

Computational design of materials for heterogeneous catalysis

Sergey Levchenko

Skoltech



Russian
Science
Foundation

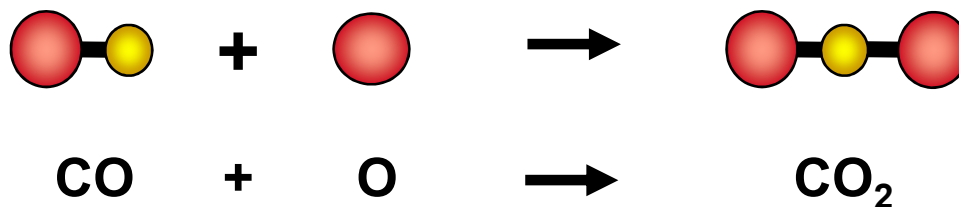
Supported by the Russian Scientific Foundation grant
No. 19072-30043 “Computational materials design
laboratory”

Did you know?

Accepted definition of a catalyst (Wilhelm Ostwald, 1902):

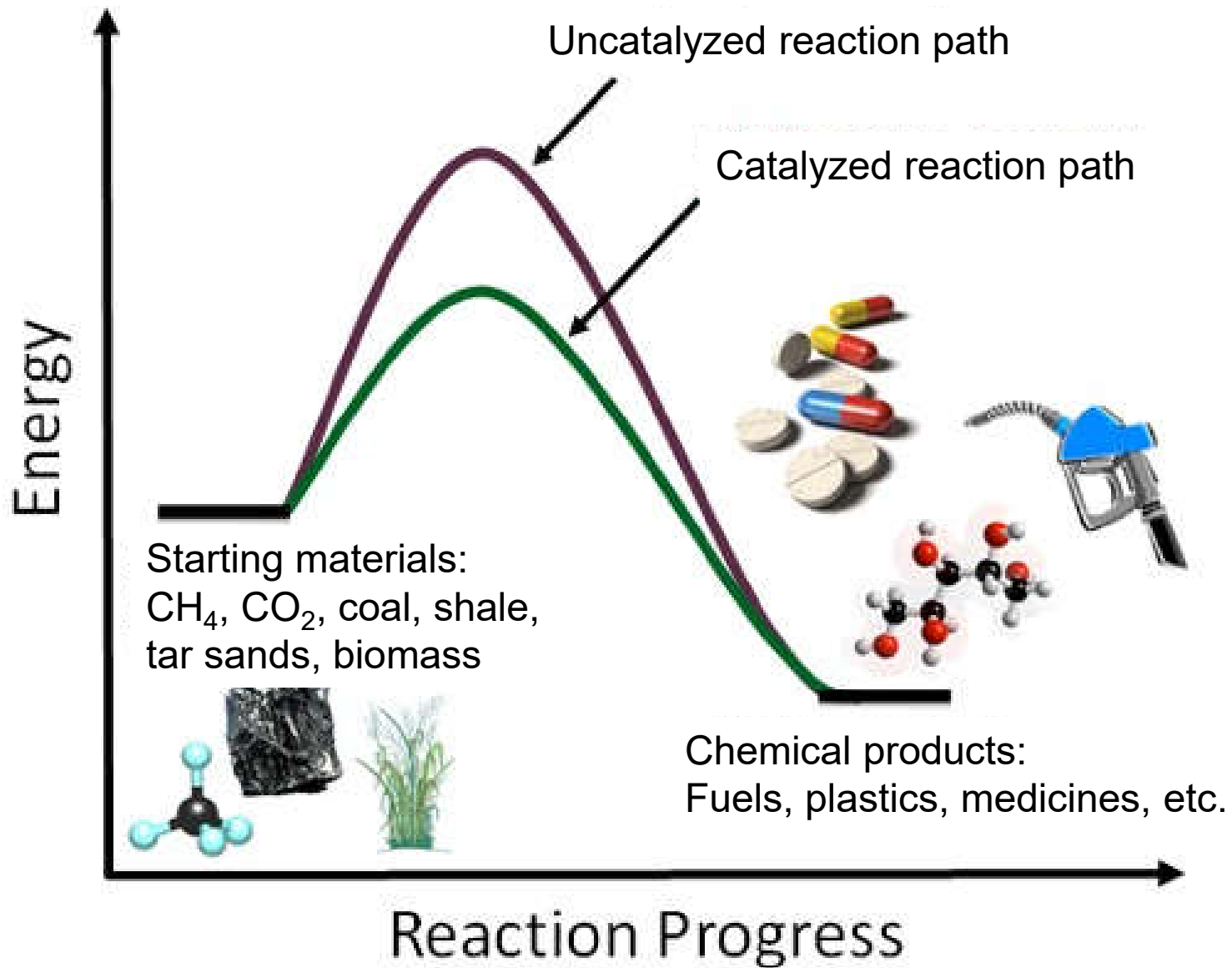
“A catalyst is a substance that changes the speed of a chemical reaction without itself undergoing any permanent chemical change.”

A catalyst provides an alternative pathway by which the reaction comes to equilibrium, although it does not alter the position of the equilibrium.



But there is no O () in the air, only O₂ () !

Catalysis



Heterogeneous catalysis

solid | gas

solid | liquid

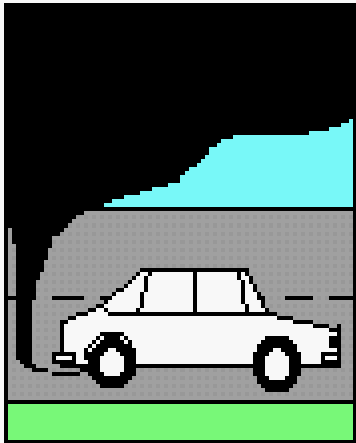
liquid | gas

The different phases allow for an easier extraction of the products

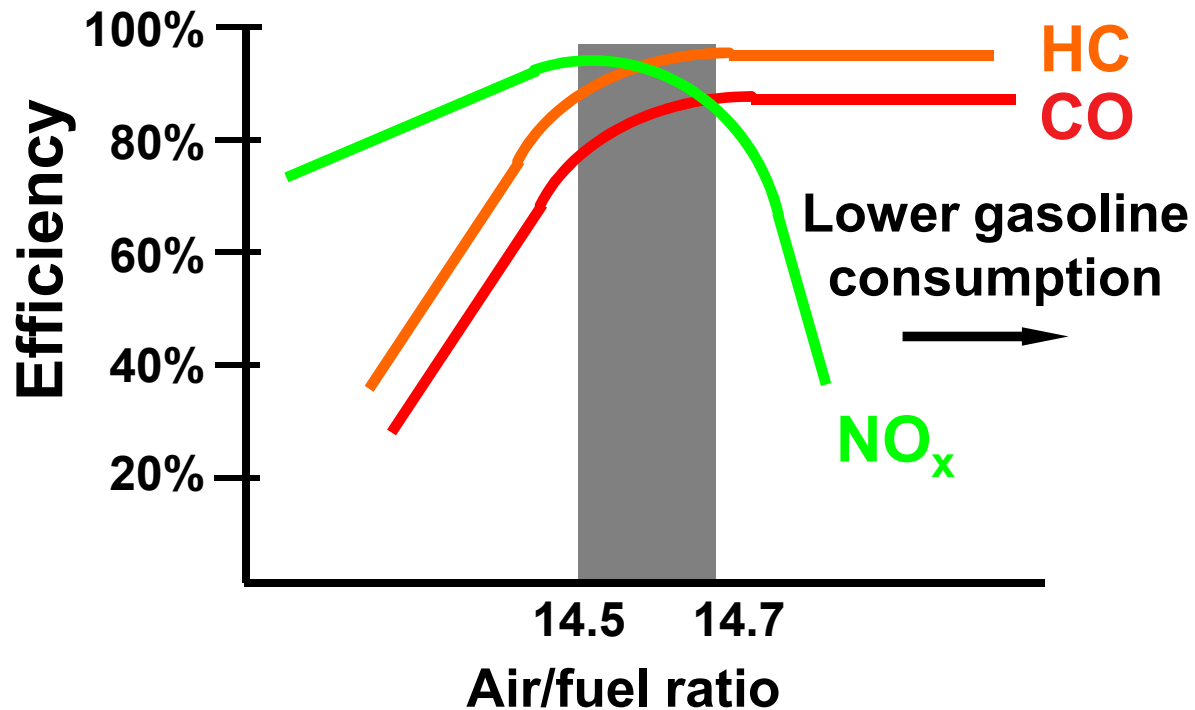
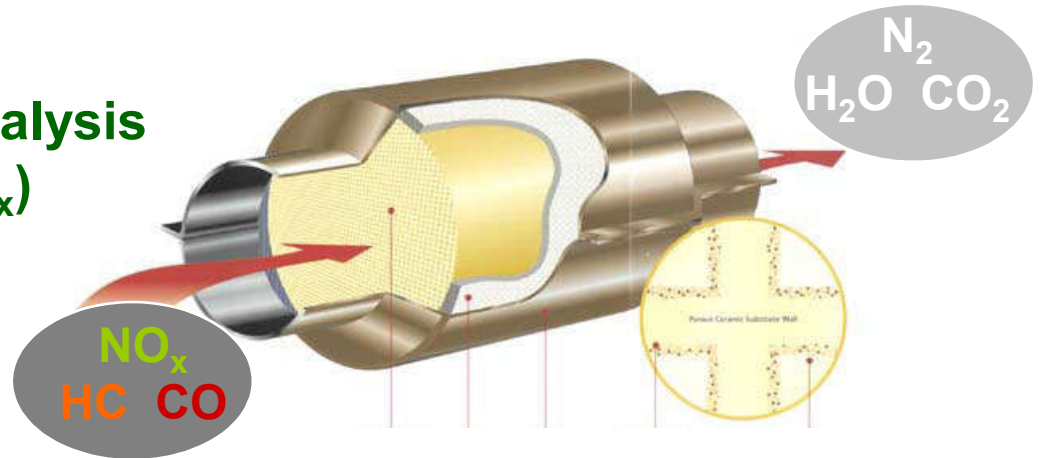
The path towards active sites is simplified

Interfaces can serve as carriers for nanostructures or active molecules

Applications of catalysis I: pollutants removal



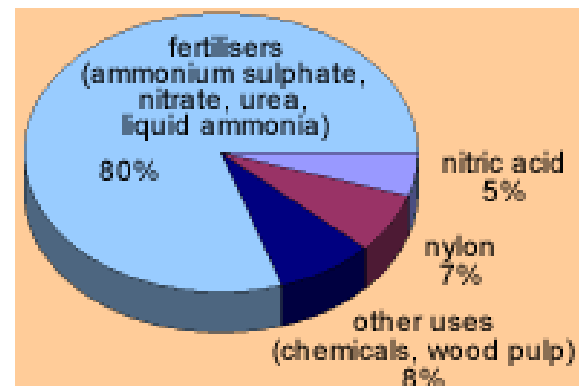
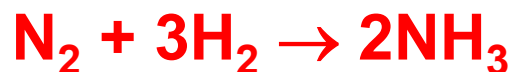
Three-way catalysis
(CO, C_xH_y, NO_x)



Applications of catalysis II: Large-scale production of chemicals

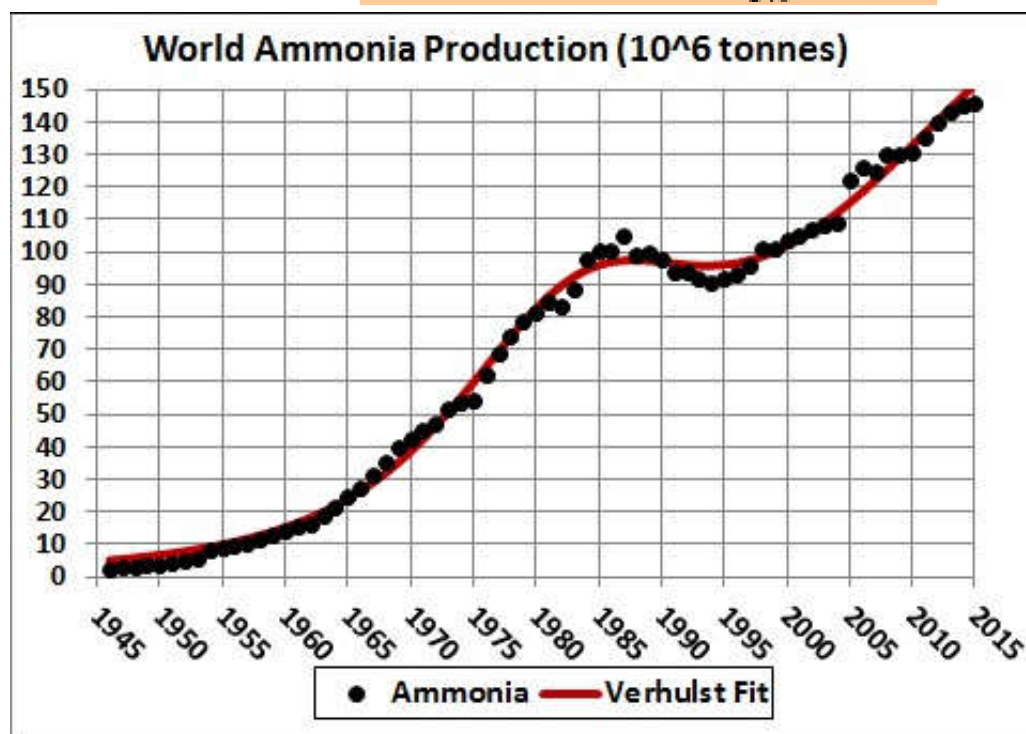


Ammonia synthesis plant

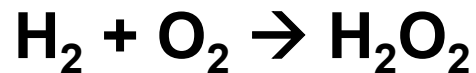
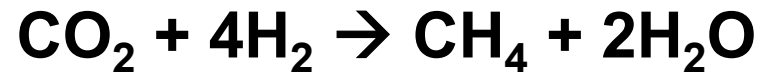
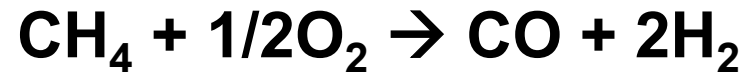
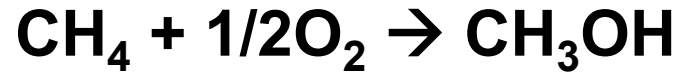


Haber-Bosch Process

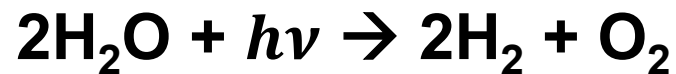
Discovered in 1909 by testing about 20.000 catalysts (BASF, Alwin Mittasch)



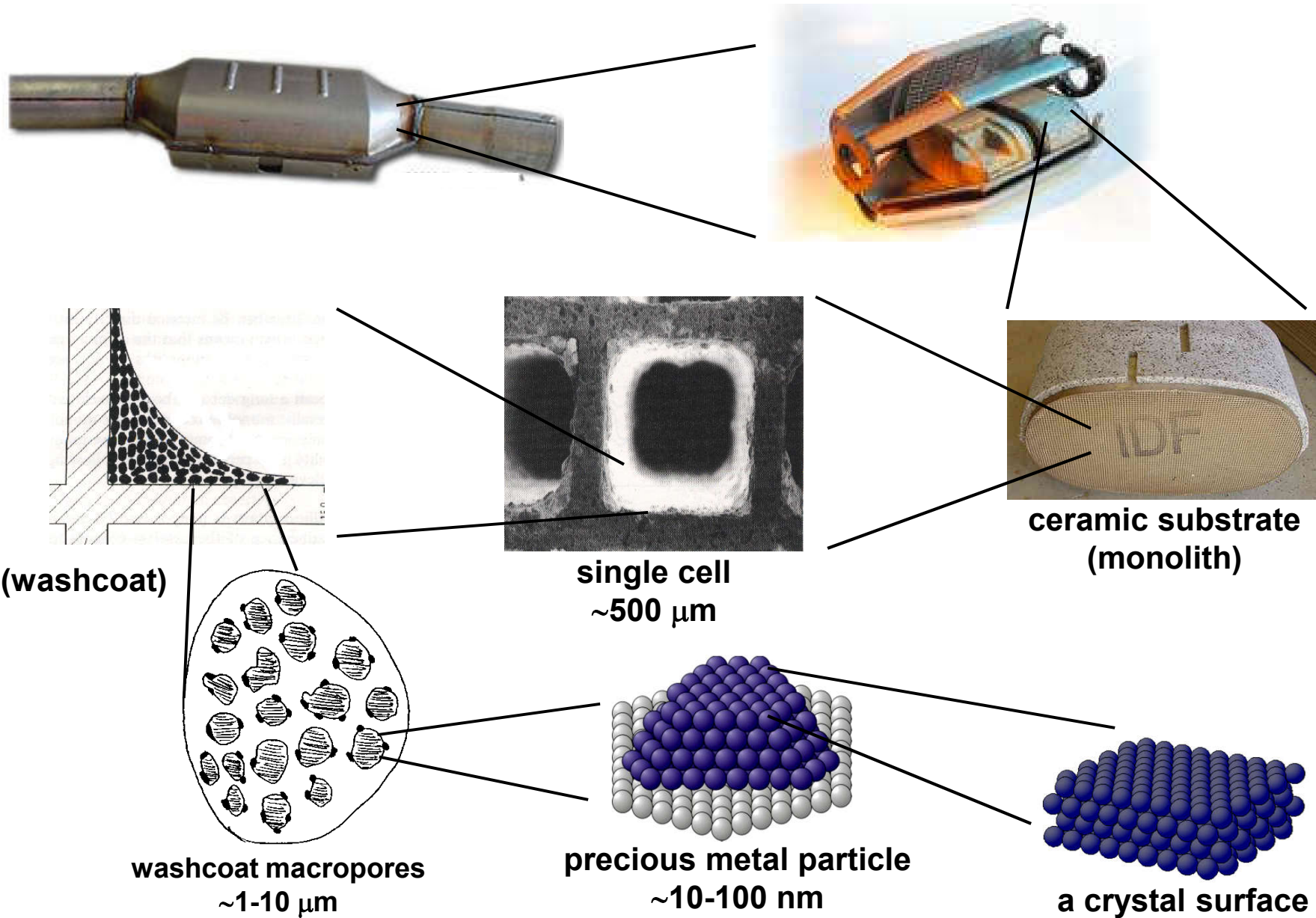
“Dream reactions”



(“green” oxidant, paper industry, water purification)



Different spatial scales of a catalytic process



Did you know?

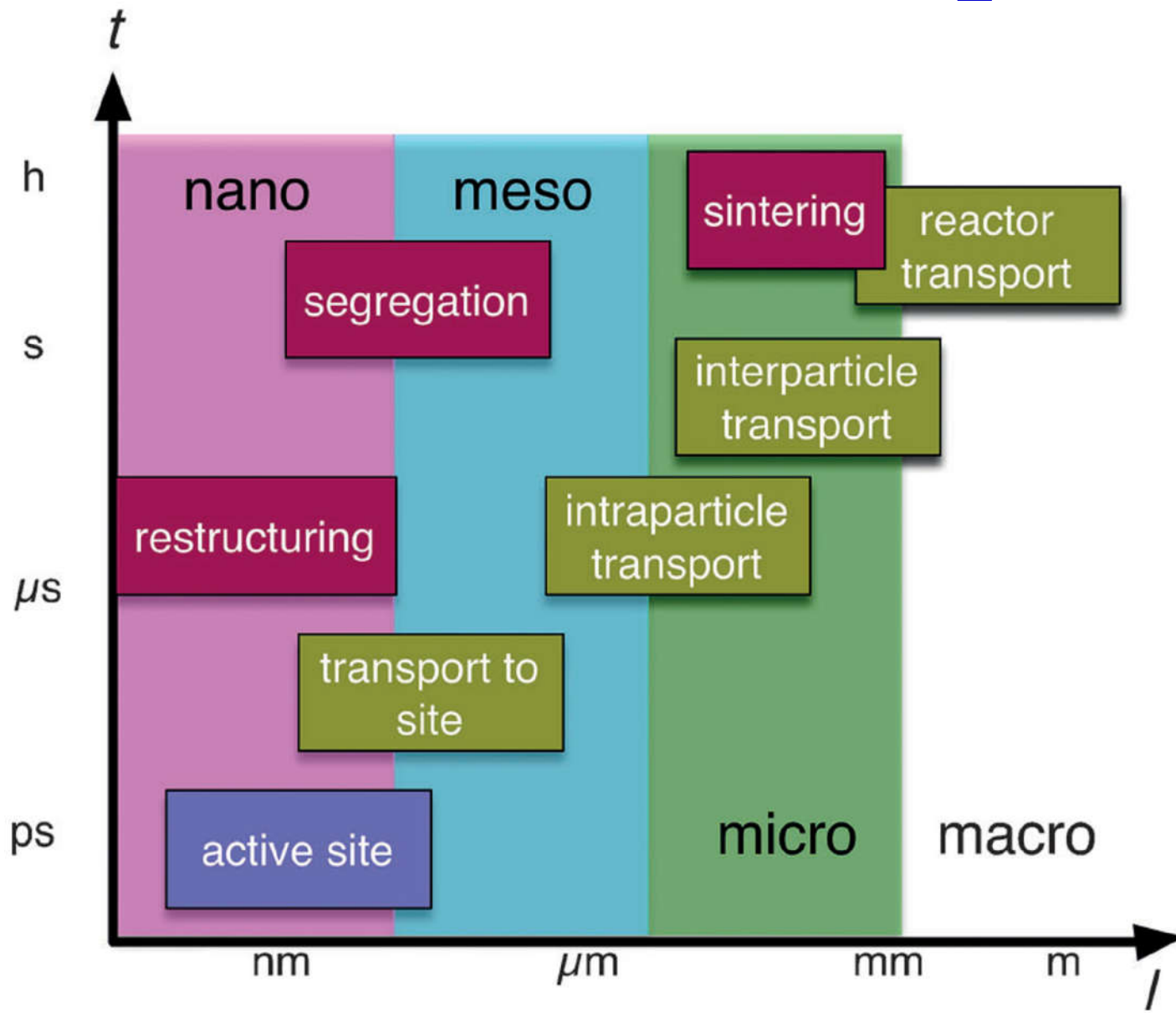
Accepted definition of a catalyst (Wilhelm Ostwald, 1902):

“A catalyst is a substance that changes the speed of a chemical reaction without itself undergoing any permanent chemical change.”

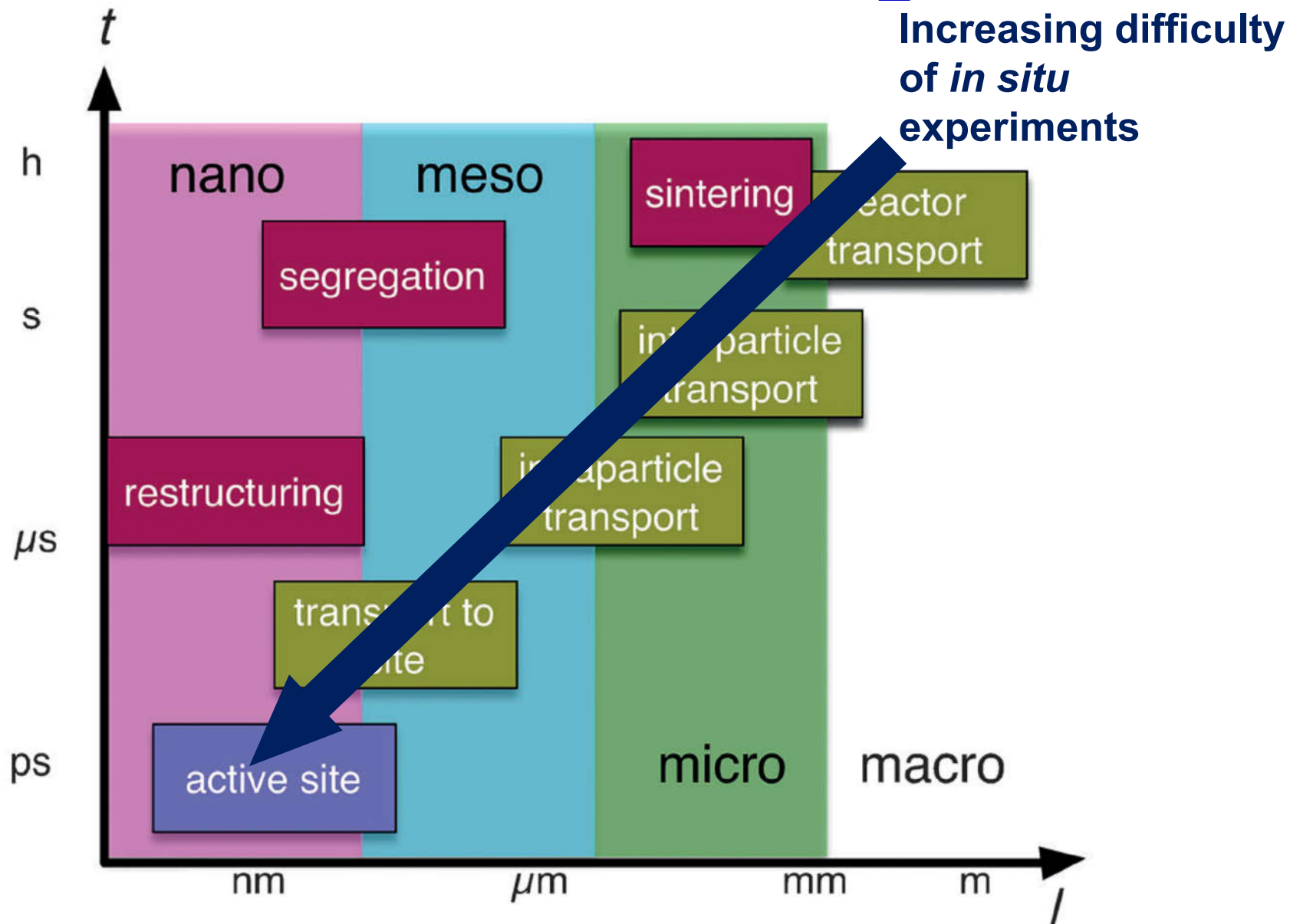


In heterogeneous catalysis, in most cases we don't know what this substance is

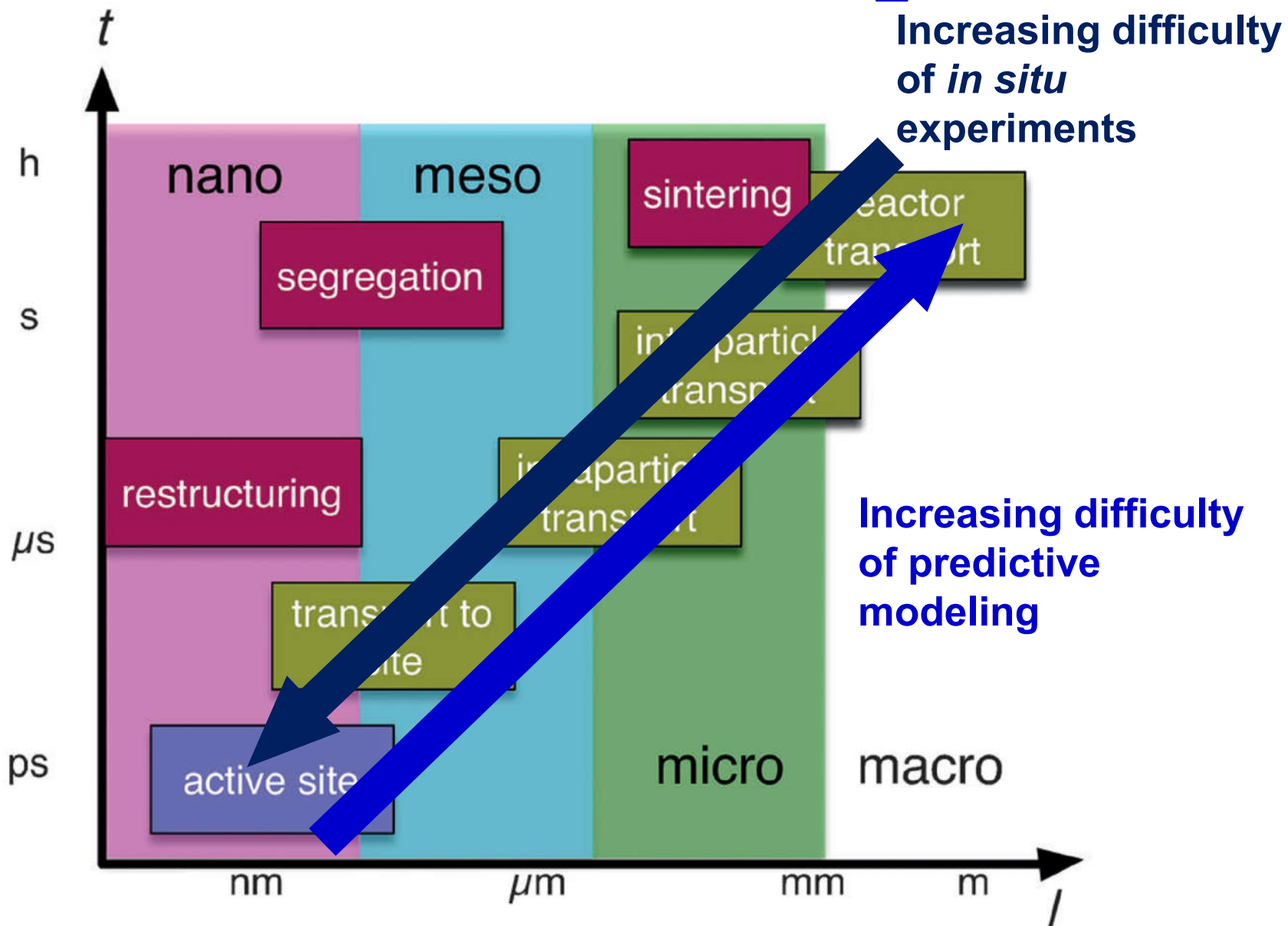
Multiscale nature of heterogeneous



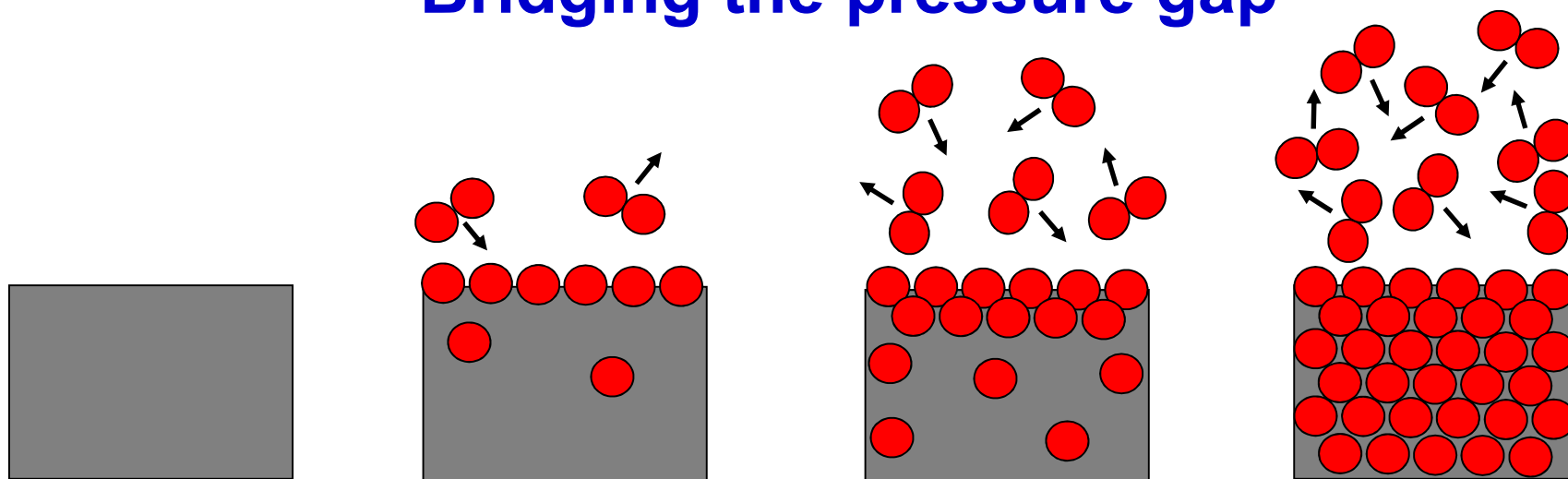
Multiscale nature of heterogeneous



Multiscale nature of heterogeneous



Bridging the pressure gap

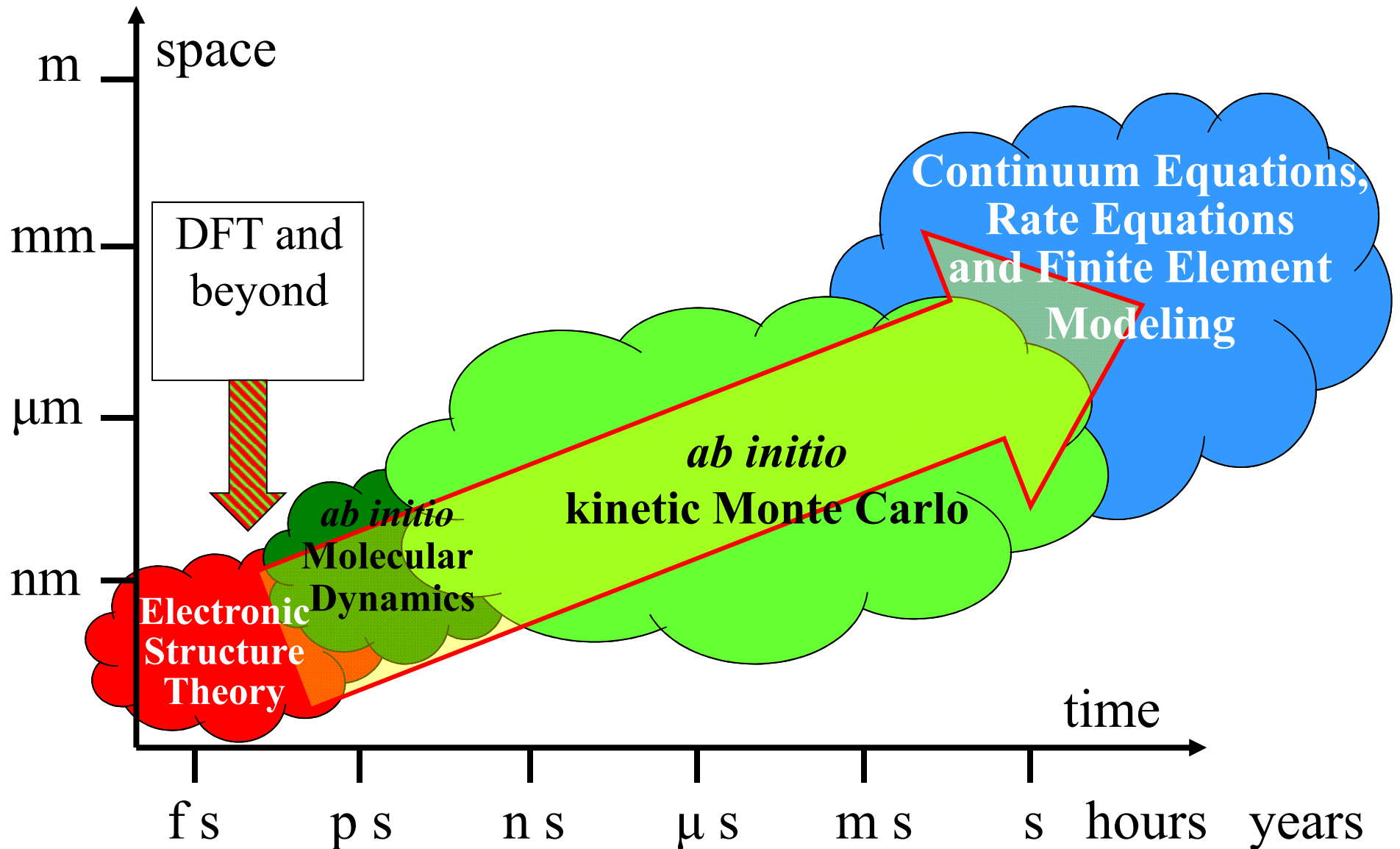


Theory: Development and use of first-principle statistical mechanics approaches

Experiment: Awareness and diligent experiments...
Development and use of „*in situ*“ techniques

Nanometer and sub-nanometer thin oxide films at surfaces of late transition metals,
K. Reuter, in: *Nanocatalysis: Principles, Methods, Case Studies,*
(Eds.) U. Heiz, H. Hakkinen, and U. Landman, Springer (Berlin, 2006)
<http://www.fhi-berlin.mpg.de/th/paper.html>

Multiscale modeling

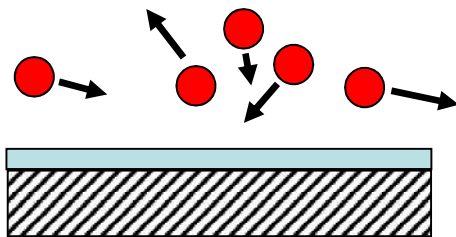


Real versus model surfaces (materials and pressure gaps)

- 1) A straightforward DFT calculation corresponds to $T = 0$ K
- 2) In literature, 1-5000 Å is called a “surface”, depending on experimental technique
- 3) Real surfaces are full of defects (vacancies, steps, dislocations,...)
- 4) Surface structure and composition can strongly depend on preparation
- 5) Impurities and adsorbates can influence the surface morphology

Special care must be taken in experiment to produce a defect-free pure surface of known termination. But even then...

- 6) A surface cannot be separated from a gas (or liquid) above it



$$\nu = \frac{p}{\sqrt{2\pi mkT}}$$

For $T = 300$ K, $p = 1$ atm $\rightarrow \nu \sim 10^8$ site⁻¹ s⁻¹



Requires $p \leq 10^{-12}$ atm to keep a “clean” surface clean; surface can also lose atoms

DFT (internal) versus free energy

At constant T a system minimizes its free energy ($-TS$), not internal energy U

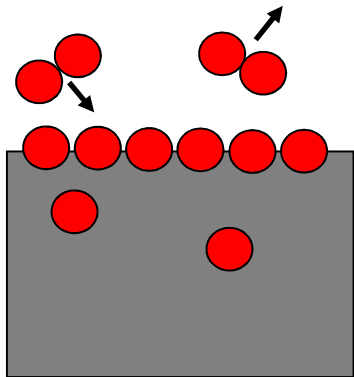
If also volume V is constant, the energy minimized is Helmholtz free energy F

$$F = U - TS$$

If (T,p) are constant, the energy minimized is Gibbs free energy G

$$G = U + pV - TS = \sum_i \mu_i N_i$$

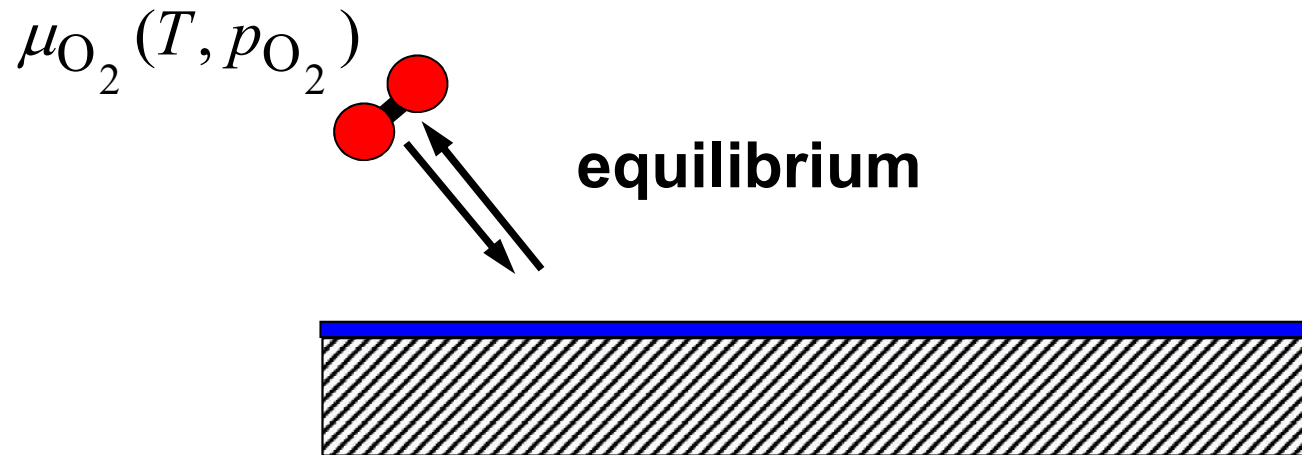
Chemical potential μ_i of the i -th atom type is the change in free energy as the number of atoms of that type in the system increases by 1



In thermodynamic equilibrium, μ_i is the same in the whole system (surface, bulk, gas)

Statistics plays a crucial role due to a macroscopically large number of particles in the system

First-principles atomistic thermodynamics



$$G(T, p) = E^{\text{tot}} + F^{\text{vib}} - TS^{\text{conf}} + pV$$

DFT

Surface free formation energy:

$$\gamma_{\text{surf}}(T, p) = \frac{1}{A} \left[G_{\text{surf}}(\{N_i\}) - \sum_i \mu_i N_i \right]$$

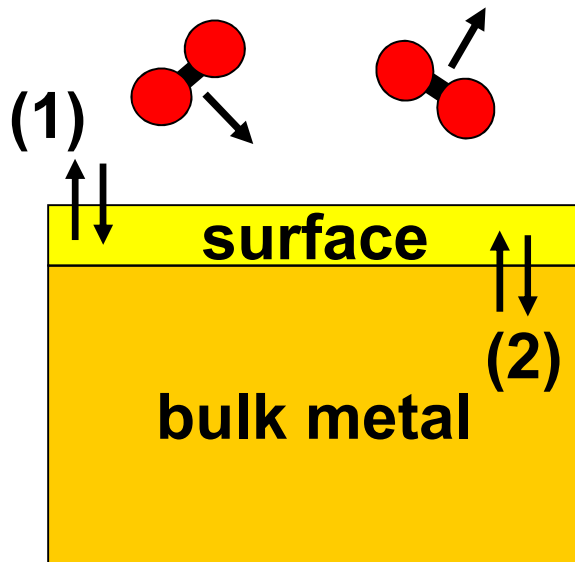
Defect free formation energy:

$$\Delta G(T, p) = \frac{1}{A} \left[G_{\text{def}}(\{N_i + \Delta N_i\}) - \sum_i \mu_i \Delta N_i - G_{\text{perf}}(\{N_i\}) \right]$$

C.M. Weinert and M. Scheffler, *Mater. Sci. Forum* 10-12, 25 (1986);
 E. Kaxiras *et al.*, *Phys. Rev. B* 35, 9625 (1987)

Example I: Metal surface in contact with O₂ gas

$$\gamma_{\text{surf}}(T, p) = \frac{1}{A} [G_{\text{surf}}(N_{\text{O}}, N_{\text{M}}) - N_{\text{O}}\mu_{\text{O}} - N_{\text{M}}\mu_{\text{M}}]$$



Reservoirs:

1) $\mu_{\text{O}}(T, p_{\text{O}_2})$ from ideal gas

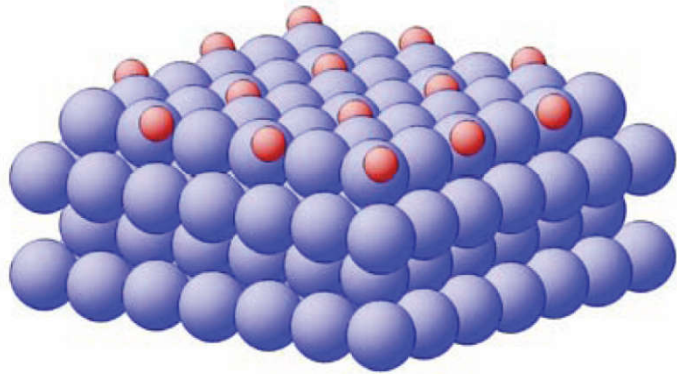
2) $\mu_{\text{M}} = g_{\text{M}}^{\text{bulk}}$

Neglect for now F^{vib} and TS^{conf} :

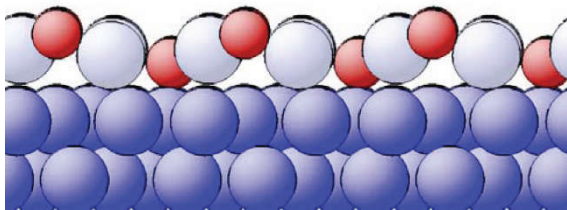
$$\gamma(T, p) \approx \left(E_{\text{surf}}^{\text{slab}} - N_{\text{M}}E_{\text{M}}^{\text{bulk}} - N_{\text{O}} \frac{1}{2} E_{\text{O}_2} \right) / A - N_{\text{O}}\Delta\mu_{\text{O}}(T, p) / A$$

Example I: Metal surface in contact with O₂ gas

$$\gamma \approx \left(E_{\text{surf}}^{\text{slab}} - N_{\text{Pd}} E_{\text{Pd}}^{\text{bulk}} - N_{\text{O}} \frac{1}{2} E_{\text{O}_2} \right) / A - N_{\text{O}} \Delta \mu_{\text{O}} / A$$

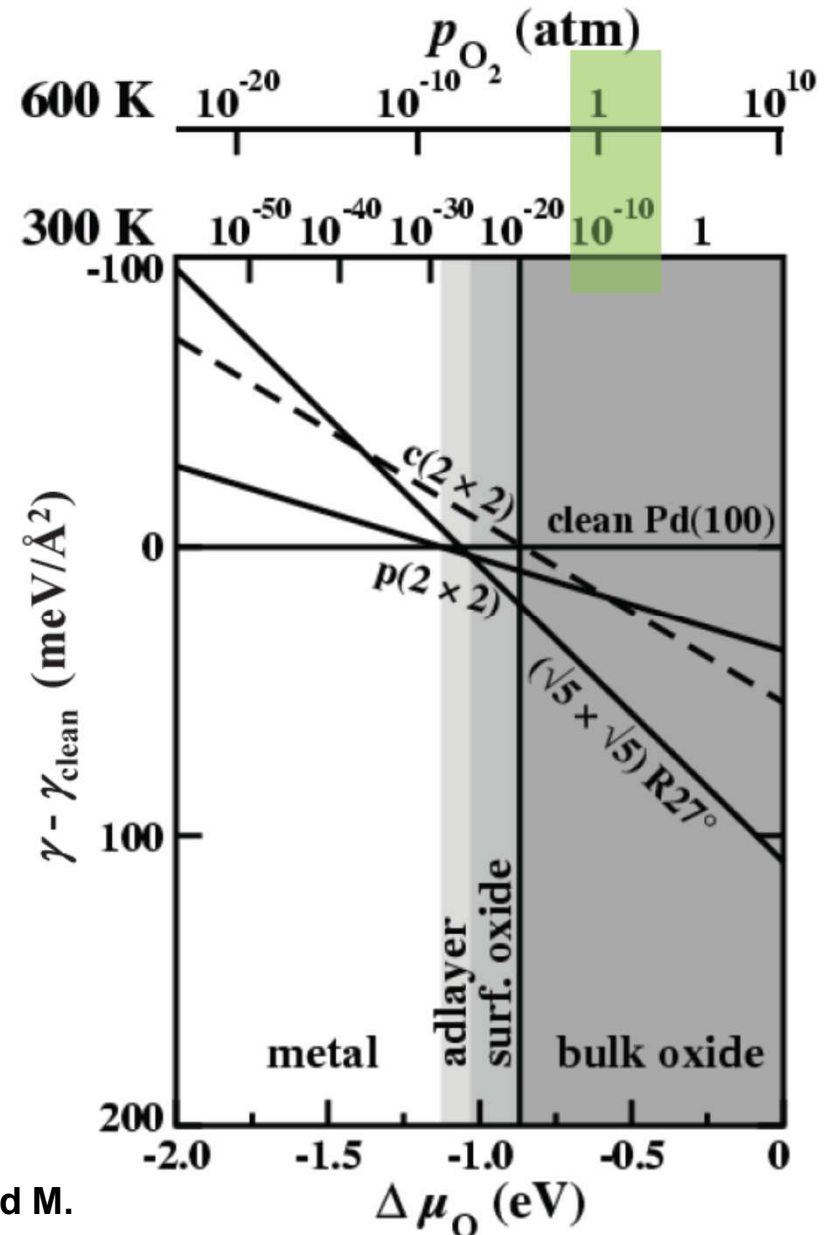


***p*(2x2) O/Pd(100)**

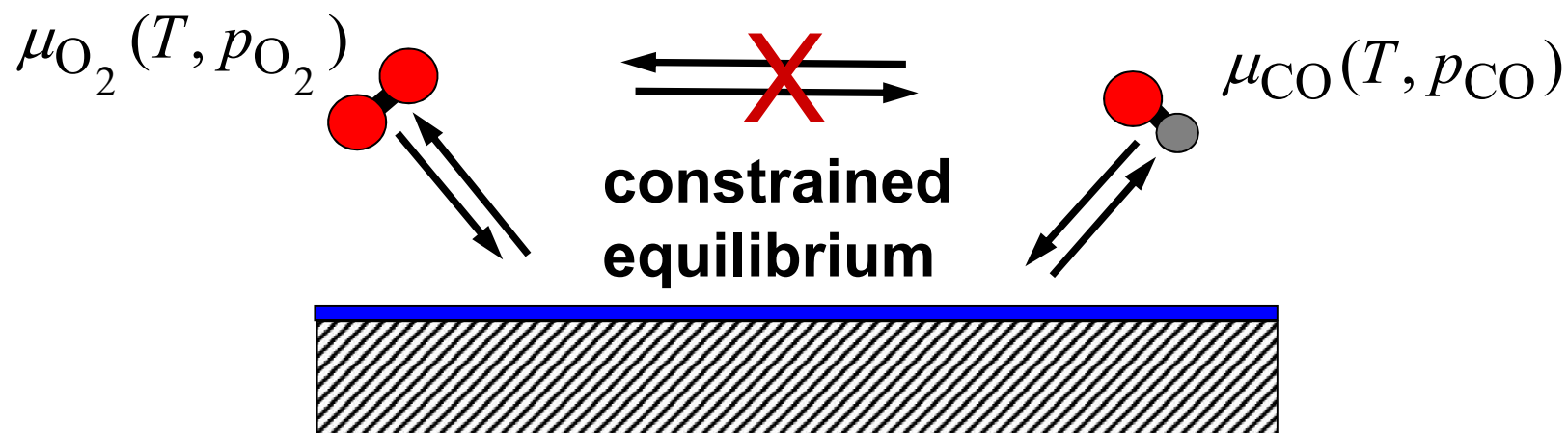


(√5x√5)R27° PdO(101)/Pd(100)

M. Todorova *et al.*, Surf. Sci. 541, 101 (2003); K. Reuter and M. Scheffler, Appl. Phys. A 78, 793 (2004)



First-principles atomistic thermodynamics: constrained equilibria



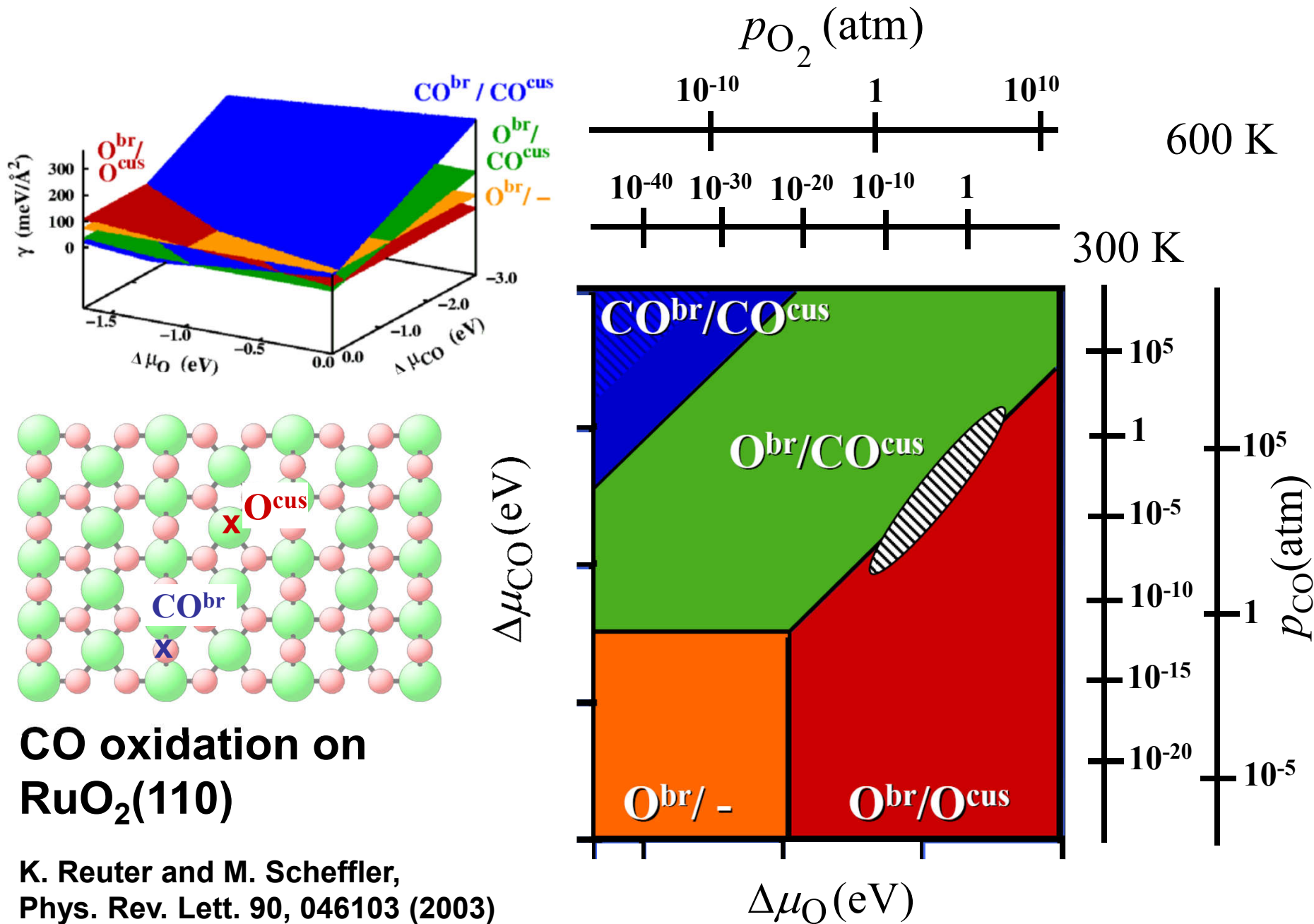
$$G(T, p) = E^{\text{tot}} + F^{\text{vib}} - TS^{\text{conf}} + pV$$

DFT

C.M. Weinert and M. Scheffler, *Mater. Sci. Forum* 10-12, 25 (1986); E. Kaxiras *et al.*, *Phys. Rev. B* 35, 9625 (1987);

K. Reuter and M. Scheffler, *Phys. Rev. B* 65, 035406 (2001); *Phys. Rev. B* 68, 045407 (2003)

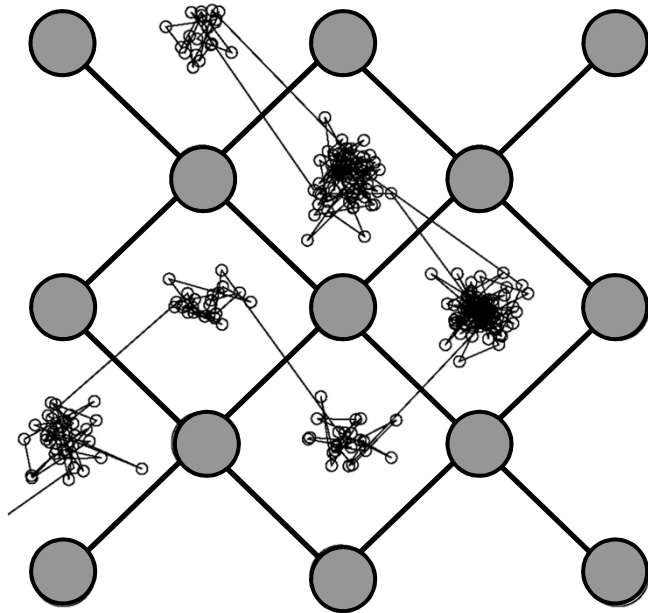
Surface phase diagrams



**CO oxidation on
RuO₂(110)**

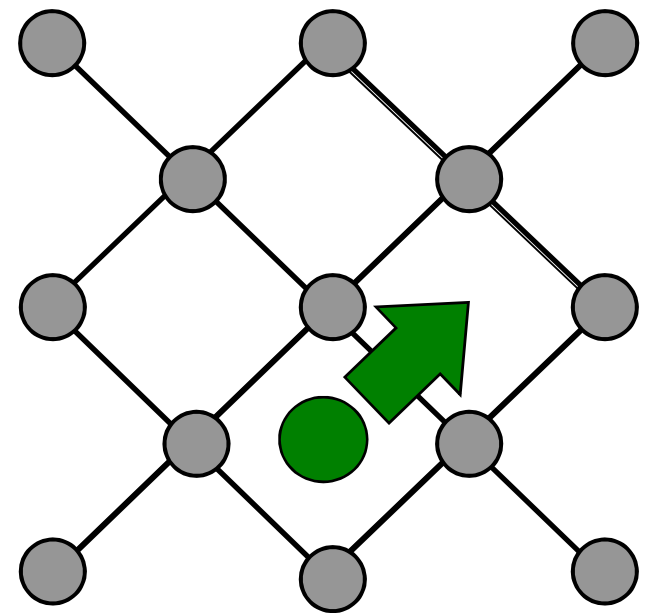
K. Reuter and M. Scheffler,
Phys. Rev. Lett. 90, 046103 (2003)

Kinetic Monte Carlo – essentially coarse-grained MD



**Molecular Dynamics:
the whole trajectory**

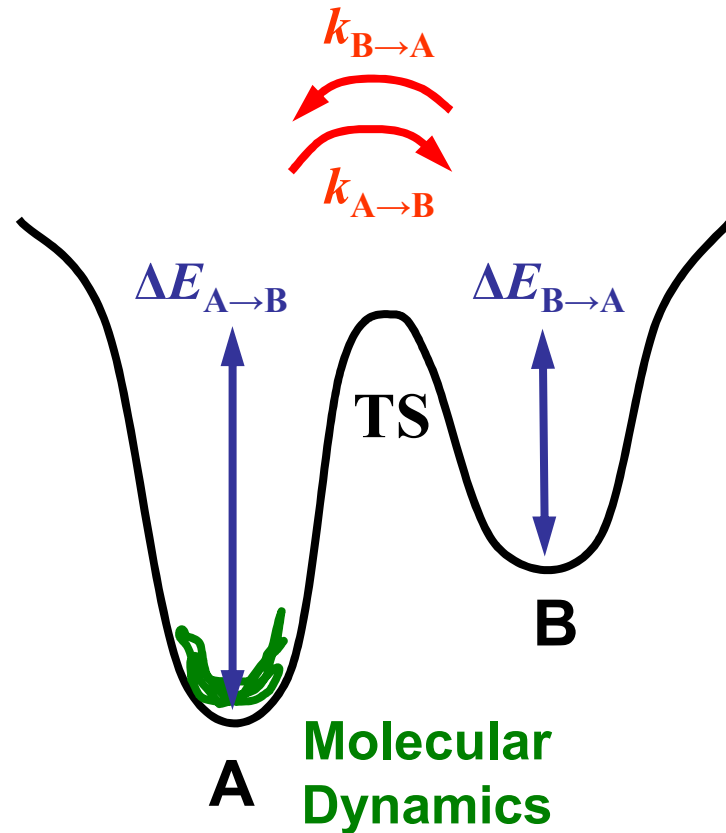
***ab initio* MD:
up to 50 ps**



**Kinetic Monte Carlo:
coarse-grained hops**

***ab initio* kMC:
up to minutes**

First-principles kinetic Monte Carlo simulations



Transition State
Theory

Molecular
Dynamics

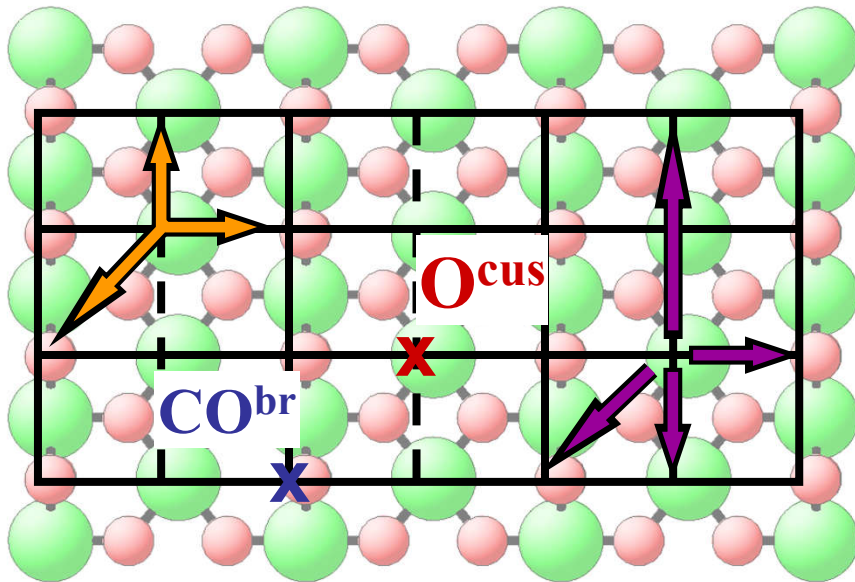
$$k_{i \rightarrow j} = \left(\frac{k_B T}{h} \right) \frac{Z_{\text{TS}(i \rightarrow j)}}{Z_i}$$
$$= \Gamma_o \exp \left(\frac{-\Delta E_{i \rightarrow j}}{k_B T} \right)$$

Master equations

$$\frac{dP_i(t)}{dt} = - \sum_j k_{i \rightarrow j} P_i(t) + \sum_j k_{j \rightarrow i} P_j(t)$$

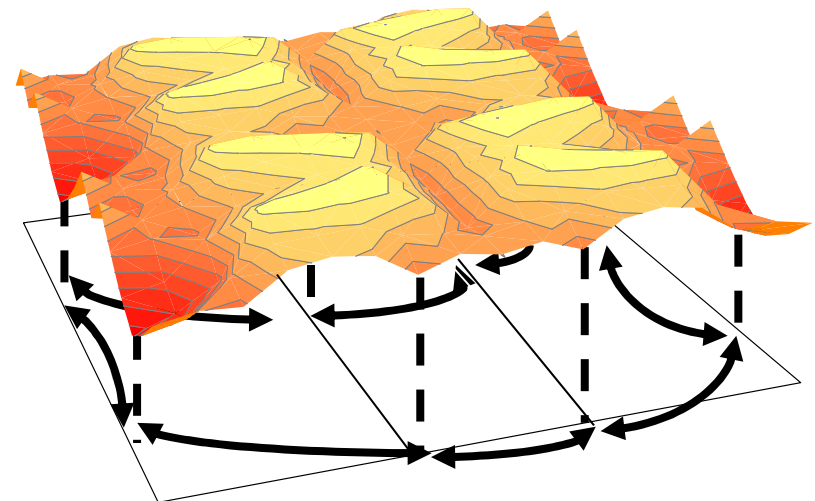
Crucial ingredients of a kMC simulation

$$\frac{dP_i(t)}{dt} = -\sum_j k_{i \rightarrow j} P_i(t) + \sum_j k_{j \rightarrow i} P_j(t)$$



1) Elementary processes

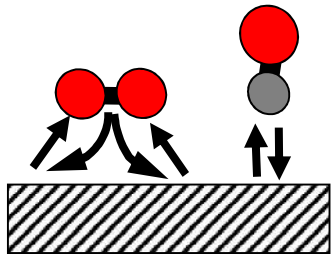
Fixed process list vs. „on-the-fly“ kMC
Lattice vs. off-lattice kMC



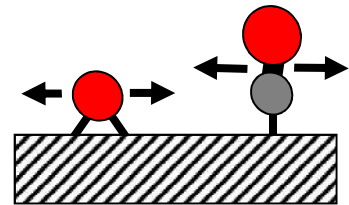
2) Process rates

PES accuracy
Reaction rate theory

kMC events for CO oxidation over RuO₂(110)

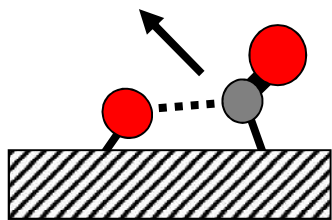


Adsorption: CO - unimolecular, O₂ – dissociative
no barrier
rate given by impingement $k \approx S_0 p / (2\pi m k_B T)$



Desorption: CO – 1st order, O₂ – 2nd order
out of DFT adsorption well (= barrier)
prefactor from detailed balance

Diffusion: hops to nearest neighbor sites
site- and element-specific
barrier from DFT (TST)
prefactor from DFT (hTST)



Reaction: site specific
immediate desorption, no readsorption
barrier from DFT (TST)
prefactor from detailed balance

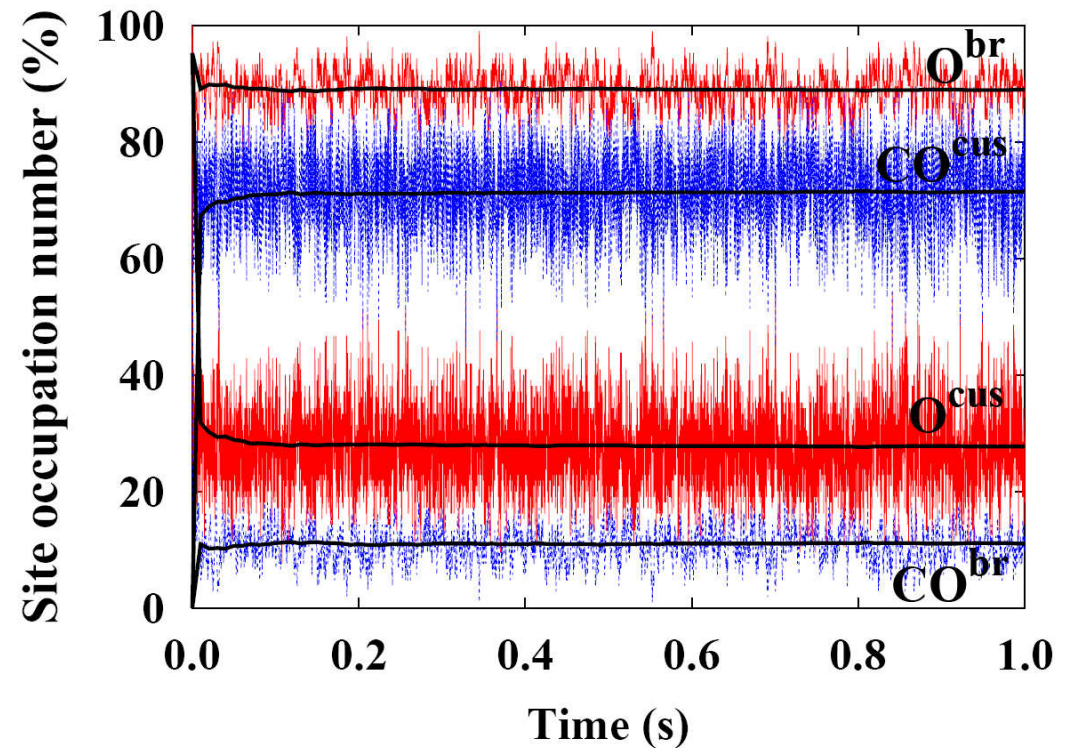
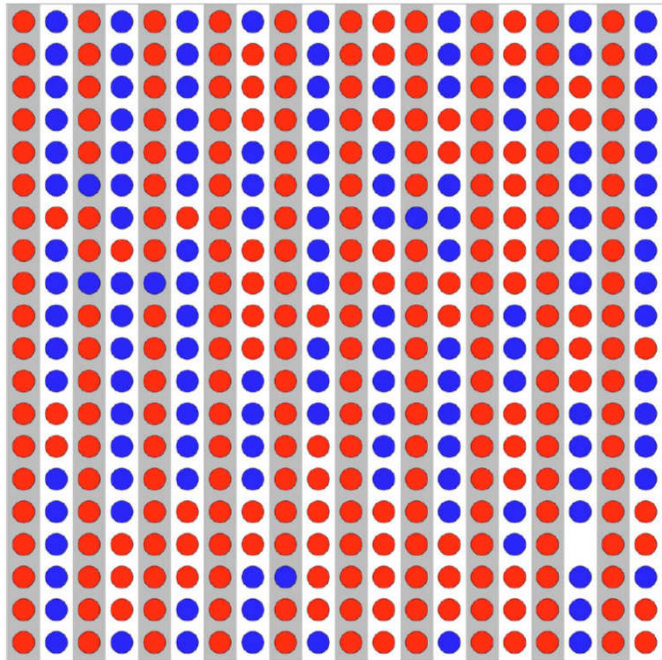
26 elementary processes
considered

The steady state of heterogeneous catalysis

$T = 600 \text{ K}$

$p_{\text{O}_2} = 1 \text{ atm}$

$p_{\text{CO}} = 7 \text{ atm}$

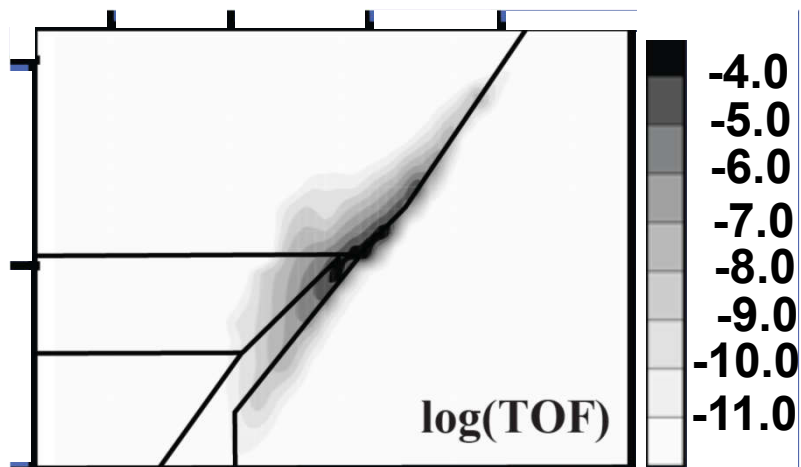
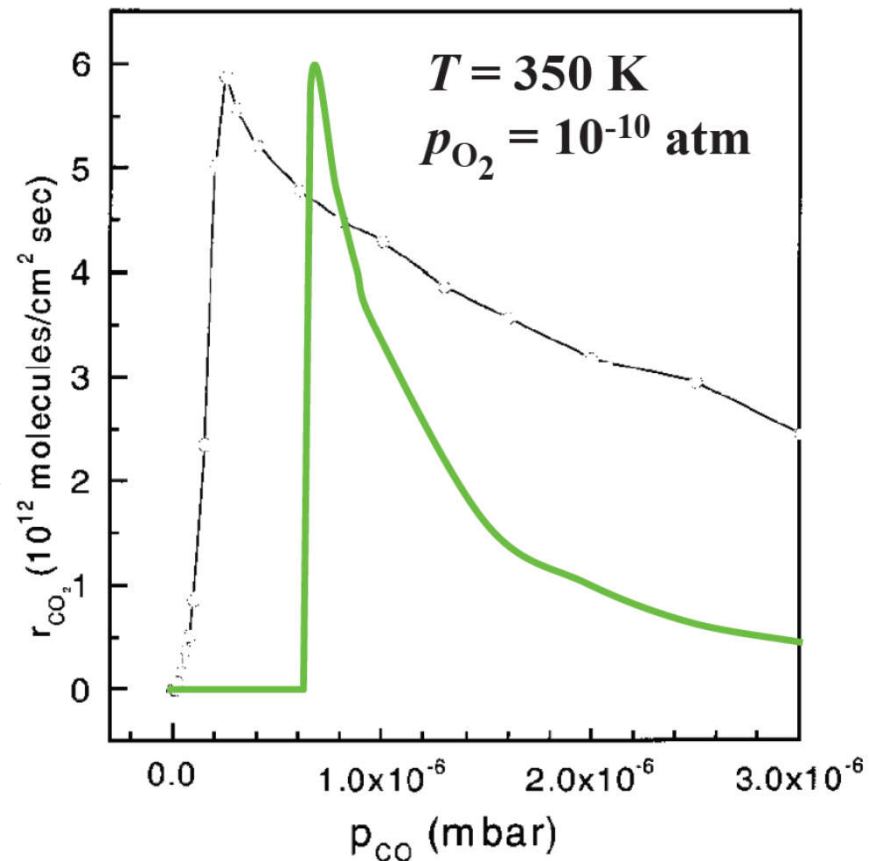
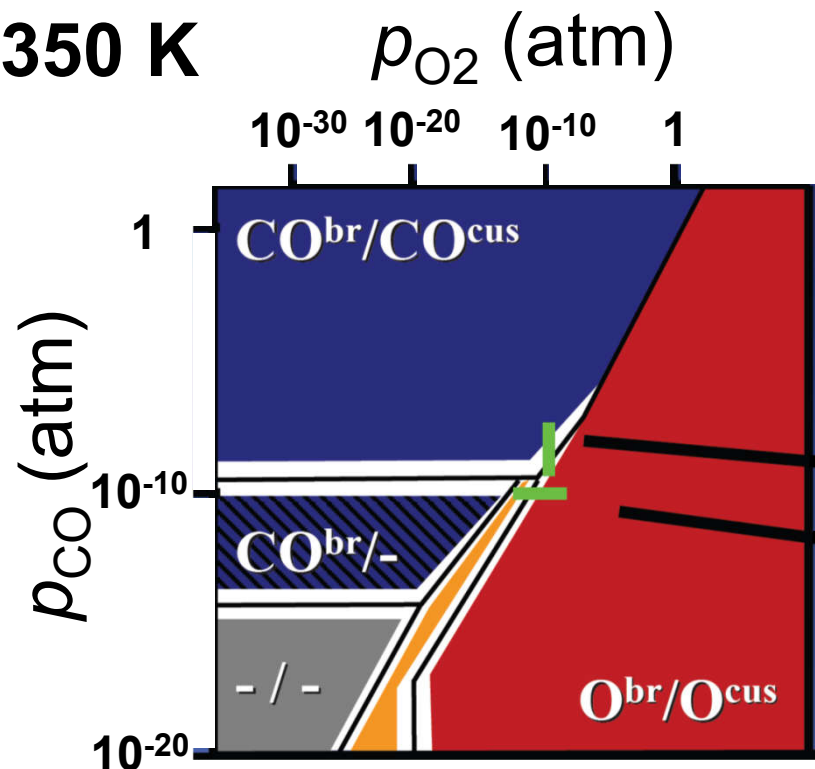


K. Reuter, D. Frenkel and M. Scheffler, Phys. Rev. Lett. 93, 116105 (2004)

K. Reuter, C. Stampfl, and M. Scheffler, Handbook of materials modeling, part A. Methods, p. 149, Springer, Berlin (2005)

Comparison to experiment

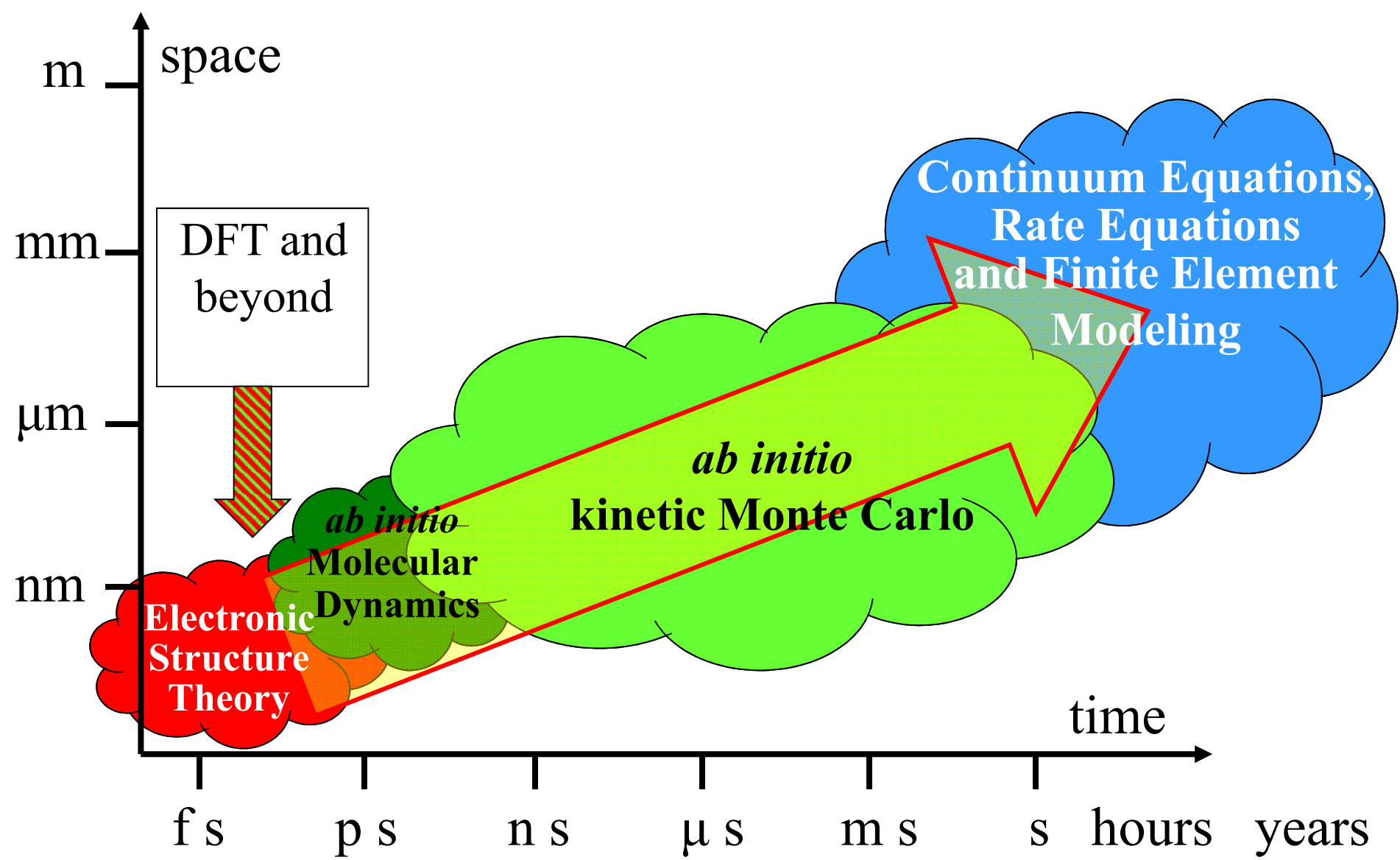
350 K



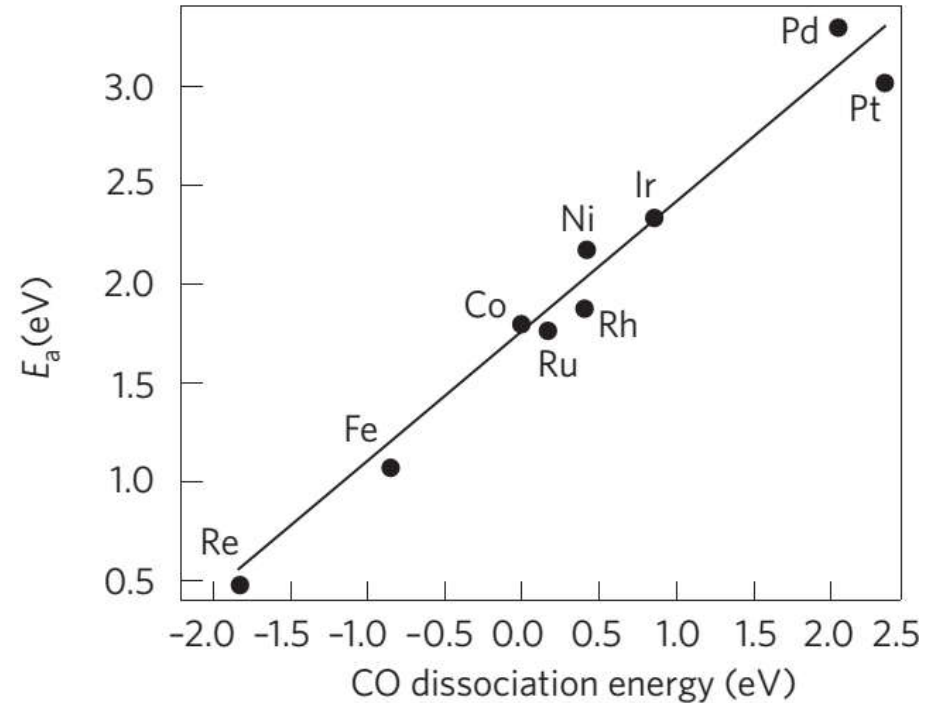
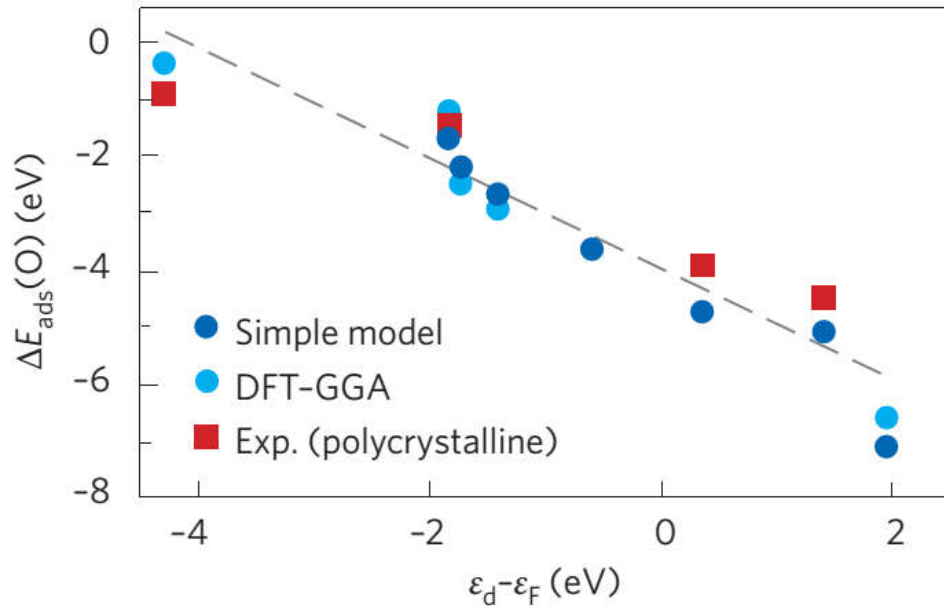
J. Wang *et al.*, *J. Phys. Chem. B*
106, 3422 (2002)

→ cf. NH₃/Ru(0001):
K. Honkala *et al.*, *Science* 307, 555 (2005)

Catalyst design



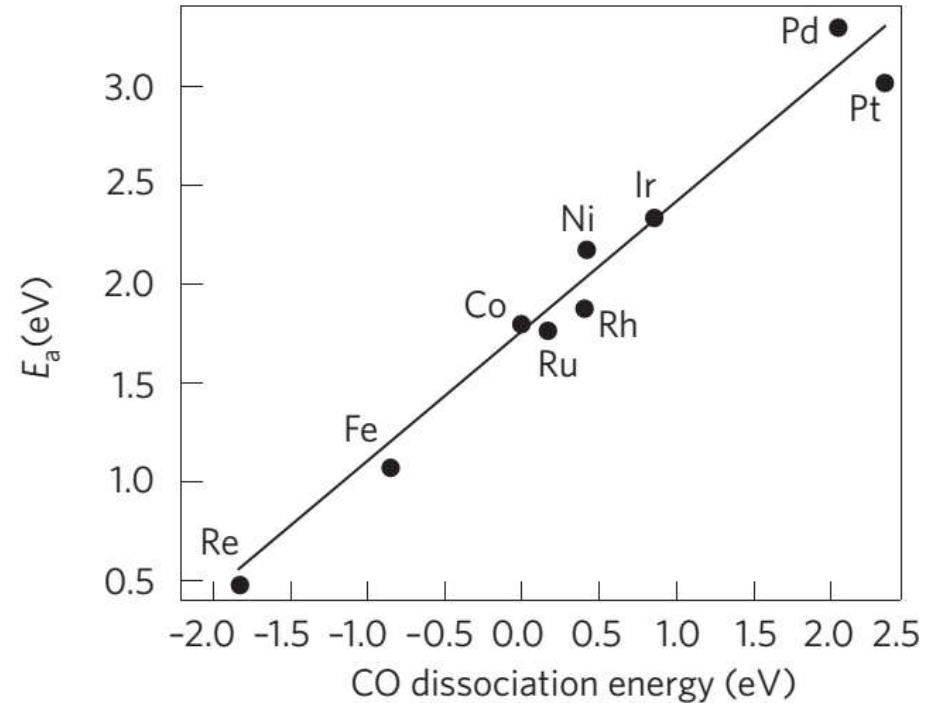
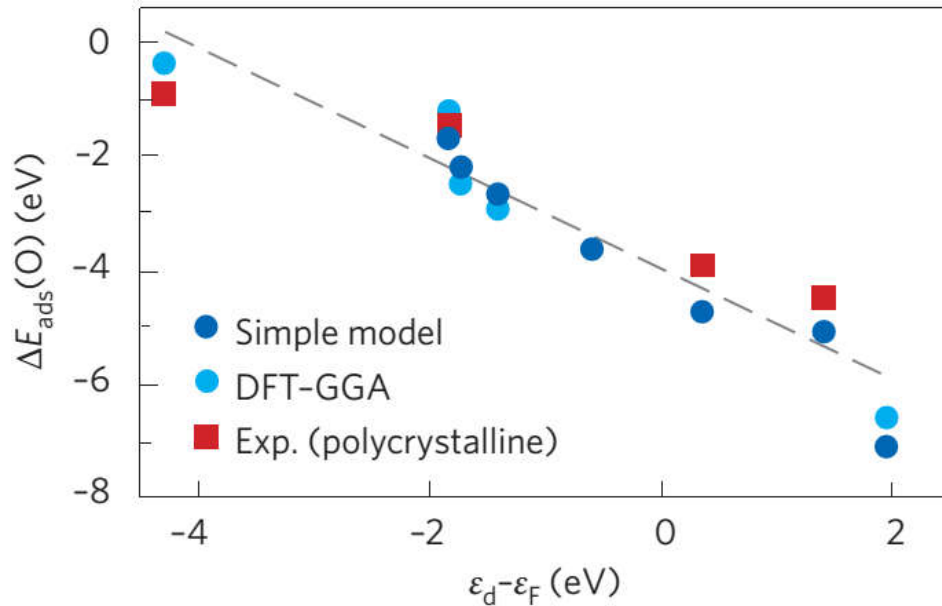
Descriptors



Simple(r) properties (bulk d-band center position and CO dissociation energy) are correlated to more complex properties (adsorption energy and reaction barrier)

The simpler quantities are called *descriptive parameters* (a *descriptor*)

Descriptors



A simple physical model (Newns-Anderson) motivates the *d*-band center descriptor

What if we don't know such a model, or we need a more accurate and more widely applicable model?

Artificial intelligence (including machine learning)

Descriptors

- 1) A descriptor d_i uniquely characterizes the material i as well as property-relevant elementary processes
- 2) The determination of the descriptor must not involve calculations as intensive as those needed for the evaluation of the property to be predicted
- 3) The dimension Ω of the descriptor should be as low as possible (for a certain accuracy demand)

Idea: calculate many *physically motivated* quantities (features), and use these features as a basis for the physical model under compactness constraints

L. M. Ghiringhelli, J. Vybiral, S. V. Levchenko, C. Draxl, and M. Scheffler, Phys. Rev. Lett. **114**, 105503 (2015)

Target property model: Kernel ridge regression versus feature selection

Regression models: Basis set expansion in materials space

kernel ridge regression

linear

$$P(\mathbf{d}) = \sum_{i=1}^N c_i \exp\left(-\|\mathbf{d}_i - \mathbf{d}\|_2^2 / 2\sigma^2\right)$$

minimize

$$P(\mathbf{d}) = \mathbf{d}\mathbf{c}$$

$$\sum_{i=1}^N (P(\mathbf{d}_i) - P_i)^2 + \lambda \sum_{i,j=1}^{N,N} c_i c_j \exp\left(-\|\mathbf{d}_i - \mathbf{d}_j\|_2^2 / 2\sigma^2\right)$$

$$\sum_{i=1}^N (P(\mathbf{d}_i) - P_i)^2 +$$

$$\lambda \|\mathbf{c}\|_0$$

$$\|\mathbf{d}_i - \mathbf{d}_j\|_2^2 = \sum_{\alpha=1}^{\Omega} (d_{i,\alpha} - d_{j,\alpha})^2$$

Target property model: Kernel ridge regression versus feature selection

kernel (Gaussian, Laplacian, linear ($d_i \cdot d_j$))

kernel ridge regression

linear

$$P(\mathbf{d}) = \sum_{i=1}^N c_i \exp(-\|\mathbf{d}_i - \mathbf{d}\|_2^2 / 2\sigma^2)$$

minimize

$$P(\mathbf{d}) = \mathbf{d}\mathbf{c}$$

$$\sum_{i=1}^N (P(\mathbf{d}_i) - P_i)^2 + \lambda \sum_{i,j=1}^{N,N} c_i c_j \exp(-\|\mathbf{d}_i - \mathbf{d}_j\|_2^2 / 2\sigma^2)$$

$$\sum_{i=1}^N (P(\mathbf{d}_i) - P_i)^2 + \lambda \|\mathbf{c}\|_0$$

penalty on similar data points

penalty on the number of non-zero coefficients $\|\mathbf{c}\|_0$

Mathematical formulation of the problem

P_j -- property value for material j (a function in materials space)

$d_{j,l}$ -- value of feature l related to material j (a basis function in materials space)

c_l -- coefficient of the expansion of the property function in terms of basis functions:

$$P_j = \sum_l d_{j,l} c_l \quad \text{How to find } c_l?$$

$$\sum_j \left(P_j - \sum_l d_{j,l} c_l \right)^2 + \lambda \|c\|_n \rightarrow \operatorname{argmin}(c)$$

$\|c\|_0$ -- number of non-zero coefficients \rightarrow NP hard! (need to try all combinations)

$\|c\|_2 = \sum_l |c_l|^2$ -- ridge regression \rightarrow not most compact!

$\|c\|_1 = \sum_l |c_l|$ -- LASSO (Least Absolute Shrinkage and Selection Operator) \rightarrow convex problem, equivalent to the NP-hard if features (columns of d) are uncorrelated

Compressed (compressive?) sensing



Raw: 15MB

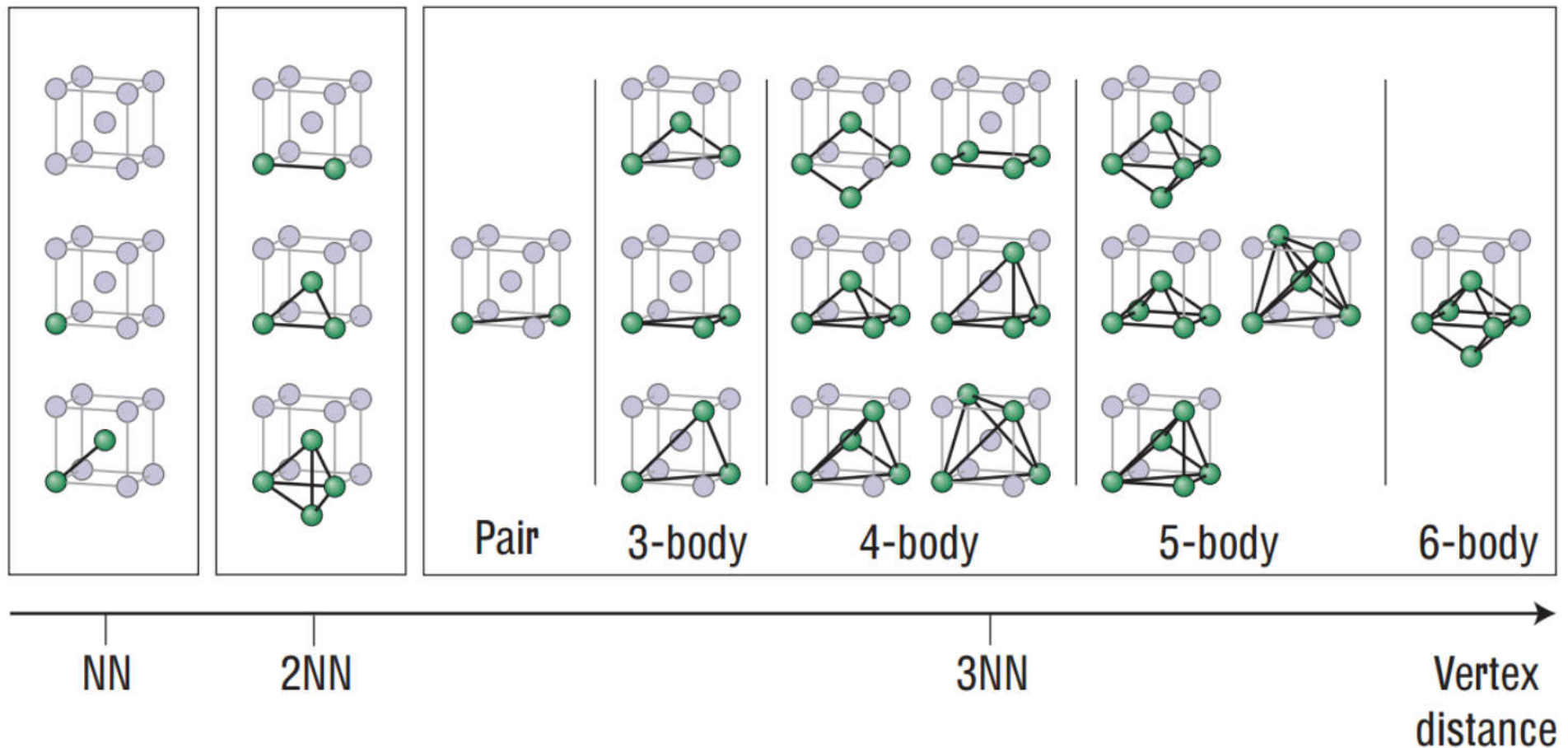


JPEG: 150KB

Expand in a basis (wavelets) → use LASSO to select most important basis functions → store compressed image

Compressive Sensing for Cluster Expansion

$$E(\sigma) = E_0 + \sum_f \Pi_f(\sigma) J_f \quad \min_{J_f} \left(\lambda \sum_f |J_f| + \sum_i (E^{DFT}(\sigma_i) - E^{CE}(\sigma_i))^2 \right) \rightarrow J_f$$



Enabling Feature Spaces with Billions of Elements by Sure Independence Screening

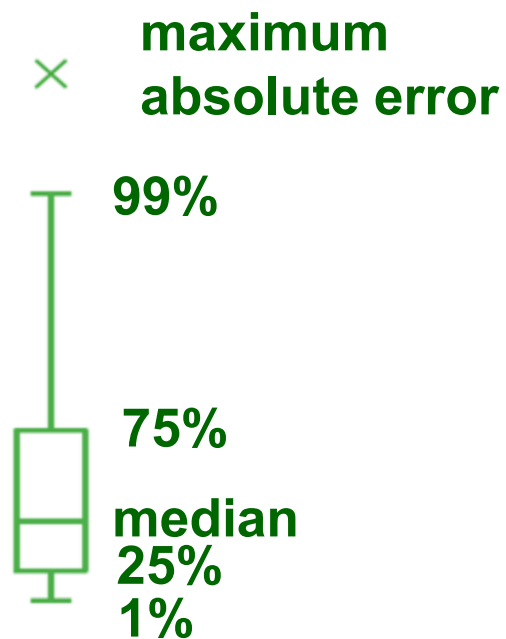
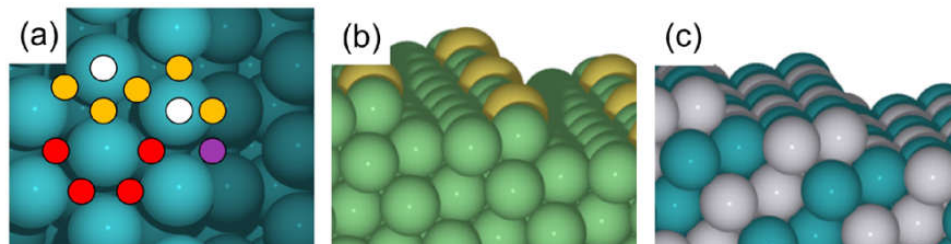
$\|c\|_1 = \sum_l |c_l|$ -- LASSO \rightarrow convex problem, equivalent to the NP-hard if features are uncorrelated \rightarrow not the case for many features \rightarrow Sure Independence Screening plus Selection Operator (SISSO)

1. Systematically construct a huge feature space (10^{11}) from primary features: $\hat{R} = \{+, -, \cdot, ^{-1}, ^2, ^3, \sqrt{\quad}, \exp, \log, |-|\}$ (use physically meaningful combinations!)
2. Select top ranked features using *Sure Independence Screening (SIS)*^[1] (correlation learning). Select n features corresponding to the n largest projection on the target property, i.e. largest components of the vector $(D^T y)$
 y : vector with the target property (e.g., catalytic activity)
 D : matrix of the feature space (e.g., 100 x 100 billion elements)
3. Apply a sparsifying operator (l_0 regularization) to the selected features to determine 1D, 2D,... descriptors

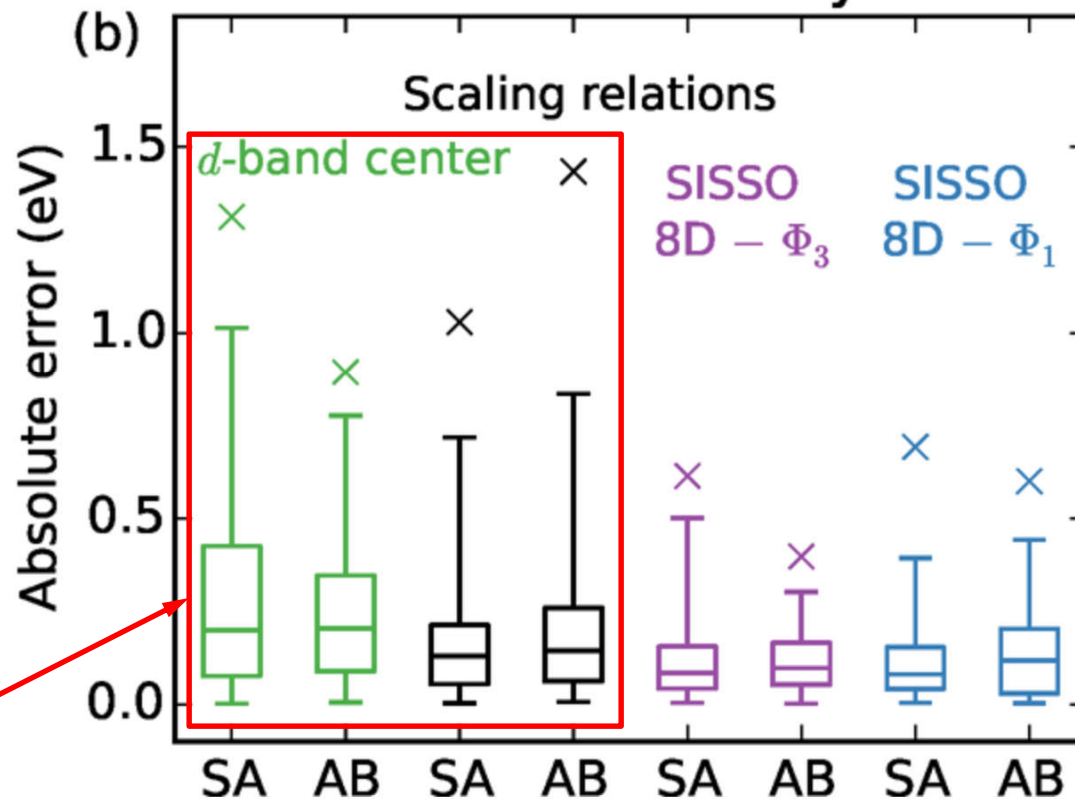
SISSO: Example

- Adsorption of molecules on metal surfaces

Adsorption of C, CH, CO, H, O, OH)



Test results alloys



previous state of the art

Conclusions

- **Catalyst design is an important and challenging problem**
- **Catalysis is intrinsically multiscale phenomenon**
- **Catalytic materials are dynamically changing during the reaction**
- **Searching for (meta)stable atomic configurations (e.g., with USPEX) at given temperature and pressure is very important**
- **Complexity can be bridged by artificial intelligence**