PES from ML or Physics

János Daru

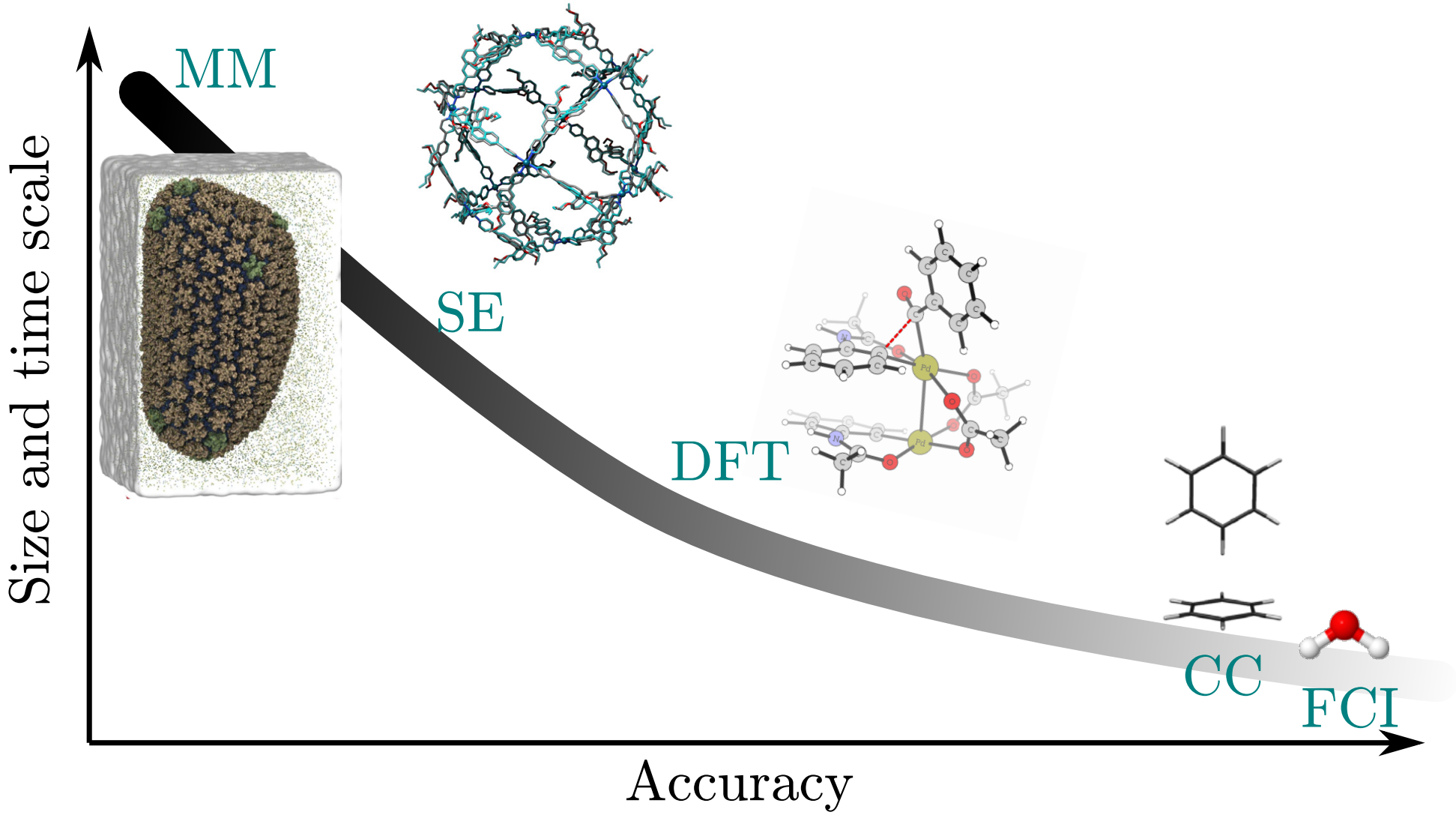
Eötvös Loránd University, Budapest
Department of Organic Chemistry

Occam vs. ML 2024, Szeged

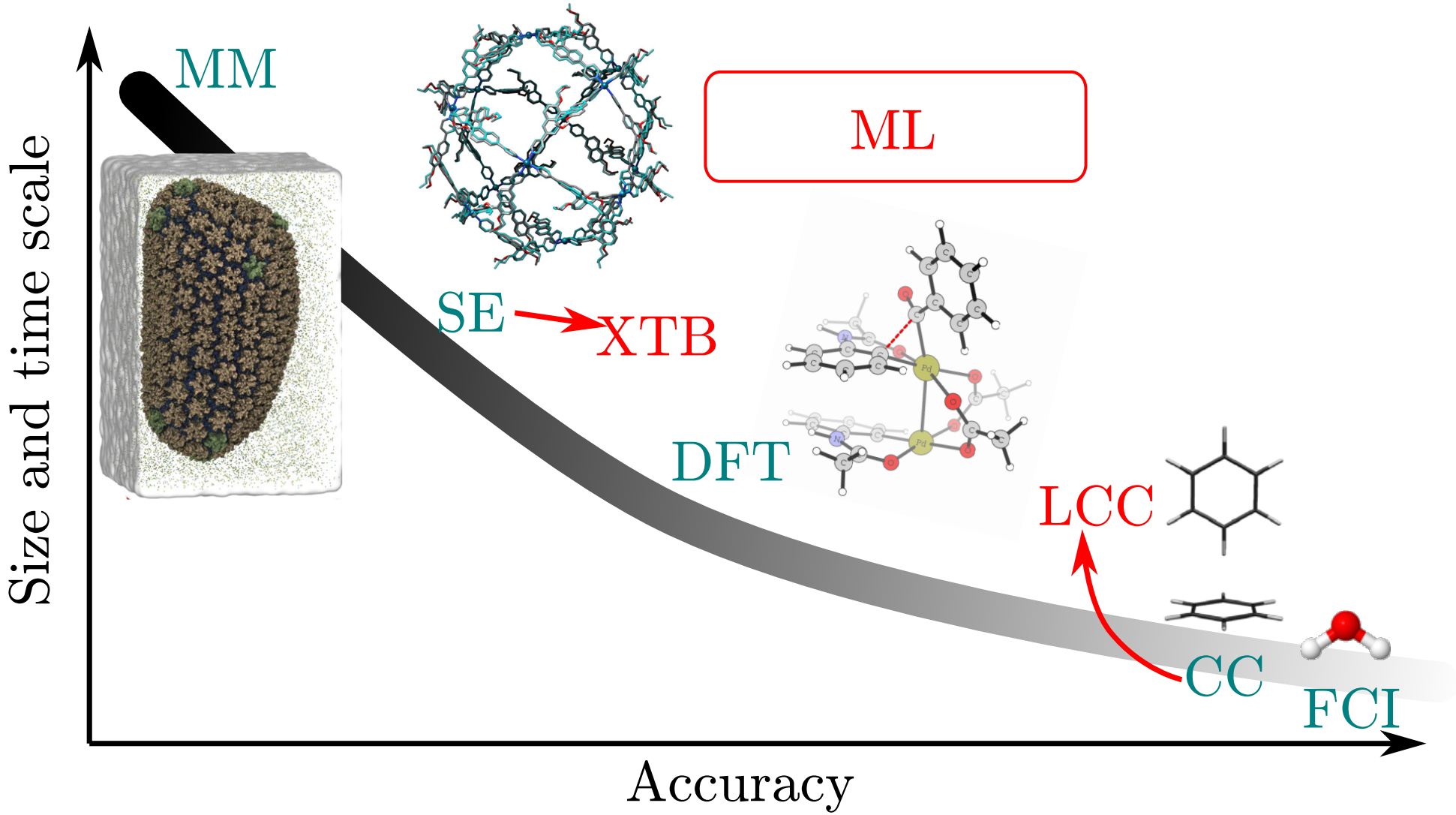
Content

Content

Electron structure → forcefields

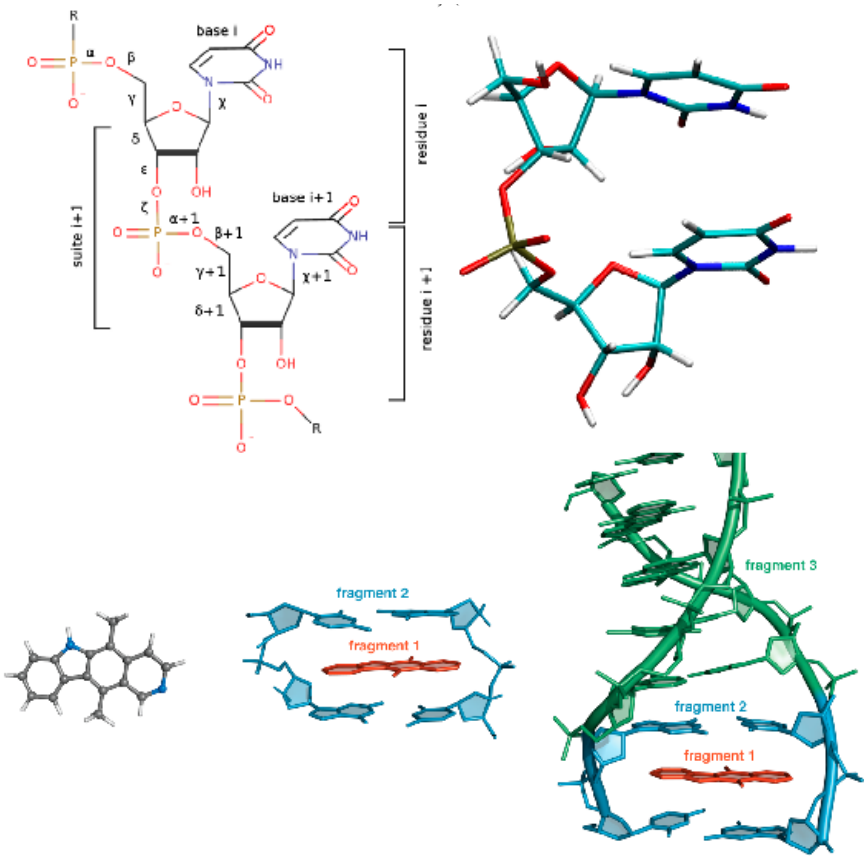


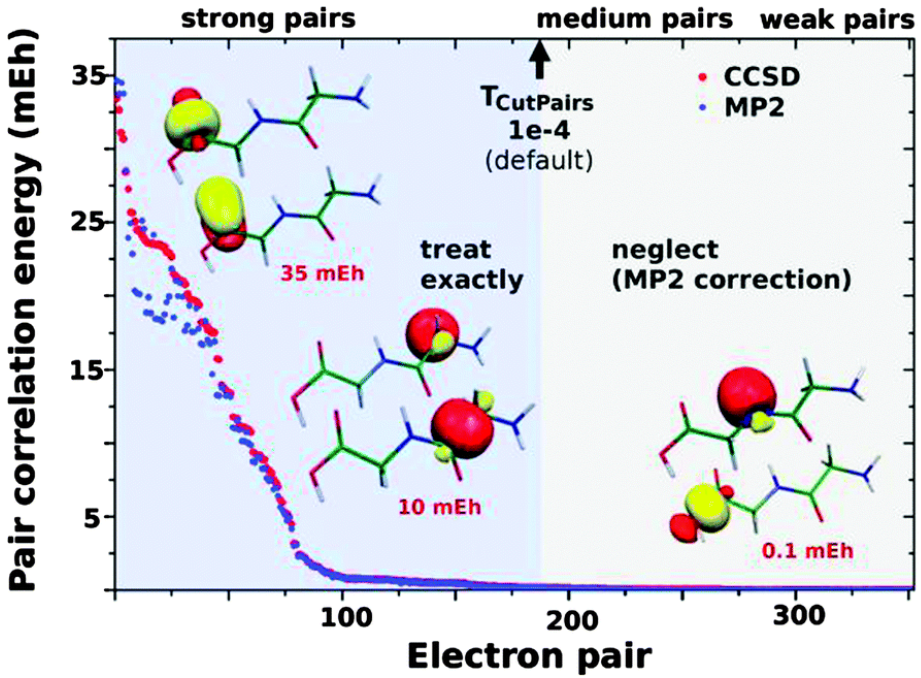
Electron structure → forcefields



Local correlation calculations

Not linear-scaling, not yet Full CI, but a huge step forward!

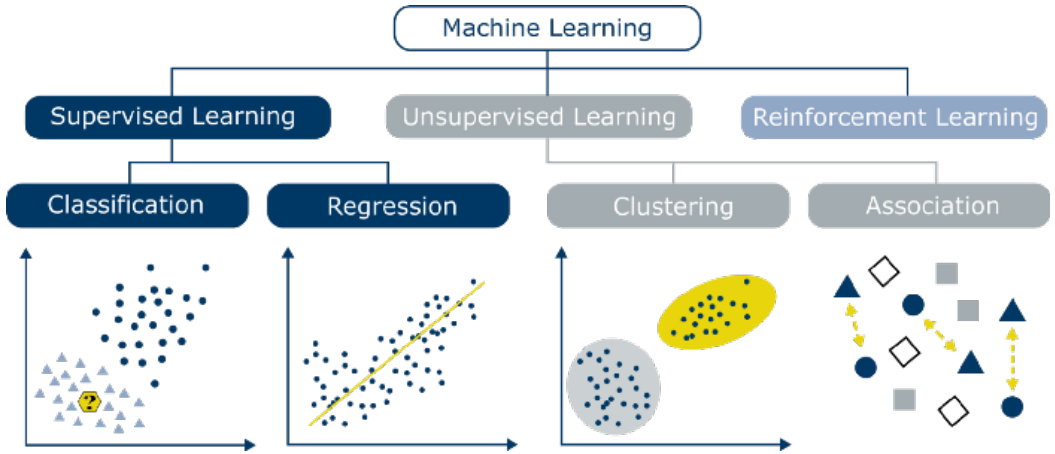


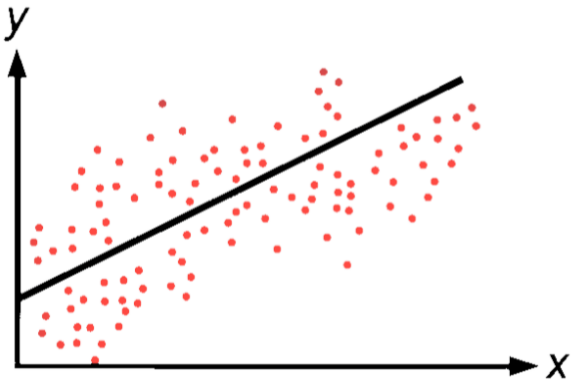


JCTC (2015), **11**: 4972 | *Chem. Soc. Rev.*, (2014), **43**: 5032
JCTC (2017), **13**: 3198

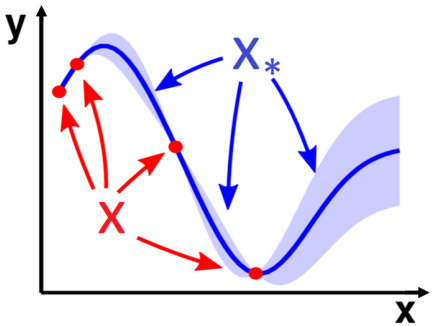
Content

Basics

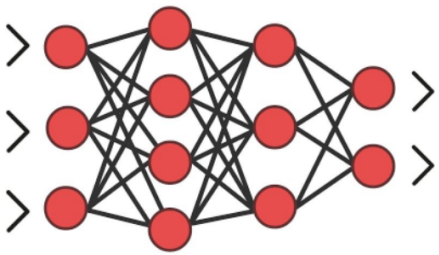




Linear



Kernel



NN

Linear Regression

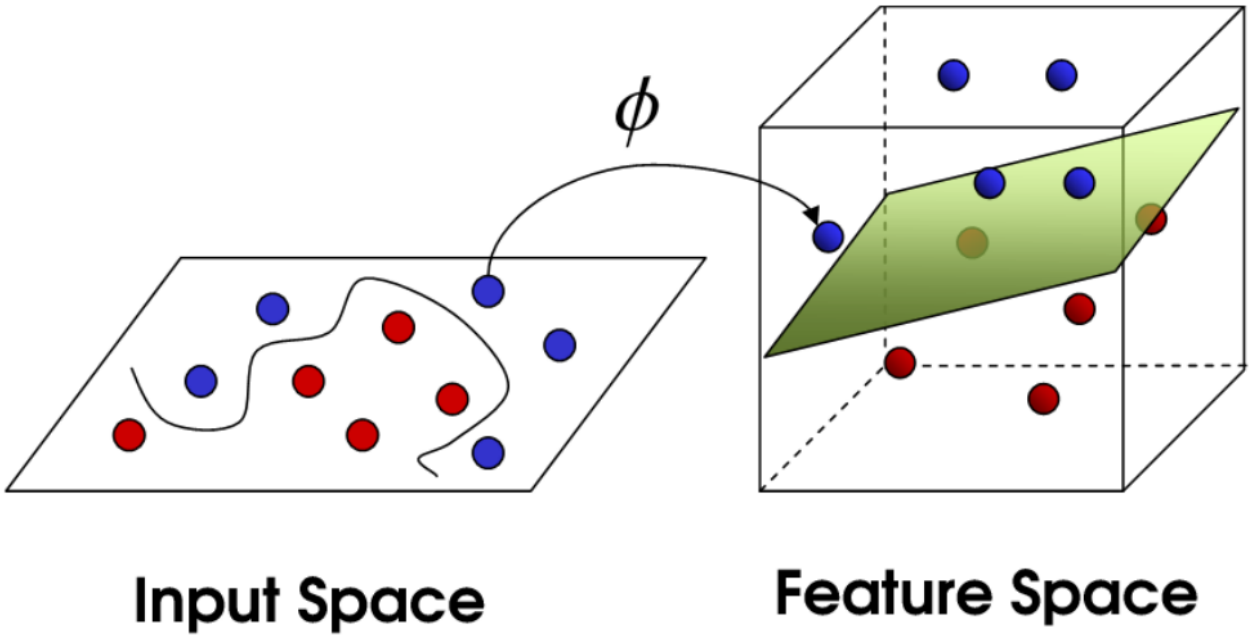
$$y_n = \mathbf{x}_n^\top \boldsymbol{\beta} + \varepsilon_n. \quad (1)$$

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y}. \quad (2)$$

Regularized case (Ridge Regression):

$$\hat{\boldsymbol{\beta}} = (\lambda \mathbf{I} + \mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y}. \quad (3)$$

Kernel Ridge Regression



Kernel Ridge Regression

$$
\begin{aligned}
y_n &= [\phi(\mathbf{x}_n)]^\top \boldsymbol{\beta} + \boldsymbol{\varepsilon_n} && (4)\\
\hat{\boldsymbol{\beta}} &= (\lambda \mathbf{I} + \boldsymbol{\Phi}^\top \boldsymbol{\Phi})^{-1} \boldsymbol{\Phi}^\top \mathbf{y}. && (5)\\
\hat{\boldsymbol{\beta}} &= \boldsymbol{\Phi}(\lambda \mathbf{I} + \boldsymbol{\Phi}\boldsymbol{\Phi}^\top)^{-1} \mathbf{y}. && (6)\\
y &= [\phi(\mathbf{x})]^\top (\lambda \mathbf{I} + \boldsymbol{\Phi}\boldsymbol{\Phi}^\top)^{-1} \mathbf{y}. && (7)
\end{aligned}
$$

Kernel Trick (kernels instead of monomials):
$k(\mathbf{x}, \mathbf{y}) = \langle \phi(\mathbf{x}), \phi(\mathbf{y}) \rangle$

$$
\begin{aligned}
\mathbf{K}_{ij} &= k(\mathbf{x}_i, \mathbf{x}_j), && (8)\\
\mathbf{k}(\mathbf{x}) &= [k(\mathbf{x}, \mathbf{x}_1), \ldots, k(\mathbf{x}, \mathbf{x}_N)]^\top. && (9)\\
y &= [\mathbf{k}(\mathbf{x})]^\top (\lambda \mathbf{I} + \mathbf{K})^{-1} \mathbf{y} && (10)
\end{aligned}
$$

GPR

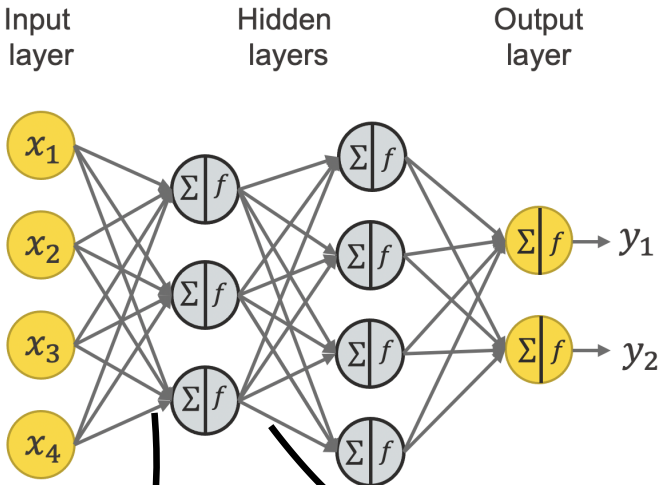
We assume a Gaussian distribution both for the measured quantity and its error:

$$
\begin{aligned}
y_n &= f(\mathbf{x}_n) + \varepsilon_n, && (11)\\
f &\sim \mathcal{N}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x})). && (12)\\
\mathbb{E}[\mathbf{f}_*] &= K(\mathbf{X}_*, \mathbf{X})[\sigma^2\mathbf{I} + K(\mathbf{X}, \mathbf{X})]^{-1}\mathbf{y} && (13)
\end{aligned}
$$

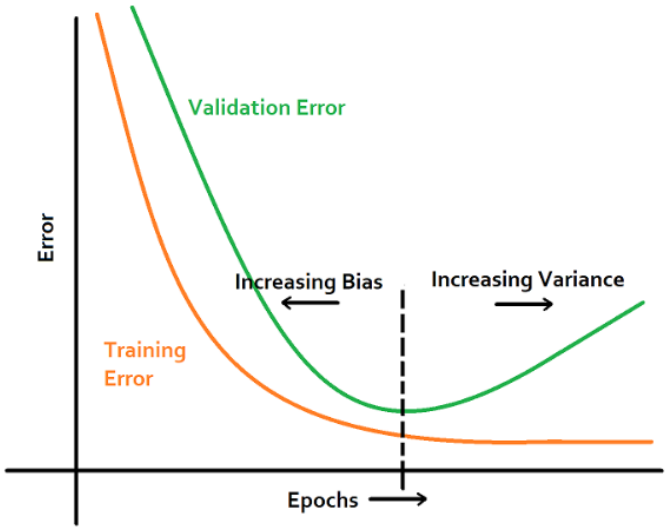
For comparison a KRR expression:

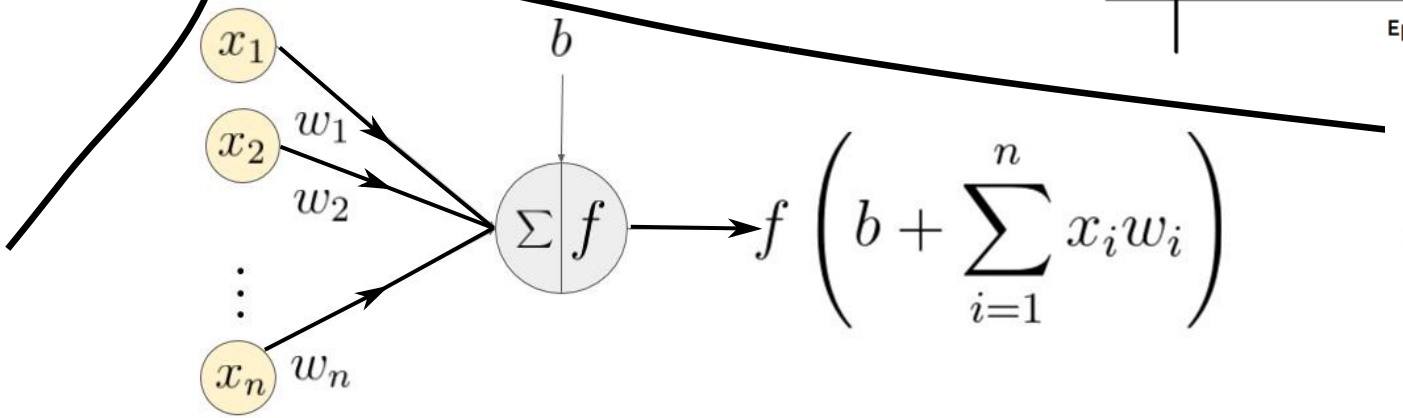
$$
y = [\mathbf{k}(\mathbf{x})]^\top(\lambda\mathbf{I} + \mathbf{K})^{-1}\mathbf{y} \qquad (14)
$$

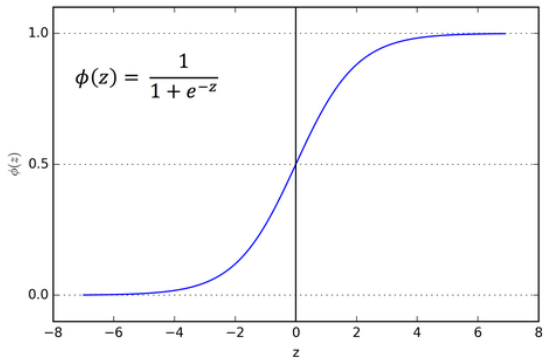
Neural Network



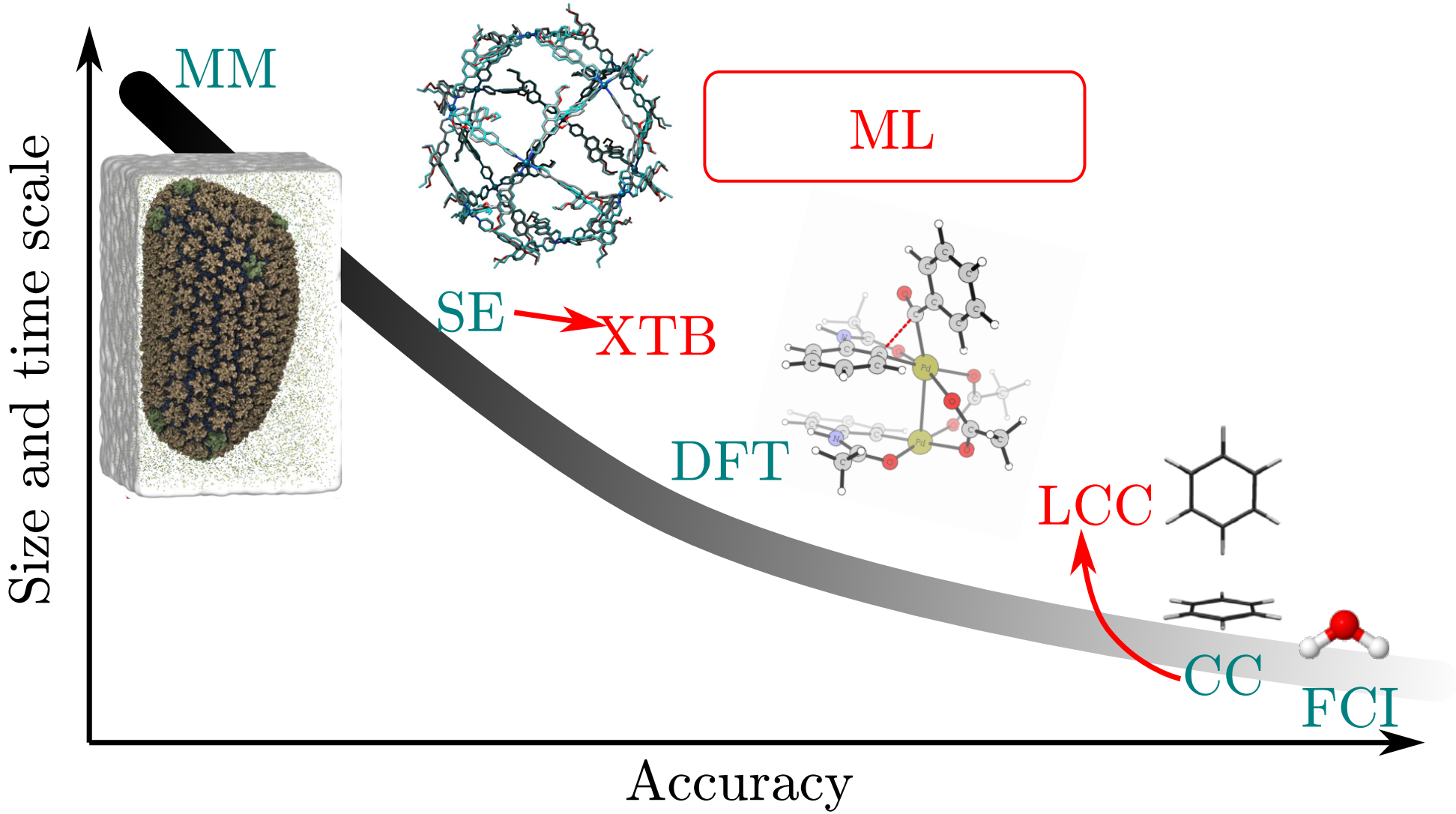
Instead of Regularization: Early Stopping







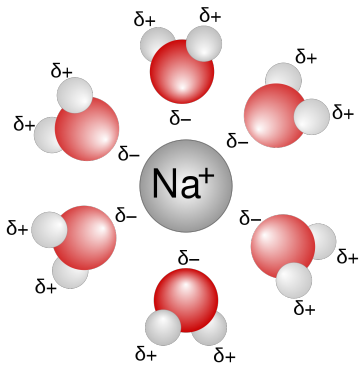
Electron structure → forcefields



Classical Force Fields

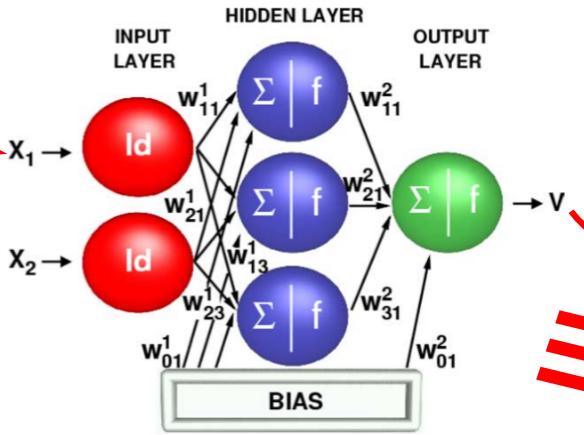
$$V(r) = \sum_{\text{bonds}} k_b(b-b_0)^2 + \sum_{\text{angles}} k_\theta(\theta-\theta_0)^2 + \sum_{\text{improp.}} k_\omega(\omega-\omega_0)^2 + \sum_{\text{Urey–Bradley}} k_u(u-u_0)^2$$

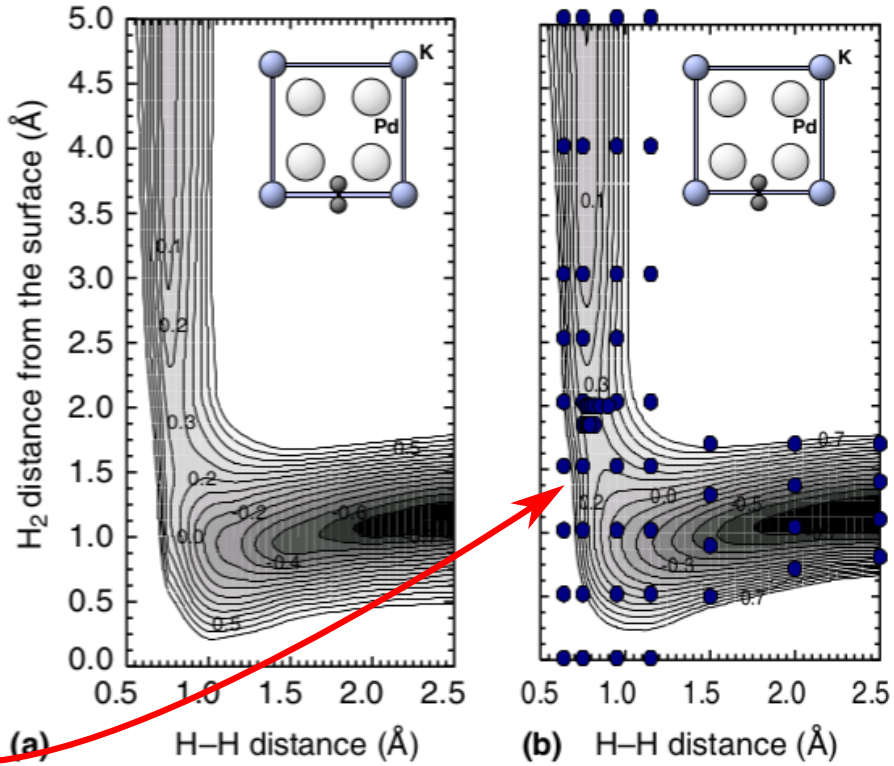
$$+ \sum_{\text{dihedrals}} k_\phi\left[1+\cos(n\phi-\delta)\right] + \sum_{\text{nonbond.}} 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \sum_{\text{nonbond.}} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}$$



The first ML forcefield

$$
\begin{aligned}
x_1 &= d, \\
x_2 &= e^{Z_c/2}, \\
x_3 &= \sin^2\theta \cos 2\phi [\cos GX_c - \cos GY_c] e^{-Z_c/2}, \\
x_4 &= \sin^2\theta \cos 2\phi [\cos 2GX_c - \cos 2GY_c] e^{-Z_c/2}, \\
x_5 &= \cos^2\theta e^{-Z_c/2}, \\
x_6 &= [\cos GX_c + \cos GY_c] e^{-Z_c/2}, \\
x_7 &= [\cos 2GX_c + \cos 2GY_c] e^{-Z_c/2}, \\
x_8 &= \sin^4\theta \cos 4\phi [\cos 2GX_c + \cos 2GY_c] e^{-Z_c/2}.
\end{aligned}
$$

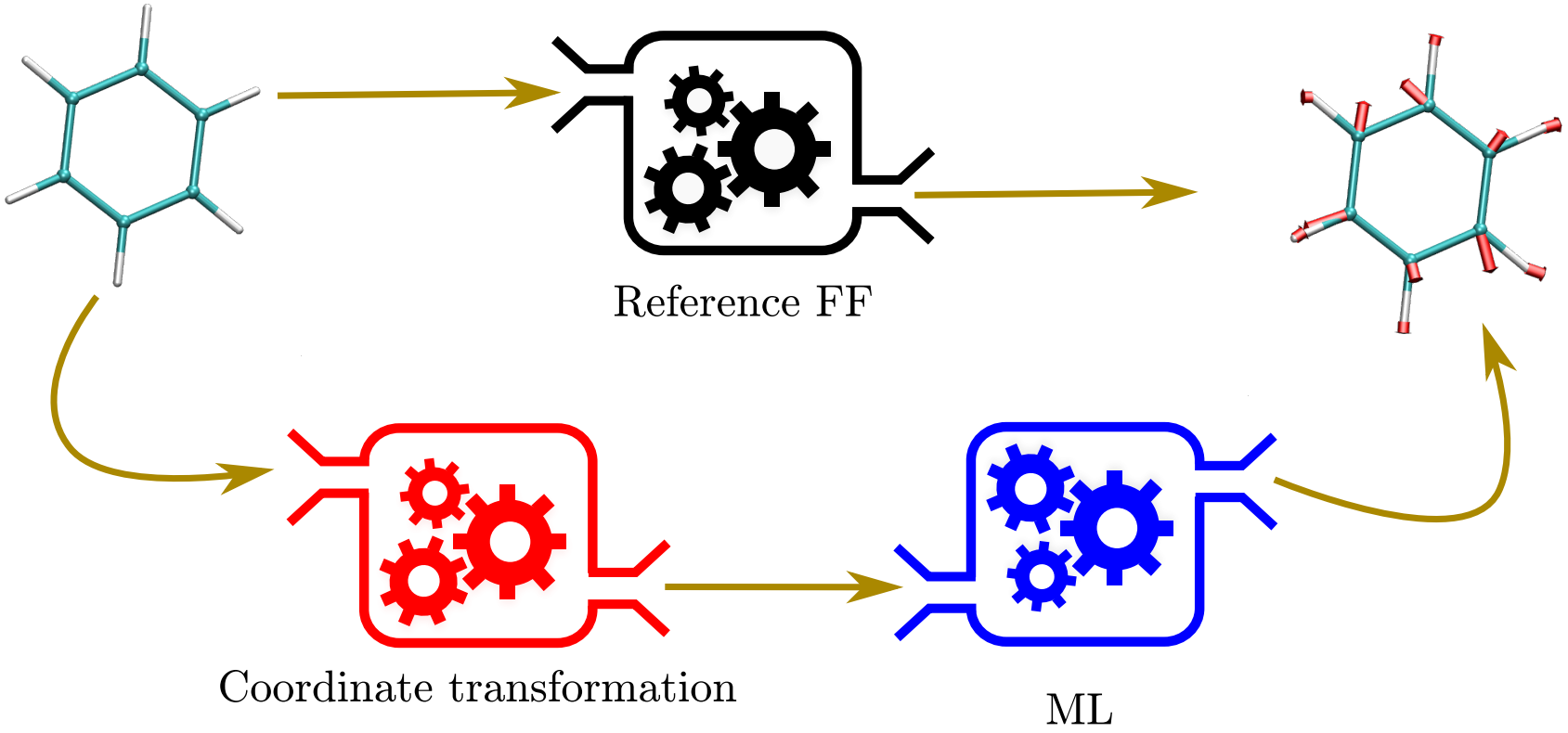




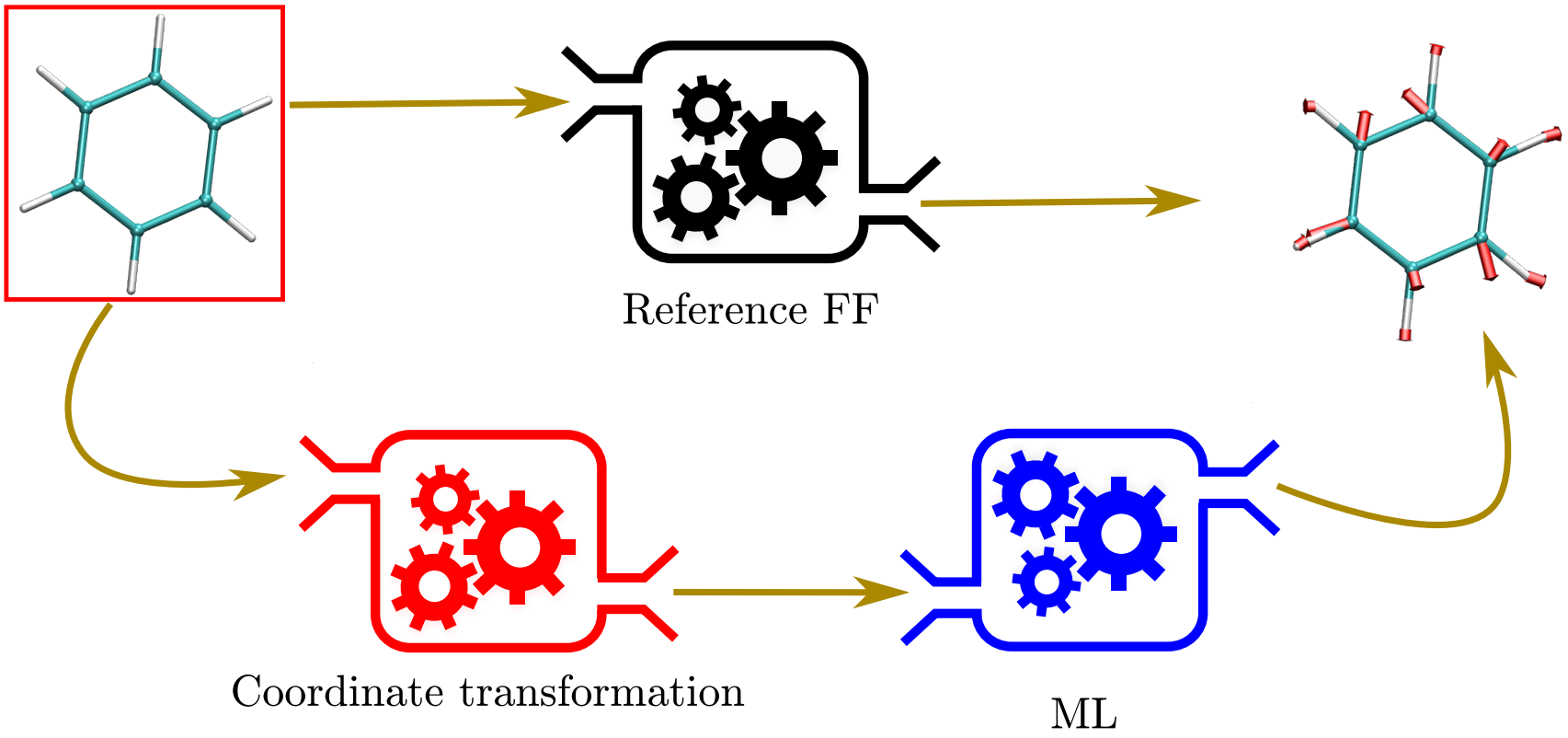
$$
V_{nn}(\mathbf{x}) = f_2\left(w_{01}^2 + \sum_j w_{j1}^2 \, f_1\left(w_{0j}^1 + \sum_{i=1}^{6} w_{ij}^1 x_i\right)\right),
$$

M. Scheffler: *Chem.Phys.Lett.* (2004), **395**: 210

General ML-FF scheme



Genearal Global ML-FF scheme

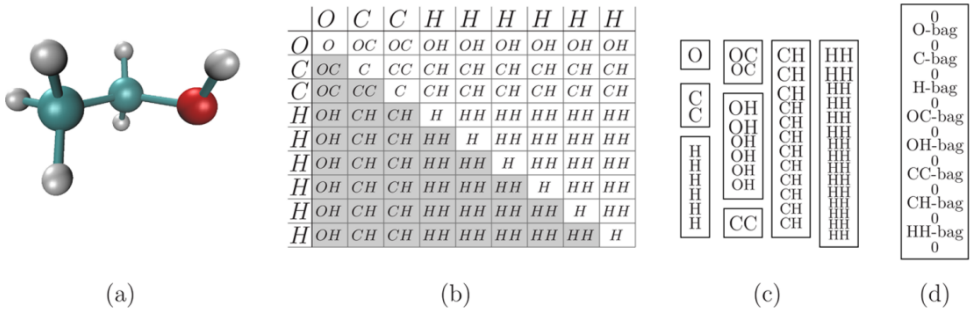


Eigenvales of the Overlap matrix

Coulomb Mátrix:

$$M_{ij} = \begin{cases} 0.5Z_i^{2.4}, & \text{if } i = j, \\ \frac{Z_i Z_j}{|r_i - r_j|}, & \text{if } i \neq j, \end{cases}$$

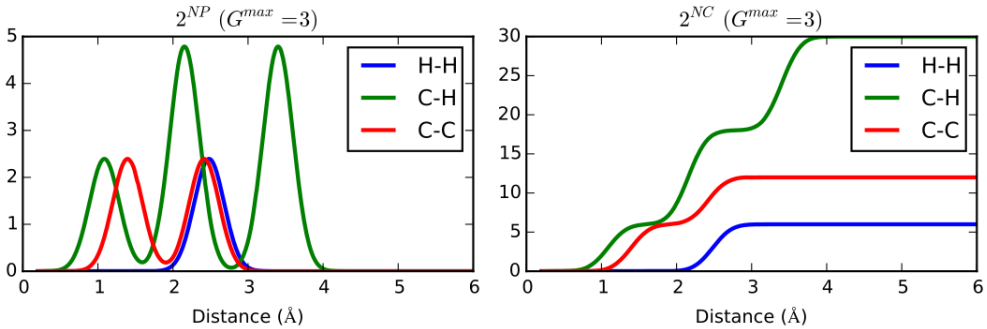
Bag of Bonds:



Encoded Bonds:

$$v_i = \sum_{j<k} f(r_{jk}, d_i, \beta).$$

$$v_i = \sum_{j<k} f(r_{jk}, d_i, \beta) \mathbb{1}(G_{jk} \leq G^{max}),$$



A. v. Lilienfeld: *JCP* (2018), **148**: 241718

Eigenvales of the Overlap matrix

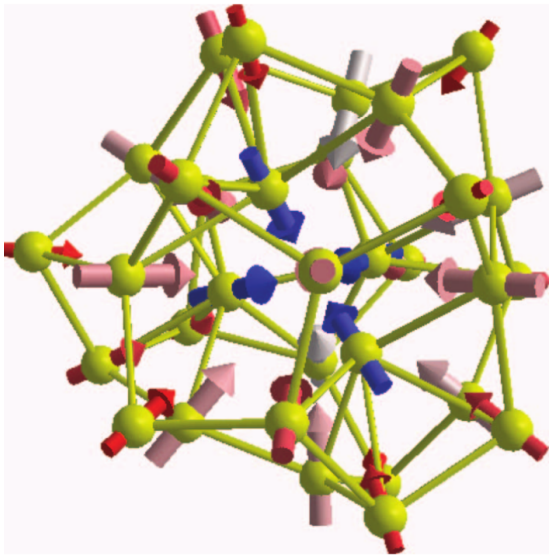
$$\left\langle \phi_i^s \middle| \phi_j^s \right\rangle = S_{ij} \qquad S_{ij} = S_{ji} = \left(\frac{2\sqrt{\alpha_i \alpha_j}}{\alpha_i + \alpha_j} \right)^{3/2} \exp\left[\frac{-\alpha_i \alpha_j}{\alpha_i + \alpha_j} r_{ij}^2 \right]$$

$$\left\langle \phi_i^{p_x} \middle| \phi_j^s \right\rangle = \frac{1}{\sqrt{\alpha_i}} \frac{\partial S_{ij}}{\partial x_i}$$

$$= -\left(\frac{2\sqrt{\alpha_i}\alpha_j}{\alpha_i + \alpha_j} \right)(x_i - x_j) S_{ij}$$

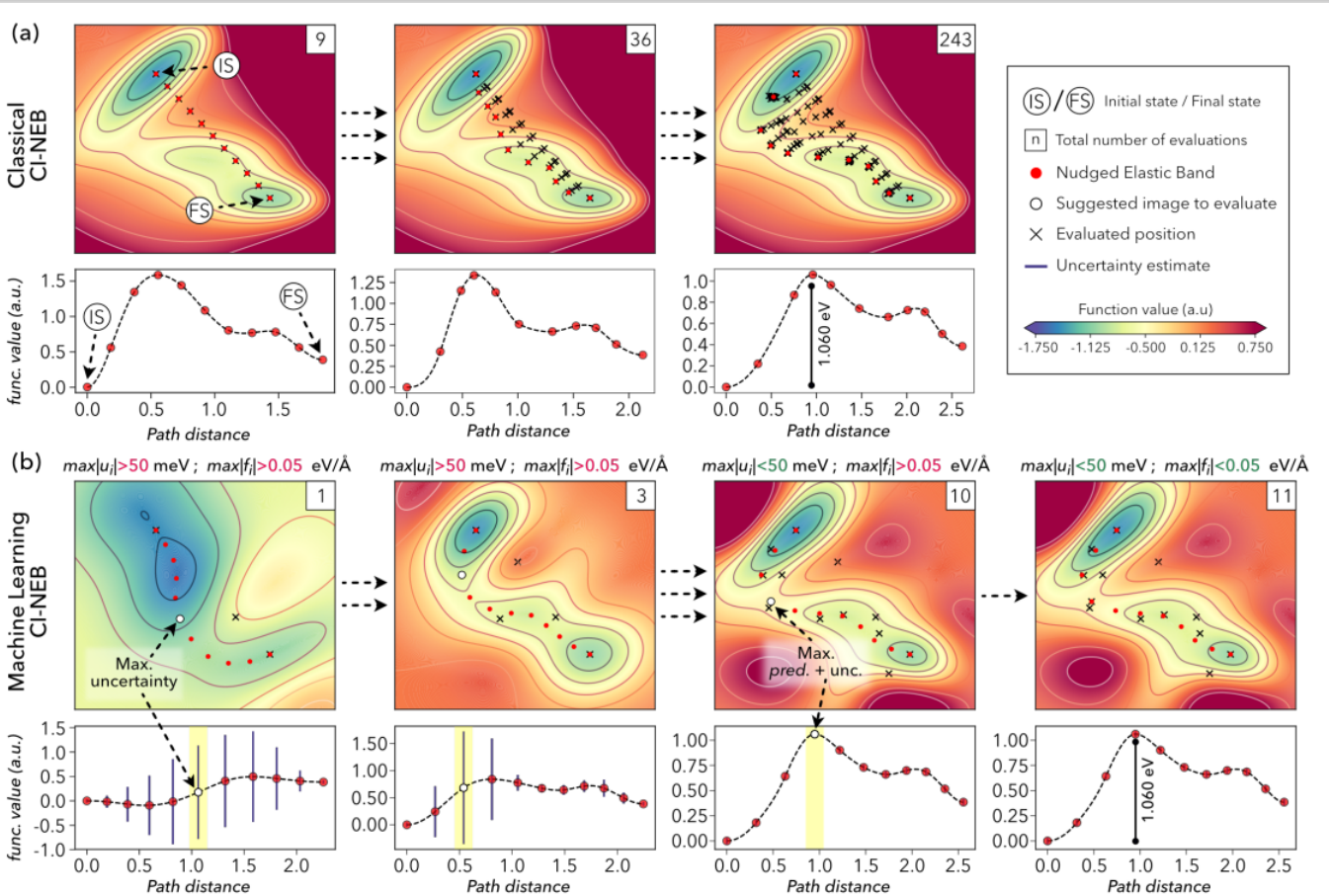
$$\left\langle \phi_i^{p_x} \middle| \phi_j^{p_{x'}} \right\rangle = \left(\frac{2\sqrt{\alpha_i \alpha_j}}{\alpha_i + \alpha_j} \right) S_{ij}$$

$$\times \left[\delta_{x,x'} - \frac{2\alpha_i \alpha_j}{\alpha_i + \alpha_j} (x_i - x_j)(x_i' - x_j') \right],$$



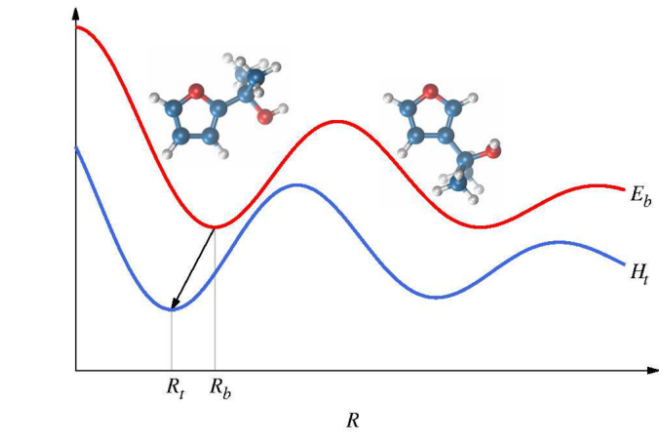
S. Goedecker: *JCP* (2013), **139**: 184118

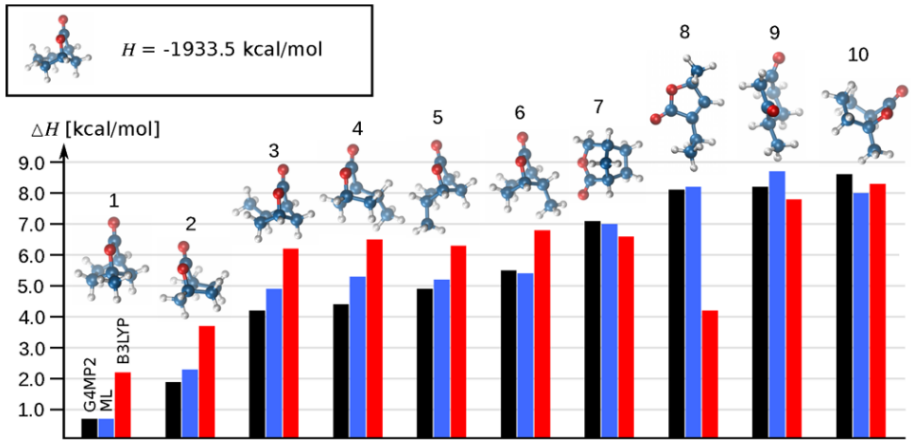
Applications

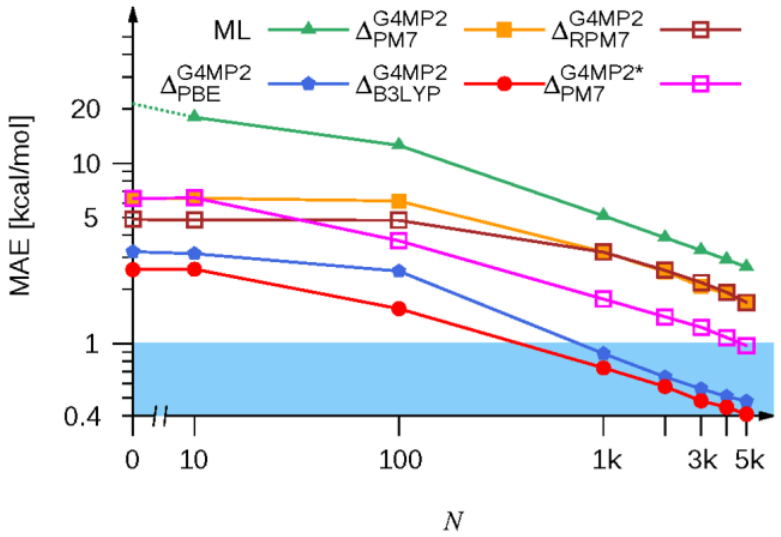


T. Bligaard: *PRL* (2019), **122**: 156001

Δ-Learning

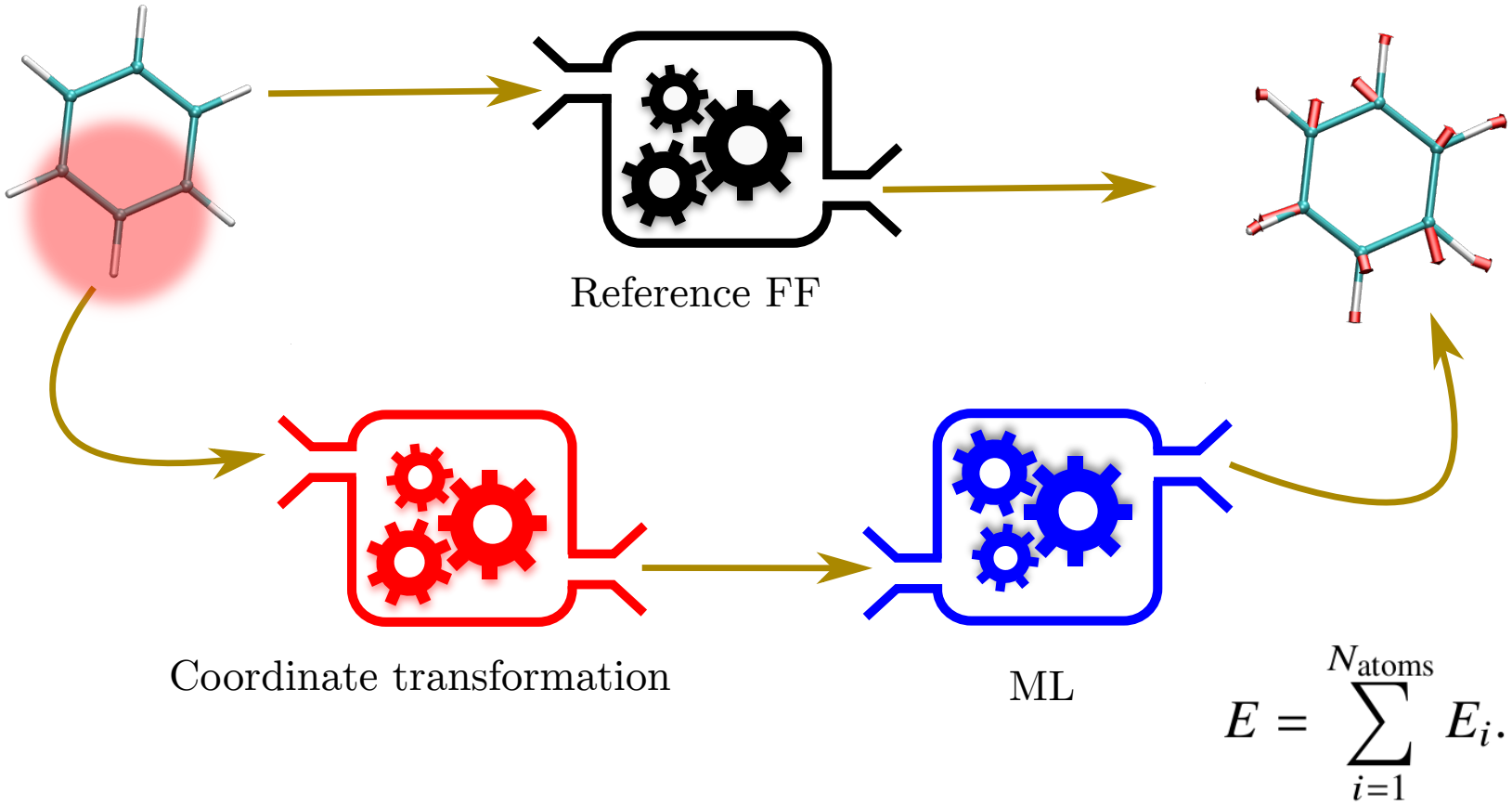






A. v. Lilienfeld: *JCTC* (2015), **11**: 2087

General Local ML-FF scheme

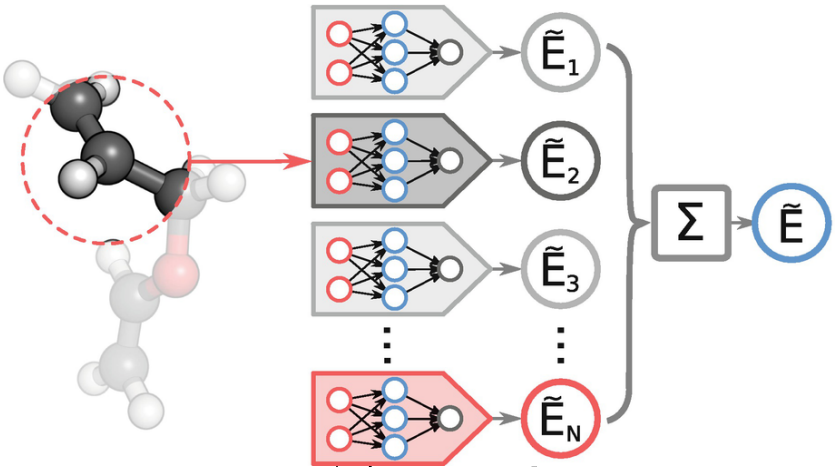


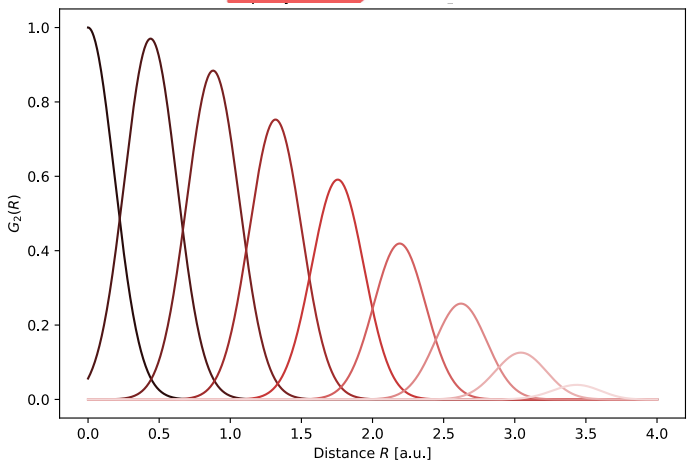
$$E = \sum_{i=1}^{N_{\text{atoms}}} E_i.$$

Pros: PBC, scaling

Cons: missing long-range interactions

Parrinello & Behler: HDNNP





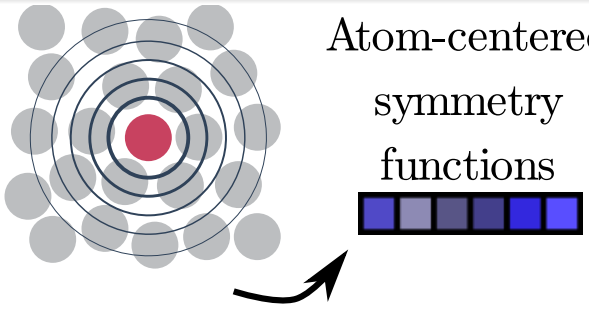
$$G^R_{a,s} = \sum_{j \neq i}^{\text{all atoms}} e^{\eta(R_{ij}-R_s)^2} f_C(R_{ij})$$

$$G^{A_{mod}}_{a,b,m,n} = 2^{1-\zeta} \sum_{j,k \neq i}^{\text{all atoms}} (1 + \cos(\theta_{ijk} - \theta_m))^\zeta \exp\left[-\eta\left(\frac{R_{ij} + R_{ik}}{2} - R_n\right)^2\right] f_C(R_{ij}) f_C(R_{ik})$$

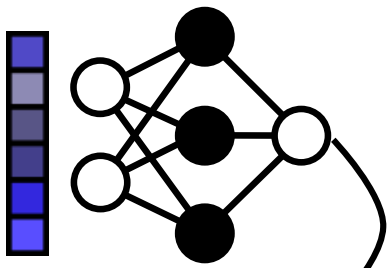
$$f_C(R_{ij}) = \begin{cases} 0.5 \times \left[1 + \cos\left(\frac{\pi R_{ij}}{R_C}\right)\right] & \text{for } R_{ij} \leq R_C \\ 0.0 & \text{for } R_{ij} > R_C \end{cases}$$

Chem. Sci., (2017), **8**: 6924 | *PRL* (2007), **98**: 146401

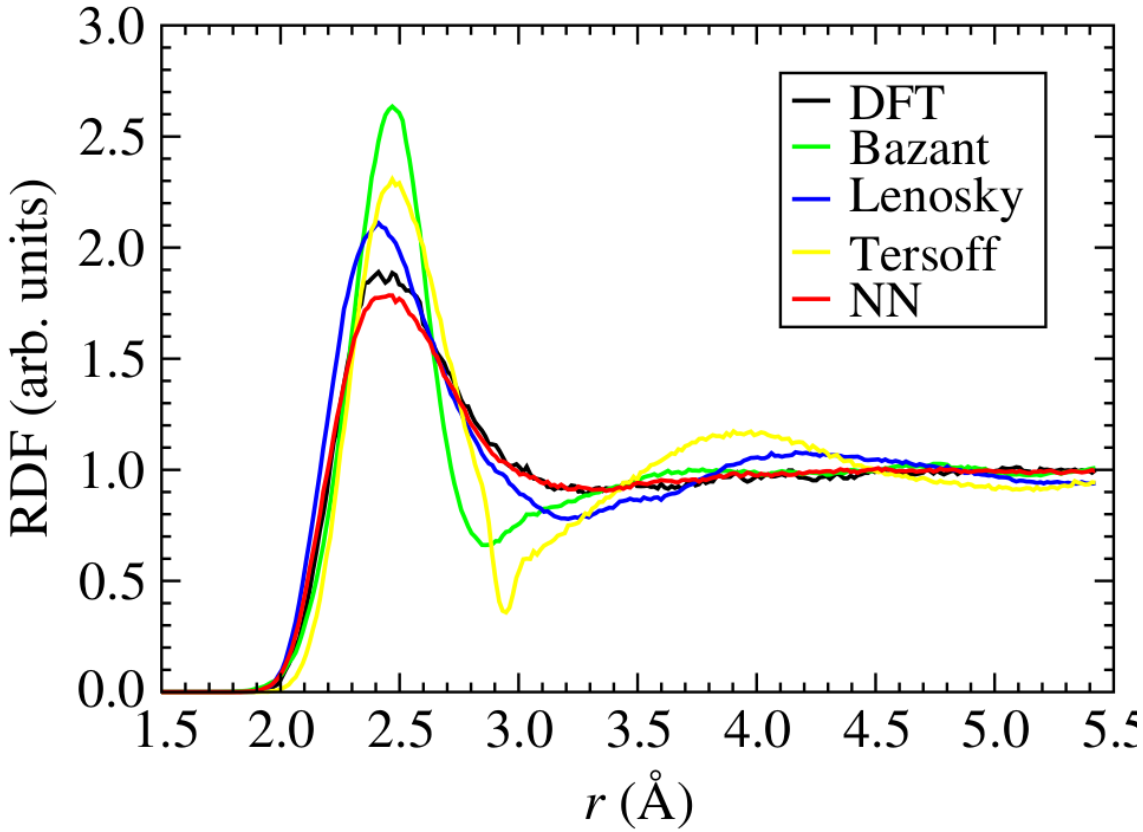
Parrinello & Behler: HDNNP



Atom-centered symmetry functions



$$E = \sum_i E_i$$

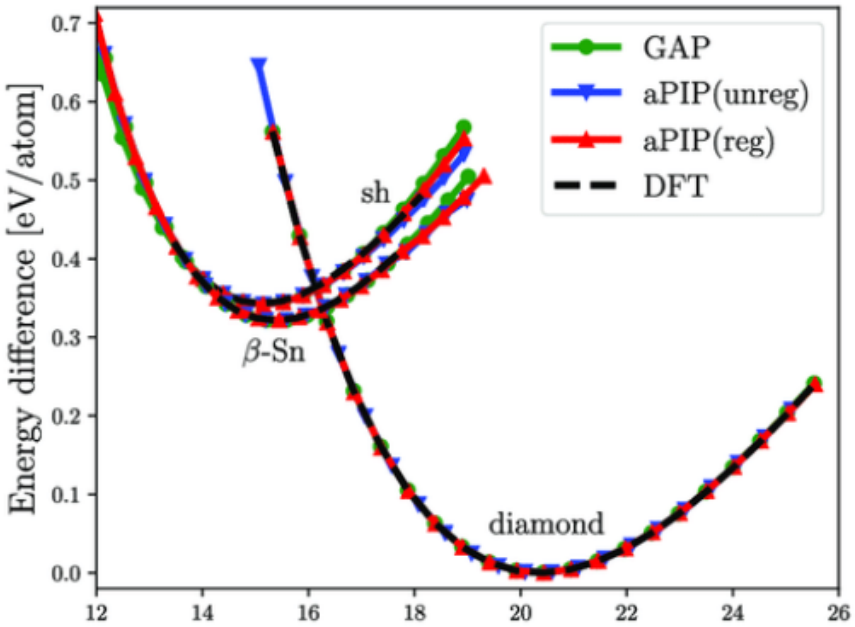


PRL (2007), **98**: 146401

Csányi & Bartók: SOAP+GAP

$$\rho_{\mathrm{SOAP}}(\mathbf{R}) = \sum_{i=1}^{N_{\mathrm{env}}} \exp(-\alpha|\mathbf{R} - \mathbf{R_i}|^2).$$

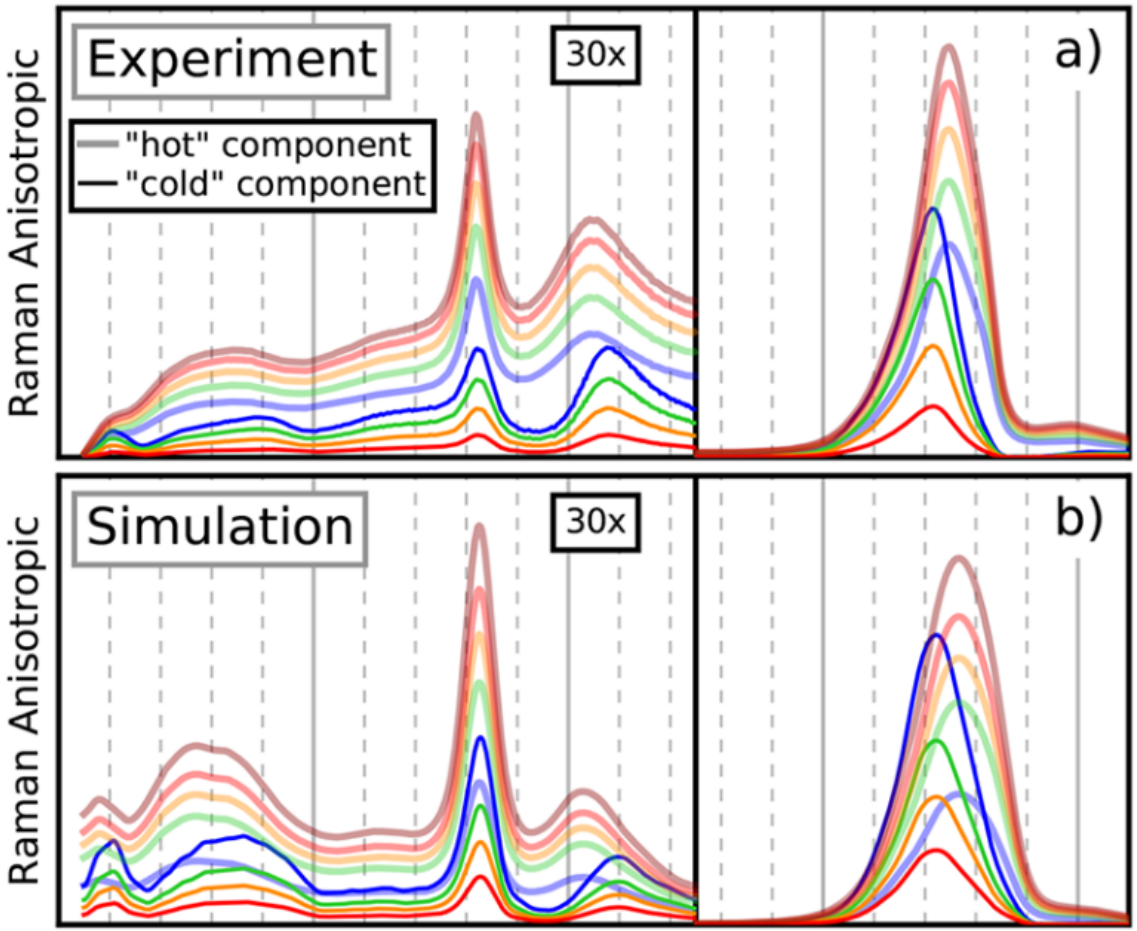
$$k(\rho_{\mathrm{SOAP}}, \rho'_{\mathrm{SOAP}}) = \sum_{n,n',l} p_{n,n',l} p'_{n,n',l}$$



$$k(\rho_{\mathrm{SOAP}}, \rho'_{\mathrm{SOAP}}) = \int d\hat{R} \left| \int \rho_{\rho_{\mathrm{SOAP}}}(\mathbf{r}) \rho_{\rho'_{\mathrm{SOAP}}}(\hat{\mathbf{R}}\mathbf{r}) d\mathbf{r} \right|^{n_{\mathrm{SOAP}}}. \quad (14)$$

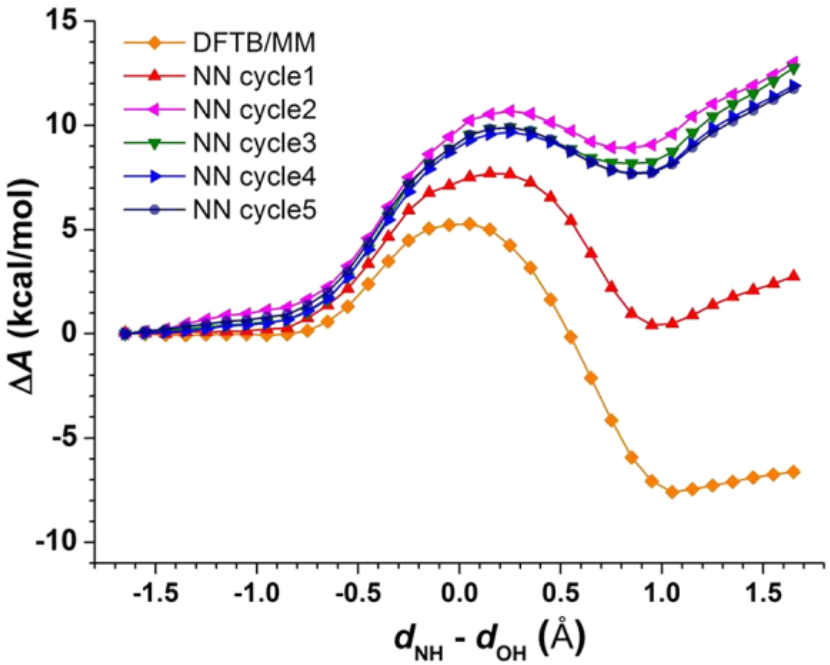
Phys. Rev. B. (2017), **87**: 219902

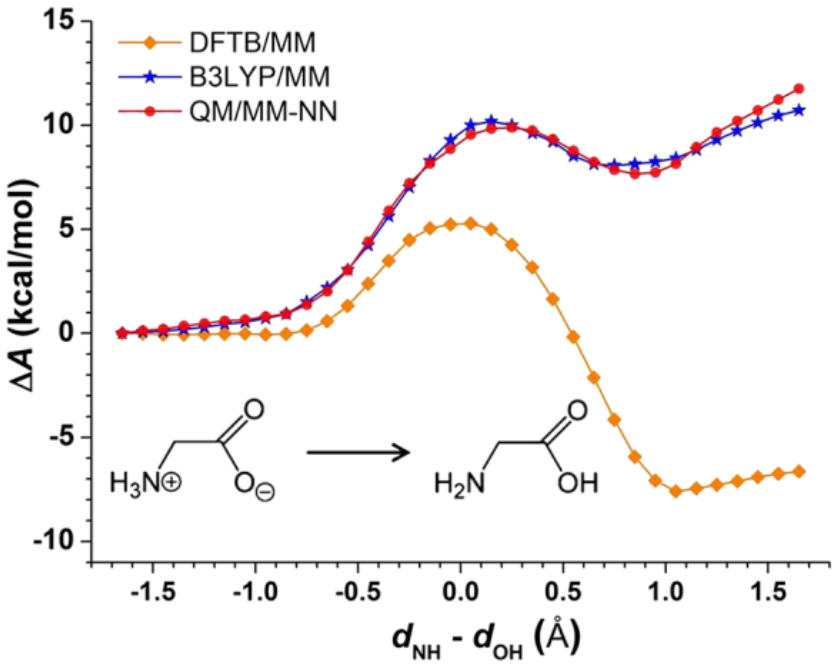
T > 0 K



T. E. Markland: *JPCL* (2018), **9**: 851

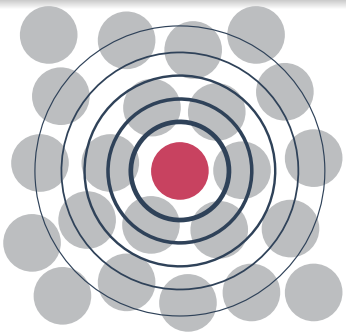
T>0 K, Δ-Learning



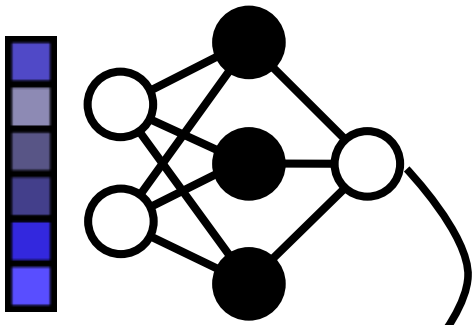


W. Yang: *JCTC* (2018), **14**: 1442

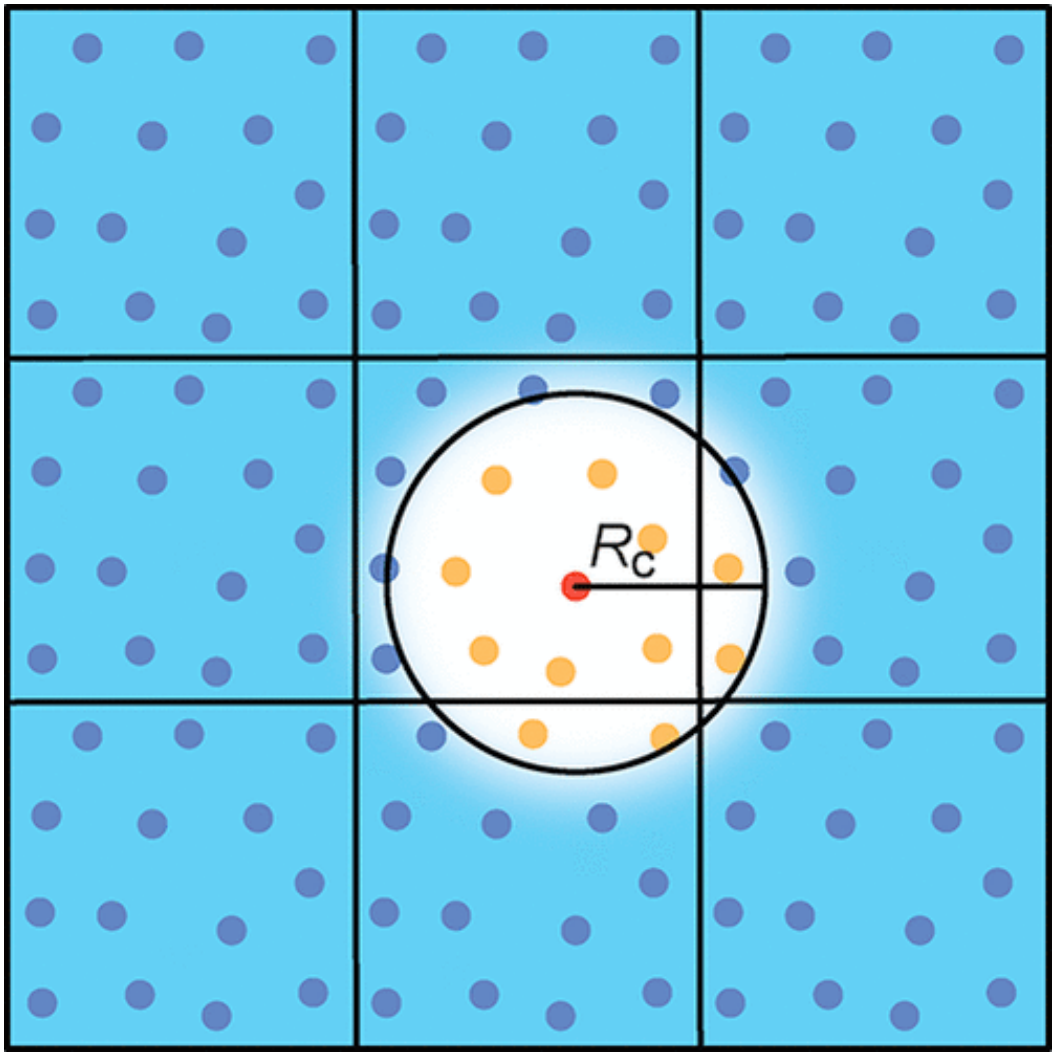
Long-range interactions



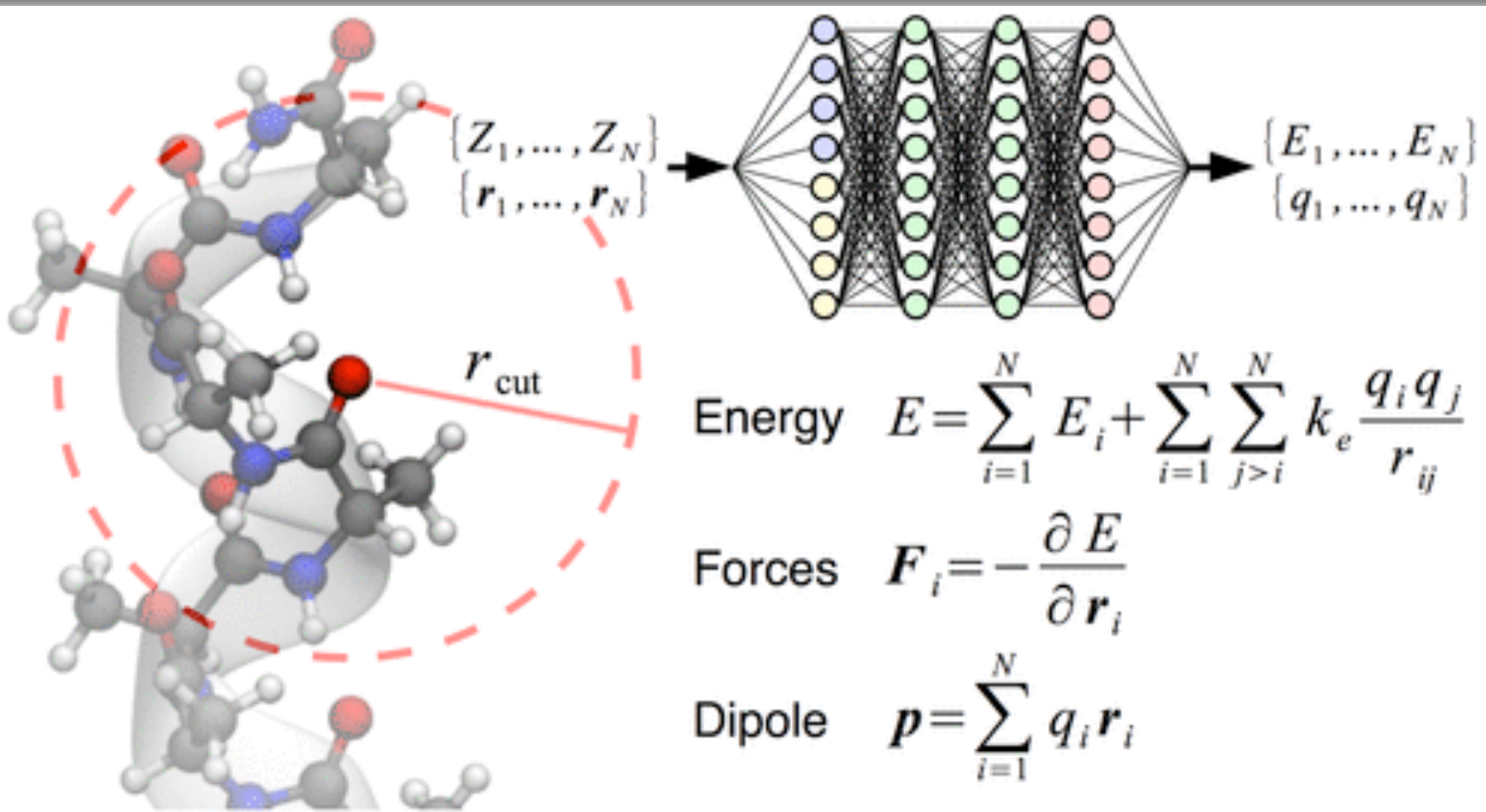
Atom-centered symmetry functions



$$E = \sum_i E_i$$



PhysNet



M. Meuwly: *JCTC* (2019), **15**: 3678

Electronegativity Equilibration

$$U_{\text{tot}}(\{q_i\}) = \sum_{i=1}^{N}\left(E_i^0 + \chi_i q_i + \frac{1}{2}J_{ii}q_i^2\right) + \frac{1}{2}\iint \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,d\mathbf{r}\,d\mathbf{r}',$$

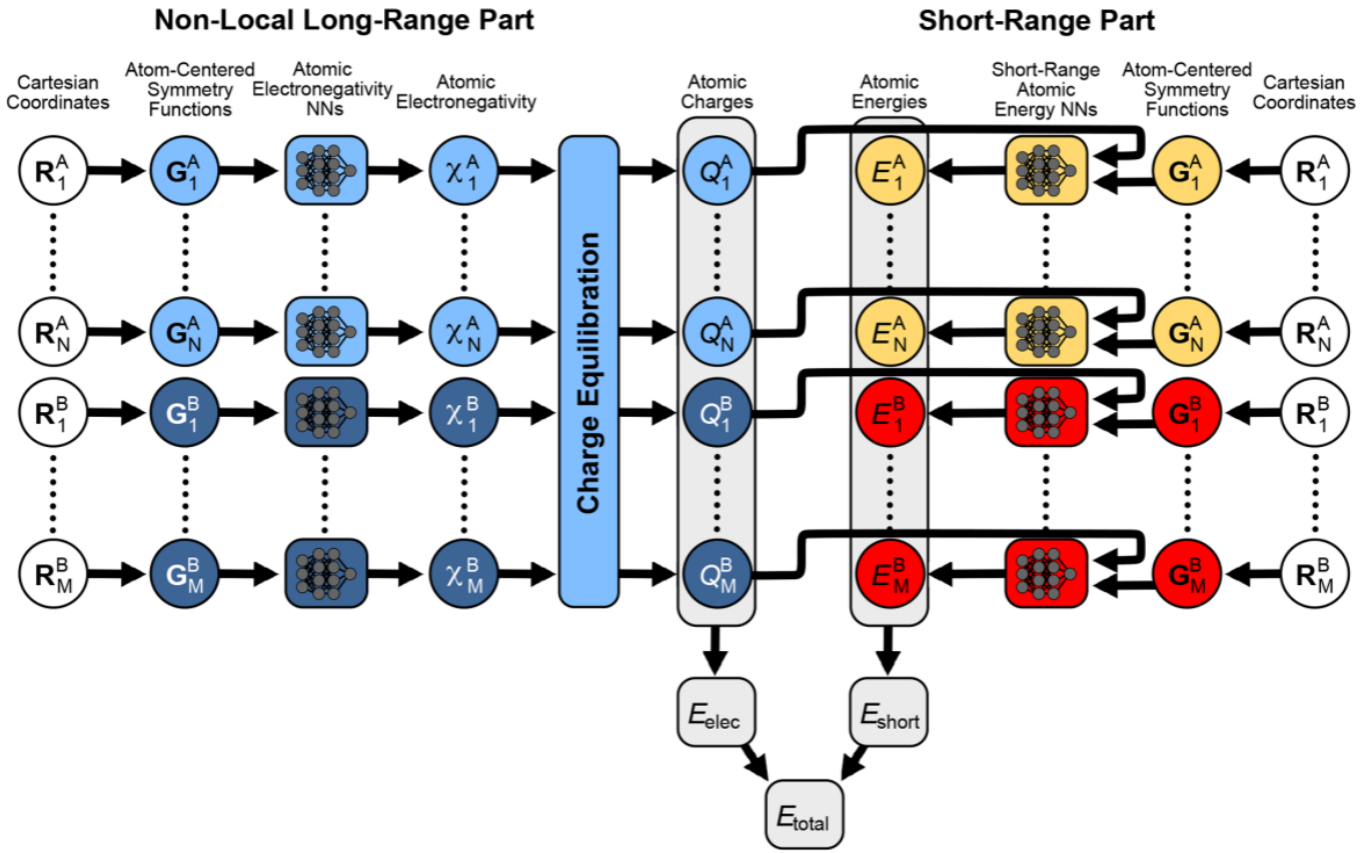
$$\rho_i(\mathbf{r}) = \frac{q_i}{\alpha_i^3\pi^{\frac{3}{2}}}\exp\left(-\frac{|\mathbf{r}-\mathbf{r}_i|^2}{\alpha_i^2}\right).$$

$$U_{\text{tot}}(\{q_i\},\{\mathbf{r}_i\}) = \sum_{i=1}^{N}\left[E_i^0 + \chi_i q_i + \frac{1}{2}\left(J_{ii} + \frac{2\gamma_{ii}}{\sqrt{\pi}}\right)q_i^2\right] + \sum_{i>j}^{N} q_i q_j \frac{\operatorname{erf}(\gamma_{ij}r_{ij})}{r_{ij}},$$

$$\left(\begin{array}{ccc|c} & A_{i,j} & & \begin{matrix}1\\ \vdots\\ 1\end{matrix} \\ \hline 1 & \dots & 1 & 0 \end{array}\right)\left(\begin{matrix} q_1 \\ \vdots \\ \frac{q_N}{\lambda}\end{matrix}\right) = \left(\begin{matrix} -\chi_1 \\ \vdots \\ \frac{-\chi_N}{q_{\text{tot}}}\end{matrix}\right)$$

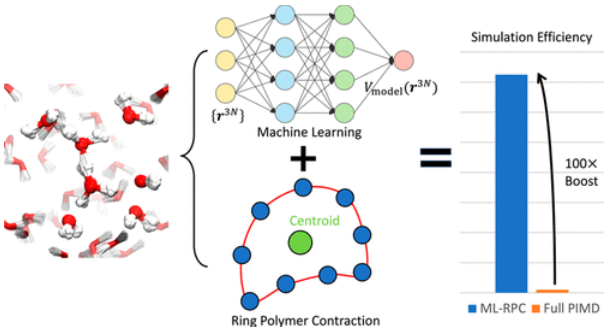
S. Goedecker: *Phys.Rev.B.* (2015), **92**: 045131

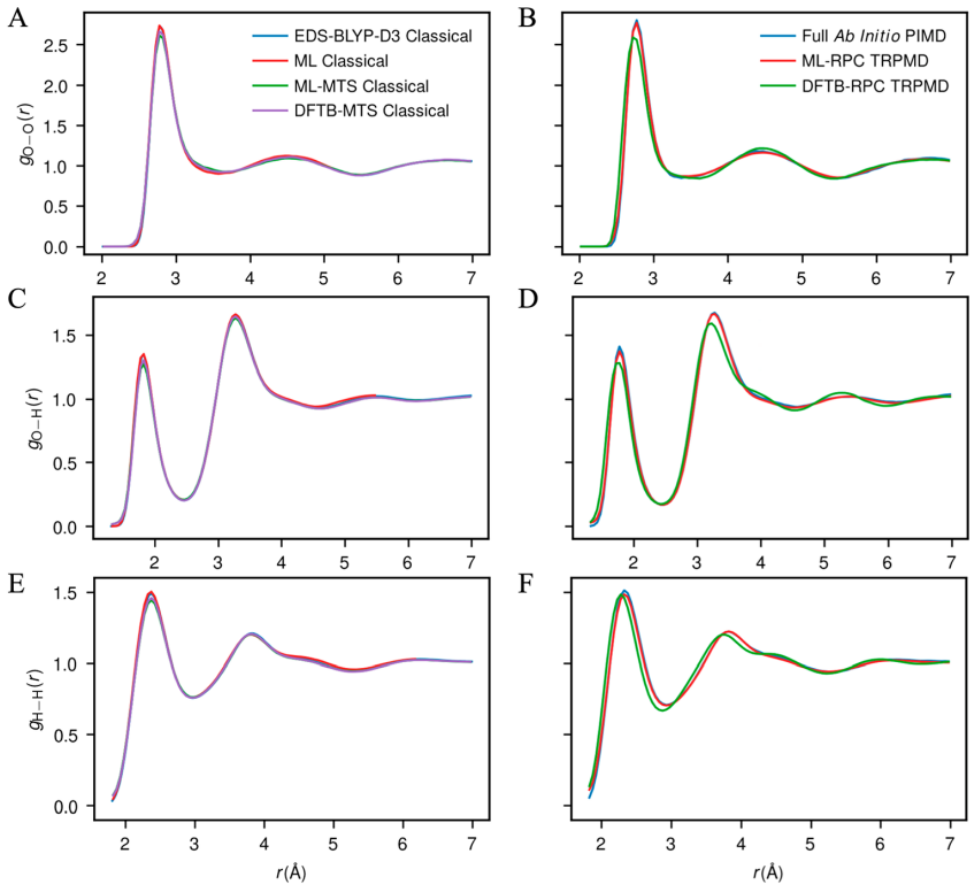
4. Generation NN:



S. Goedecker & Behler: *Nat. Commun.* (2021), **12**: 398

T > 0 K

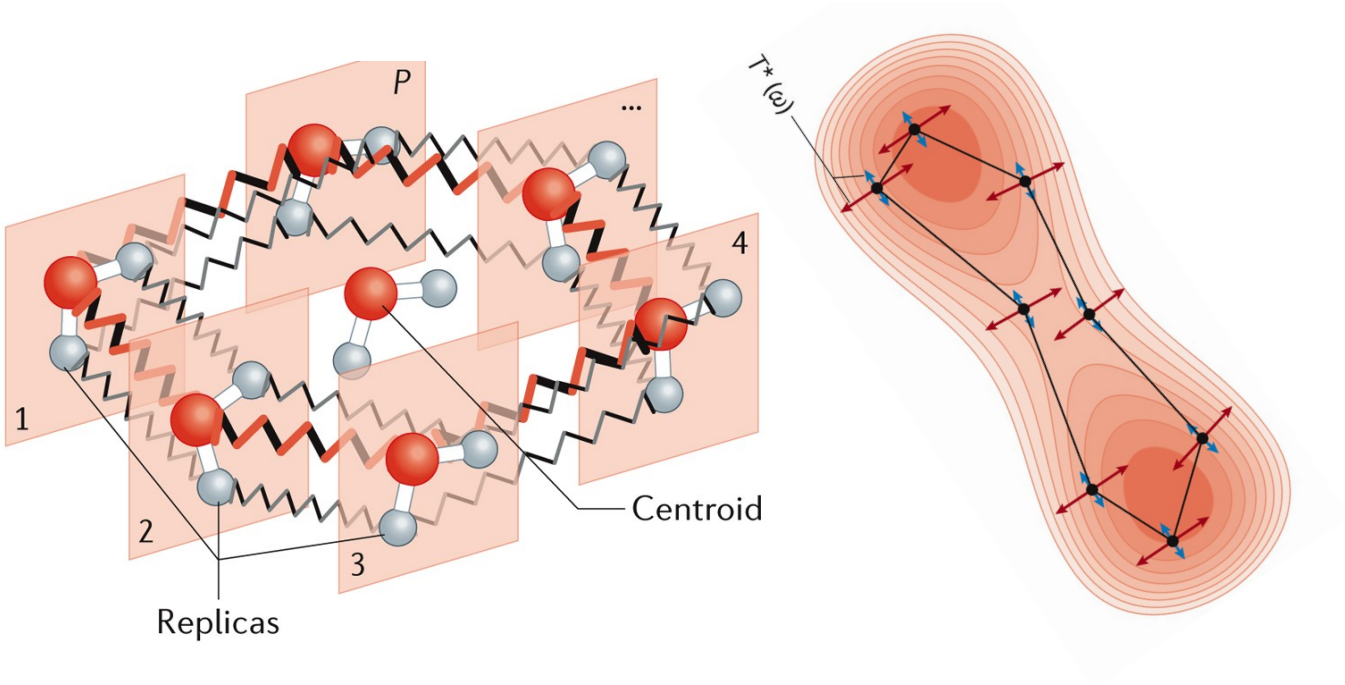




G. A. Voth: *JCTC* (2022), **18**: 599

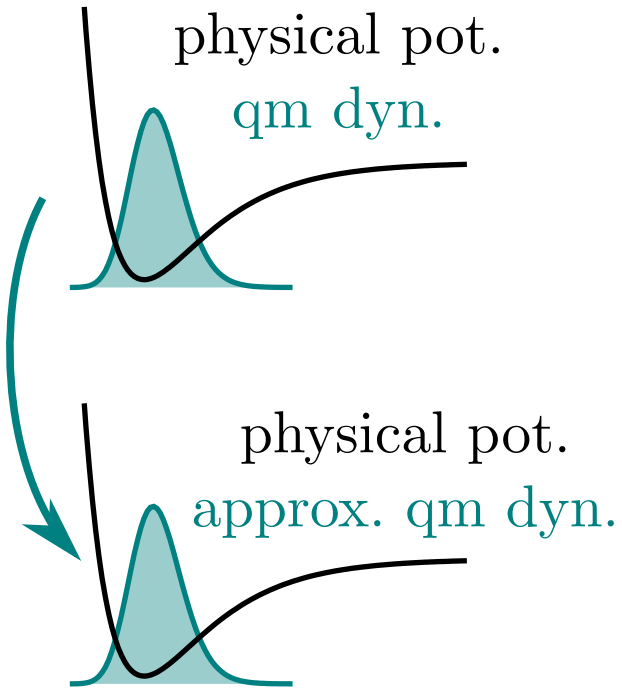
Content

ab initio PIMD



Nat Rev Chem 2, (2018) 0109

ab initio PIMD



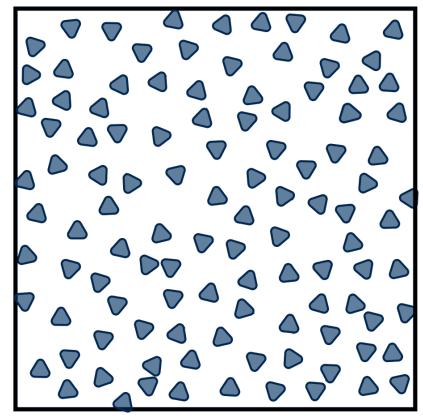
pros:

- PBC explicit solvent
- PBC interfaces
- anharmonicity
- dynamical quantities
- decent NQE

issues:

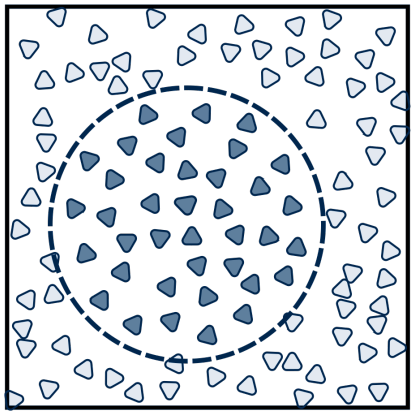
- extreme expe$$$ive
- cost limitted level of theory, scale and time scale

CCMD - LocalCC as reference

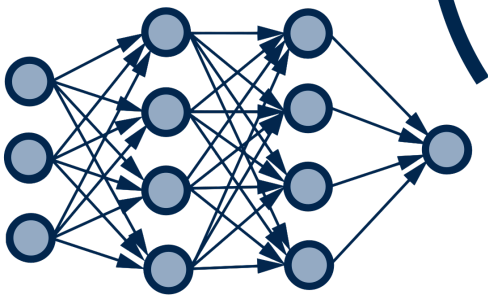


PBC simulation

Iterative Refinement



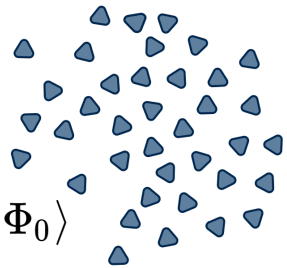
Selection of clusters



ML training

Reference calculations:

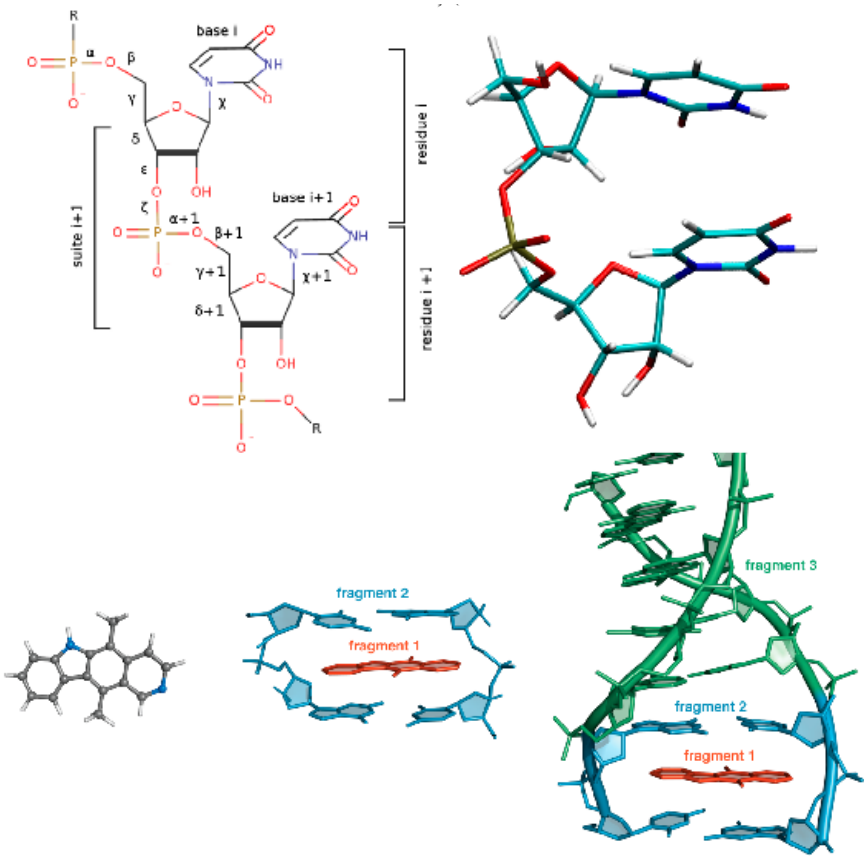
$$|\Psi\rangle = e^T |\Phi_0\rangle$$

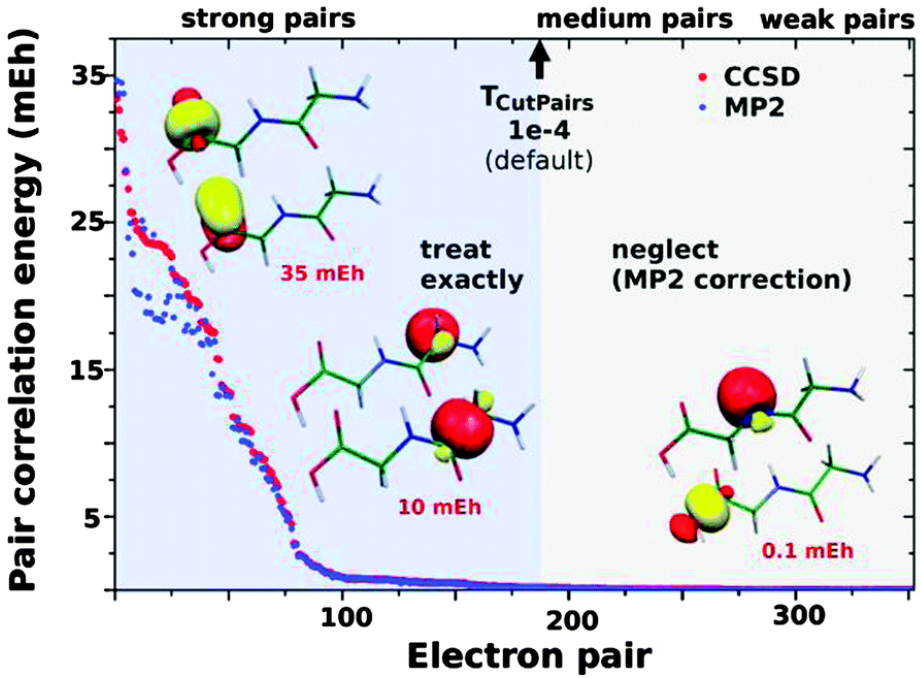


Daru, J., Forbert, H., Behler, J., Marx D., PRL, (2022), **129**, 226001

Local correlation calculations

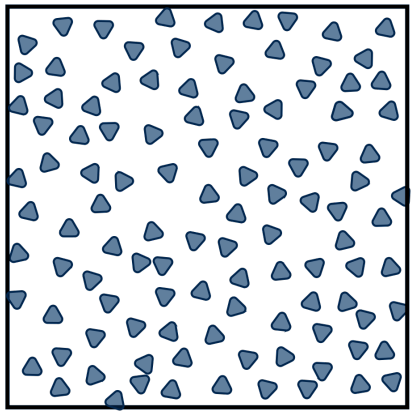
Not linear-scaling, not yet Full CI, but a huge step forward!





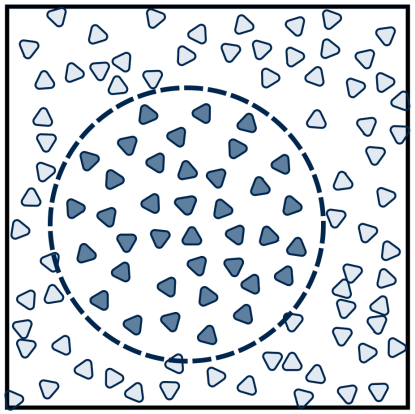
JCTC (2015), **11**: 4972 | *Chem. Soc. Rev.*, (2014), **43**: 5032
JCTC (2017), **13**: 3198

CCMD - LocalCC as reference

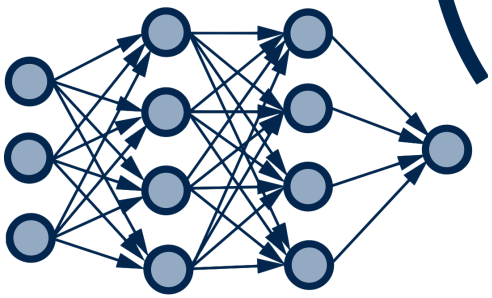


PBC simulation

Iterative Refinement



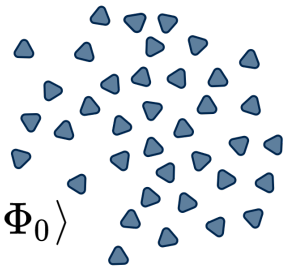
Selection of clusters



ML training

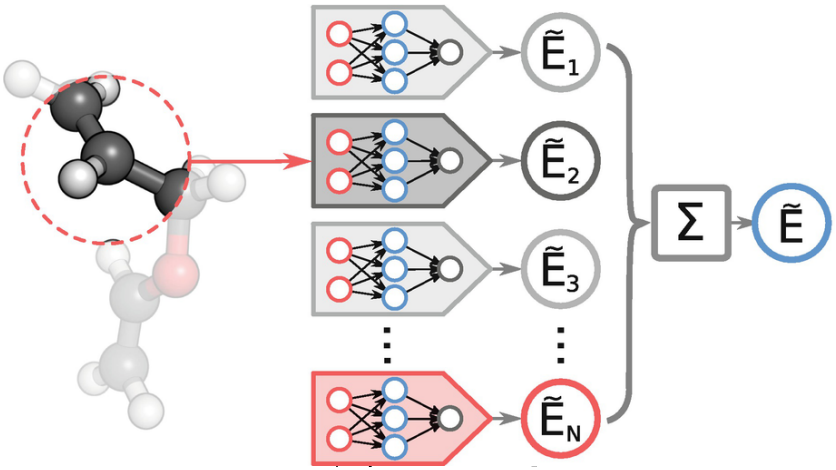
Reference calculations:

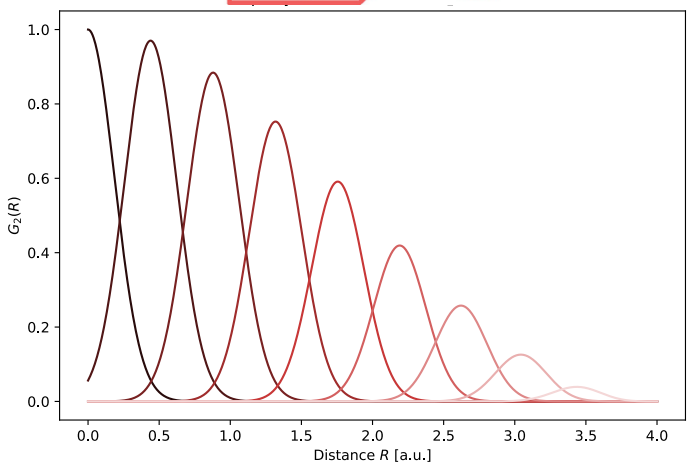
$$|\Psi\rangle = e^T |\Phi_0\rangle$$



Daru, J., Forbert, H., Behler, J., Marx D., PRL, (2022), **129**, 226001

ML: Machine Learning Potentials





$$G^{R}_{a,s} = \sum_{j \neq i}^{\text{all atoms}} e^{\eta(R_{ij}-R_s)^2} f_C(R_{ij})$$

$$G^{A_{mod}}_{a,b,m,n} = 2^{1-\zeta} \sum_{j,k \neq i}^{\text{all atoms}} (1 + \cos(\theta_{ijk} - \theta_m))^{\zeta} \exp\left[-\eta\left(\frac{R_{ij} + R_{ik}}{2} - R_n\right)^2\right] f_C(R_{ij}) f_C(R_{ik})$$

$$f_C(R_{ij}) = \begin{cases} 0.5 \times \left(1 + \cos\left(\frac{\pi R_{ij}}{R_C}\right)\right) & \text{for } R_{ij} \leq R_C \\ 0.0 & \text{for } R_{ij} > R_C \end{cases}$$

Chem. Sci., (2017), **8**: 6924 | *PRL* (2007), **98**: 146401

Electrostatics

$$E^{\mathrm{lr}} = \sum_{A<B} \frac{Q_A Q_B}{r_{AB}} g_{AB}(r_{AB}), \tag{15}$$

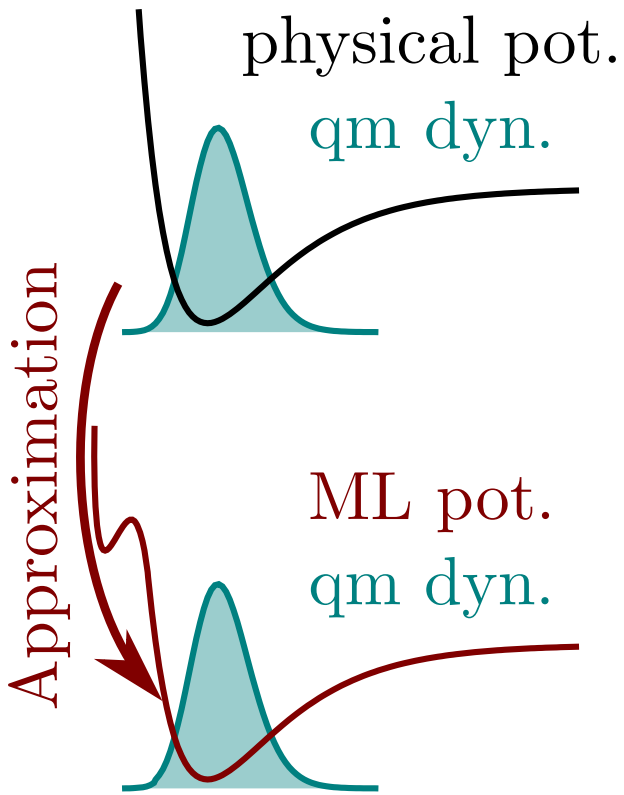
where Q_A and Q_B are the partial atomic charge of atom A and B, i.e. O and H, while $g_{AB}(r_{AB})$ is the Gauss-Gauss Coulomb integral attenuation factor at the interatomic distance r_{AB}

$$g_{AB}(r_{AB}) = \mathrm{erf}\left(\sqrt{\frac{a_A a_B}{a_A + a_B}} r_{AB}\right). \tag{16}$$

The short-range core-core repulsions within the system are represented using the pairwise Yukawa potential,

$$E^{\mathrm{sr}} = \sum_{A<B} \frac{Z_A Z_B}{r_{AB}} \exp(-\beta_{AB} r_{AB}), \tag{17}$$

ML PIMD

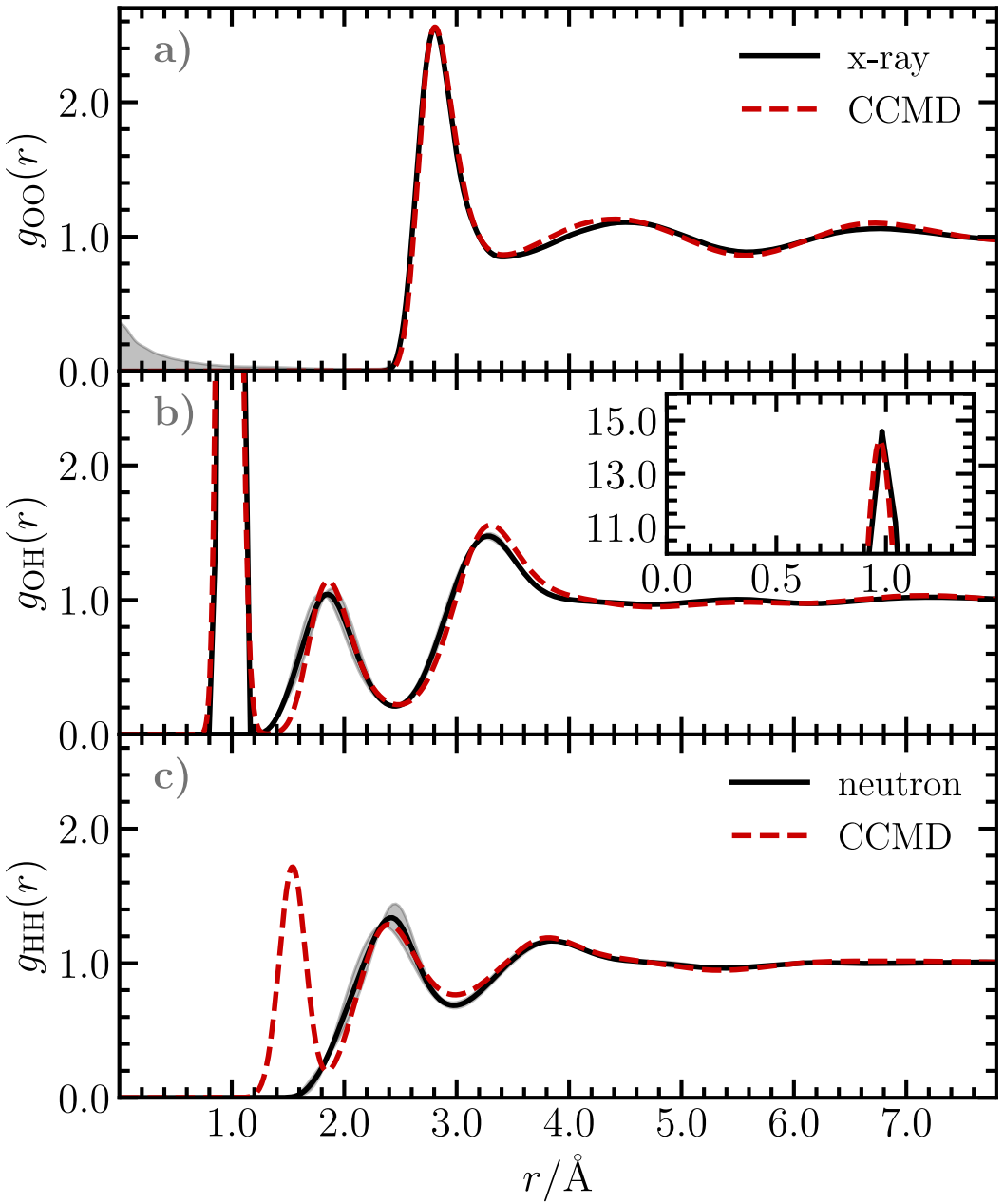


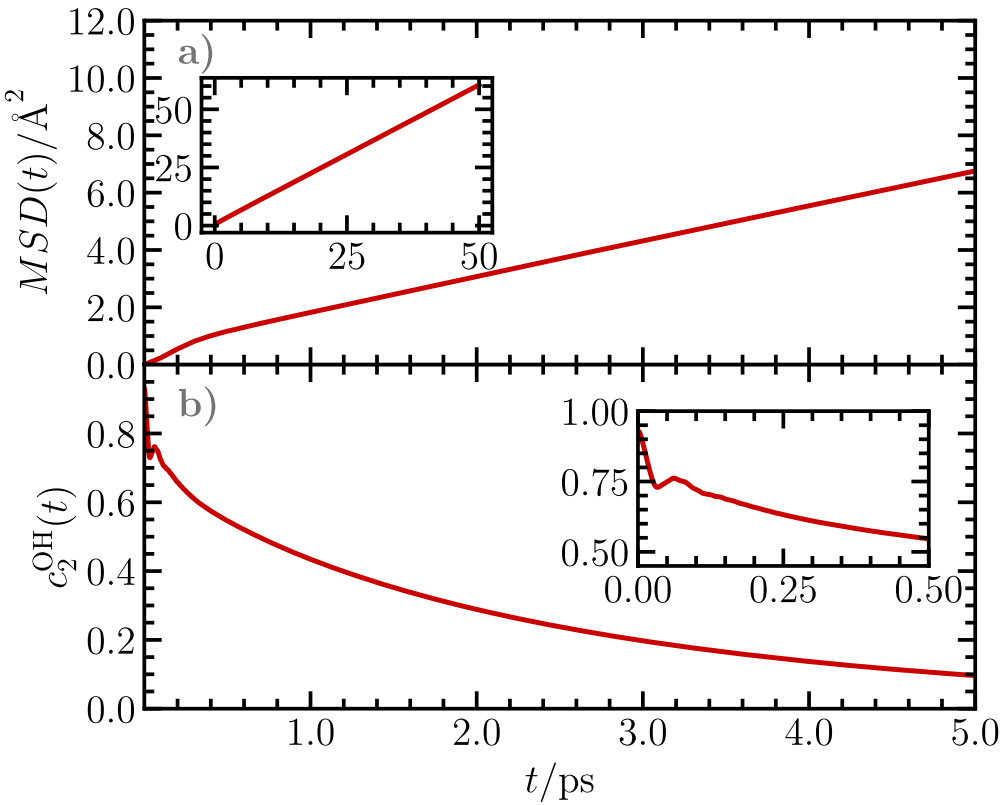
pros:

- PBC explicit solvent
- PBC interfaces
- anharmonicity
- NQE
- approx. dynamical quantities

issues:

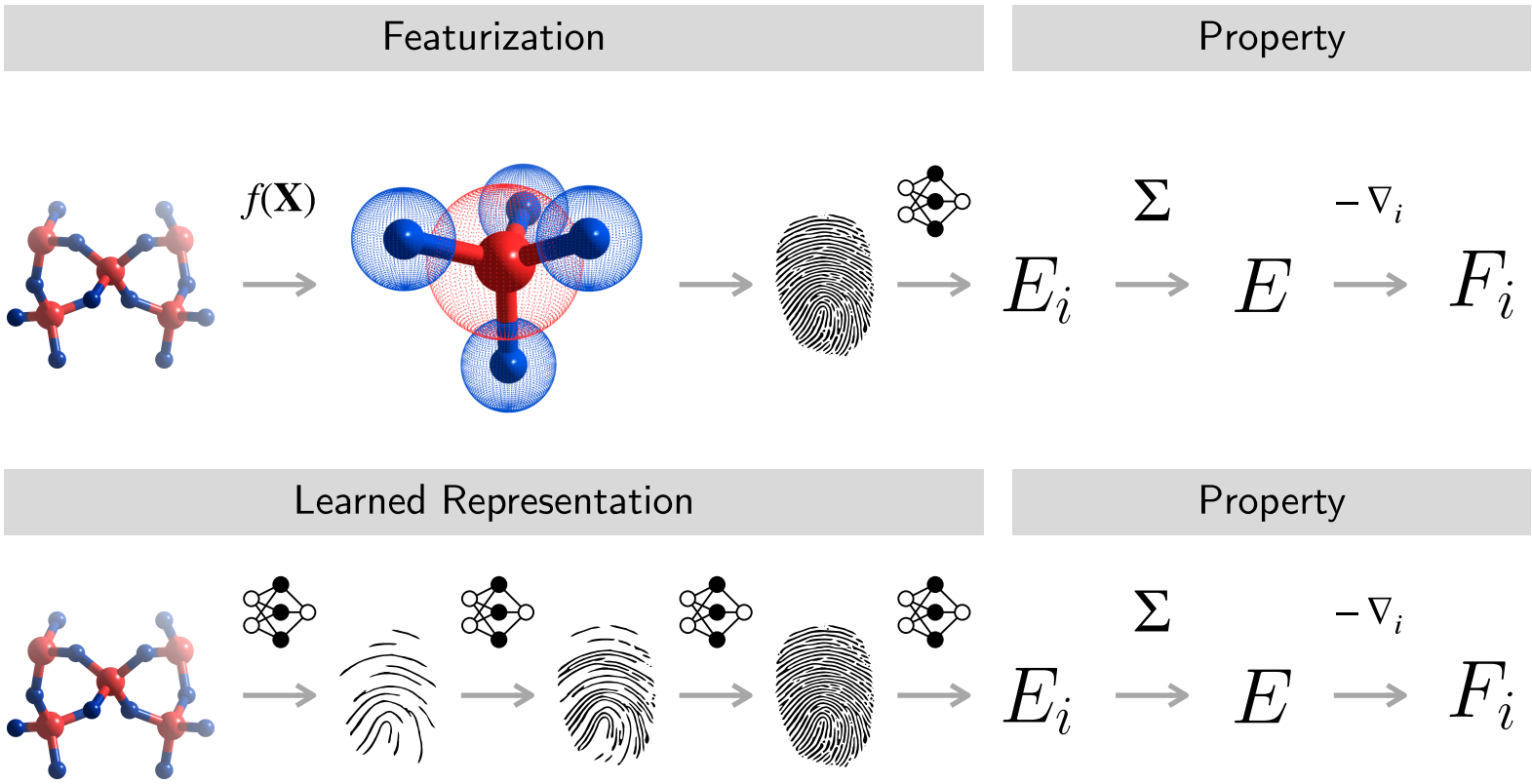
- PES can be unphysical
- iterative refinement is needed
- labor intensive


a)
x-ray
CCMD
$g_{OO}(r)$
2.0
1.0
0.0
b)
$g_{OH}(r)$
15.0
13.0
11.0
0.0
0.5
1.0
c)
neutron
CCMD
$g_{HH}(r)$
1.0
2.0
3.0
4.0
5.0
6.0
7.0
$r/\text{Å}$



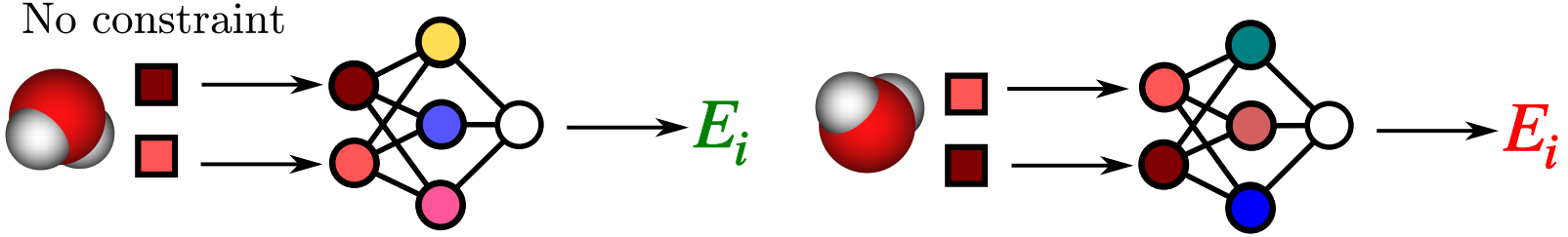
	D / Å^2·ps^{-1}	τ_{OH} / ps
Experimental	0.23 Å^2	1.71–1.96 to 2.5–3.0
Calculated	0.244 ± 0.002	1.733 ± 0.029 to 2.992 ± 0.065

Learned representations

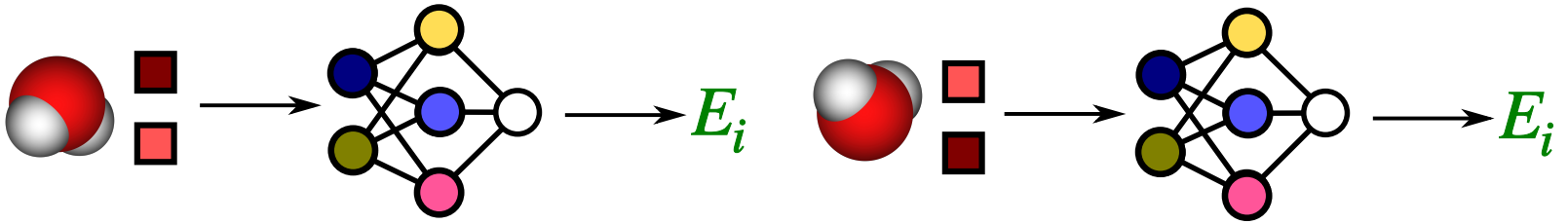


Equivariance

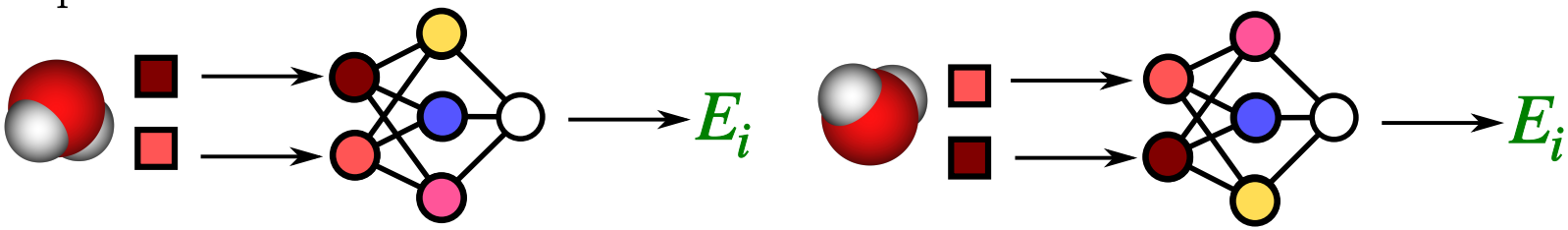
No constraint



Invariants



Equivariants



Equivariance

Allegro

$$\begin{aligned}\mathbf{V}^{ij,L}_{n,(\ell_1,p_1,\ell_2,p_2)\to(\ell_{\text{out}},p_{\text{out}})} &= \sum_{k\in\mathcal{N}(i)} w^{ik,L}_{n,\ell_2,p_2}\left(\mathbf{V}^{ij,L-1}_{n,\ell_1,p_1}\otimes\vec{Y}^{ik}_{\ell_2,p_2}\right)\\ &= \sum_{k\in\mathcal{N}(i)} \mathbf{V}^{ij,L-1}_{n,\ell_1,p_1}\otimes\left(w^{ik,L}_{n,\ell_2,p_2}\vec{Y}^{ik}_{\ell_2,p_2}\right)\\ &= \mathbf{V}^{ij,L-1}_{n,\ell_1,p_1}\otimes\left(\sum_{k\in\mathcal{N}(i)} w^{ik,L}_{n,\ell_2,p_2}\vec{Y}^{ik}_{\ell_2,p_2}\right)\end{aligned}$$

A. Musaelian et al. *arXiv:2204.05249 (2022)*

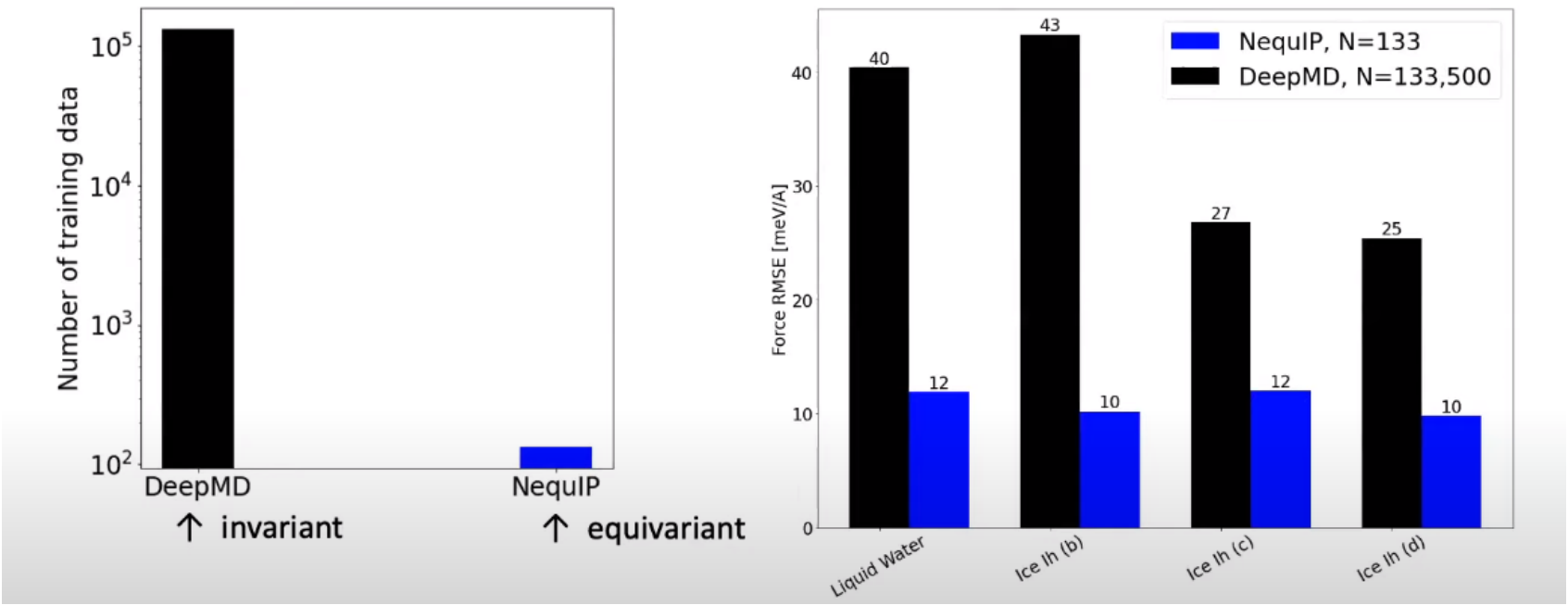
MACE

$$\boldsymbol{m}^{(t)}_i = \sum_j \boldsymbol{u}_1\left(\sigma^{(t)}_i;\sigma^{(t)}_j\right) + \sum_{j_1,j_2}\boldsymbol{u}_2\left(\sigma^{(t)}_i;\sigma^{(t)}_{j_1},\sigma^{(t)}_{j_2}\right) + \cdots + \sum_{j_1,\ldots,j_\nu}\boldsymbol{u}_\nu\left(\sigma^{(t)}_i;\sigma^{(t)}_{j_1},\ldots,\sigma^{(t)}_{j_\nu}\right),$$

I. Batatia et al. *arXiv:2206.07697* (2022)

$$A^{(1)}_{i,kl_1m_1} = \sum_{j\in\mathcal{N}(i)} R^{(1)}_{kl_1}(r_{ji})Y^{m_1}_{l_1}(\hat{\boldsymbol{r}}_{ji})\mathcal{W}^{(1)}_{kz_j}.$$

Symmetry constrains → improved gerneralization



slide from: Albert Musaelian

Thank you for ...

... founds:

OTKA-FK147031

nkfih

... cpu time:

KIFÜ

HPC @hu
Kompetencia Központ

... the attention!

... collaboration:

Andreas W. Hauser
Dominik Marx
Jörg Behler
Zoltán Novák
Gergely Tolnai
Tibor Soós
Szilárd Varga

daru.web.elte.hu
janos.daru@ttk.elte.hu