

Atoms and computers

MICHELE PARRINELLO

USI, Faculty of Informatics, Institute of Computational Sciences, Lugano
ETH, Department of Chemistry and Biotechnologies, Zurich

A grim outlook



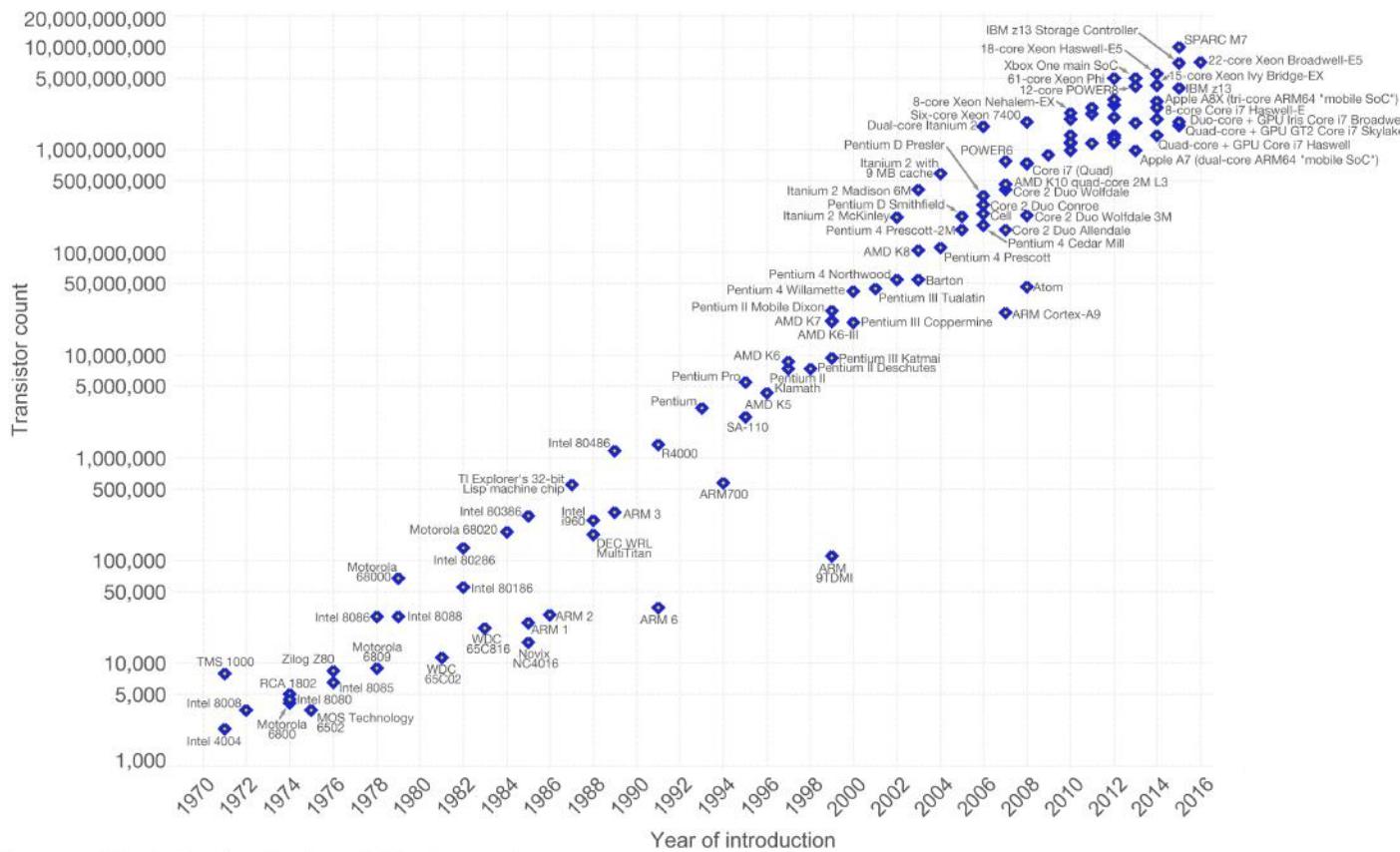
P. A. M. Dirac
Proc. Roy. Soc.
Ser. A, 123, 714 (1929)

The fundamental laws necessary for the mathematical treatment of a large part of physics and the whole of chemistry are thus completely known, and the difficulty lies only in the fact that application of these laws leads to equations that are too complex to be solved.

Moore's law

Moore's Law – The number of transistors on integrated circuit chips (1971-2016) Our World is Shrinking

Moore's law describes the empirical regularity that the number of transistors on integrated circuits doubles approximately every two years. This advancement is important as other aspects of technological progress – such as processing speed or the price of electronic products – are strongly linked to Moore's law.



Data source: Wikipedia (https://en.wikipedia.org/wiki/Transistor_count)

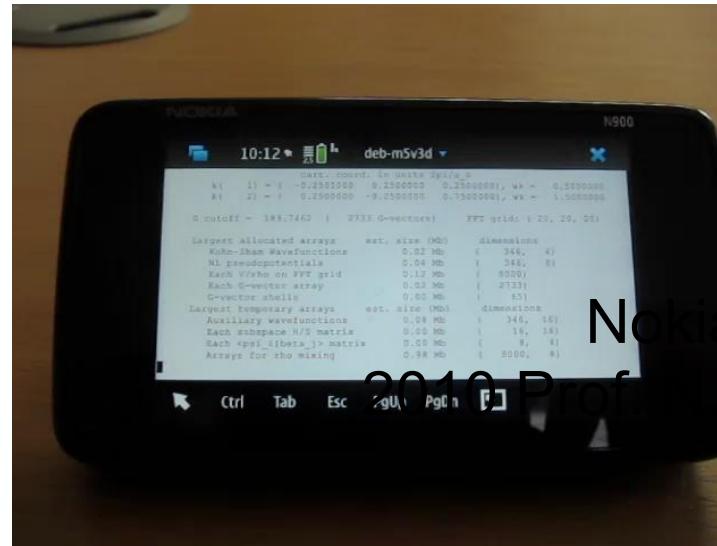
The data visualization is available at OurWorldInData.org. There you find more visualizations and research on this topic.

Licensed under CC-BY-SA by the author Max Rose

Computer evolution during my career



CDC CYBER 170
Trieste 1984/85



Nokia N900
2010 Prof. Marzari, EPFL

The genius of Fermi

266.

STUDIES OF NON LINEAR PROBLEMS

E. FERMI, J. PASTA, and S. ULAM
Document LA-1940 (May 1955).

ABSTRACT.

A one-dimensional dynamical system of 64 particles with forces between neighbors containing nonlinear terms has been studied on the Los Alamos computer MANIAC I. The nonlinear terms considered are quadratic, cubic, and broken linear types. The results are analyzed into Fourier components and plotted as a function of time.

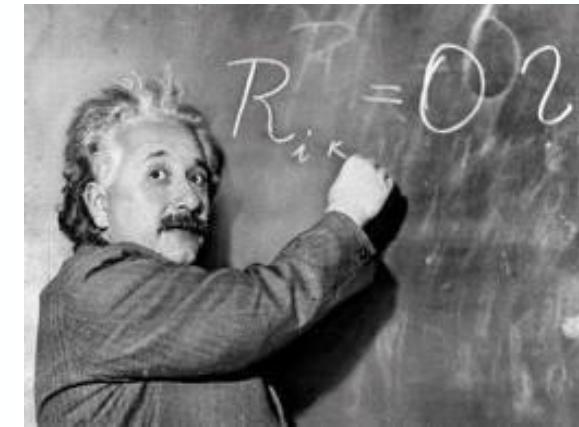


The triangle of science

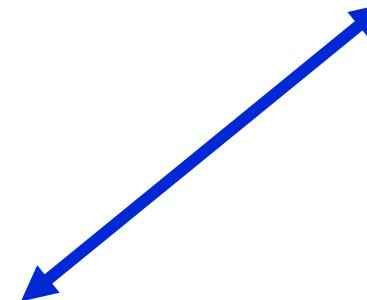
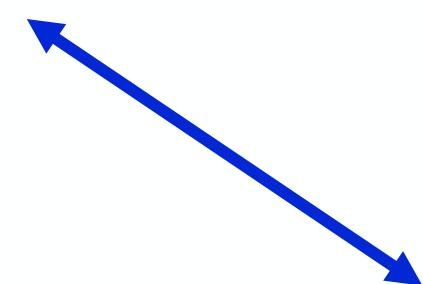
Experiment



Theory



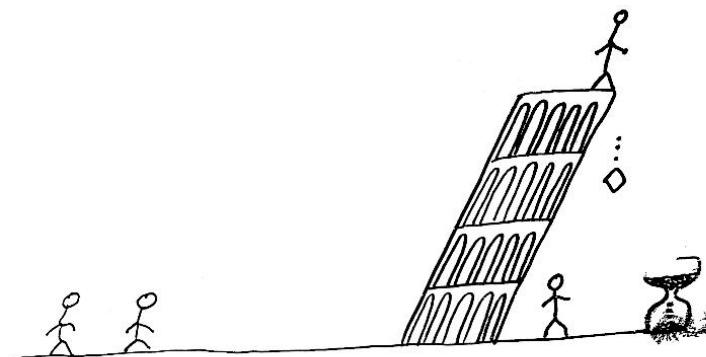
Simulation



Galileo Galilei and Computational Physics



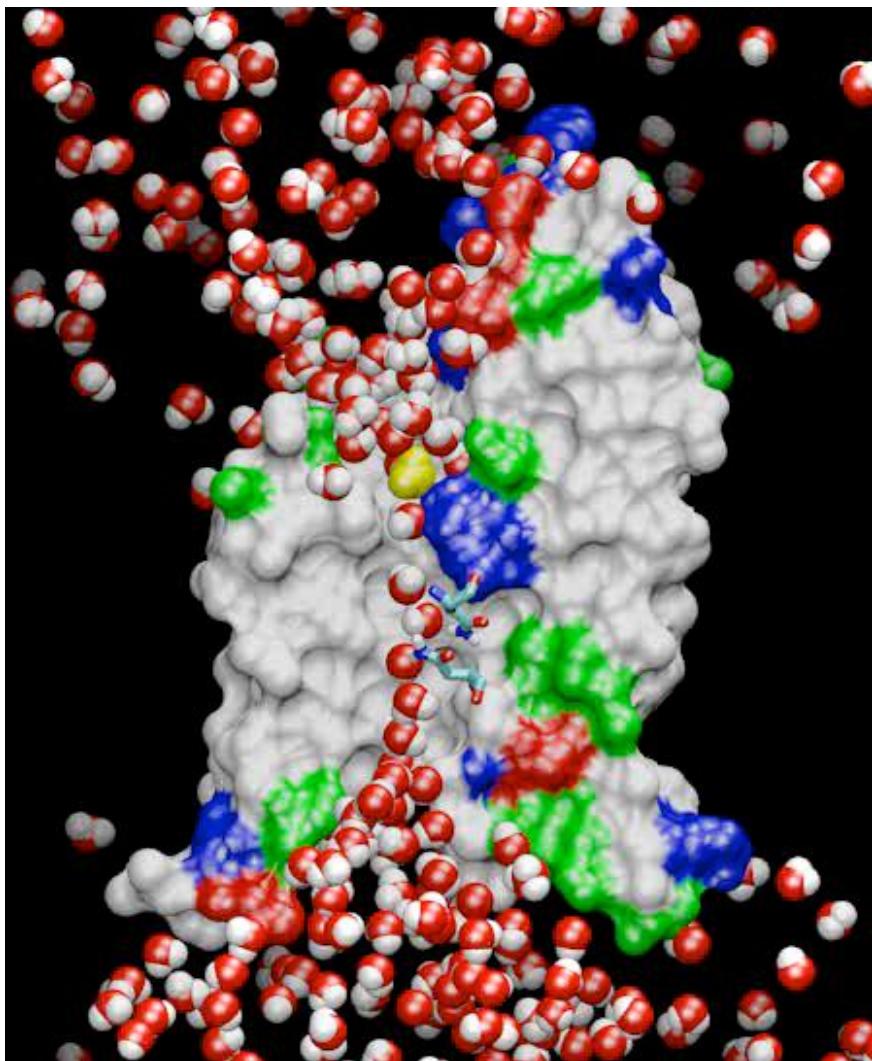
When EXPERIMENTAL PHYSICS was young



“Mr. Galileo!
Why are your students
mindlessly dropping balls
and watching hourglasses?
Shouldn’t you teach them to THINK,
as Aristotle showed how to do?”

A hand written slide from Ken Wilson
Physics Nobel Prize, 1982

Becoming respectable



Acquaporine is a protein that regulates the flux of water across the cell membrane. For resolving this structure Peter Agre got the 2003 Nobel prize.

The movie is downloaded from the Nobel Prize site. The simulation is by K. Shulten, and is presented as a supporting evidence of the correctness of the experimental structure.

What is molecular dynamics?

Molecular dynamics is a set of numerical techniques that allows the behaviour of complex assemblies of molecules such as liquids, solids, surfaces and so on to be simulated.

These simulations:

- Help explain experiments,
- Replace experiments,
- Predict new phenomena,
- Provide invaluable insight,
- Are a kind of virtual microscopy.

The fundamental equation

$$M_I \ddot{R}_I = -\nabla U(R_1, R_2, \dots, R_N)$$

Mass time Acceleration= Force



Is molecular dynamics of any practical use?

The world about us, and biology itself, can be described as resulting from a set of complex physico-chemical reactions.

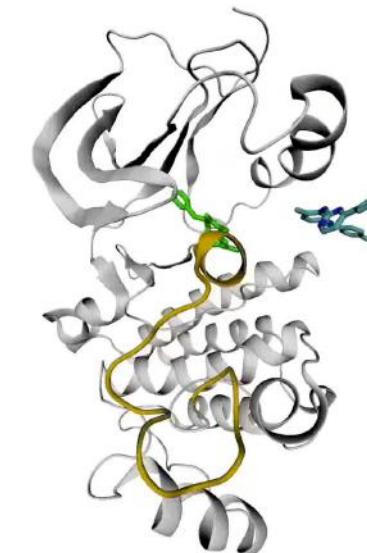
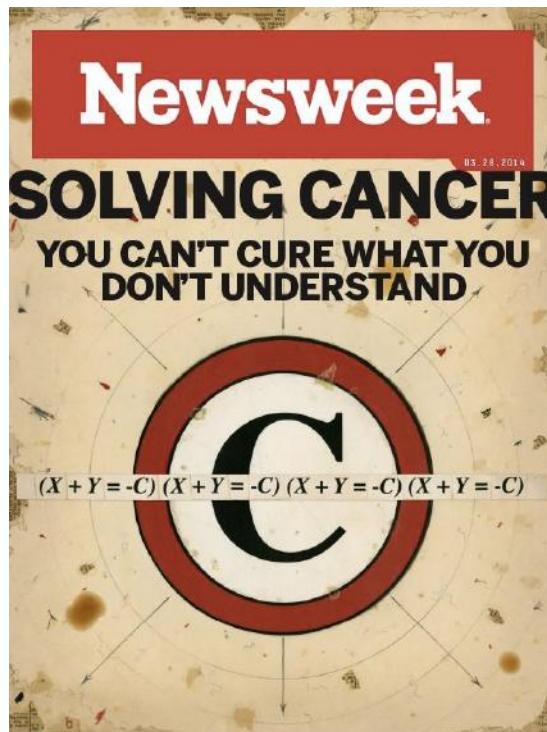
Together with experiments, simulations are an indispensable tool to understand these phenomena.

This understanding can be used to solve many of mankind's problems.

We shall present three representative examples that address, with the help of molecular dynamics, three areas contemporary societal concern.

- **Health**
- **Energy**
- **Environment**

Drug design



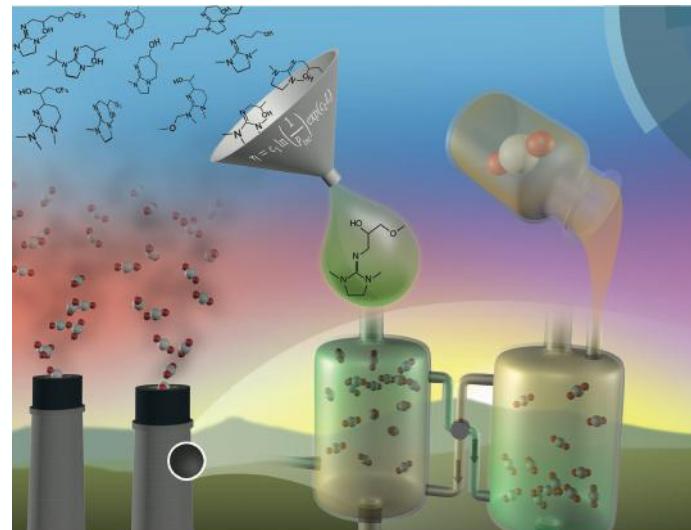
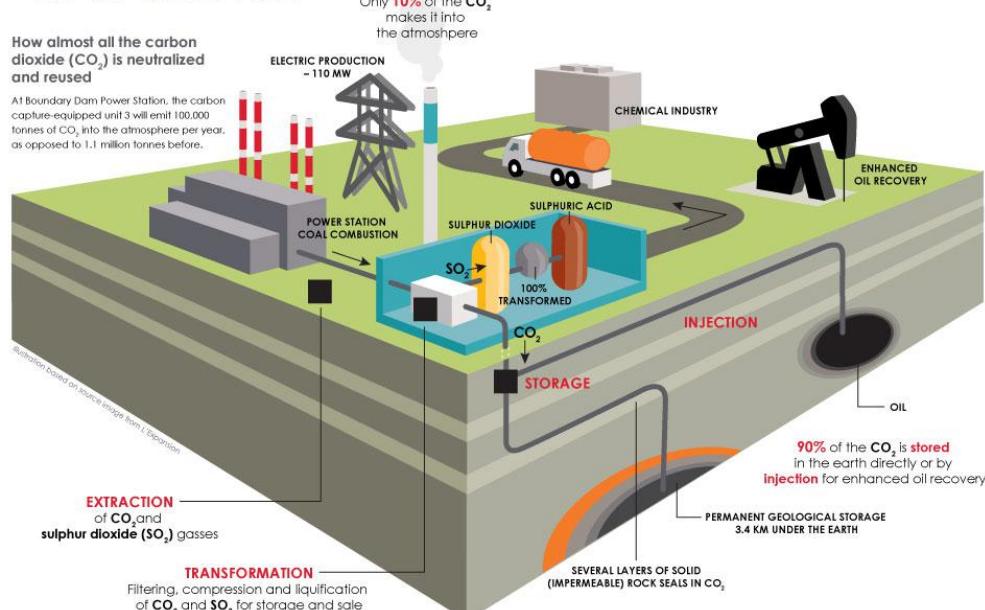
Courtesy F. Gervasio

Sound track G. Piccini

Carbon capture

Carbon capture and storage

AT A GLANCE



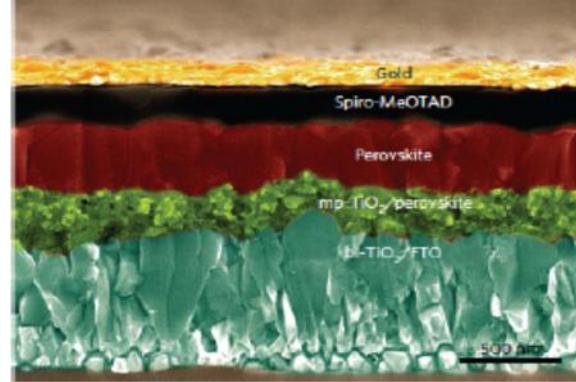
Courtesy V. Glezakou and R. Rousseau

New, cheaper photovoltaic cells

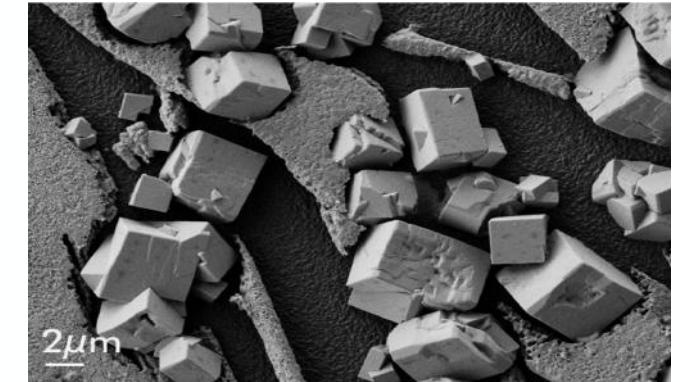


Silicon

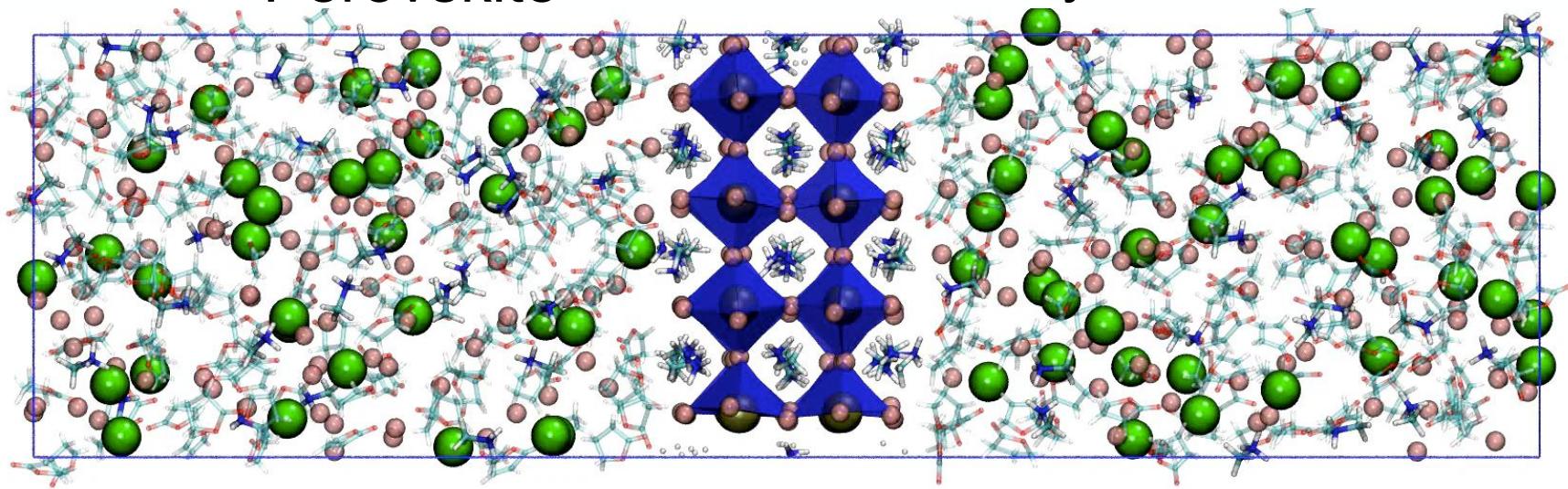
Controlling the quality of the perovskite crystals is essential for efficiency and durability.



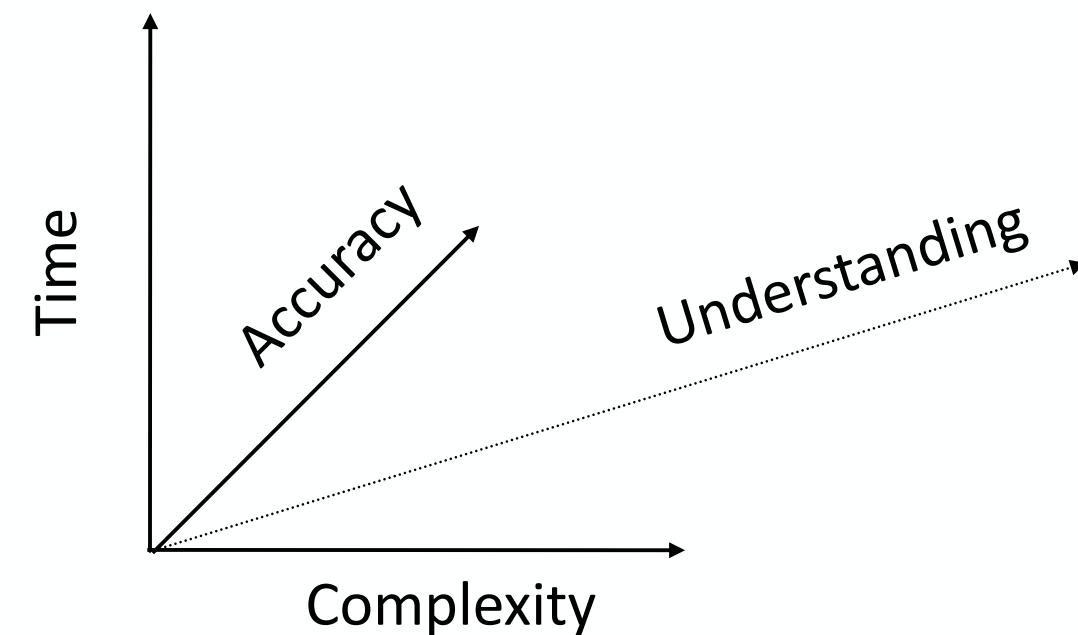
Perovskite



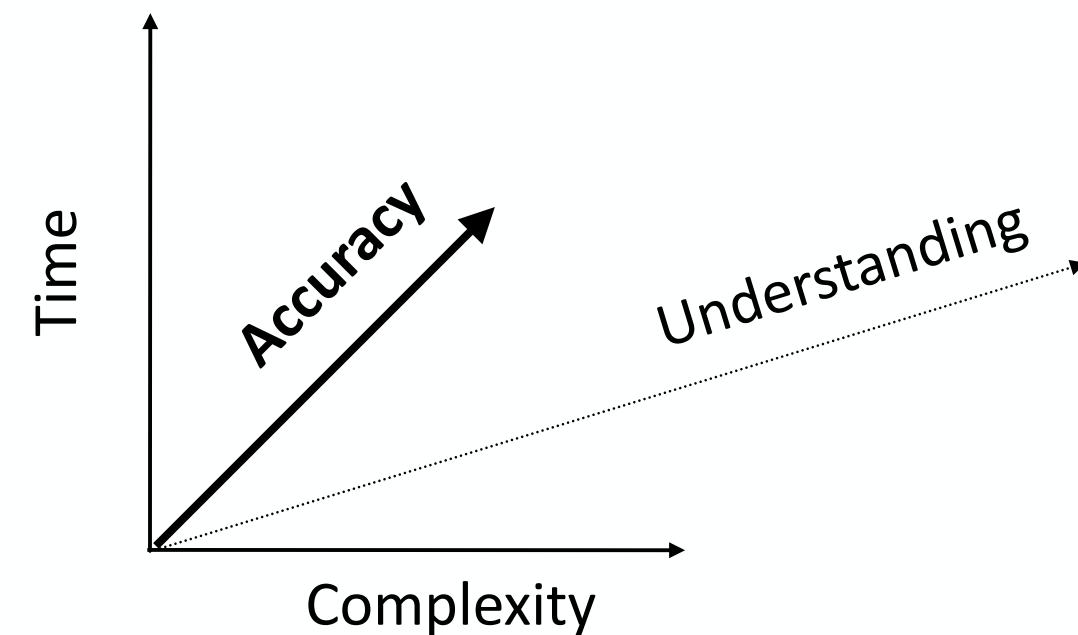
Crystals



The challenges



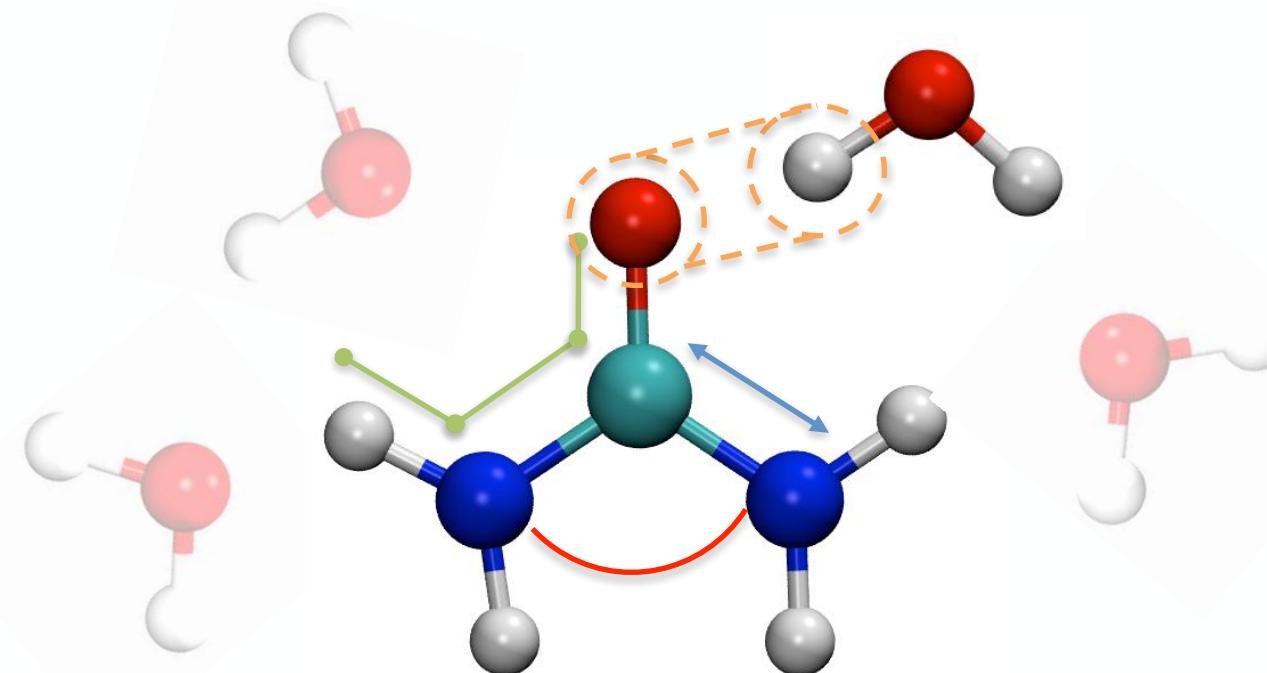
The challenges



How do the forces look like

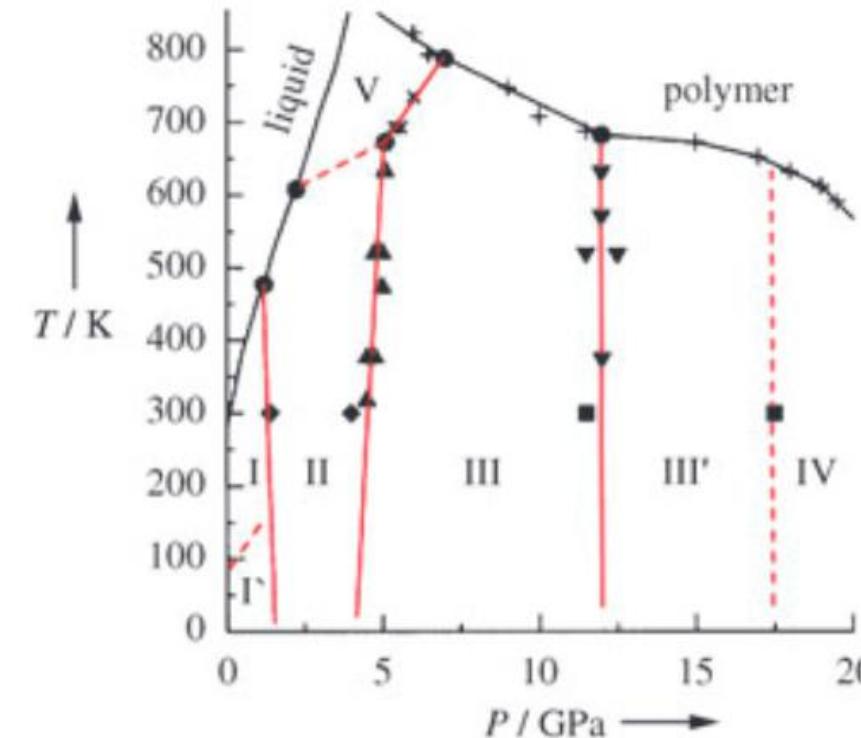
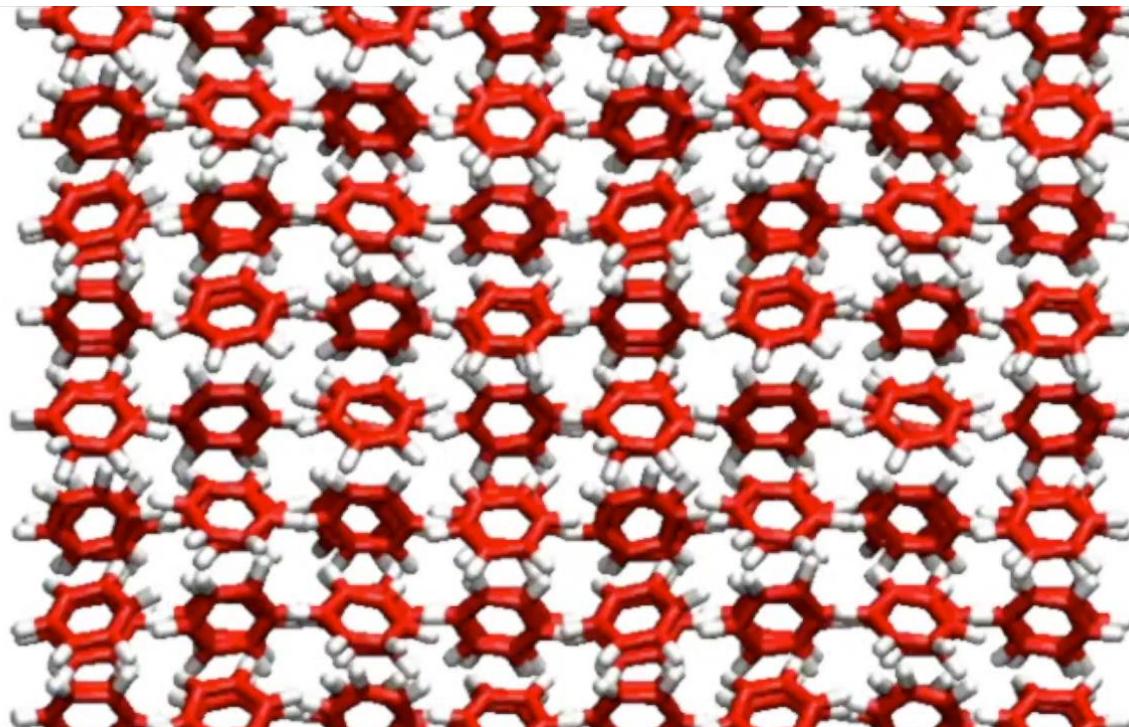
$$E = \sum_{bonds} K_r (r - r_{eq})^2 + \sum_{angles} K_\theta (\theta - \theta_{eq})^2 + \sum_{dihedrals} E_t + \sum_{i < j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} + \frac{q_i q_j}{\epsilon R_{ij}} \right]$$

bond stretching angle bending torsions VdW + Electrostatic interactions

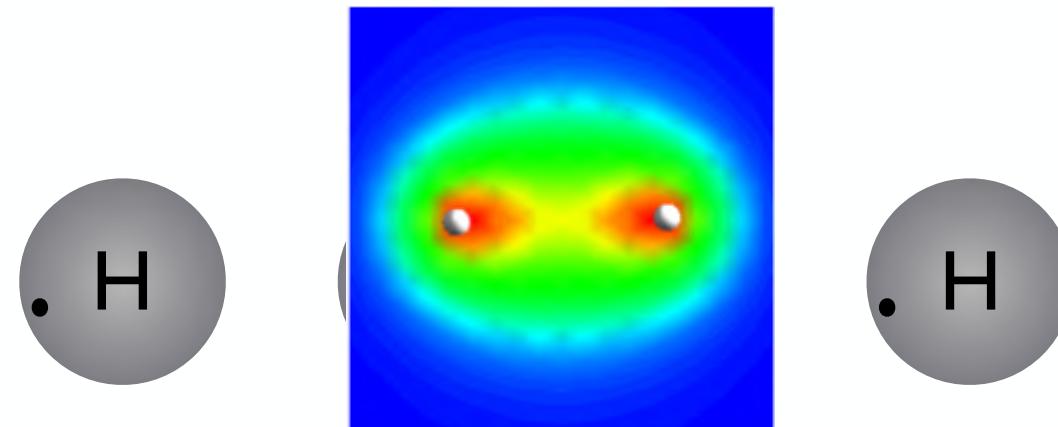
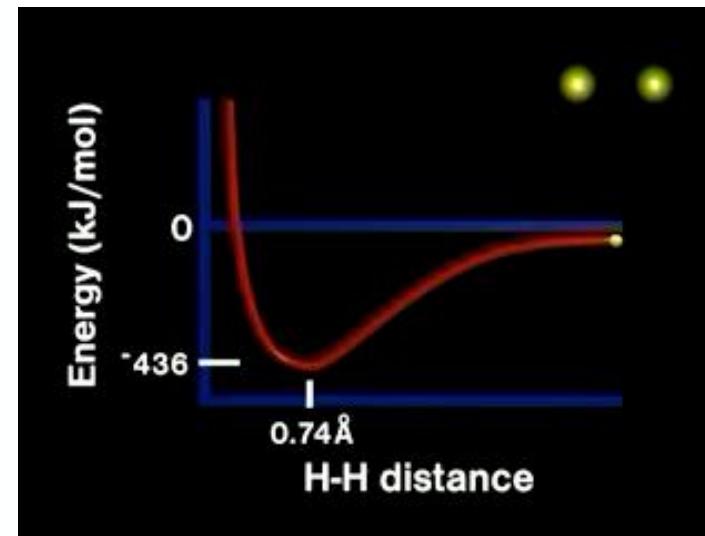
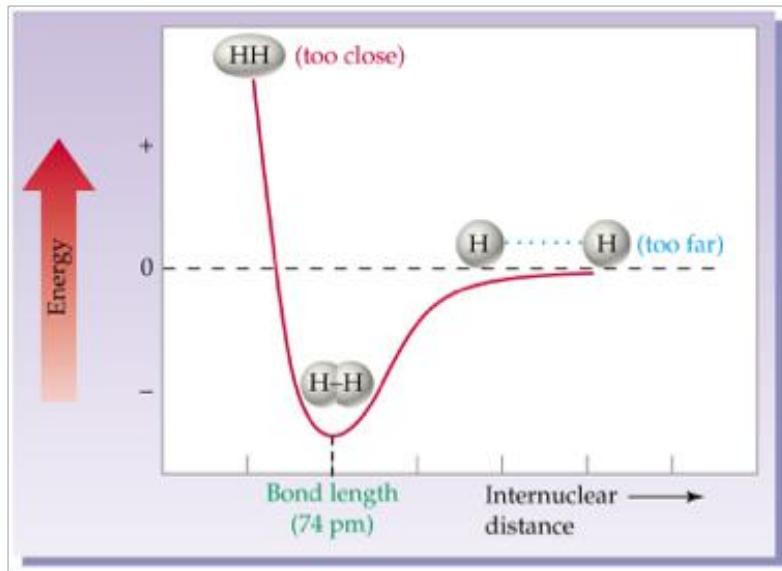


The dance of the atoms

Making benzene molecules dance



Chemical bonds

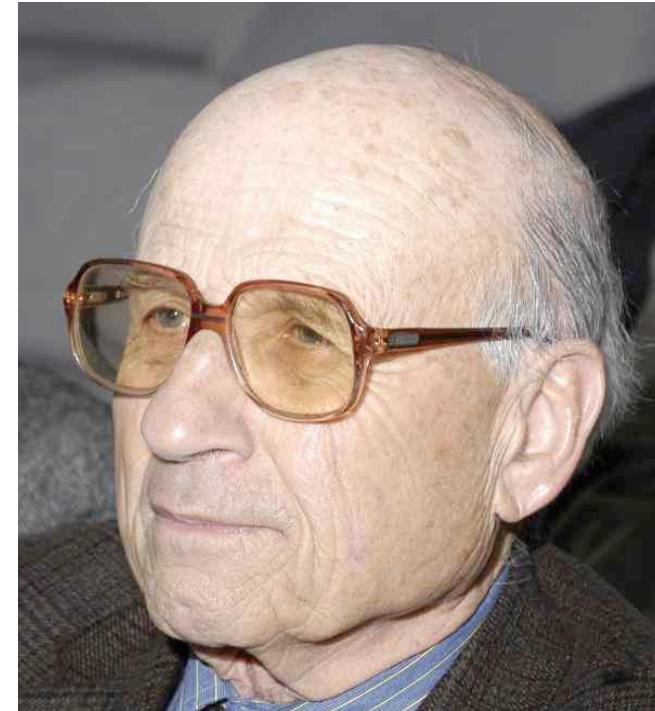


Quantum equations



$$H\psi = E_0\psi$$

Schöredinger equation



$$E_0 = E[\rho_0(r)]$$

Density functional theory

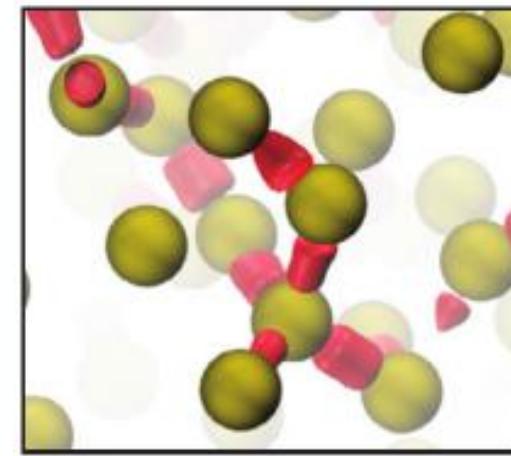
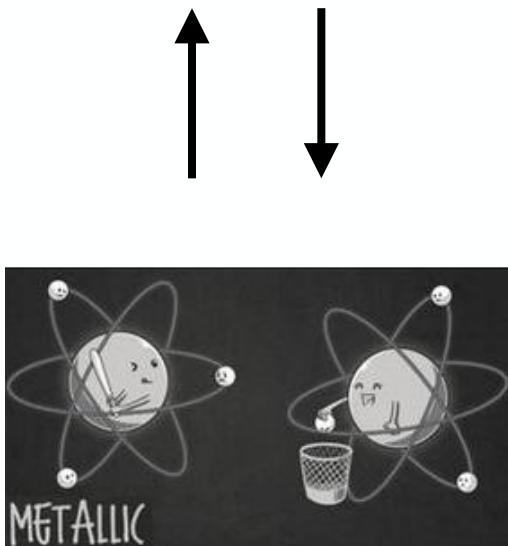
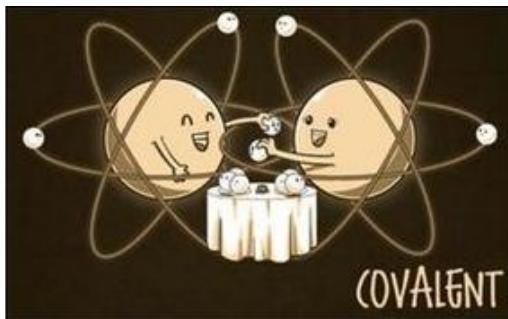
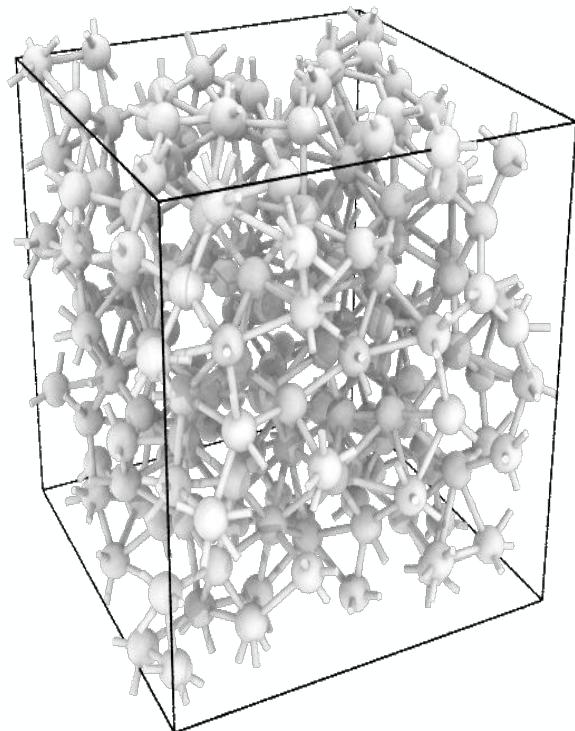
The marriage of two worlds

Molecular dynamics
can describe the complex and dynamic
environment of real life chemistry.

Electronic structure theory
provides the ability to describe the formation
and breaking of chemical bonds.

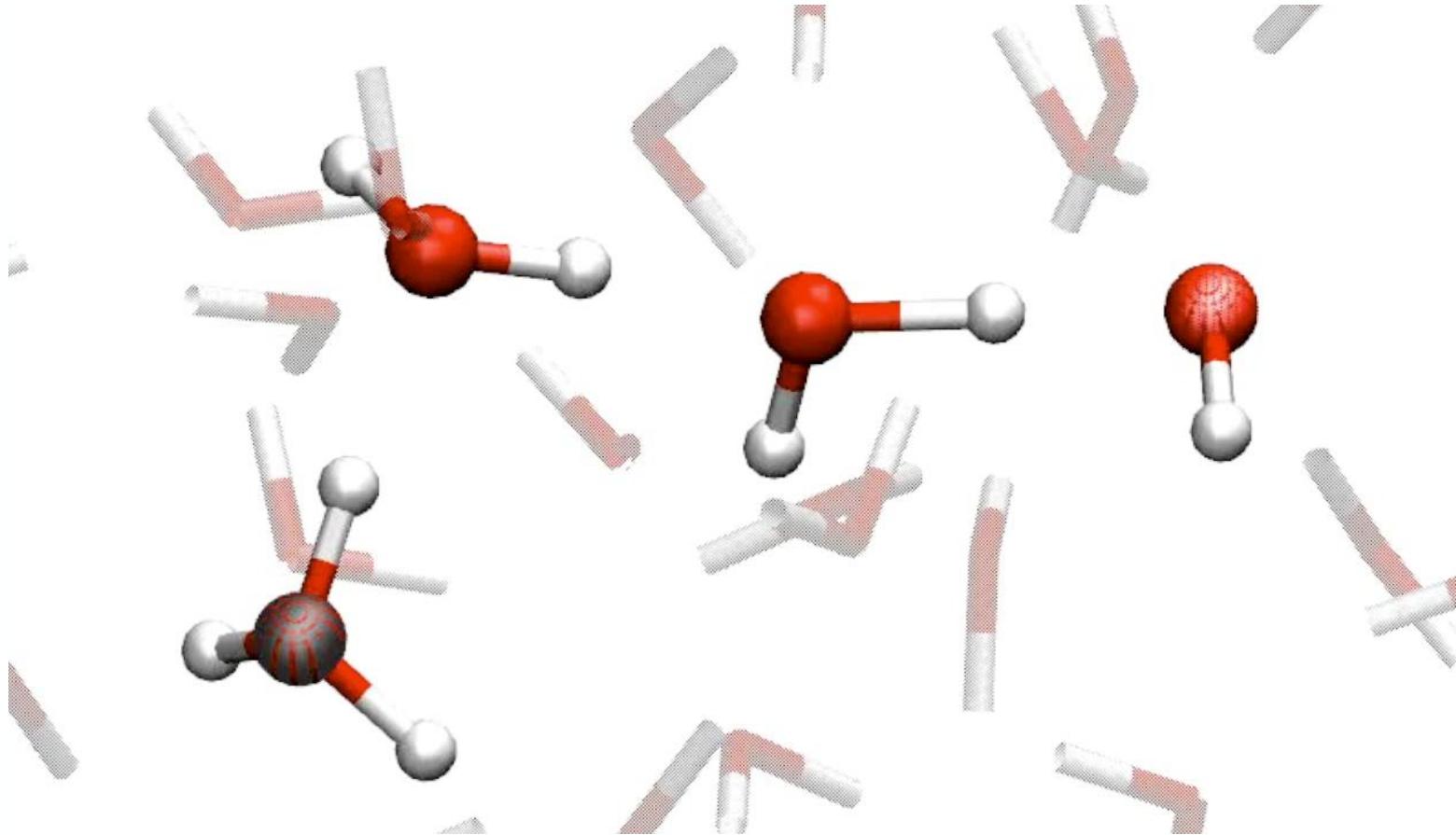


Silicon crystallisation



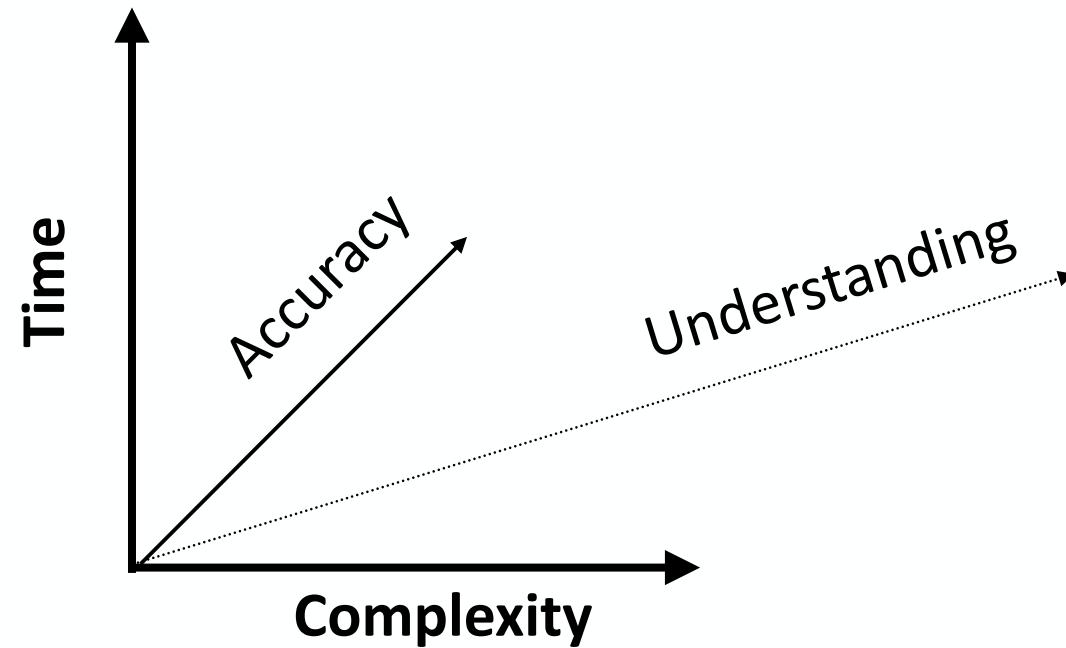
Non local chemistry

Proton diffusion



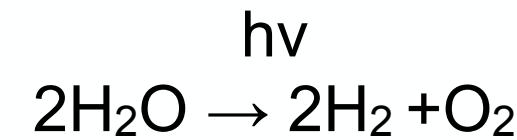
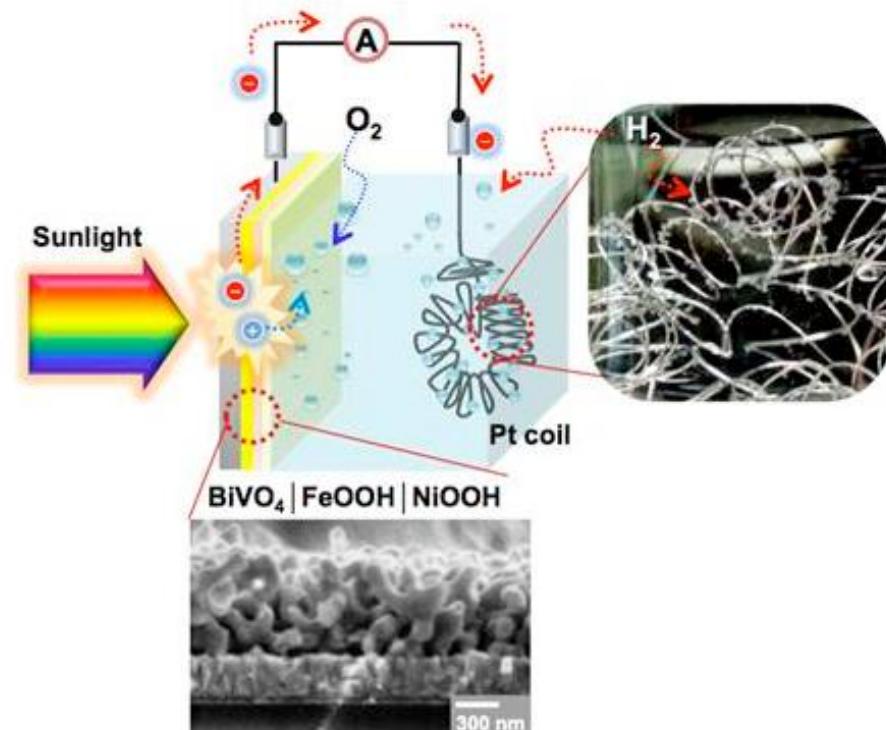
Courtesy Ali Hassanali

The challenges



A complex system

Photo-catalytic water splitting



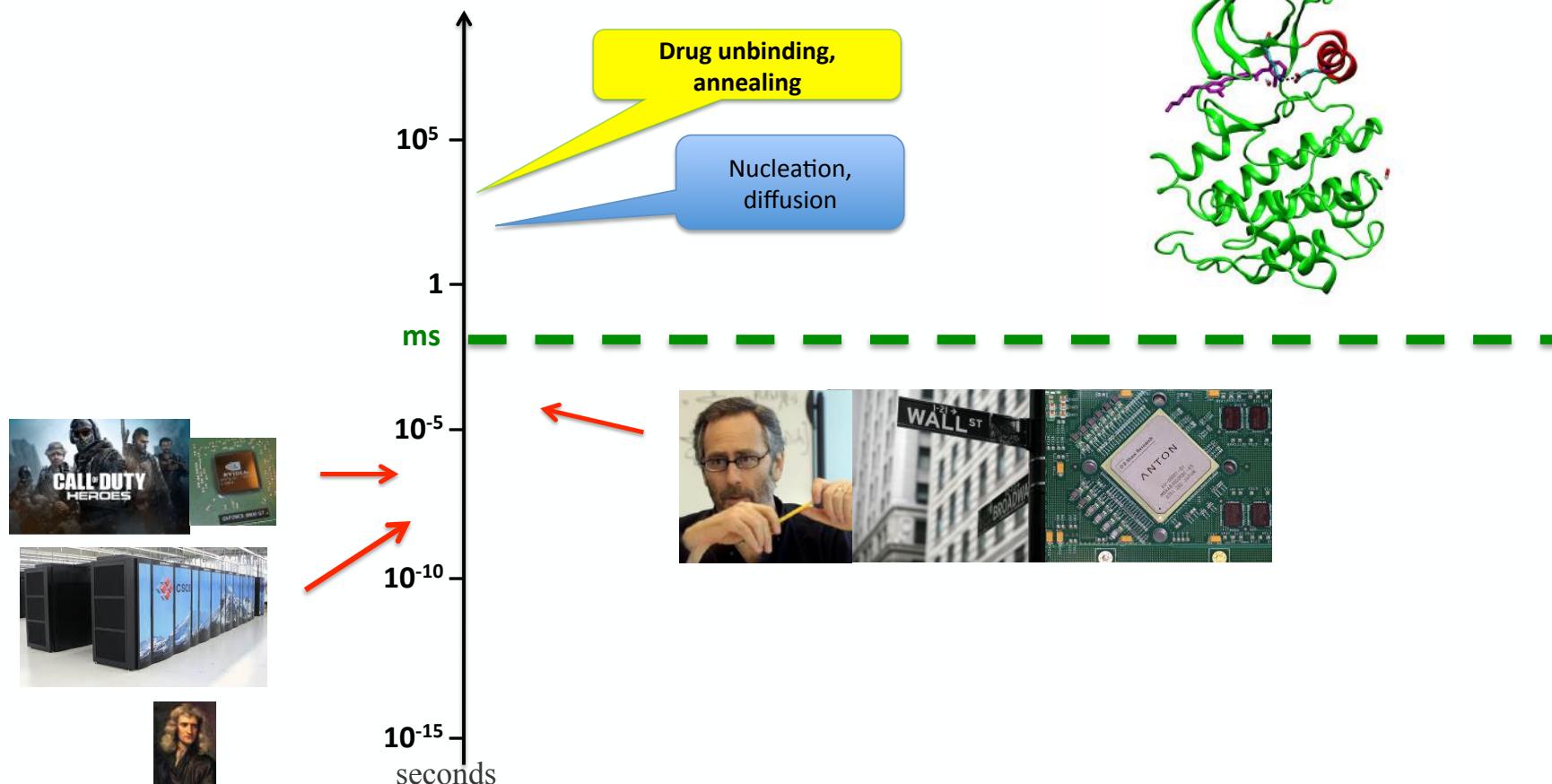
Absorb light

Transport electrons and holes from the solid absorber to the liquid

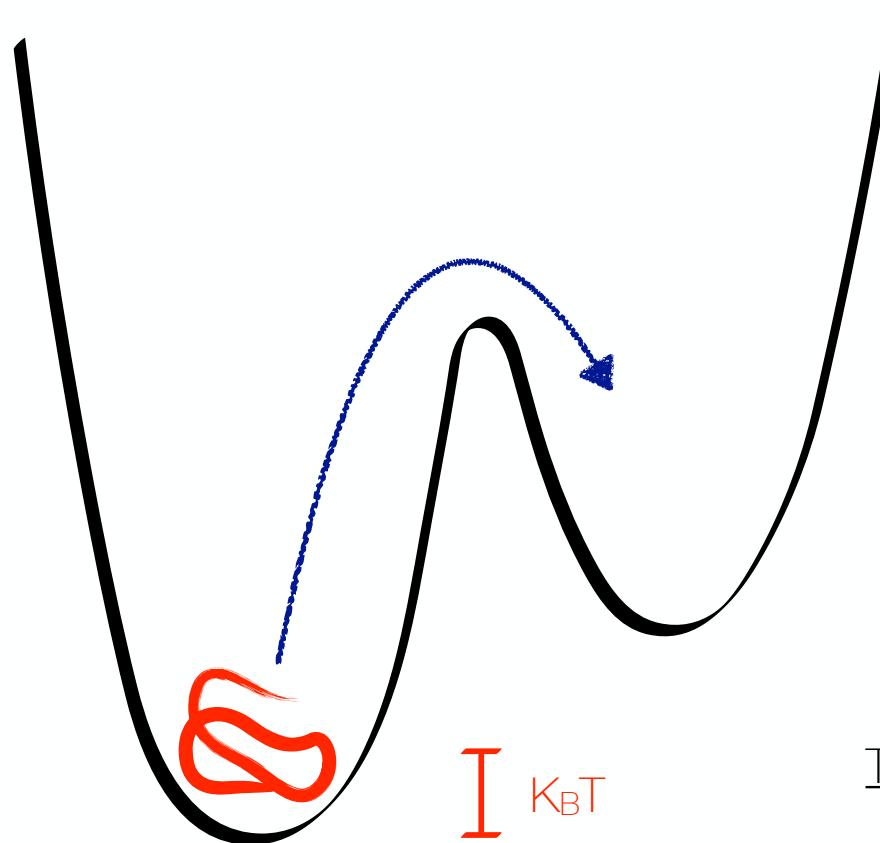
Harvest charges for chemical reaction

Courtesy G. Galli

The time challenge



Energy Barriers and Rare Events



- Large barriers imply long time scales
- Thermal excitation not sufficient in MD

Example:

$$\Delta G = 150 \text{ kJ/mol}$$

$$T = 300 \text{ K}$$

$$k = 4.78 \cdot 10^{-14} \text{ s}^{-1}$$

$$t_{1/2} = 459824 \text{ s} = 5.3 \text{ days}$$

The Higher the barrier the less frequent the transition

A complex problem



Back to the classics for inspiration



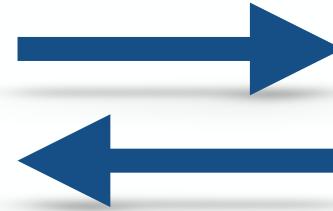
Isaiah 40:4

Every valley shall be raised up, every mountain and hill made low; the rough ground shall become level, the rugged places a plain.

The research program



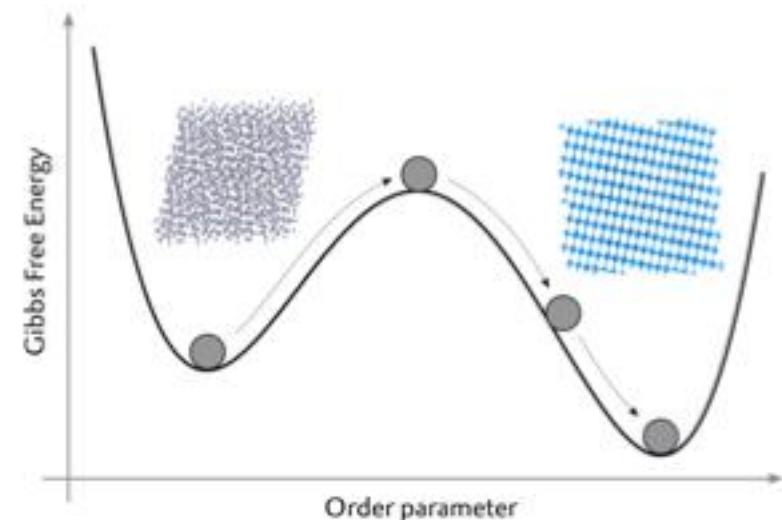
Switzerland



Tuscany

Learning from crystallisation

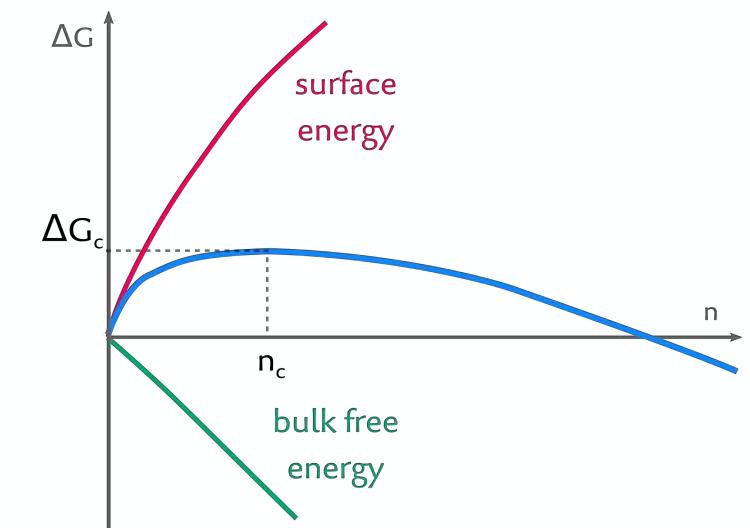
Fluctuations form clusters of the new phase.



Use the cluster size n as order parameter

Free energy cost

$$\Delta G(n) = -\Delta\mu n + c\gamma n^{2/3}$$



Describe the system in a low dimensional space

The collective variables

$$\mathbf{s}(\mathbf{R}) = (s_1(\mathbf{R}), \dots, s_M(\mathbf{R}))$$

The probability distribution

$$P(\mathbf{s}) = \int d\mathbf{R} \delta(\mathbf{s} - \mathbf{s}(\mathbf{R})) P(\mathbf{R})$$

The free energy surface

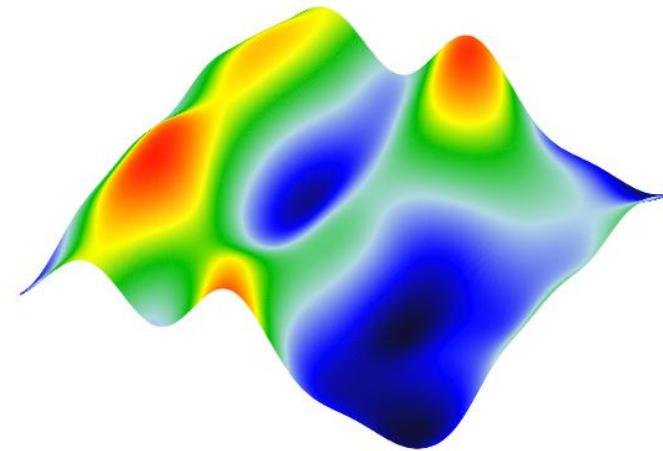
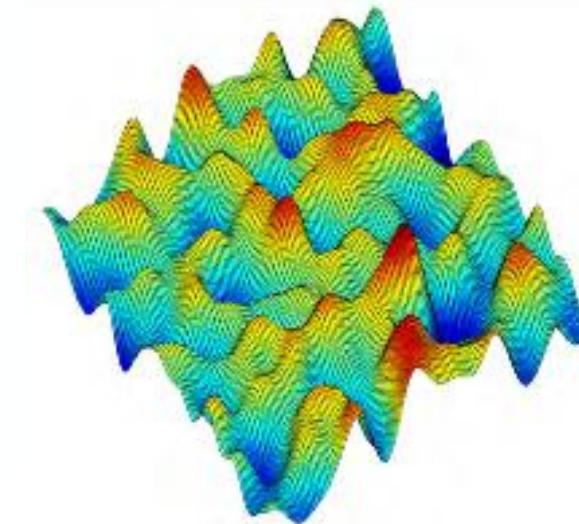
$$F(\mathbf{s}) = -\frac{1}{\beta} \log P(\mathbf{s})$$

A dimensional reduction

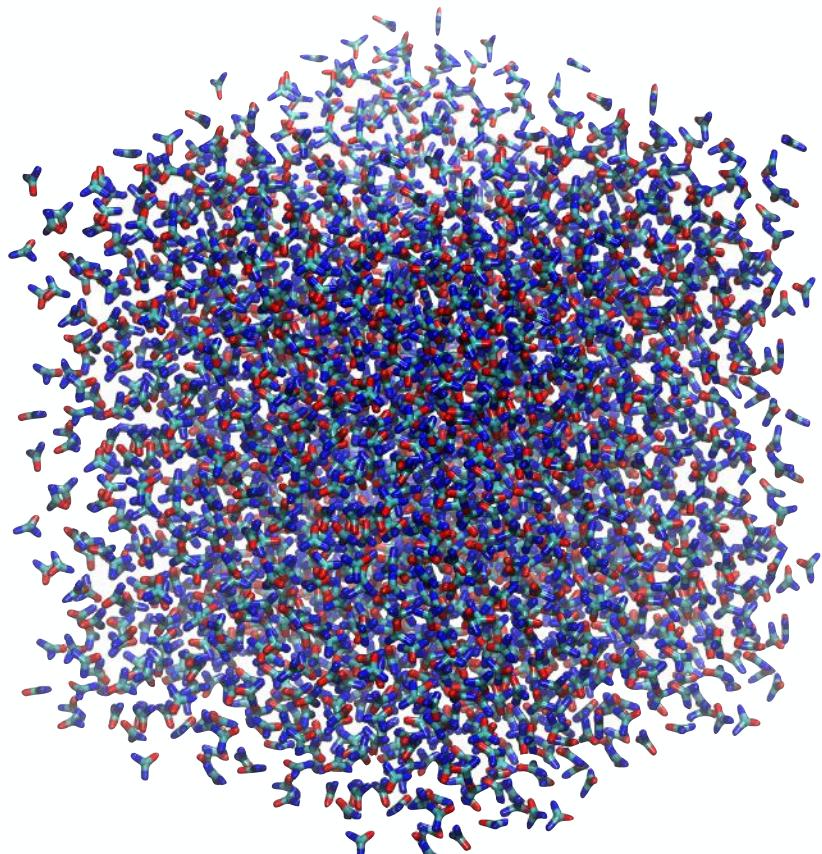
From a high dimensional and rugged Potential Energy Surface



To a low dimensional and smooth Free Energy Surface

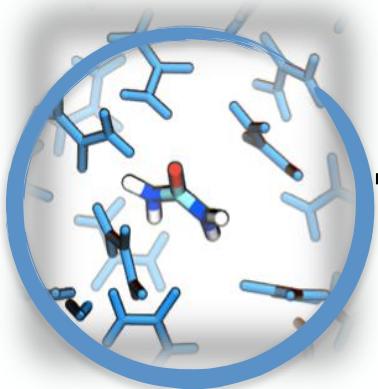


A dimensional reduction

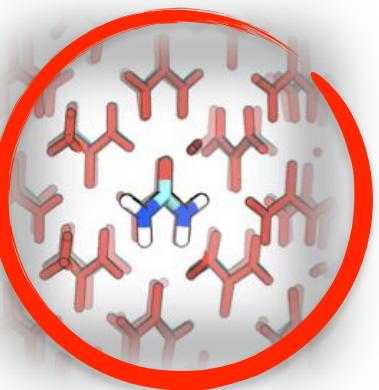


Fully atomistic description

local density



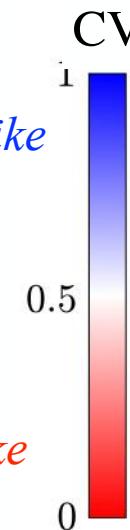
local order



Collective
Variables

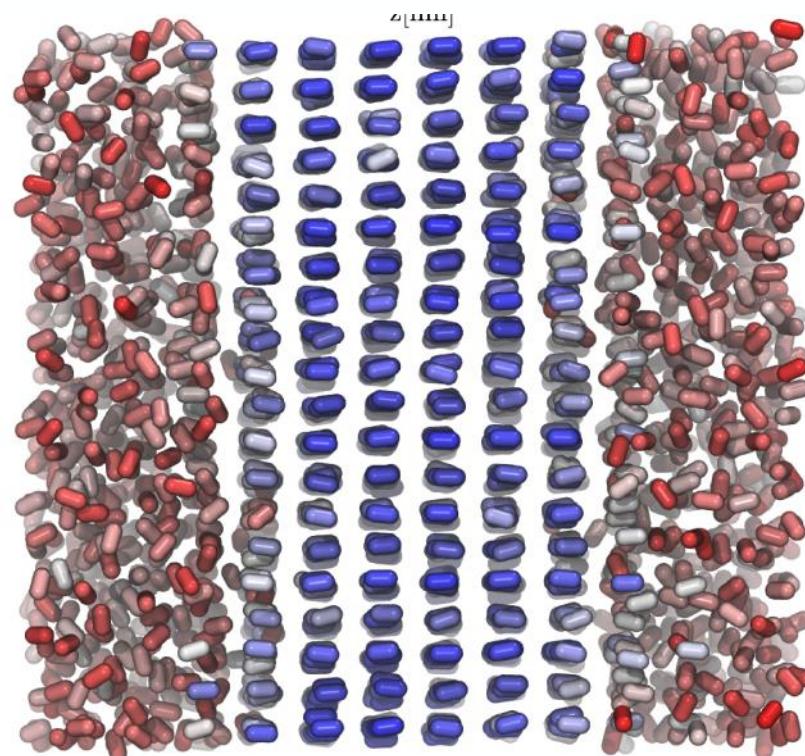


Crystal-like



Liquid-like

CV description



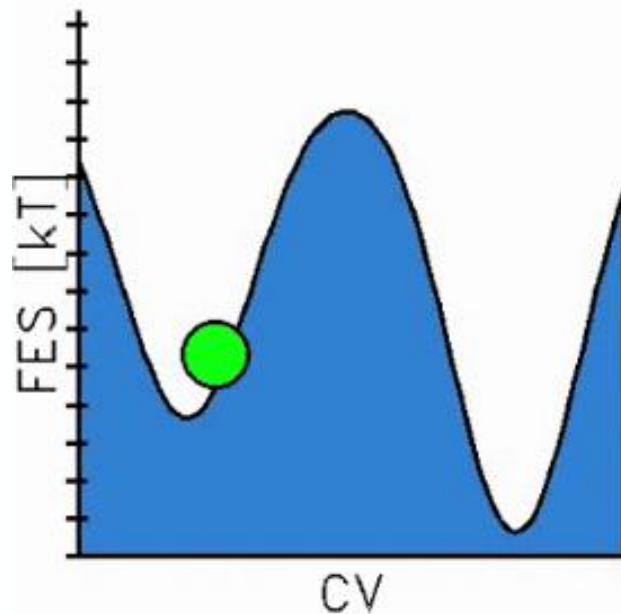
Sampling methods

We have developed two collective-coordinates-based enhanced sampling methods

- **Metadynamics**
- **Variationally enhanced sampling**

Metadynamics

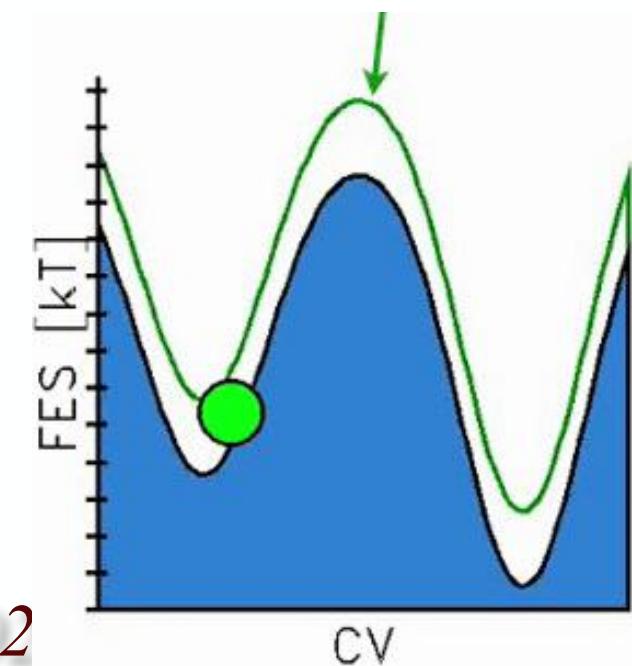
Standard dynamics



The bias potential is built iteratively by adding a local repulsive potential that discourages revisiting regions already explored.

$$V_n(\mathbf{s}) = V_{n-1}(\mathbf{s}) + w_g e^{-\frac{\beta V_{n-1}(\mathbf{s})}{\gamma-1}} G(\mathbf{s}, \mathbf{s}_n)$$

Metadynamics

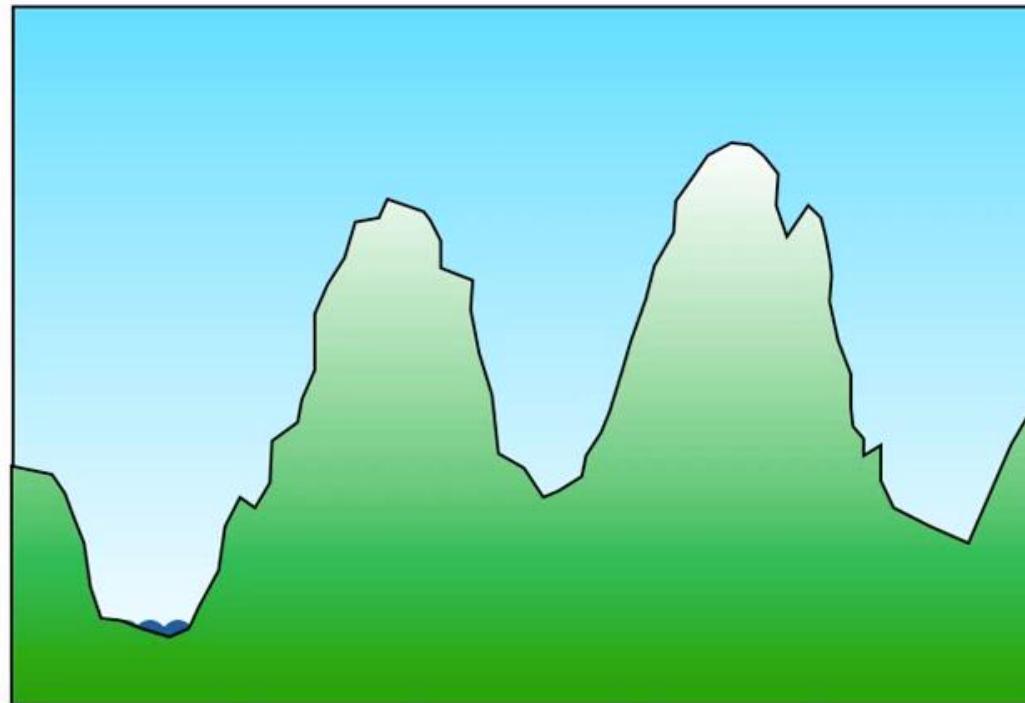


*Laio and Parrinello PNAS (2002)
Barducci, Bussi and Parrinello PRL (2008)*

Can't help showing this too

A Dutch perspective on
escaping free energy minima

Metadynamics, Laio and Parrinello, PNAS (2002)



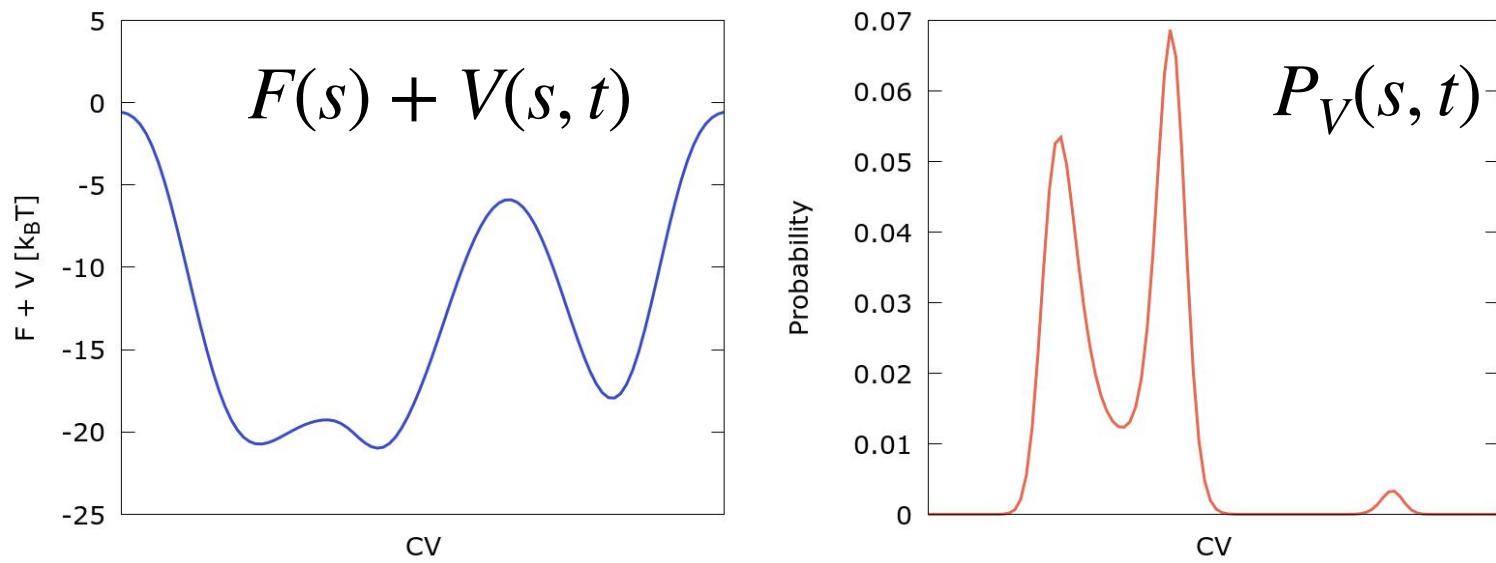
A rigorous result

The procedure amounts at solving the ordinary differential equation

$$\frac{dV(s, t)}{dt} = \int ds' G(s - s') e^{-\frac{V(s', t)}{\gamma-1}} P_V(s', t)$$

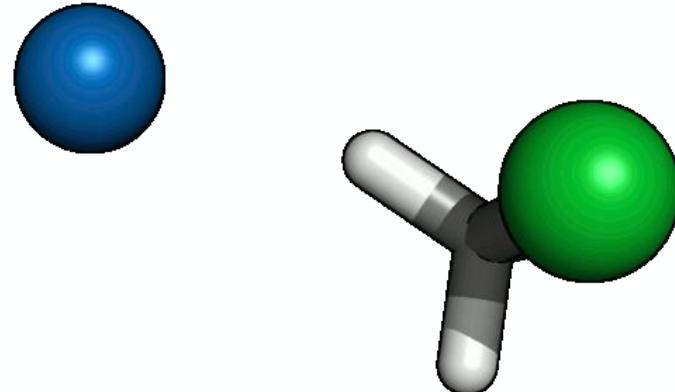
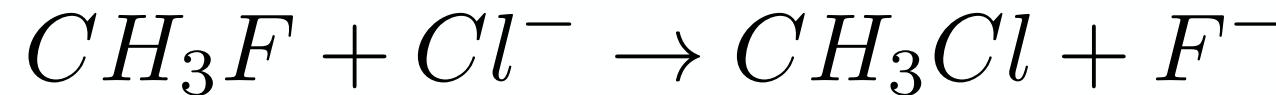
and at enhancing the fluctuations in a controlled way using the parameter γ .

$$P(s) \rightarrow P(s)^{\frac{1}{\gamma}}$$

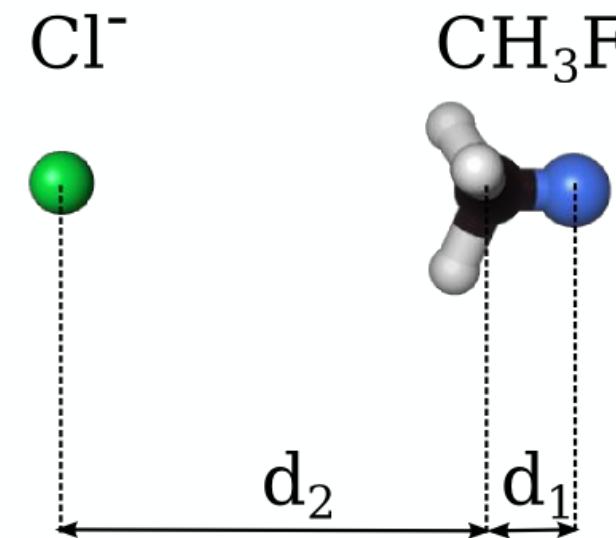
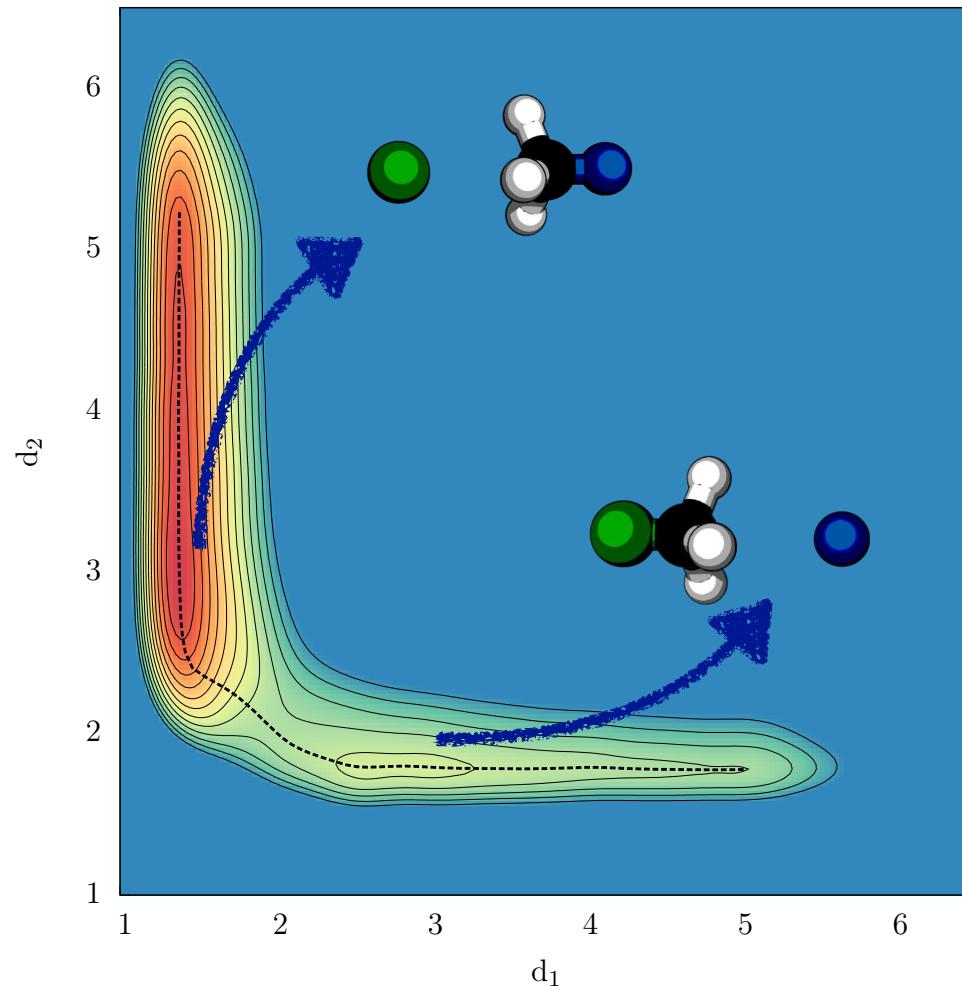


Simple collective coordinates for chemistry

Let us start from the simple S_N2 reaction

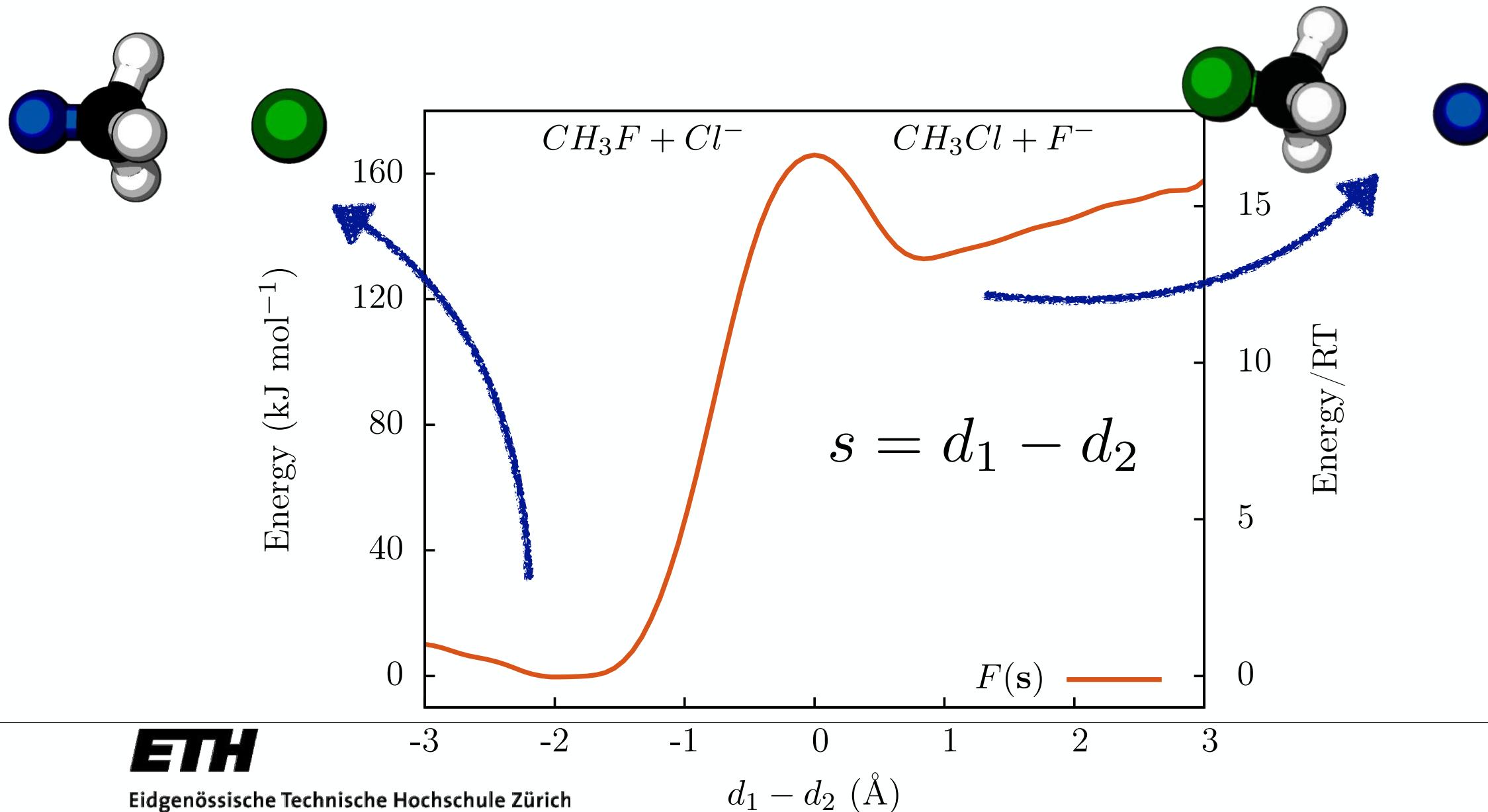


The free energy surface



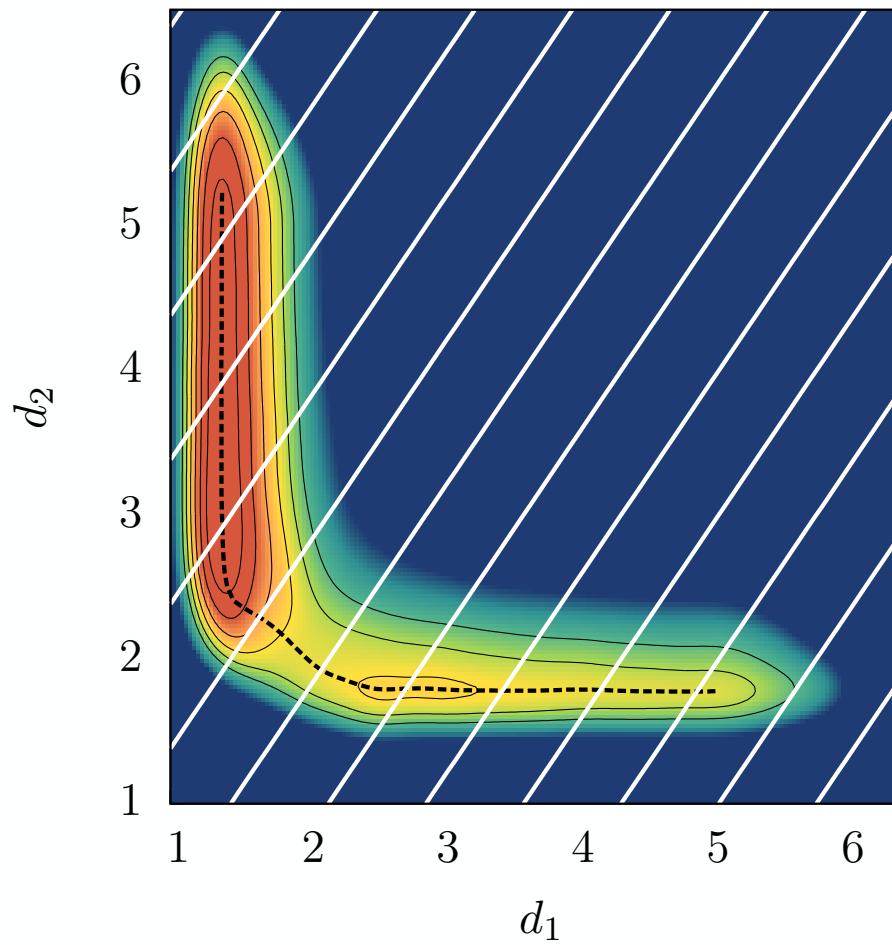
The standard approach one looks for the minimum free energy path and/or the transition state.

A heuristic CV



Surfaces of constant collective variable value

$$s = d_1 - d_2$$

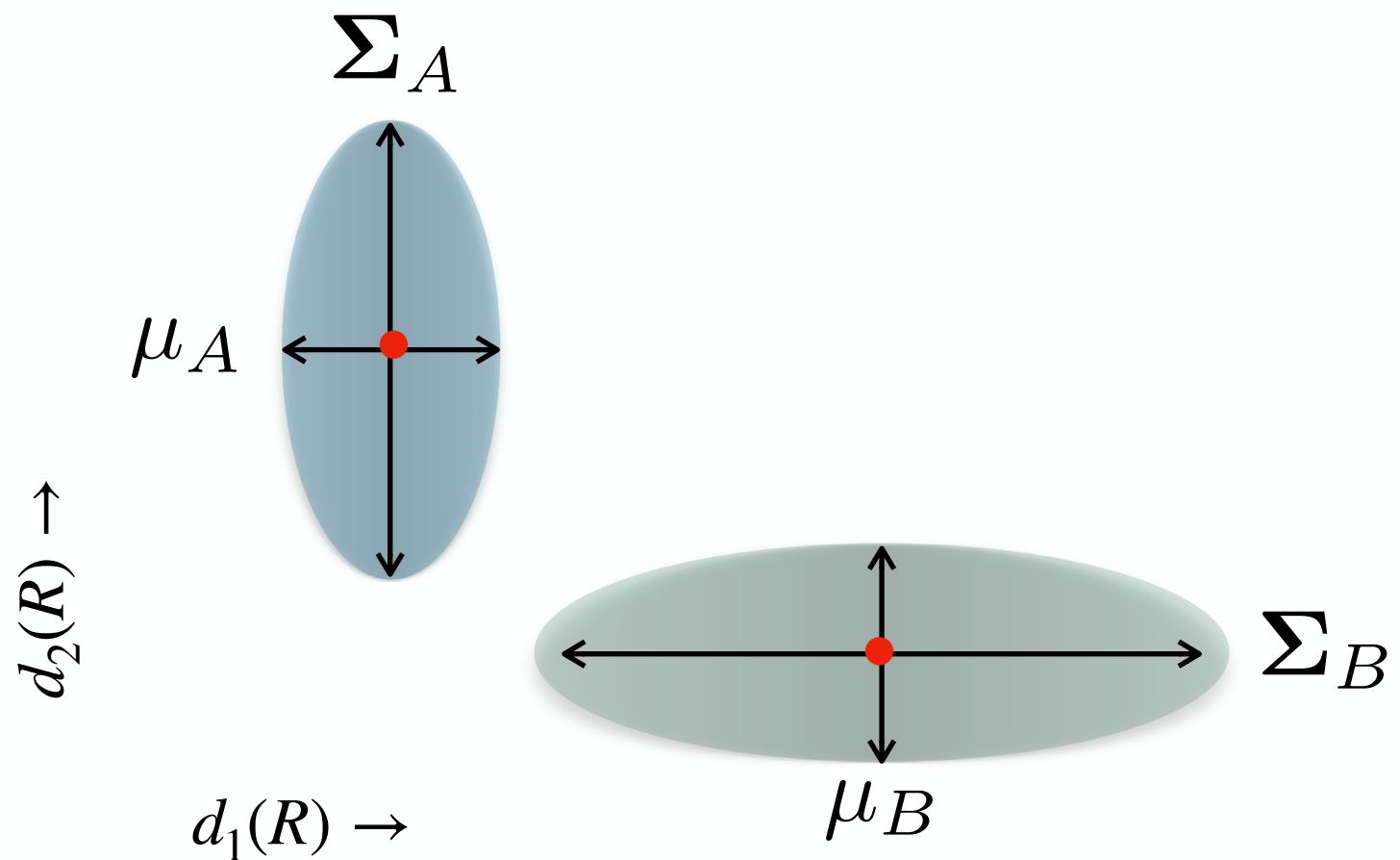


Consider a two state system

The two states are identified with a set of descriptors $d(R)$

Each metastable states has its own expectation value and covariance matrix.

Can we get a good one dimensional collective variable from this information alone?

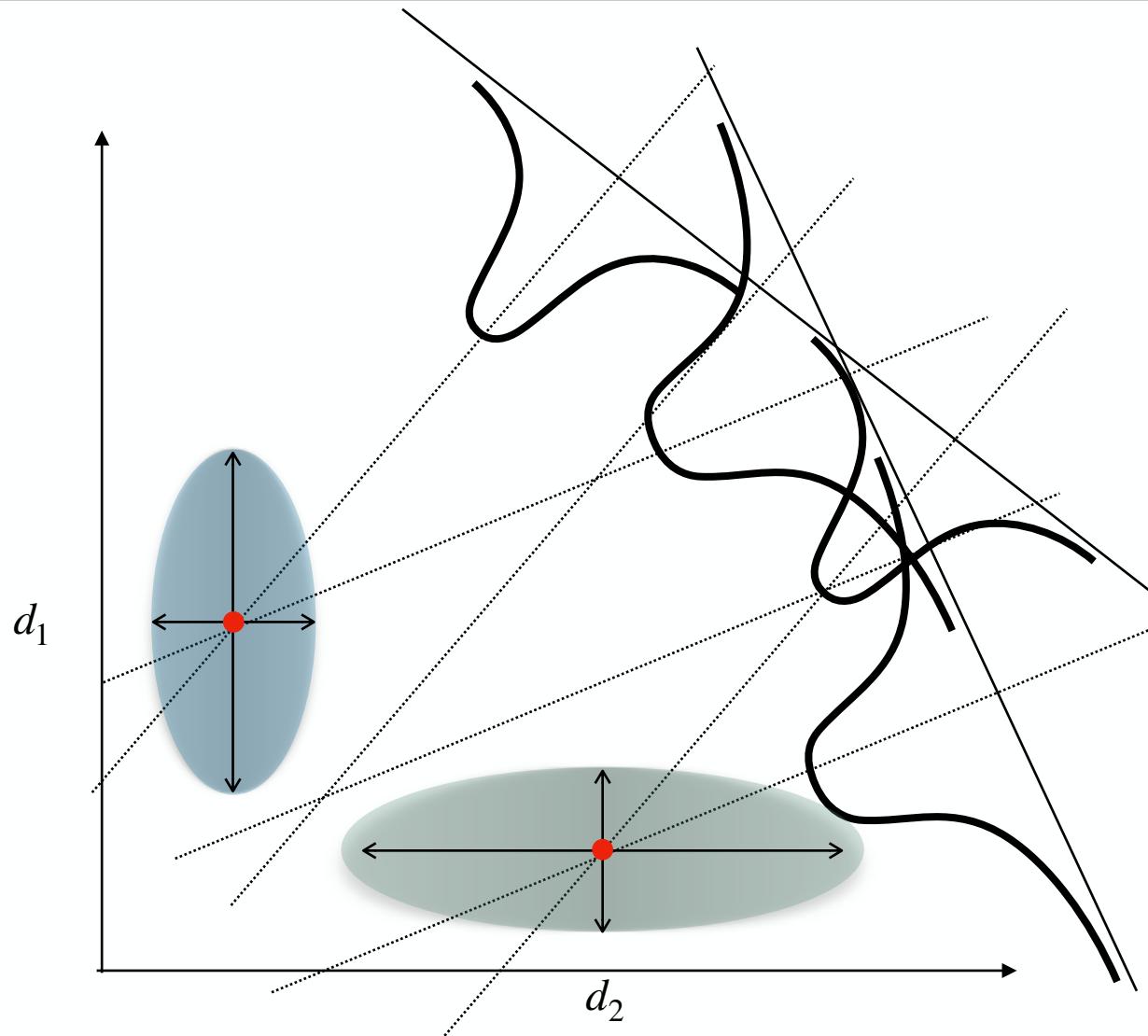


Linear discriminant analysis

Search for the one dimensional projection that best separates the two different set of data.

$$z = \mathbf{w} \mathbf{d}$$

The number of descriptors can be very large!

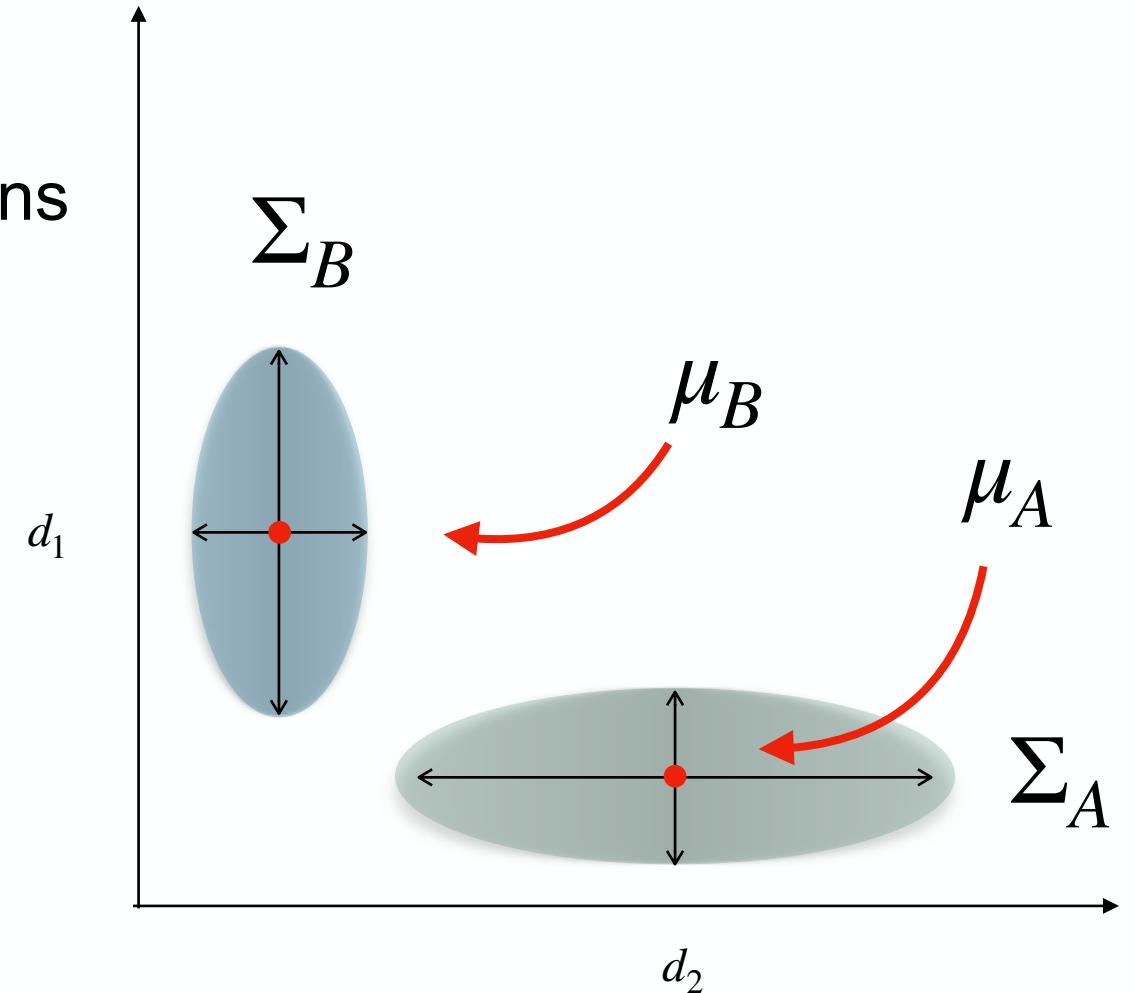


Harmonic Linear Discriminant Analysis

For the purpose of studying chemical reactions we introduce a variant that we call Harmonic Linear Discriminant Analysis that leads to:

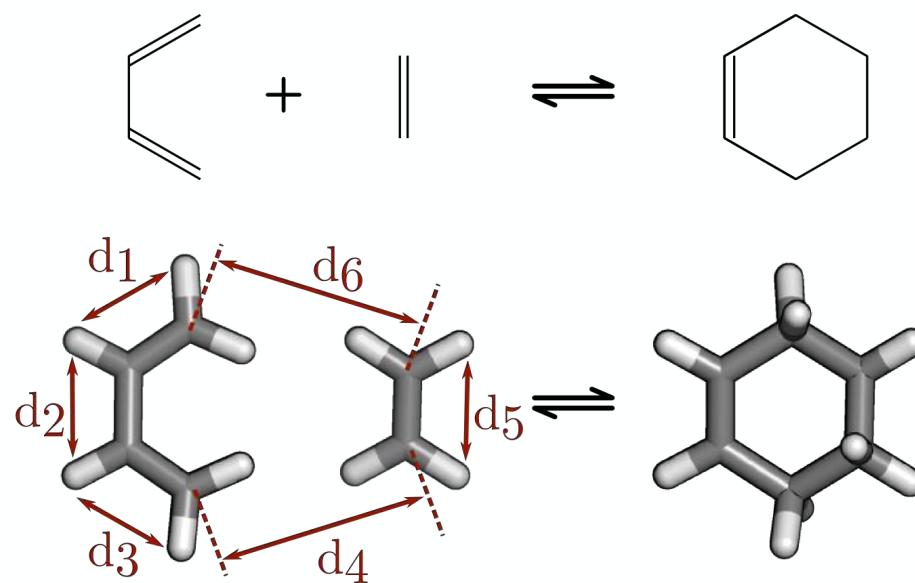
$$s(\mathbf{R}) = (\boldsymbol{\mu}_A - \boldsymbol{\mu}_B)^T \left(\frac{1}{\Sigma_A} + \frac{1}{\Sigma_B} \right) \mathbf{d}(\mathbf{R}).$$

No information on the transition state or reaction path is needed!
It is all encrypted in the fluctuations!



Ideal for chemistry!

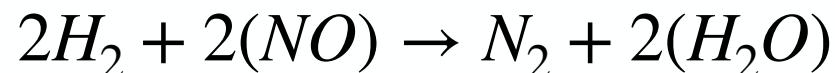
Lets study the classical Diels Alder reaction



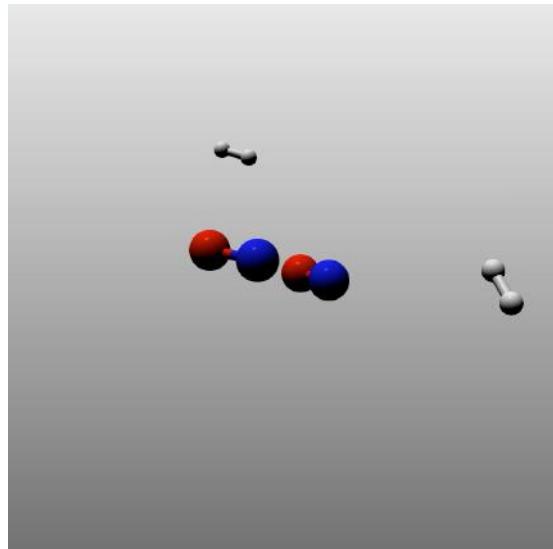
d_1	d_2	d_3	d_4	d_5	d_6
0.18	-0.46	0.13	-0.60	0.05	-0.61

A multistep process

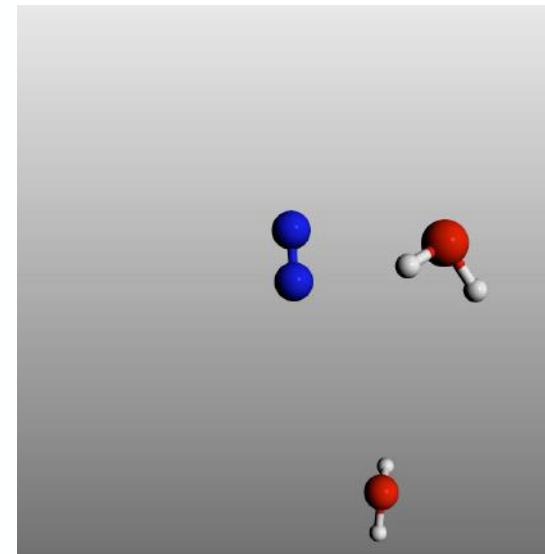
We want to simulate the reaction:



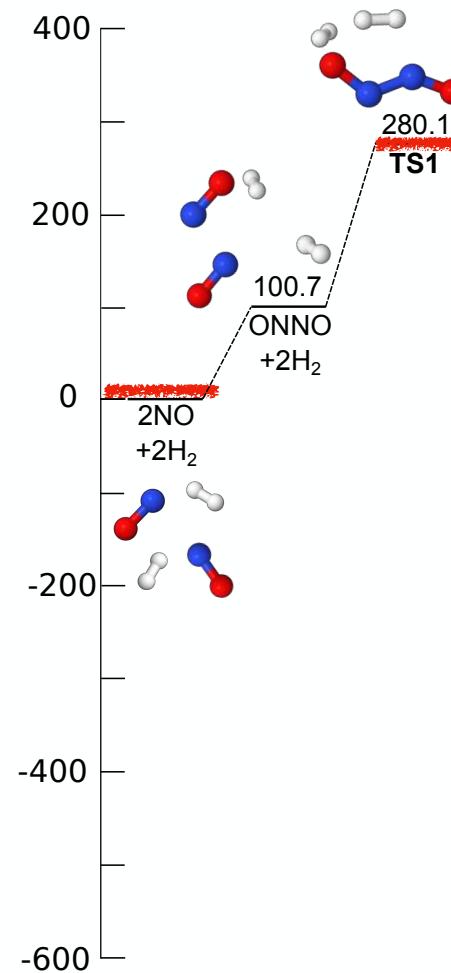
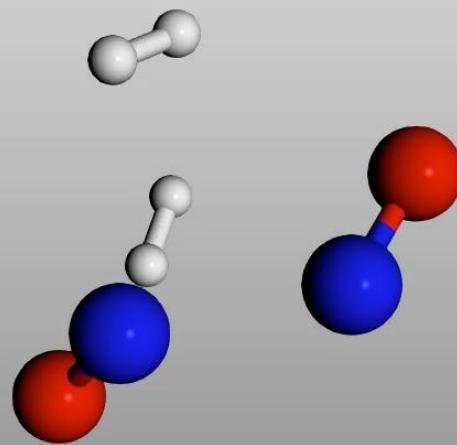
We run two independent simulations in the initial and final state and compute the fluctuation of 14 permutation invariant descriptors, to find the HLDA reaction coordinate.



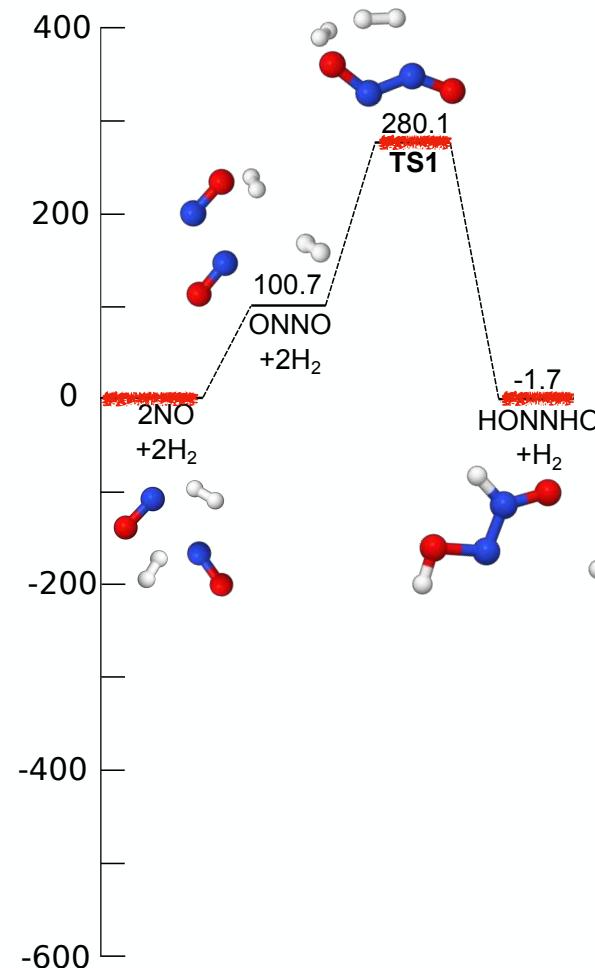
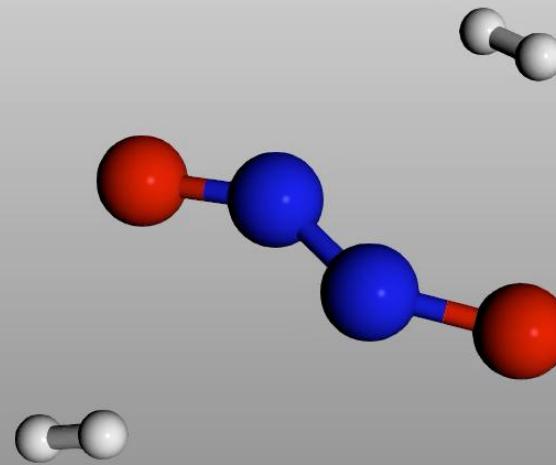
In collaboration with
Emilia Sicilia



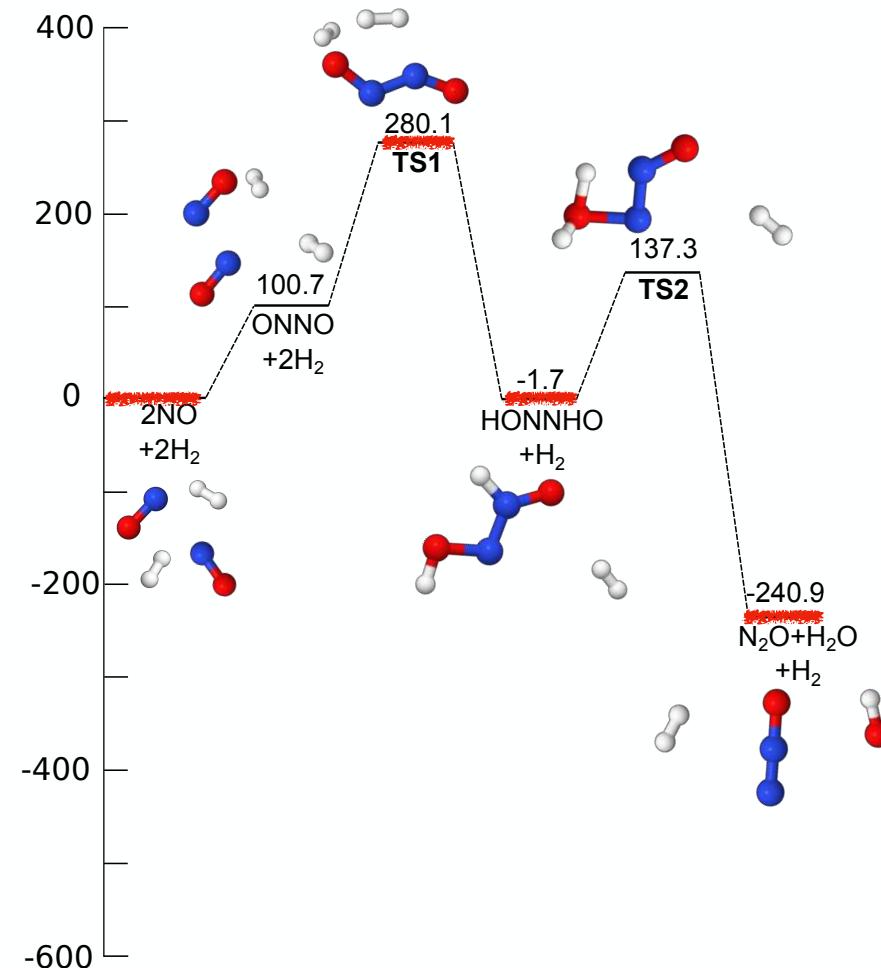
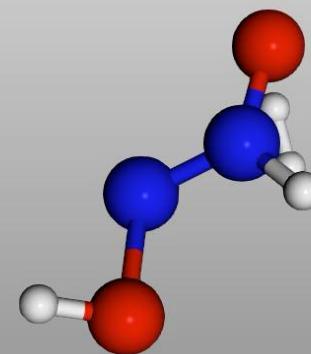
The main hurdle



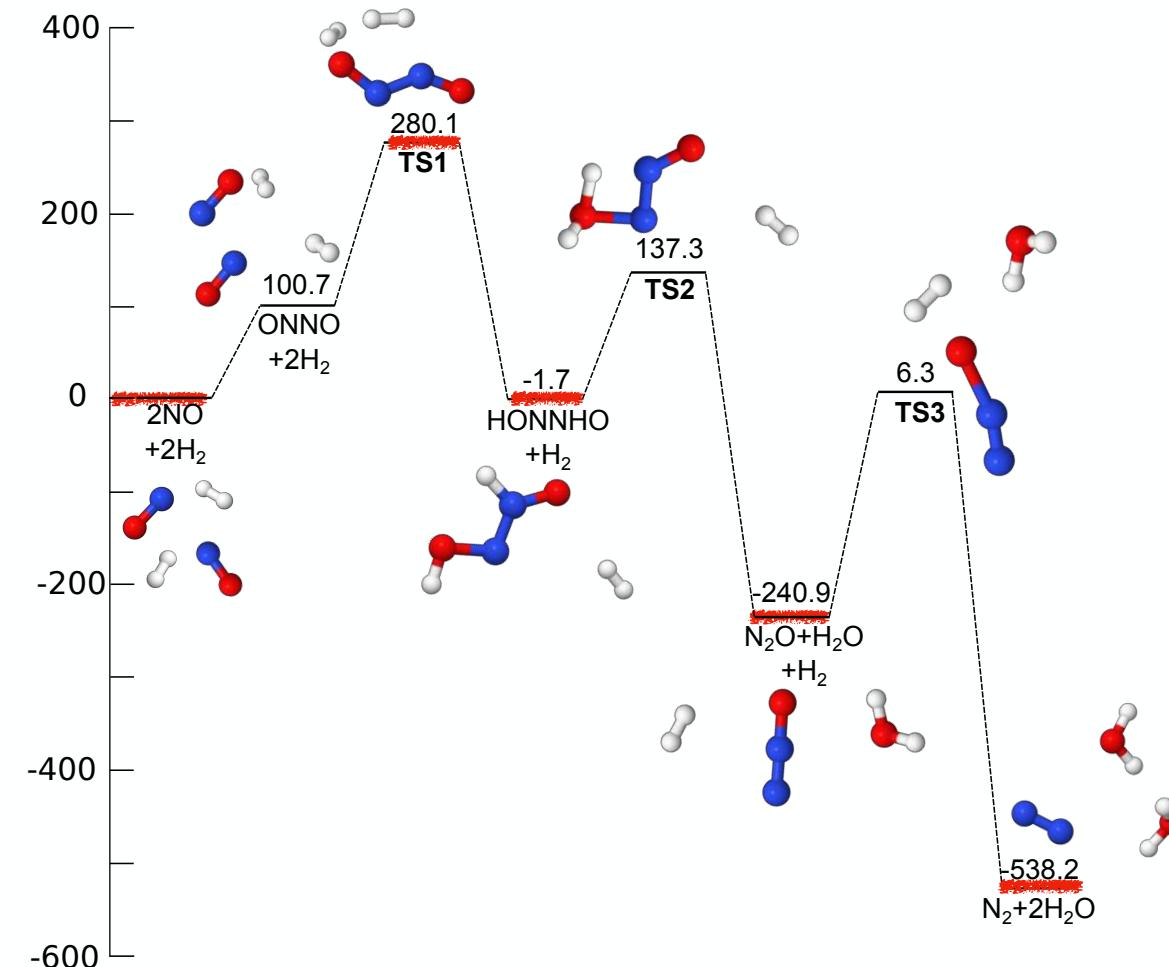
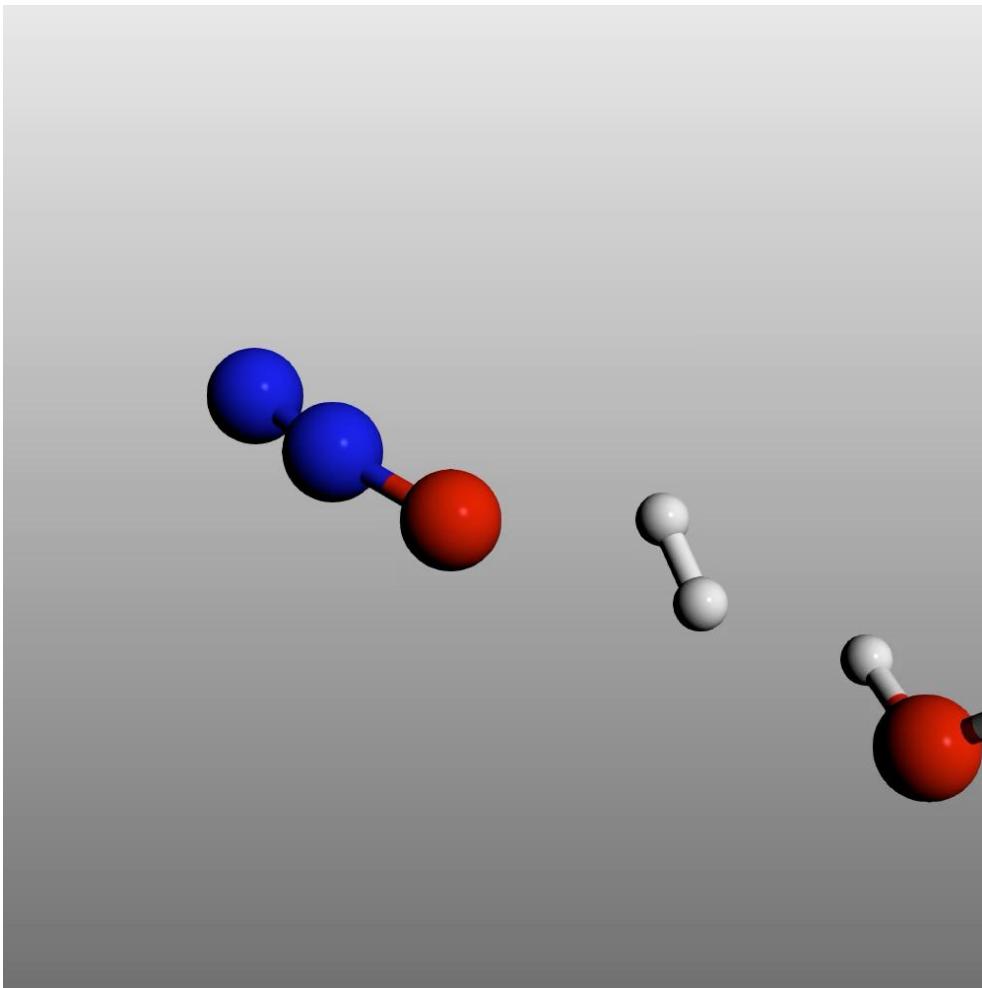
First intermediate



Second intermediate



The final product



A variational alternative

The bias is constructed by minimising the convex functional

$$\Omega(V(\mathbf{s})) = \frac{1}{\beta} \log \frac{\int d\mathbf{s} e^{-\beta(F(\mathbf{s})+V(\mathbf{s}))}}{\int d\mathbf{s} e^{-\beta F(\mathbf{s})}} + \int d\mathbf{s} p(\mathbf{s}) V(\mathbf{s})$$

At the minimum:

$$V(\mathbf{s}) = -F(\mathbf{s}) - \frac{1}{\beta} \log p(\mathbf{s}) \quad \text{i.e.} \quad P_V(\mathbf{s}) = p(\mathbf{s})$$

In the practice

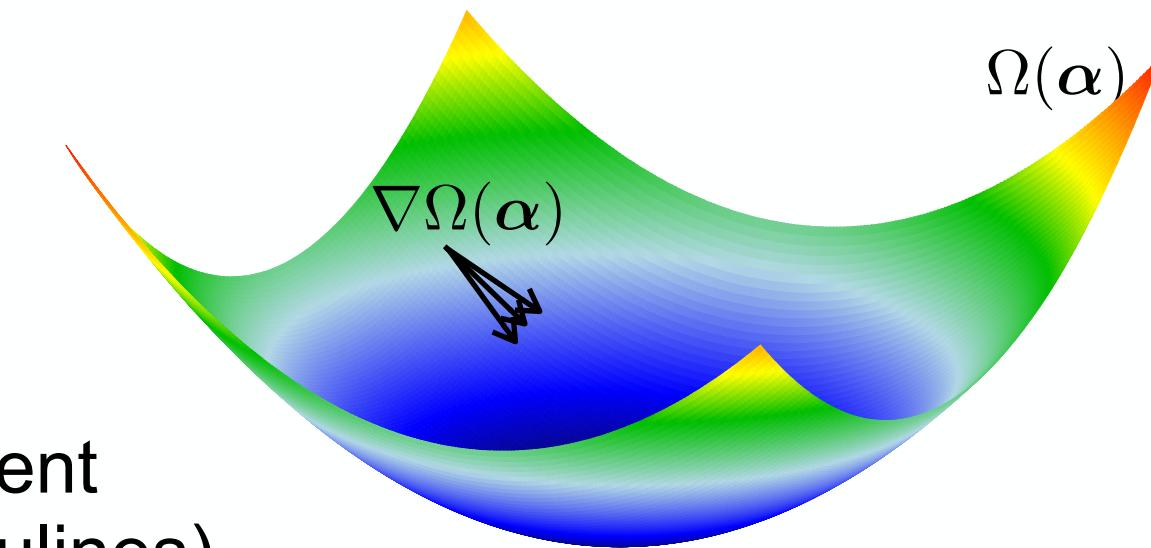
$$V(\mathbf{s}; \boldsymbol{\alpha}) = \sum_i \alpha_i f_i(\mathbf{s})$$

$$\Omega(V) \rightarrow \Omega(\boldsymbol{\alpha})$$

The convex function

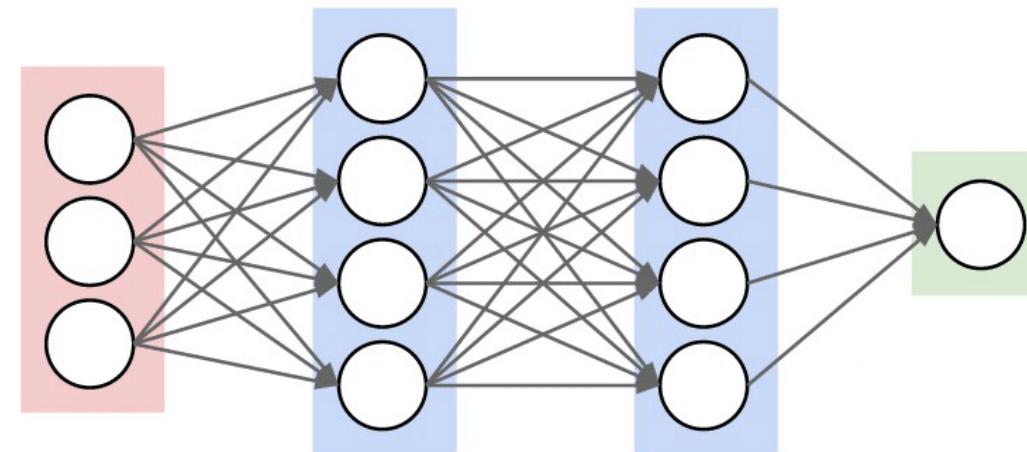
$$\Omega(\boldsymbol{\alpha})$$

is minimised using a
stochastic steepest descent
algorithm. (Bach and Moulines)



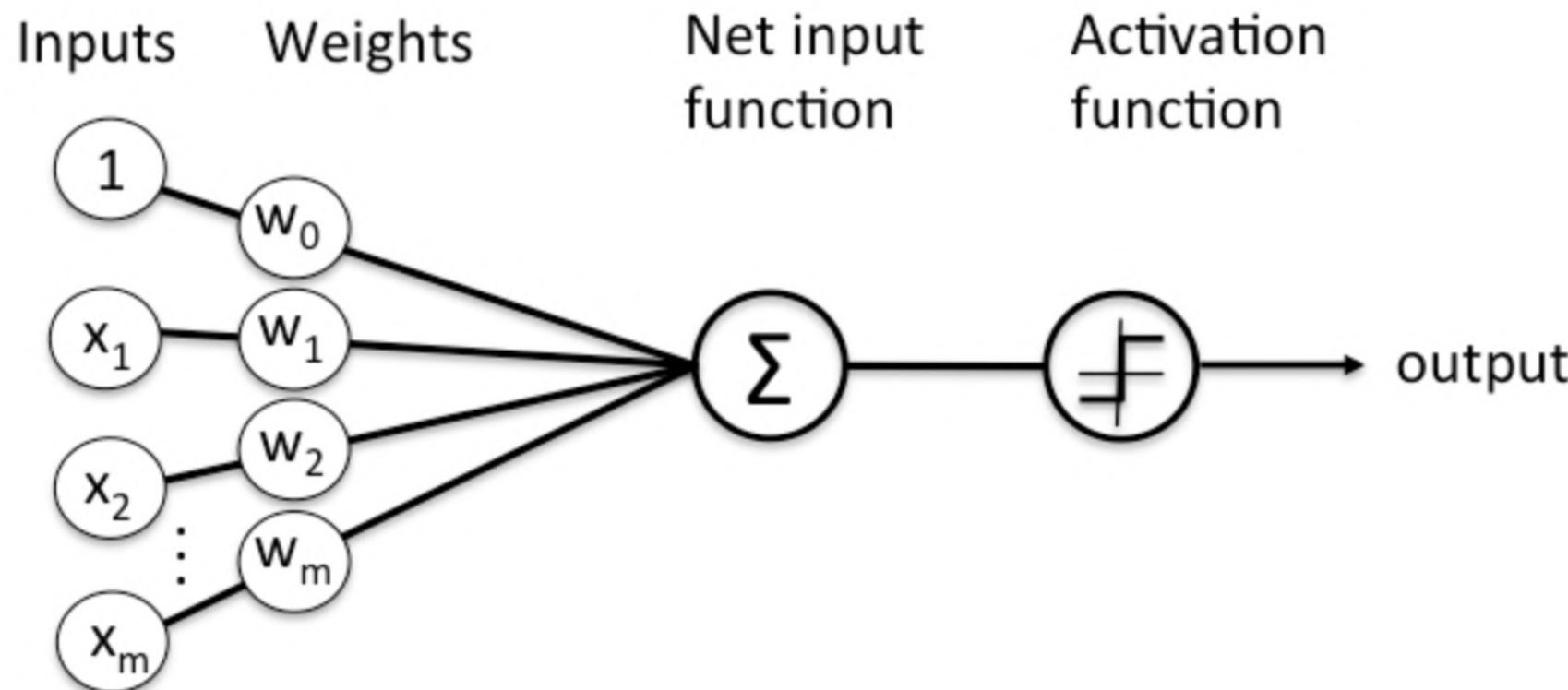
Represent the bias $V(s)$ as a neural network.

Input= collective
variables

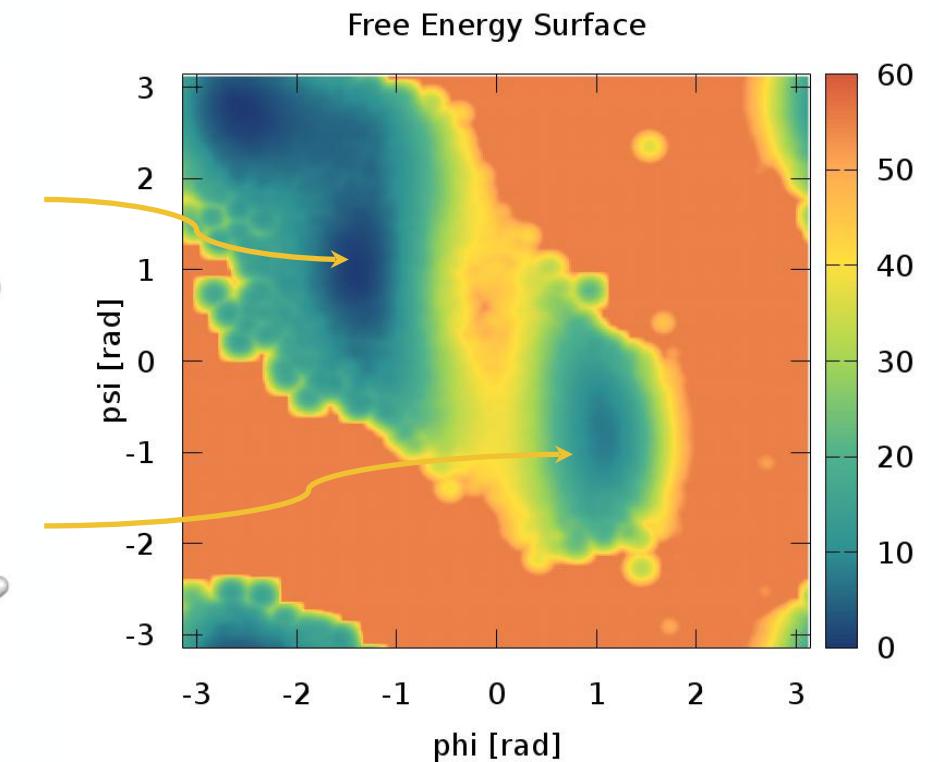
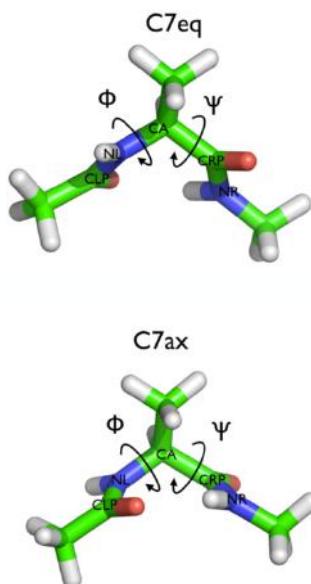
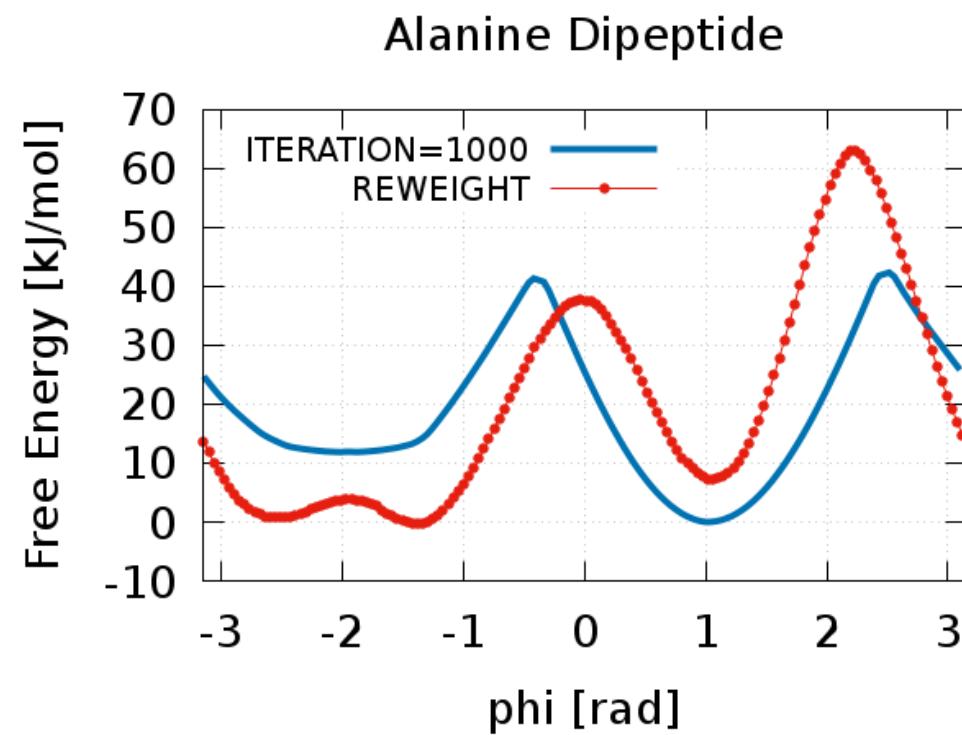


Output= bias
potential

The structure of a node



It does work



The group



Movies by
Jean Favre
and
Valerio Rizzi

Fine