UNIVERSITY OF CAMBRIDGE

International Centre for Advanced Materials

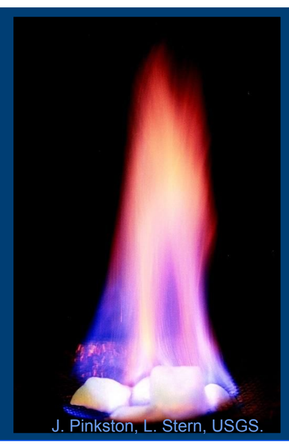
bp

GAP-models for small molecules

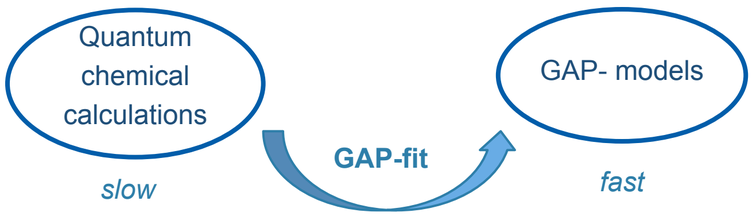
Eszter Székely, Gábor Csányi

Dept. of Engineering, University of Cambridge

(E-mail: es732@cam.ac.uk)

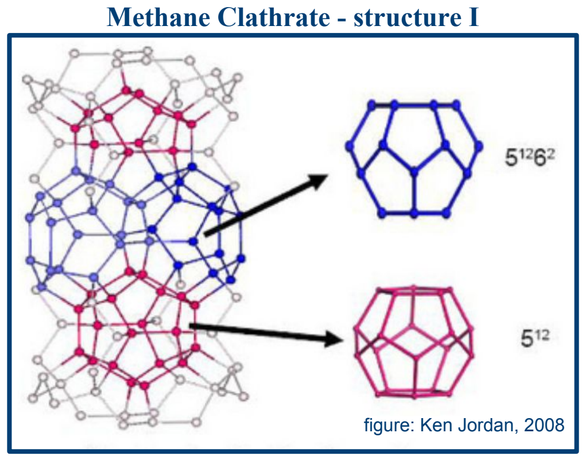


Outline of the project



- Fit PES for methane-water systems combining:
  - high-level quantum chemistry: CCSD(T)-F12
  - machine learning: GAP [1]
- GAP (*Gaussian Approximation Potentials*):
  - fast (1,000,000x than input method)
  - transferable
  - accurate
  - systematically improvable

Methane Clathrate - structure I



The molecular crystalline structure is built from water cages connected to each other containing one methane molecule each.

Methodology

- **Many-body expansion:**

$$E_N(r_1,...,r_N) = \sum_{i=1}^{N} E_{1B}(r_i) + \sum_{i<j}^{N} E_{2B}(r_i,r_j) + \sum_{i<j<k}^{N} E_{3B}(r_i,r_j,r_k) + \\ +... + E_{NB}(r_1,r_2,...,r_N)$$

$$E_N = E_{1B} + E_{2B} + E_{3B} + E_{\text{beyond-3B}}$$

Partridge-Schwenke PES H_2O, CH_4 — **GAP** — **Force fields** eg. SCME, TTM-models

- **Gaussian Approximation Potentials (GAP)**

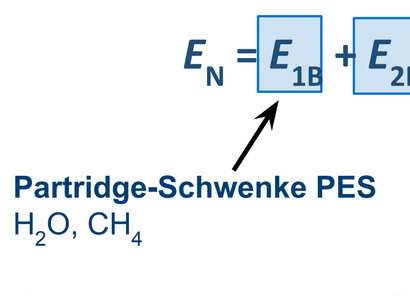
Fit the PES on a database: {geometry + energy, (forces)}
$\boldsymbol{R}_k$ $E(\boldsymbol{R}_k)$ $E'(\boldsymbol{R}_k)$

Energy function in the space of kernel basis functions: $E(\boldsymbol{R}_i) = \sum_k^n w_k K(\boldsymbol{R}_i, \boldsymbol{R}_k)$
$K(\boldsymbol{R}_i, \boldsymbol{R}_k)$

Minimise the loss function: $l(\boldsymbol{w}) = ||(K + \lambda \cdot I) \cdot \boldsymbol{w} - \boldsymbol{E}(\{\boldsymbol{R}_m\}||^2$

$$\boldsymbol{w} = (K + \lambda I)^{-1} \boldsymbol{E}(\{\boldsymbol{R}_m\})$$

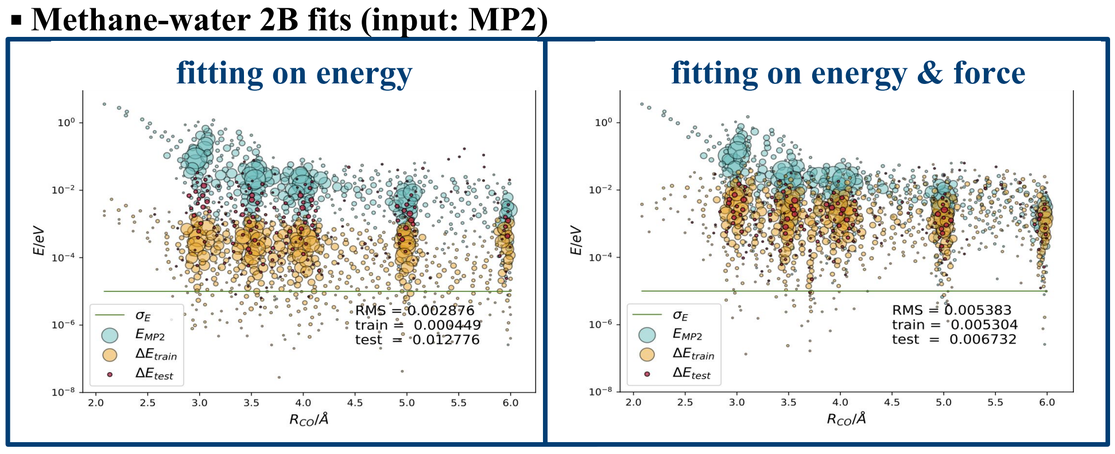
- **Quantum mechanics: CCSD(T)-F12**



CCSD(T)-F12 test calculations on CH_4-H_2O 2B:
Differences from database (orange) changing:

- basis set settings:
  - AVTZ
  - basis1: default=AVTZ,H=VTZ
  - basis2: H=VTZ,O=AVTZ,C=AVTZ
- using counterpoise (CP) correction or not (noCP)

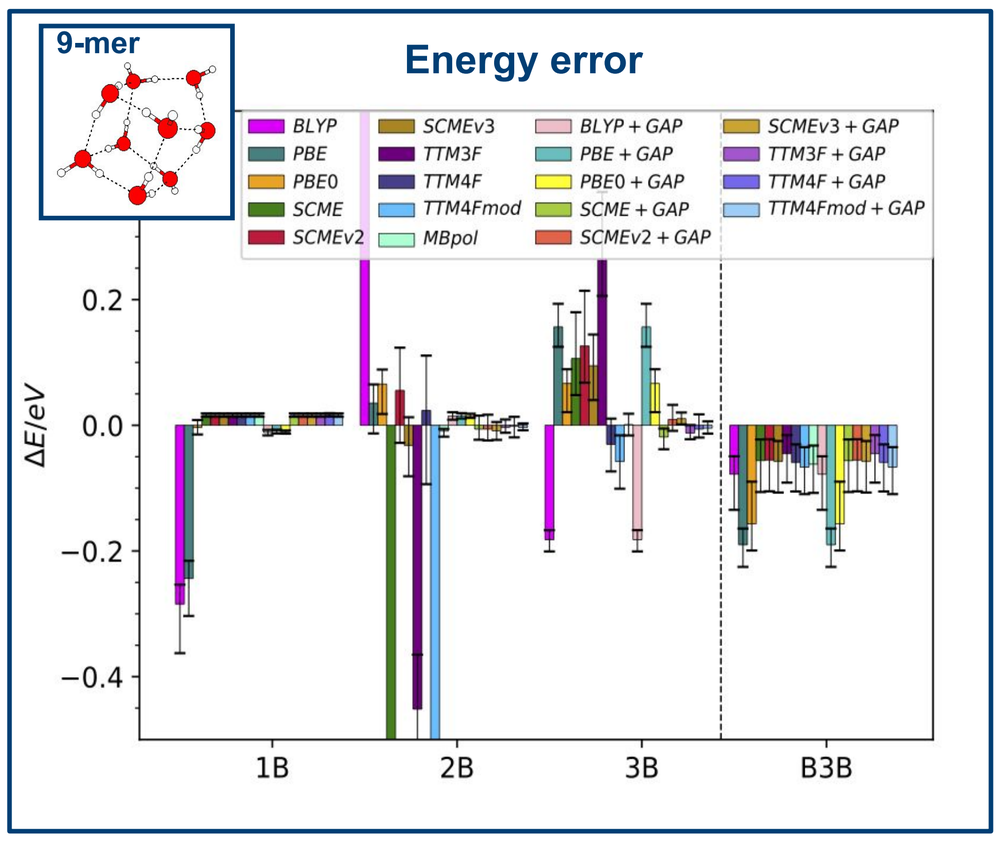
- **Methane-water 2B fits (input: MP2)**



Fitting including the forces reduces overfitting

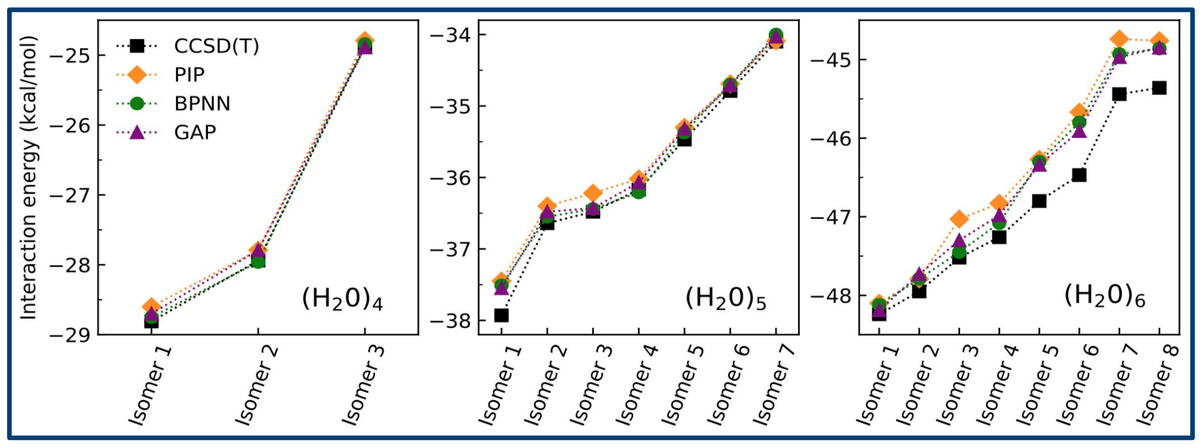
Results

- **Comparison of different models on compressed water nonamers**



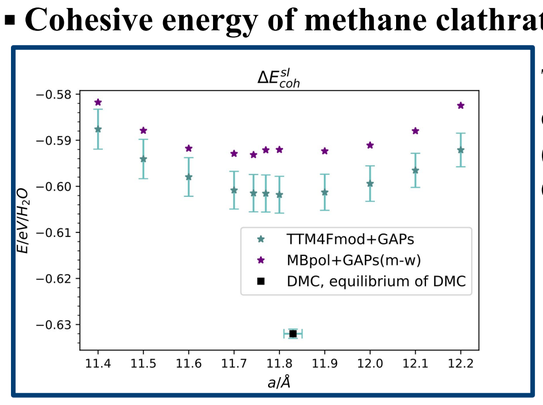
- The GAP-corrections are in the 2B and 3B terms for the classical models (SCME- and TTM*x*F -models) but only in the 2B terms for the DFT-s (BLYP, PBE, PBE0)
- The DFT+GAP-s are from Ref. [2]

- **GAP & PIP & BPNN**



As shown in Ref. [3] for the water 2B & 3B energies, GAP is comparable with other state-of-the-art fitting methods: PIP (*permutationally invariant polynomials*) and BPNN (*Behler-Parinello neural networks*) when fed the same inputs of energy differences.

- **Cohesive energy of methane clathrate**



The energy minimum is at the experimental lattice constant but there's still a difference from DMC. (DMC result is from Ref. [4])
Concerns & questions:

- whether DMC and CCSD(T)-F12 are comparable on that scale?
- different level of theory in inputs
- fits could be improved more.
- 4B effects are more important than thought?

Conclusions

- Better fit including both *E* & *E'* than only *E*.
- GAP & PIP & BPNN result in similar fits fitted only on *E*-s
- GAP-corrected water force fields achieve higher accuracy than DFTs with shorter evaluation time
- Largest errors are in >3B terms for water
- CH_4-H_2O interaction terms are not enough accurate yet.

Acknowledgements: The authors would like to acknowledge the funding from the Peterhouse Research Studentship and the funding and technical support from BP through the BP International Centre for Advanced Materials (BP-ICAM) which made this research possible.
Related publications: T.T. Nguyen et al, *J. Chem. Phys.* 148(241725), 1–11. (2018)

References:
[1] A. P. Bartók, G. Csányi, *Int. J. Quant. Chem.* 115(16) (2015), code available at: www.libatoms.org
[2] D. Alfè et al, *J. Chem. Phys.* 141(1) (2014)
[3] T. T. Nguyen et al, *J. Chem. Phys.* 148(241725), 1–11. (2018))
[4] S. J. Cox et al, *J. Chem. Phys.* 140(17) (2014)