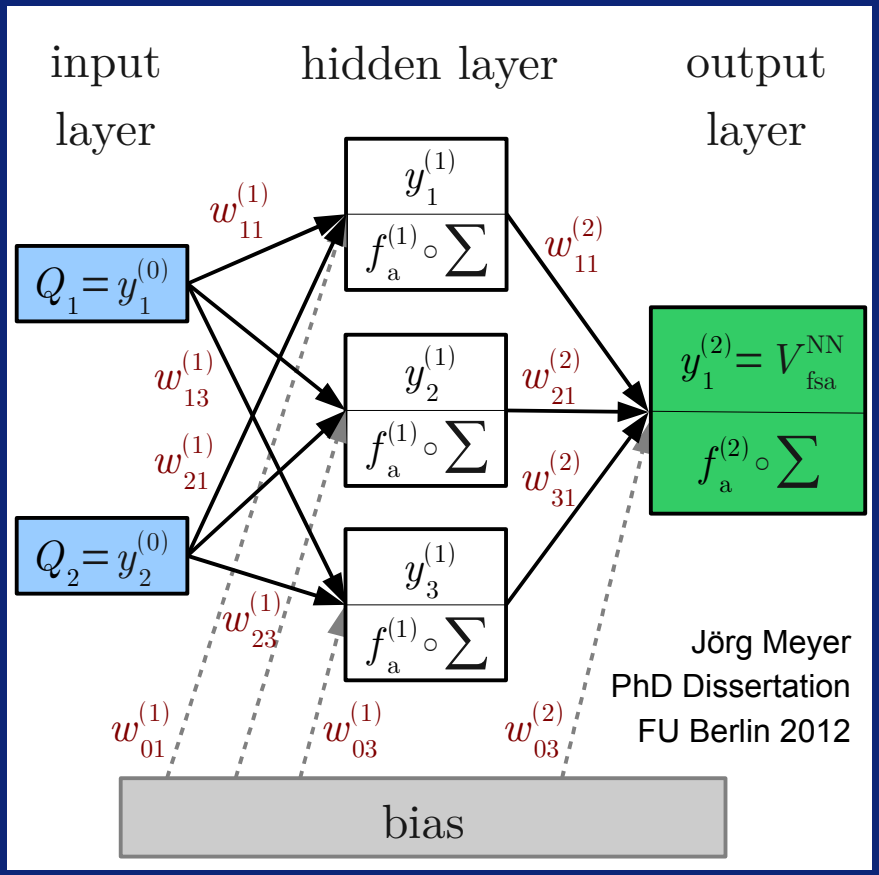
Machine learning in computational chemistry

Foundations and applications



Jörg Meyer, Theoretical Chemistry
j.meyer@chem.leidenuniv.nl

Universiteit Leiden
The Netherlands

Winter School on
Theoretical Chemistry and Spectroscopy
Han-sur-Lesse, December 10 - 14

Contents

1) Neural Networks (NNs):
   Structure & “Learning” (<1h)
   Hands-on: Training a simple NN (45 min)

2) NNs for potential energy surfaces:
Coordinate representation (~90 min)
***Hands-on:* (Re-)Fitting a potential energy surface for O_2@Pd(100) (~15 min + tonight)**

3) Applications in gas-surface dynamics (<1h)

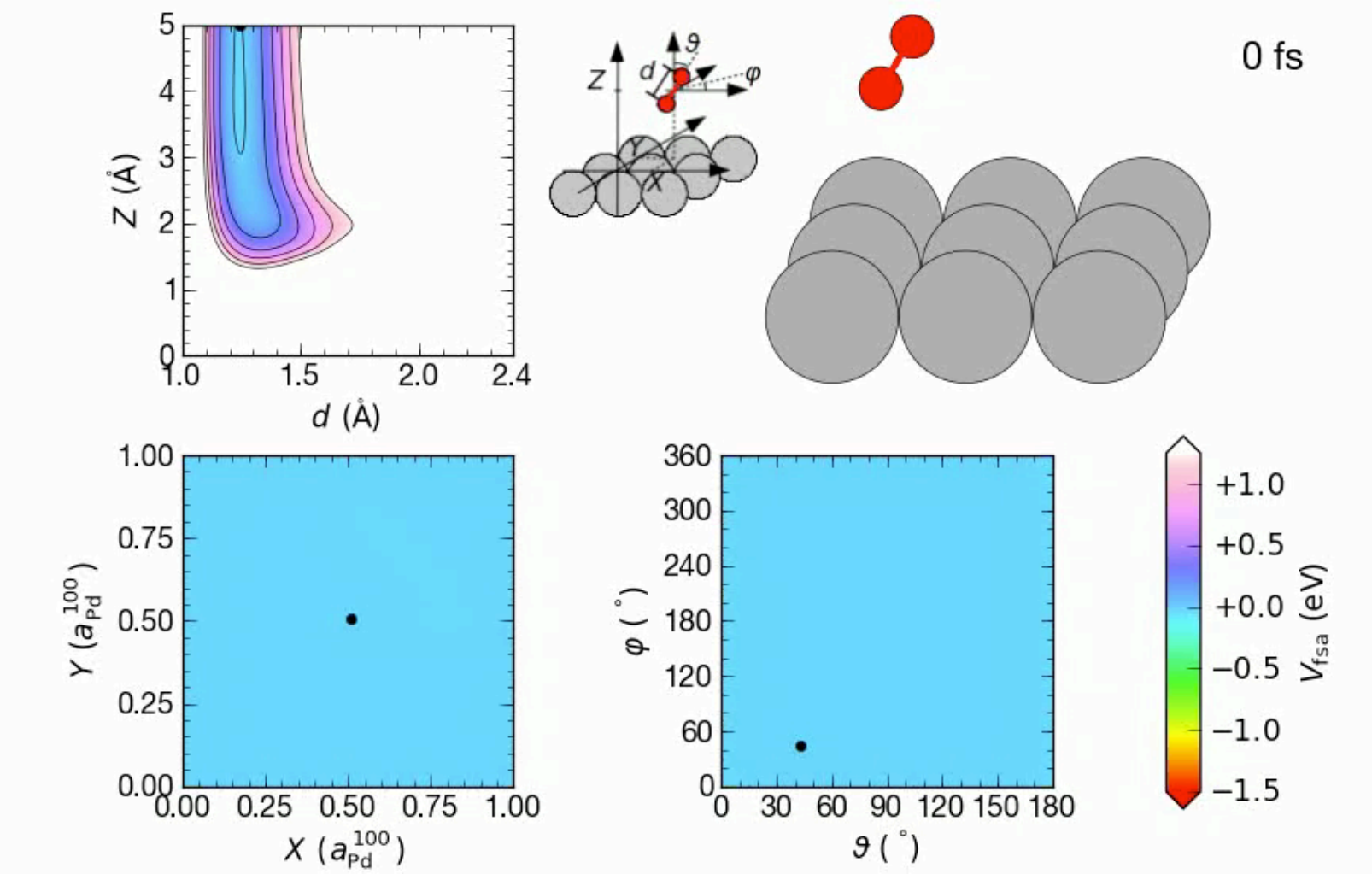
Back to the roots

Neural Network PESs for gas-surface dynamics

State of the art

High-Dimensional Neural Network Potentials

Interaction of molecules with metal surfaces



→ no “chemically intuitive interaction terms” (bonded+non-bonded) available…

“Early NN adopter”

Table 2. List of NNPs for molecule–surface interactions published to date.

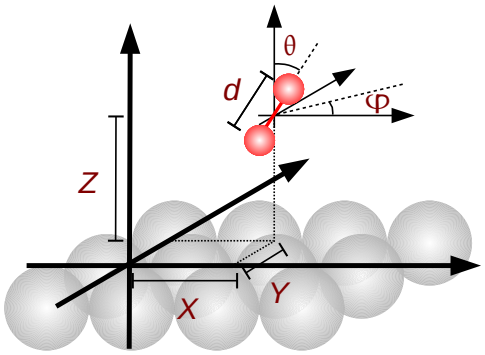
Year	Ref.	System	Reference method
1995	Blank *et al* [77]	CO @ Ni(1 1 1)	empirical PES
1995	Blank *et al* [77]	H_2 @ Si(1 0 0)-(2 × 1)	DFT (LDA)
2004	Lorenz *et al* [223]	H_2 @ K(2 × 2)/Pd(1 0 0)	DFT (PW91)
2005	Behler *et al* [123, 224, 225]	O_2 @ Al(1 1 1)	DFT (RPBE)
2006	Lorenz *et al* [226]	H_2 @ Pd(1 0 0)	empirical PES
2006	Lorenz *et al* [226]	H_2 @ (2 × 2)S/Pd(1 0 0)	empirical PES
2007	Ludwig and Vlachos [227]	H_2 @ Pt(1 1 1)	empirical PES
2007	Ludwig and Vlachos [227]	H_2 @ Pt(1 1 1)	DFT (PW91)
2008	Behler *et al* [225]	O_2 @ Al(1 1 1)	DFT (PBE)
2008	Latino *et al* [228]	ethanol @ Au(1 1 1)	DFT (B3LYP)
2008	Carbogno *et al* [229]	O_2 @ Al(1 1 1)	DFT (RPBE)
2009	Manzhos *et al* [97]	N_2O @ Cu(1 0 0)	DFT
2009	Carbogno *et al* [230]	O_2 @ Al(1 1 1)	DFT (RPBE)
2010	Latino *et al* [231]	ethanol @ Au(1 1 1)	DFT (B3LYP)
2010	Manzhos and Yamashita [232]	N_2O @ Cu(1 0 0)	DFT
2012	Goikoetxea *et al* [233]	O_2 @ Ag(1 1 1)	DFT (PBE)
2013	Liu *et al* [234]	HCl @ Au(1 1 1)	DFT (PW91)

J. Behler, J. Phys.: Condens. Matter **26**, 183001 (2014).

PESs construction

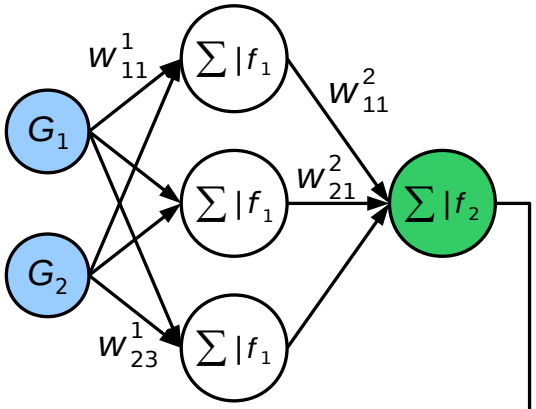
1. 6D DFT Potential Energy Surface → **2.** Interpolation: Neural Network[1] (exploiting symmetry) → **3.** Classical Molecular Dynamics

O_2 center of mass spherical coordinates



$V_{\text{fsa}}(X, Y, Z, d, \theta, \varphi)$

frozen surface approximation



$$V = f_2\left(\sum_{j=1}^{3} w^2_{j1} \cdot f_1\left(\sum_{i=1}^{2} w^1_{ij} \cdot G_i\right)\right)$$

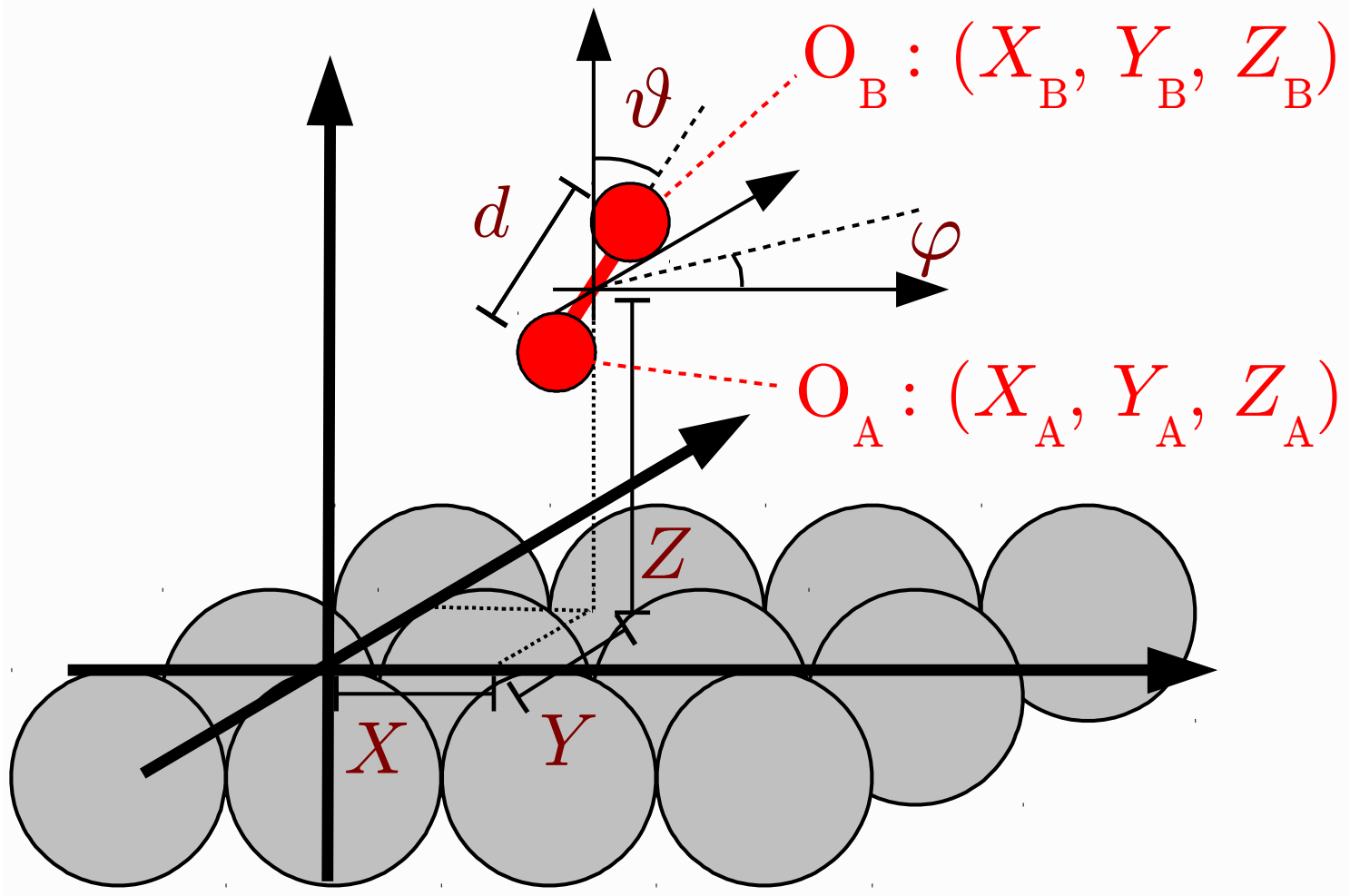
$$H = T + V_{\text{fsa}}$$

$$\dot{q}_i = \frac{\partial H}{\partial p_i}$$

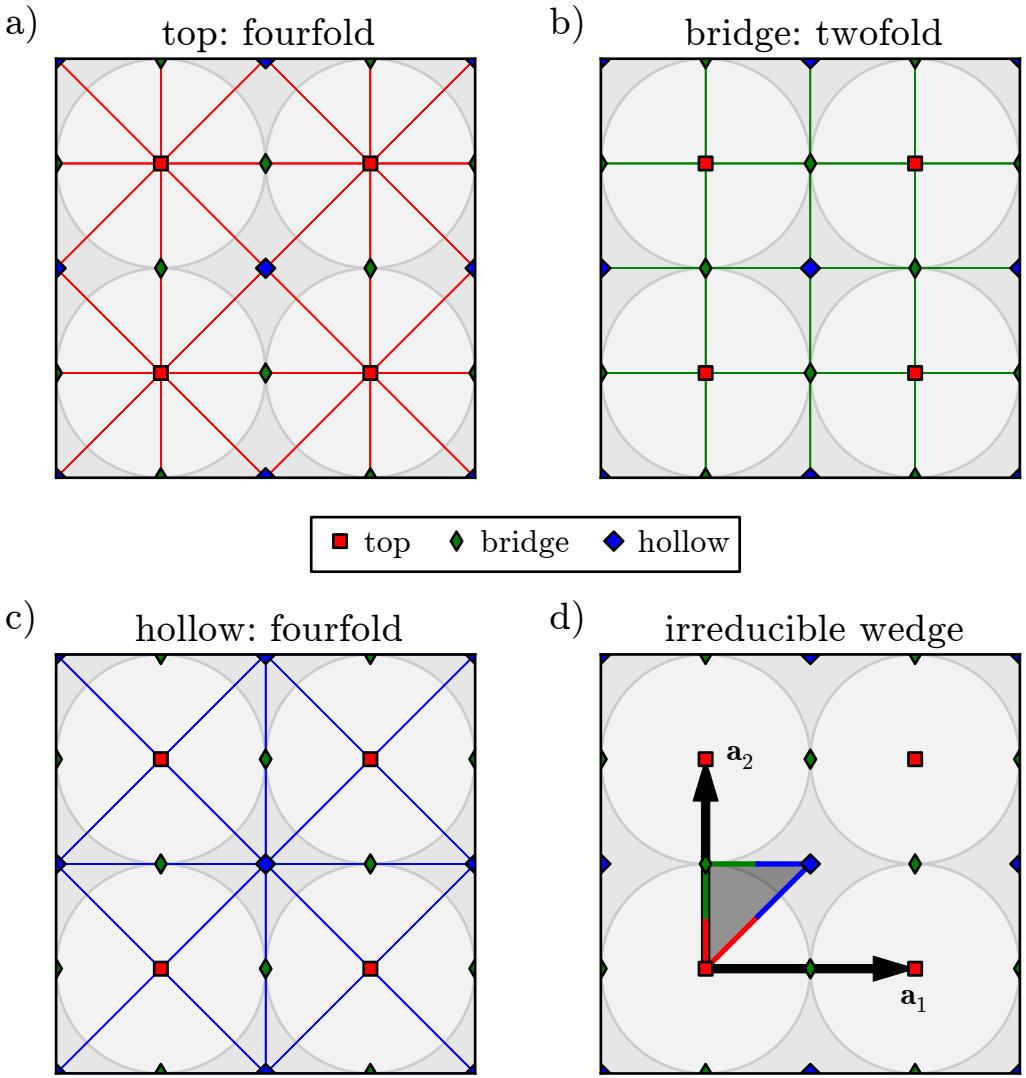
$$\dot{p}_i = -\frac{\partial H}{\partial q_i}$$

[1] S. Lorenz, A. Groß and M. Scheffler, Chem. Phys. Lett. **395** (2004) 210

Diatomic molecule on a surface

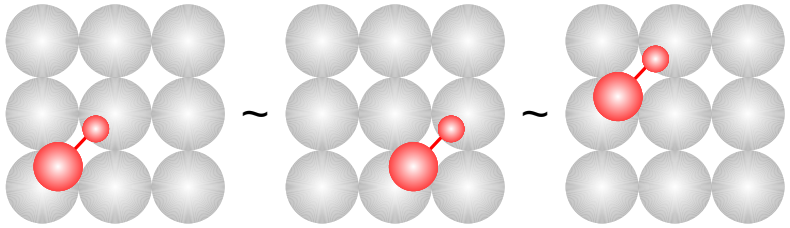


fcc(100) surface

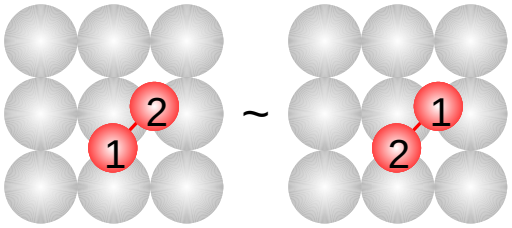


Symmetry (1)

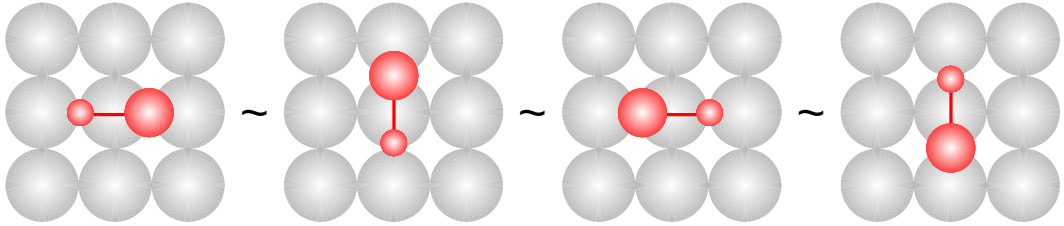
T_{100}



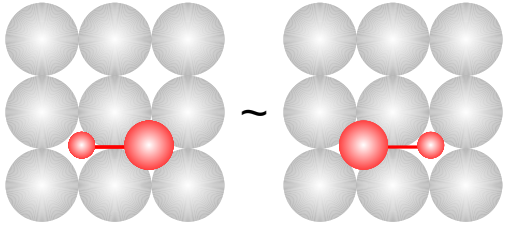
i (*X*,*Y*,*Z*)



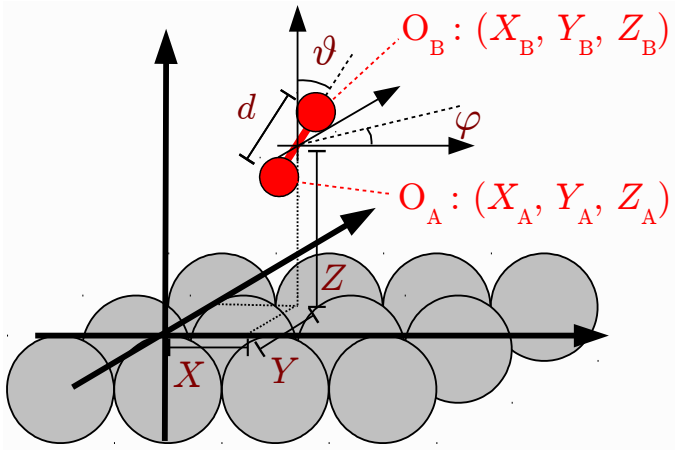
C_{4v} (*top*, hollow)



C_{2v} (*bridge*)



Symmetry (2)



$$\mathcal{Q}^{\mathrm{cart}} = \{\boldsymbol{R}^{\mathrm{cart}}\} \,/\, {\sim^{\mathrm{cart}}}$$

$$\equiv$$

$$\mathcal{Q}^{\mathrm{sph}} = \{\boldsymbol{R}^{\mathrm{sph}}\} \,/\, {\sim^{\mathrm{sph}}}$$

molecular configurations

inequivalent by symmetry

1:1

$\longrightarrow$

differentiable

?

$$\{\boldsymbol{Q}\} \subseteq \mathbb{R}^{\mathrm{n}}$$

symmetry adapted coordinates

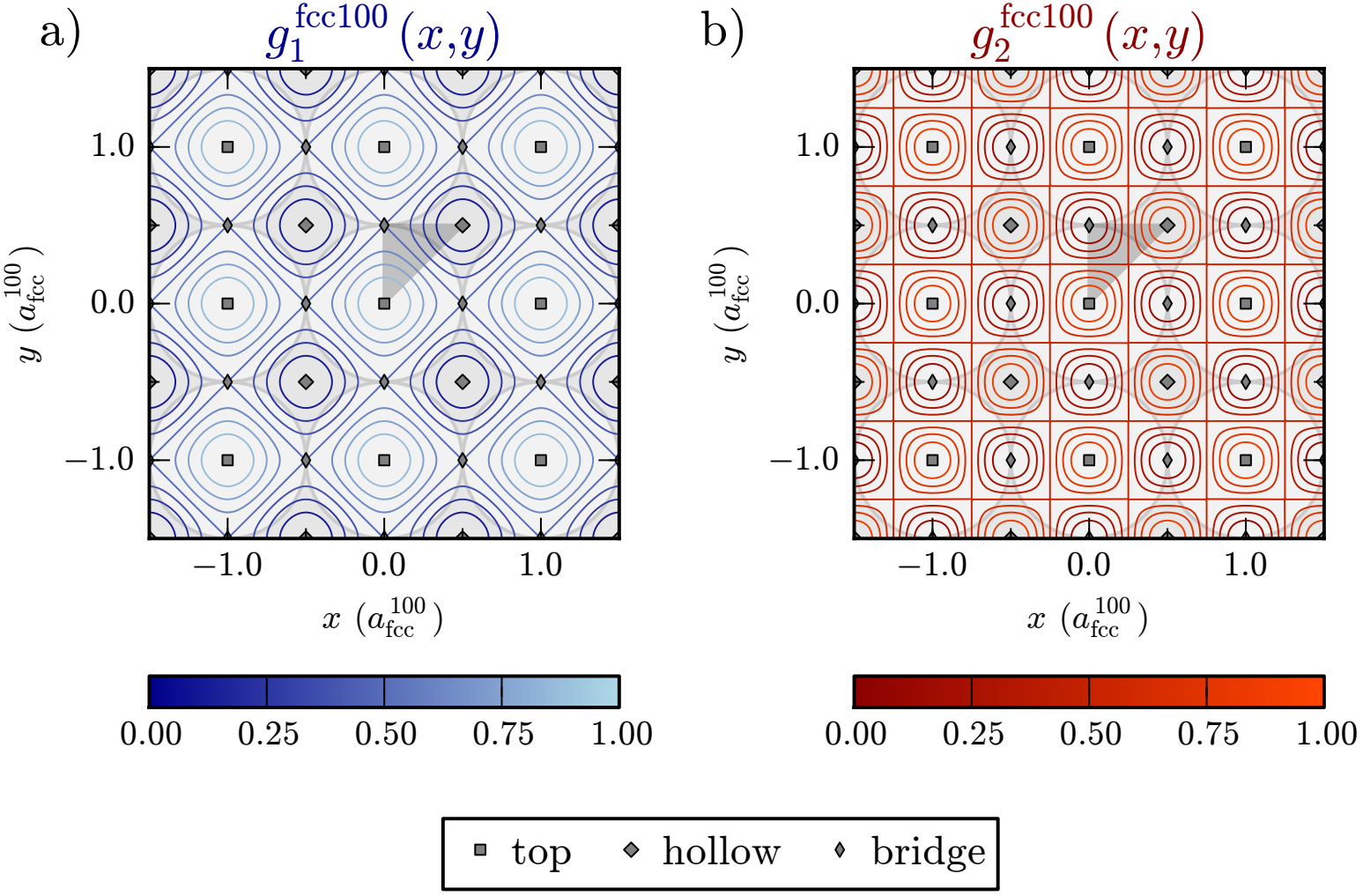
(Lateral) symmetry functions (1)

$$g_1^{\mathrm{fcc100}}(x,y) = \frac{1}{4}\left[\cos\left(\boldsymbol{G}_{01}^{\mathrm{fcc100}} \cdot \begin{pmatrix} x \\ y \end{pmatrix}\right) + \cos\left(\boldsymbol{G}_{10}^{\mathrm{fcc100}} \cdot \begin{pmatrix} x \\ y \end{pmatrix}\right)\right] + \frac{1}{2}$$
$$= \frac{1}{4}\left[\cos\left(\frac{2\pi}{a_{\mathrm{fcc}}^{100}} x\right) + \cos\left(\frac{2\pi}{a_{\mathrm{fcc}}^{100}} y\right)\right] + \frac{1}{2}$$

$$g_2^{\mathrm{fcc100}}(x,y) = \frac{1}{4}\left[\cos\left(\boldsymbol{G}_{11}^{\mathrm{fcc100}} \cdot \begin{pmatrix} x \\ y \end{pmatrix}\right) + \cos\left(\boldsymbol{G}_{1\bar{1}}^{\mathrm{fcc100}} \cdot \begin{pmatrix} x \\ y \end{pmatrix}\right)\right] + \frac{1}{2}$$
$$= \frac{1}{4}\left[\cos\left(\frac{2\pi}{a_{\mathrm{fcc}}^{100}} x\right) \cdot \cos\left(\frac{2\pi}{a_{\mathrm{fcc}}^{100}} y\right)\right] + \frac{1}{2} \quad .$$

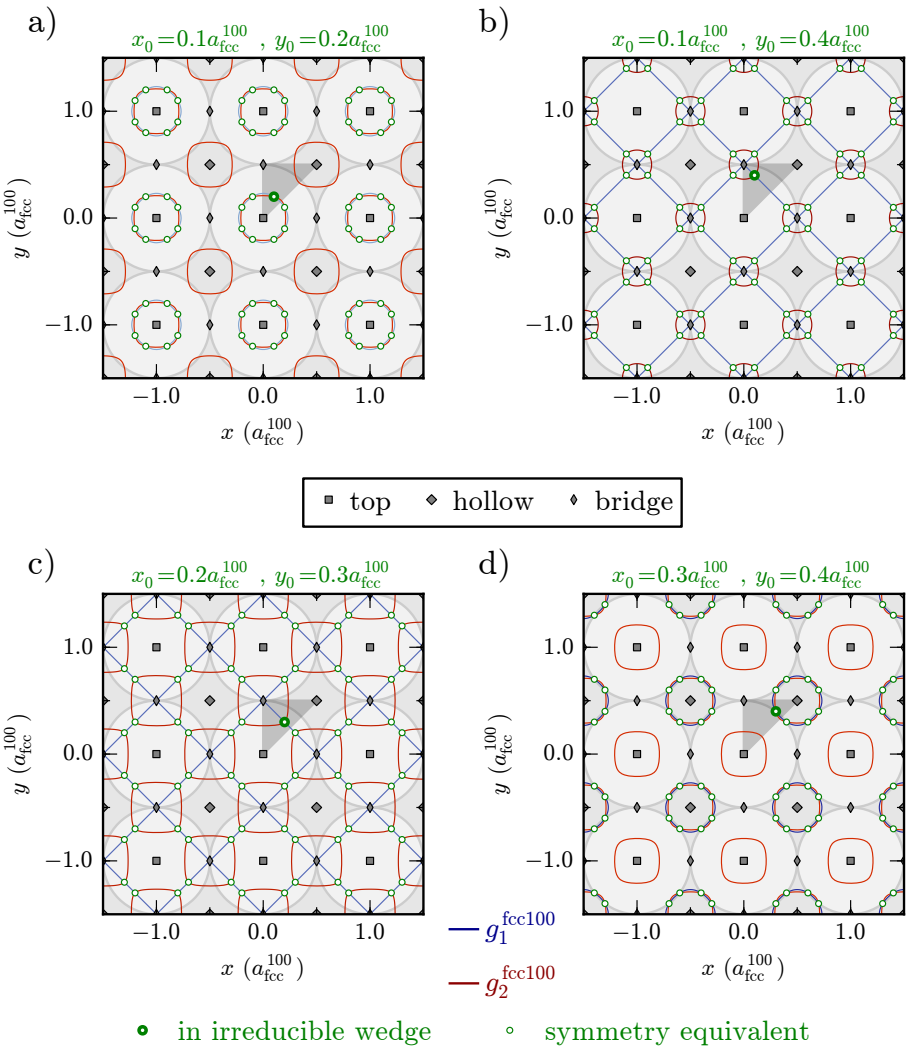
Jörg Meyer, PhD Dissertation, FU Berlin 2012

(Lateral) symmetry functions (2)

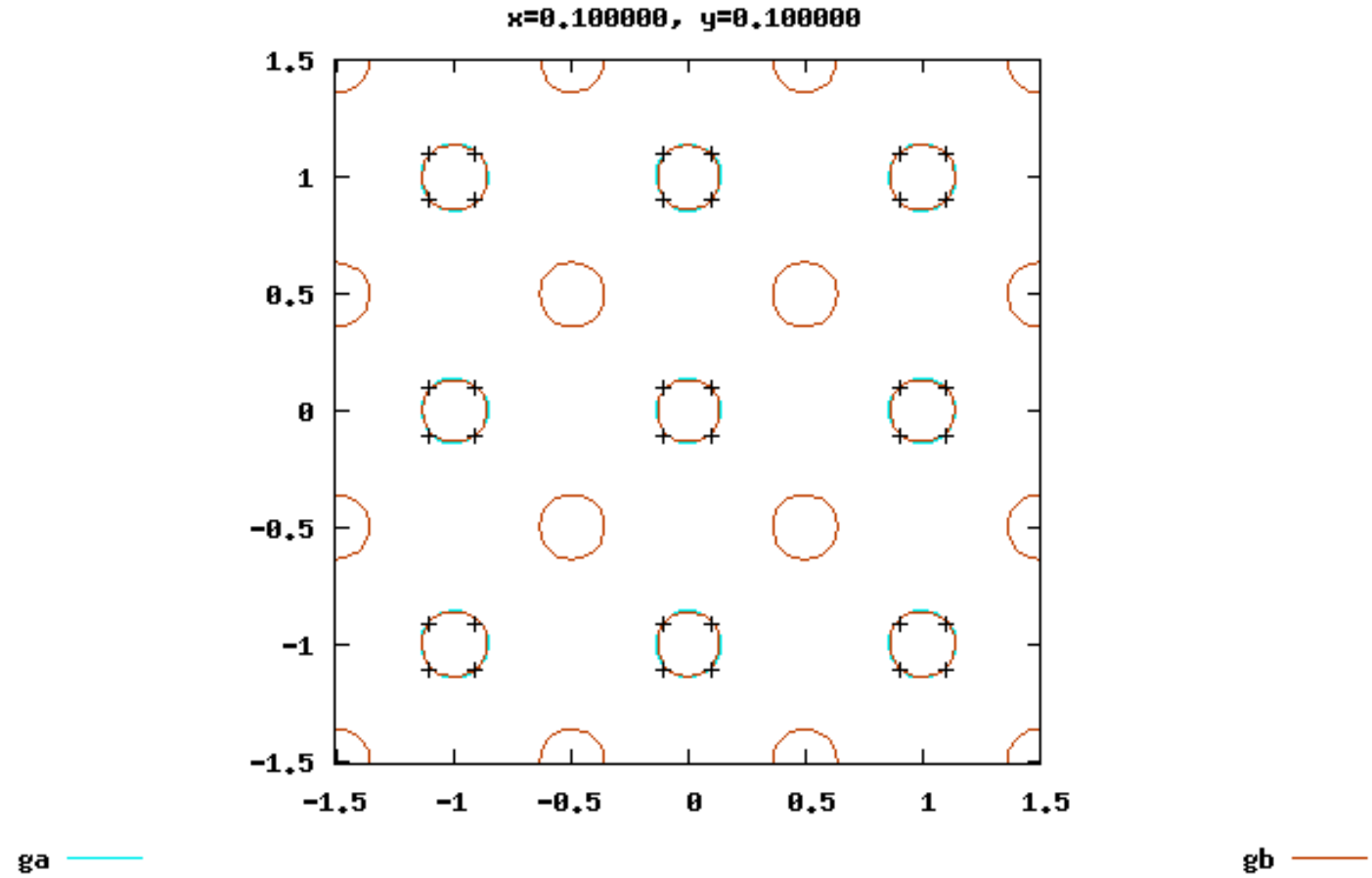


Jörg Meyer, PhD Dissertation, FU Berlin 2012

Intersections



Coordinates ←→ Input layer vector



Putting it together…

$$
\begin{aligned}
Q_1 &= \frac{1}{2}\left[f(Z_\mathrm{A})\cdot g_1^{\mathrm{fcc100}}(X_\mathrm{A},Y_\mathrm{A}) + f(Z_\mathrm{B})\cdot g_1^{\mathrm{fcc100}}(X_\mathrm{B},Y_\mathrm{B})\right]\\
Q_2 &= f(Z_\mathrm{A})\cdot g_1^{\mathrm{fcc100}}(X_\mathrm{A},Y_\mathrm{A}) \cdot f(Z_\mathrm{B})\cdot g_1^{\mathrm{fcc100}}(X_\mathrm{B},Y_\mathrm{B})\\
Q_3 &= \frac{1}{2}\left[f(Z_\mathrm{A})\cdot g_2^{\mathrm{fcc100}}(X_\mathrm{A},Y_\mathrm{A}) + f(Z_\mathrm{B})\cdot g_2^{\mathrm{fcc100}}(X_\mathrm{B},Y_\mathrm{B})\right]\\
Q_4 &= f(Z_\mathrm{A})\cdot g_2^{\mathrm{fcc100}}(X_\mathrm{A},Y_\mathrm{A}) \cdot f(Z_\mathrm{B})\cdot g_2^{\mathrm{fcc100}}(X_\mathrm{B},Y_\mathrm{B})\\
Q_5 &= f(Z)\cdot g_1^{\mathrm{fcc100}}(X,Y)\\
Q_6 &= f(Z)\cdot g_2^{\mathrm{fcc100}}(X,Y)\\
Q_7 &= f(Z)\\
Q_8 &= d\\
Q_9 &= \cos(\vartheta)^2 \quad ,
\end{aligned}
$$

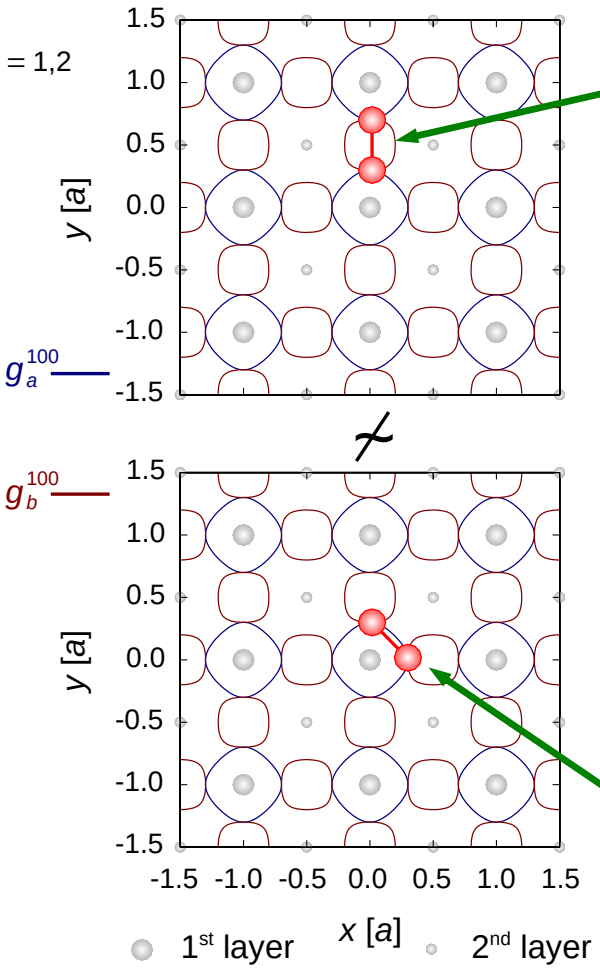
Jörg Meyer, PhD Dissertation, FU Berlin 2012

Coordinates ←→ Input layer vector

$$\left.\begin{aligned} a_i(X_i, Y_i, Z_i) &= \exp\left(-\frac{1}{2}Z_i\right)\cdot g^{a}_{100}(X_i, Y_i) \\ b_i(X_i, Y_i, Z_i) &= \exp\left(-\frac{1}{2}Z_i\right)\cdot g^{b}_{100}(X_i, Y_i) \end{aligned}\right\} \quad i = 1,2$$

$$\begin{aligned} G_1 &= \frac{1}{2}\cdot[a_1 + a_2] \\ G_2 &= a_1 \cdot a_2 \\ G_3 &= \frac{1}{2}\cdot[b_1 + b_2] \\ G_4 &= b_1 \cdot b_2 \\ G_5 &= d \\ G_6 &= [\cos(\theta)]^2 \\ \hline G_7 &= \exp\left(-\frac{1}{2}Z\right)\cdot g^{a}_{100}(X, Y) \\ G_8 &= \exp\left(-\frac{1}{2}Z\right)\cdot g^{b}_{100}(X, Y) \\ G_9 &= \exp\left(-\frac{1}{2}Z\right) \end{aligned}$$

- G_1 to G_6 are not sufficient (→ collisions!)



G1 = 0.24695544
G2 = 0.06098699
G3 = 0.12742918
G4 = 0.01623820
G5 = 1.11104452
G6 = 0.0
G7 = 0.18340583
G8 = 0.00106778
G9 = 0.36787944

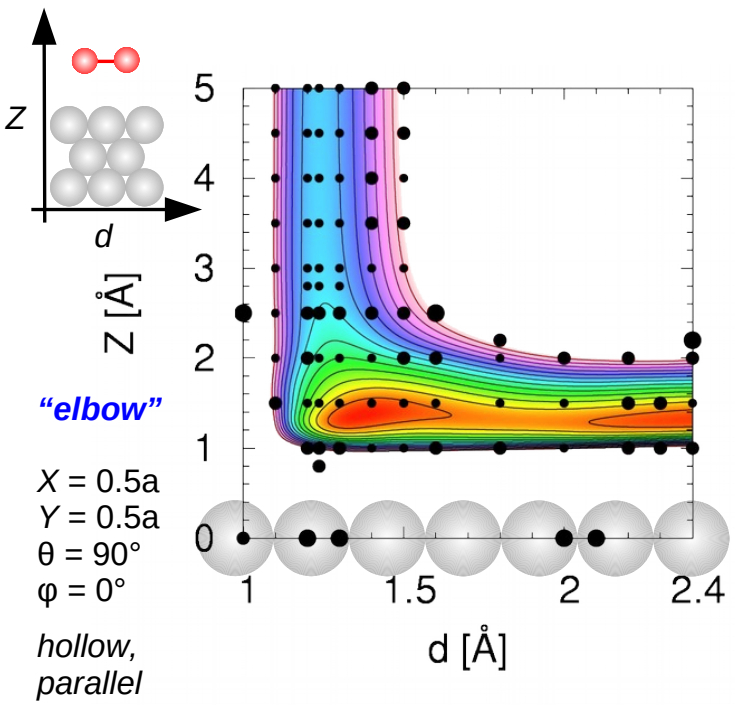
=

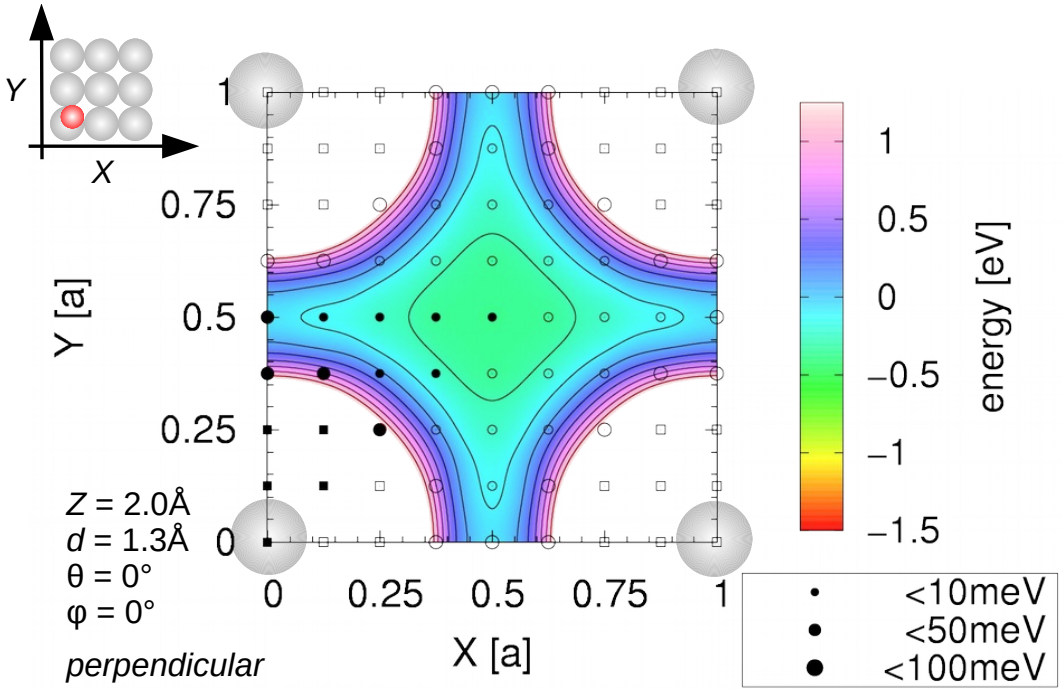
G1 = 0.24695544
G2 = 0.06098699
G3 = 0.12742918
G4 = 0.01623820
G5 = 1.11104452
G6 = 0.0
G7 = 0.28388258
G8 = 0.23824325
G9 = 0.36787944

≠

I. Goikoetxea, J. Beltrán, J. Meyer, and K. Reuter, New J. Phys. **14**, 013050 (2012).
V. J. Bukas, J. Meyer, M. Alducin, and K. Reuter, Z. Phys. Chem. **227**, 1523 (2013).

Quality of NN-PES



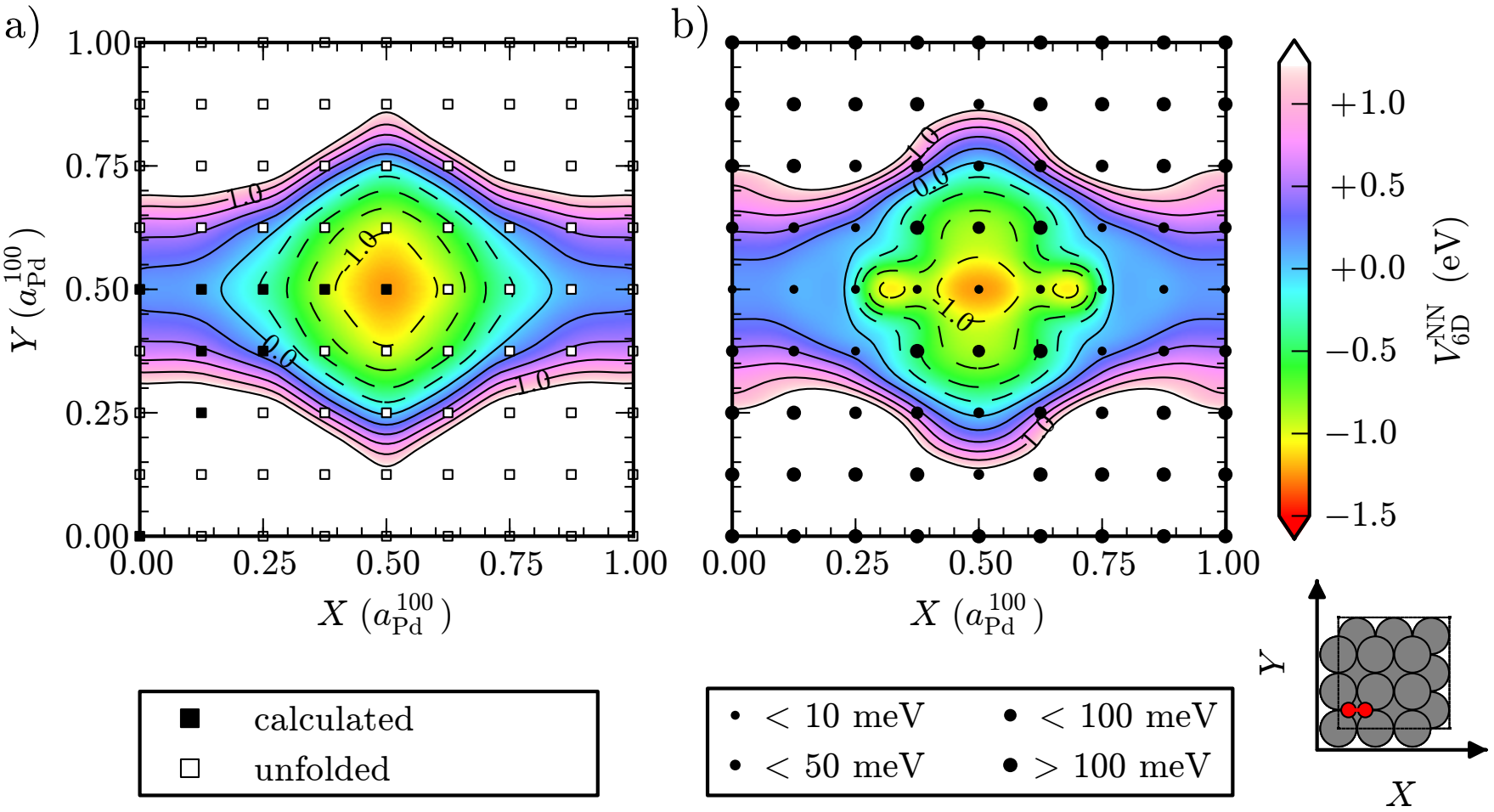


- ~4000 DFT data points in more than 30 ***"elbows"***
- best network topology (out of over 40 tries):

$$V_{\text{fsa}}^{\text{optimal}} = \tanh\left(\sum_{l=1}^{29} w_0^4 + w_l^4 \cdot \tanh\left(\sum_{k=1}^{29} w_0^3 + w_{kl}^3 \cdot \tanh\left(\sum_{j=1}^{29} w_0^2 + w_{jk}^2 \cdot \tanh\left(\sum_{i=1}^{9} w_0^1 + w_{ij}^1 \cdot G_i\right)\right)\right)\right)$$

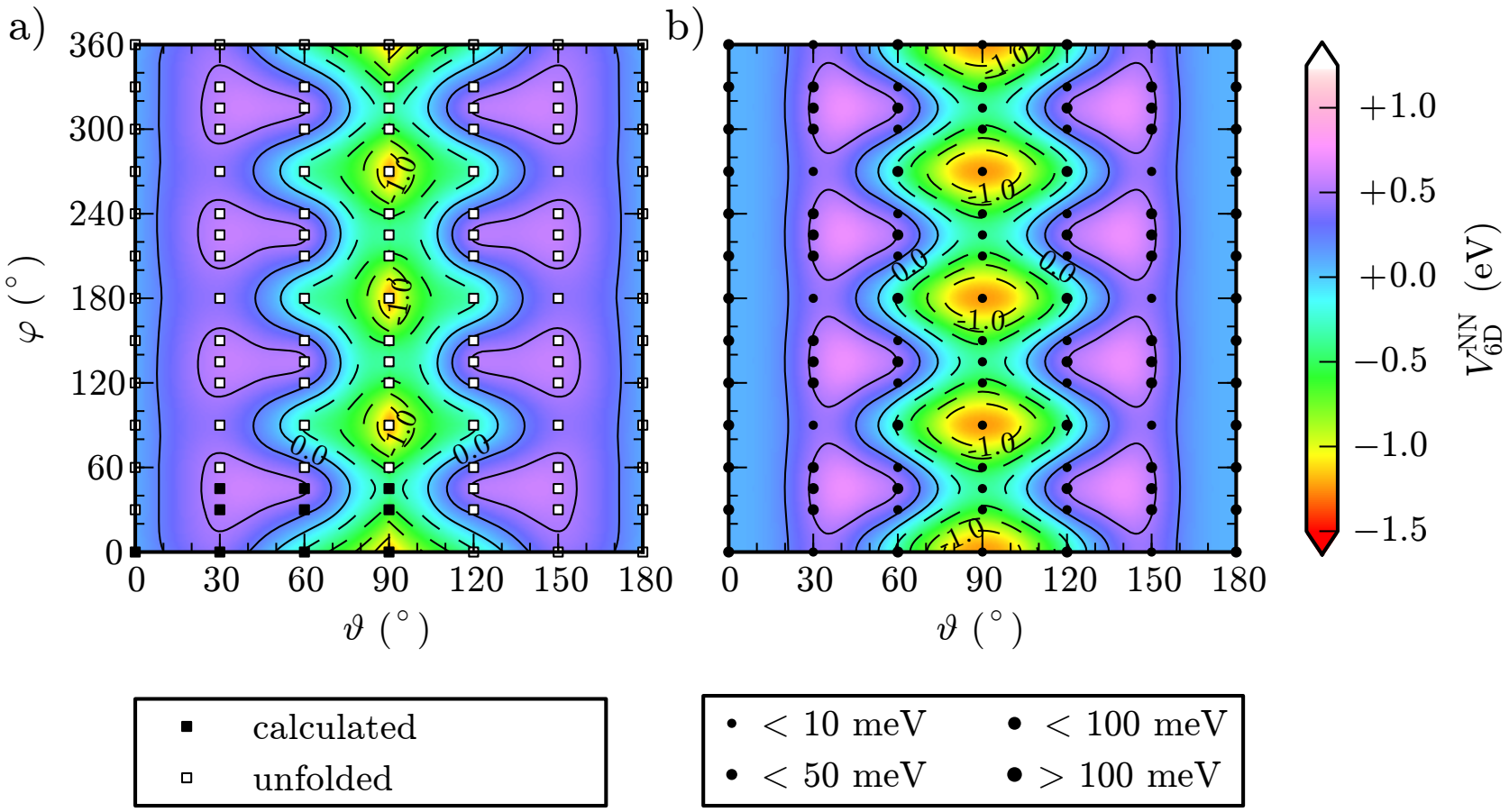
- RMSE for $V_{\text{fsa}}^{\text{optimal}} \leq 1\text{eV} : 19\text{meV}$

Quality of NN-PES (2)



Jörg Meyer, PhD Dissertation, FU Berlin 2012

Quality of NN-PES (3)



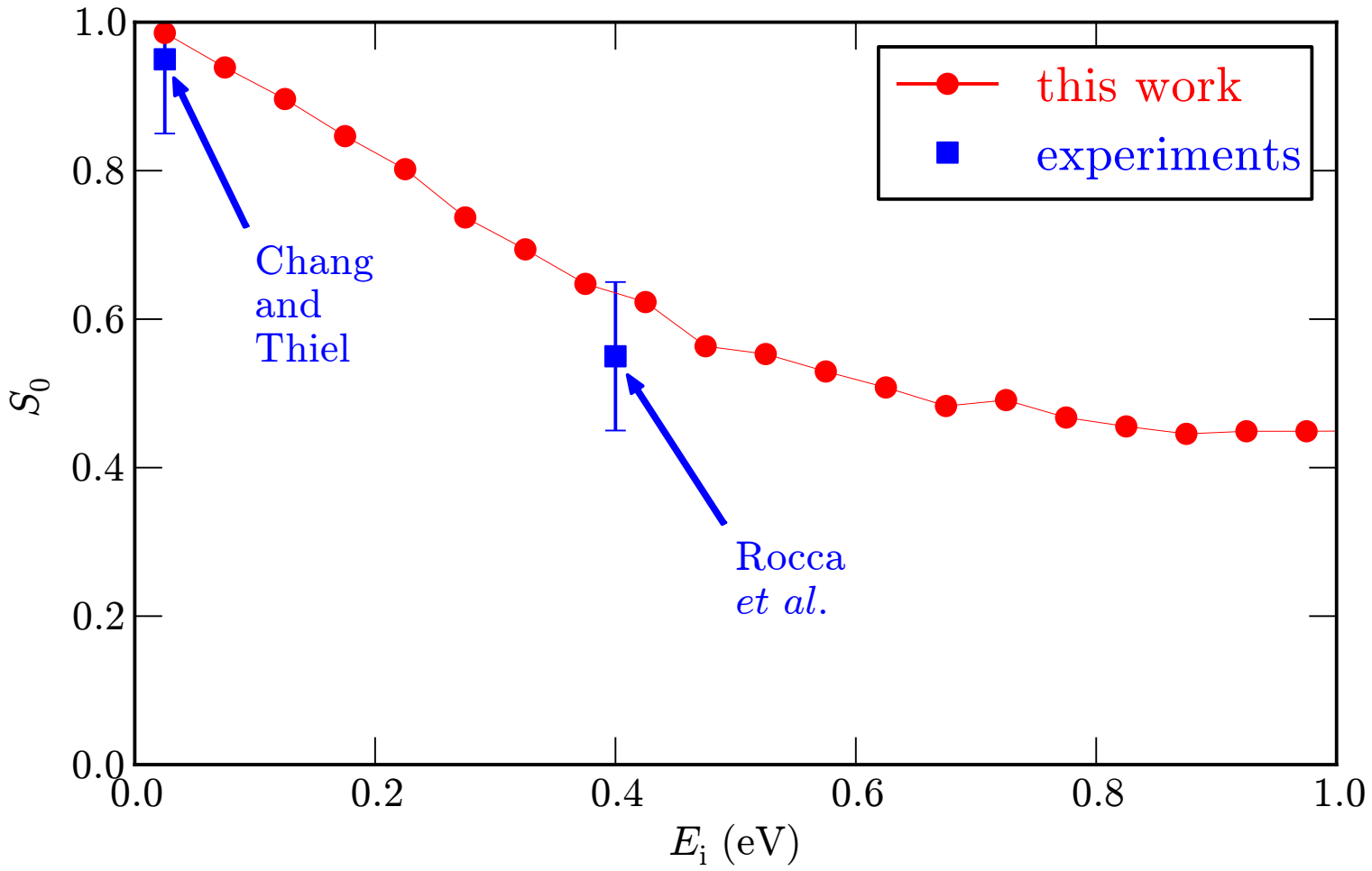
Jörg Meyer, PhD Dissertation, FU Berlin 2012

And what about forces?

$$\nabla_{\boldsymbol{R}^{\mathrm{ph}}} V_{6\mathrm{D}} \approx \nabla_{\boldsymbol{Q}} V_{6\mathrm{D}}^{\mathrm{NN}} \cdot \frac{\partial(Q_1, \ldots, Q_9)}{\partial(R_1^{\mathrm{ph}}, \ldots, R_6^{\mathrm{ph}})}$$

Need to Jacobian of symmetry functions
(with respect to “physical coordinates”)

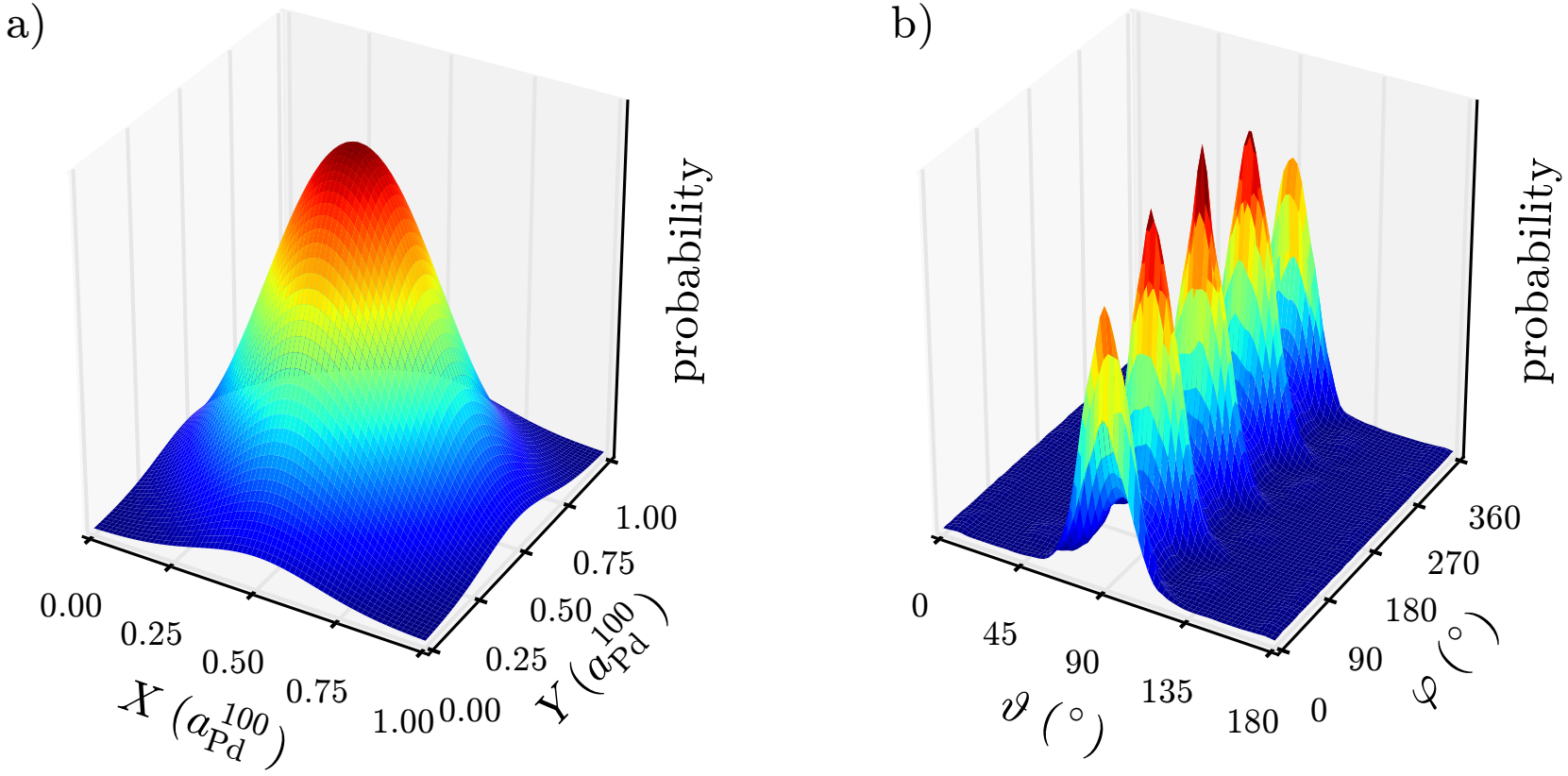
Reaction probability



Jörg Meyer, PhD Dissertation, FU Berlin 2012

V. J. Bukas, S. Mitra, J. Meyer, K. Reuter, *J. Chem. Phys.* **143**, 034705 (2015)

Dissociation channels



Jörg Meyer, PhD Dissertation, FU Berlin 2012

Efficiency

neural network
with symmetry constraints

- ~4000 DFT data points

✔ corresponds to 6D regular grid with only 4 points per direction

✔ symmetry equivalent to >**42000** DFT data points per surface unit cell

➔ **~4000** SCF cycles +

➔ several thousand trajectories per E_{kin}
(↳ CPU time insignificant!)

☺ extensive statistical sampling possible (also for other properties...)

☹ frozen surface approximation

but: $\frac{m(\mathrm{Pd})}{m(\mathrm{O_2})} \approx 3.35$

ab-initio MD
on ground-state Born-Oppenheimer surface

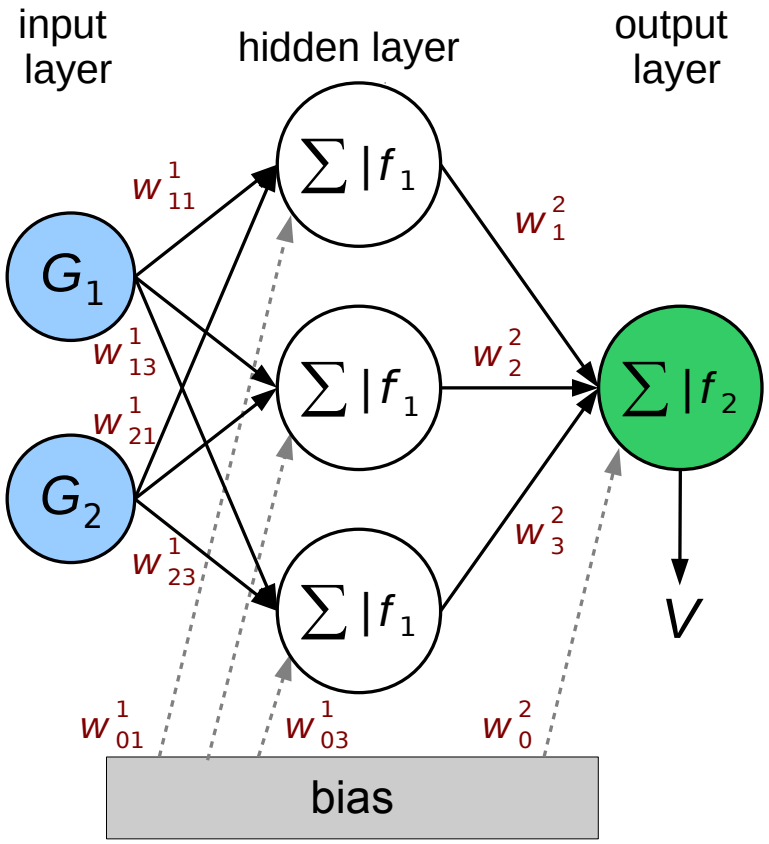
- $\geq$ 100 trajectories per $E_{kin}(t=0)$
- average integration time $\geq$ 1000fs
- time step $\leq$ 2.5fs

➔ $\geq$ **40000** SCF cycles **per** $\boldsymbol{E_{kin}}$
with properly converged forces!

☹ limited statistical sampling

☺ mobile substrate
but only within employed supercell...

Summary



$$V = f_2\left(\sum_{j=1}^{3} w_0^2 + w_j^2 \cdot f_1\left(\sum_{i=1}^{2} w_0^1 + w_{ij}^1 \cdot G_i\right)\right)$$

activation functions $f_1, f_2 \in \{0 \frown 1\}$

- ☺ very flexible: can in principle approximate any potential energy surface to any accuracy

  in practice:
  weight optimization, a.k.a. “learning”, based on training and test sets of input data for different network topologies

- ☺ analytical function with simple derivative (→ exact forces)

- ☹ no explicit knowledge about physical properties of the system

- ↳ perform training with *symmetry constraints*

- ↳ carefully monitor *interpolation quality* (in “areas of interest”)

Back to the roots

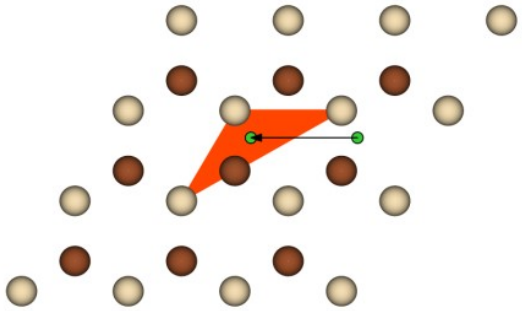
Neural Network PESs for gas-surface dynamics

State of the art

High-Dimensional Neural Network Potentials

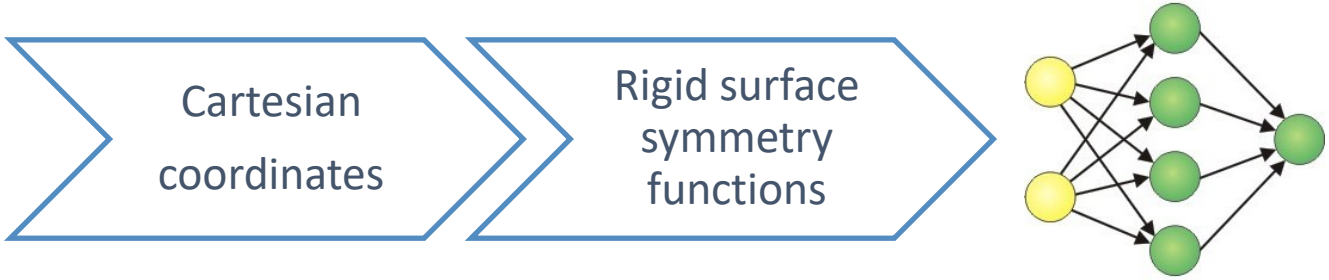
The challenge

- Permutation invariant
- Translational symmetry
- Rotational symmetry



Cartesian coordinates are not suitable!

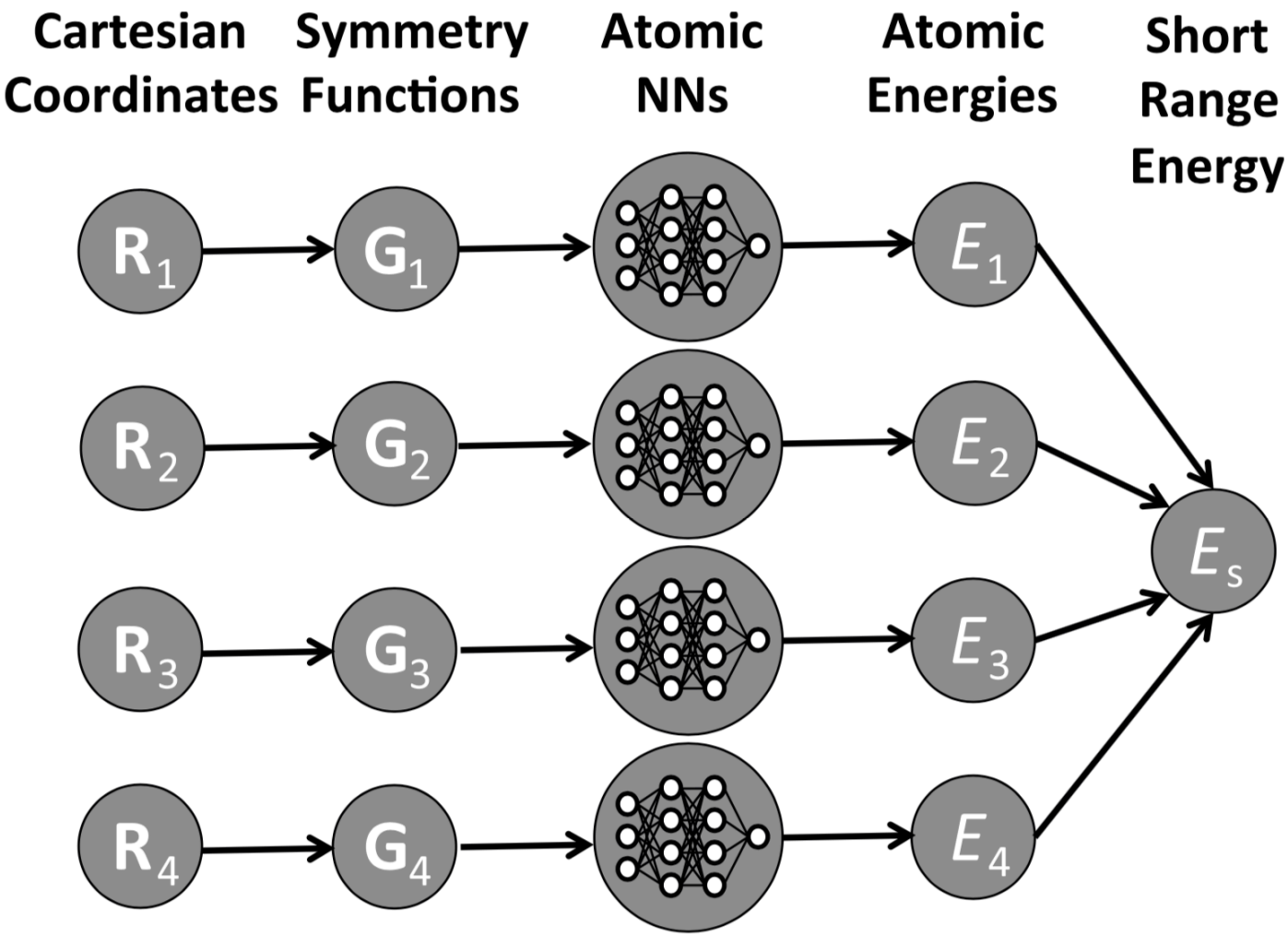
Prior NNPs for molecule-surface reactions



Only applicable to low-dimensional systems! ☹

S. Lorenz *et al.*, CPL **395**, 210 (2004), T. B. Blank *et al.*, JCP **103**, 4129 (1995), B. Jiang *et al.*, JCP **141**, 034109 (2014).

Introducing atomic NNs

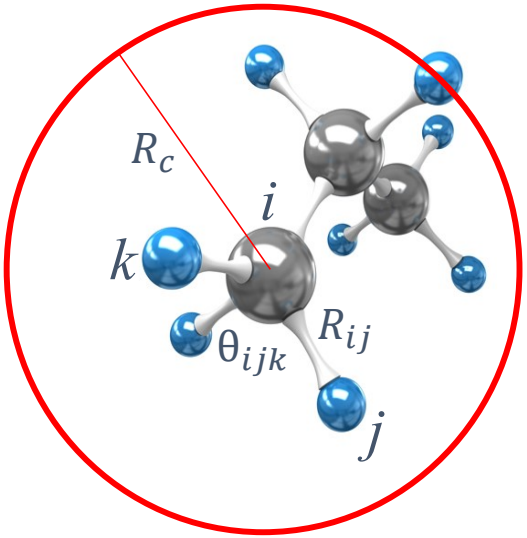


J. Behler, *J. Phys.: Condens. Matter*. **26**, 183001 (2014).

Fingerpint of local atomic environment

- Constructed based on chemical environment of the atoms

$$f_c(R) = \begin{cases} 0.5\left[\cos\left(\pi\frac{R}{R_c}\right)+1\right] & R < R_c \\ 0 & R \geq R_c \end{cases}$$



R_c= 6 Å

- Permutation invariance ✓
- Translational symmetry ✓
- Rotational symmetry ✓

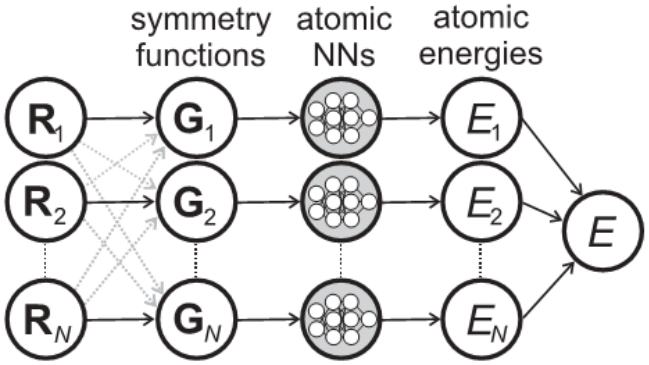
Behler and Parrinello, PRL **98**, 146401 (2007).

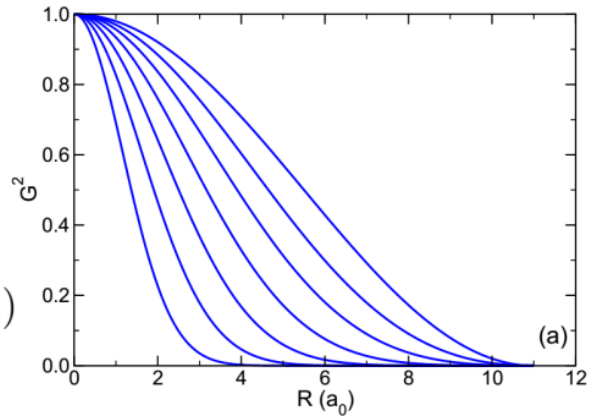
Many-body symmetry functions

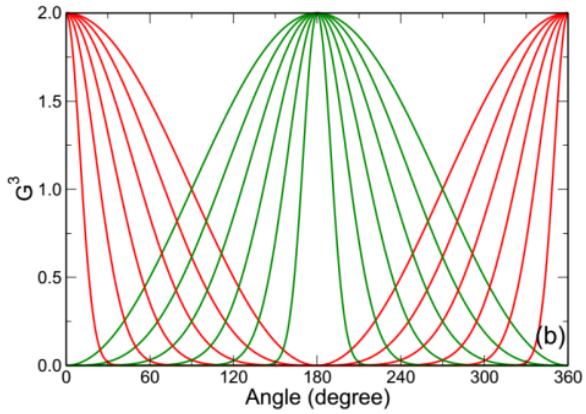
Description of interatomic interactions:

$$G_i^2 = \sum_{j \neq i} e^{-\eta(R_{ij}-R_s)^2} f_c(R_{ij})$$

$$G_i^3 = 2^{1-\zeta} \sum_{j,k \neq i} (1 + \lambda \cos\theta_{ijk})^{\zeta} e^{-\eta(R_{ij}^2+R_{ik}^2+R_{jk}^2)} f_c(R_{ij}) f_c(R_{ik}) f_c(R_{jk})$$







Applicable to mobile surfaces!

J. Behler, M. Parrinello, Phys. Rev. Lett. **98**, 146401 (2007).