of structural representations as correlations of the atom-density
 - Invariant and equivariant feature construction
- Equivariant models for covariant properties and electronic structure
 - Equivariant models of tensorial properties
 - Extension of atom-centered descriptors to N -center features, and message passing
 - Equivariant learning and electronic-structure modeling

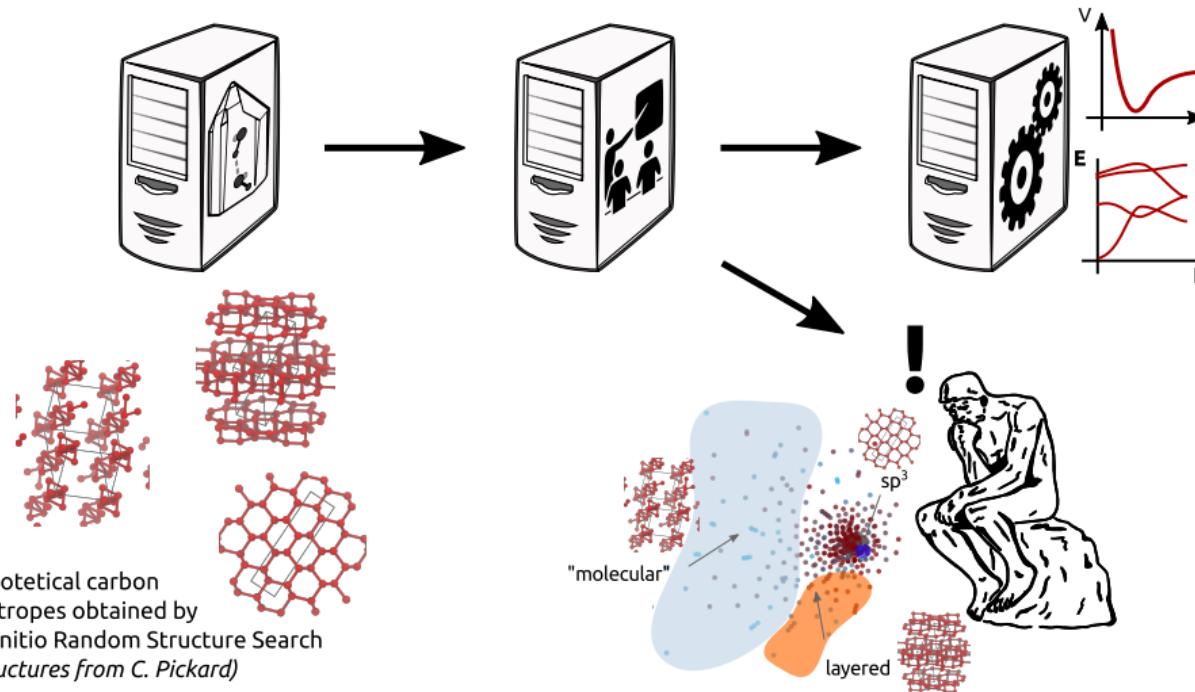
Bottom-up modeling for chemistry

- Most widespread application of ML targets prediction of the end property
- This course focuses on the application as part of “bottom-up”, physics-based modeling



Black-box predictions and intuitive understanding

- Machine learning can be used to predict molecules or materials' properties or to process large amounts of complex simulation data to produce a human-understandable form
- Focus on interpretability helps achieve both accurate predictions and intuitive understanding



Hands-on examples

- Will be showing examples of how to compute many of these models.
- Jupyter notebooks available for download on <https://github.com/ceriottm/ale-notebooks>
- These are *not* practical examples: over-simplified models and over-verbose implementations, purpose is to clarify concepts

```
if chemiscope.jupyter._is_running_in_notebook():
    from IPython.display import display
    display(cs)
```

-0.05 0 0.05
PCA[1]

structure 1 atom 2

2. A structure-property map for azaphenacene organic conductors

This examples uses a small dataset of 156 molecular crystals (Yang et al., Chem. Mater. (2018)), assembled from 28 different structural isomers of azaphenacene derivatives. Lattice energy and electronic mobility have been computed for each structure, as well as a favorable classification of the resulting materials as conductors or insulators based on the mean number of hydrogen bonds per molecule.