Fostering accuracy in modelling materials and molecular complexes with quantum Monte Carlo

Andrea Zen

Università di Napoli Federico II

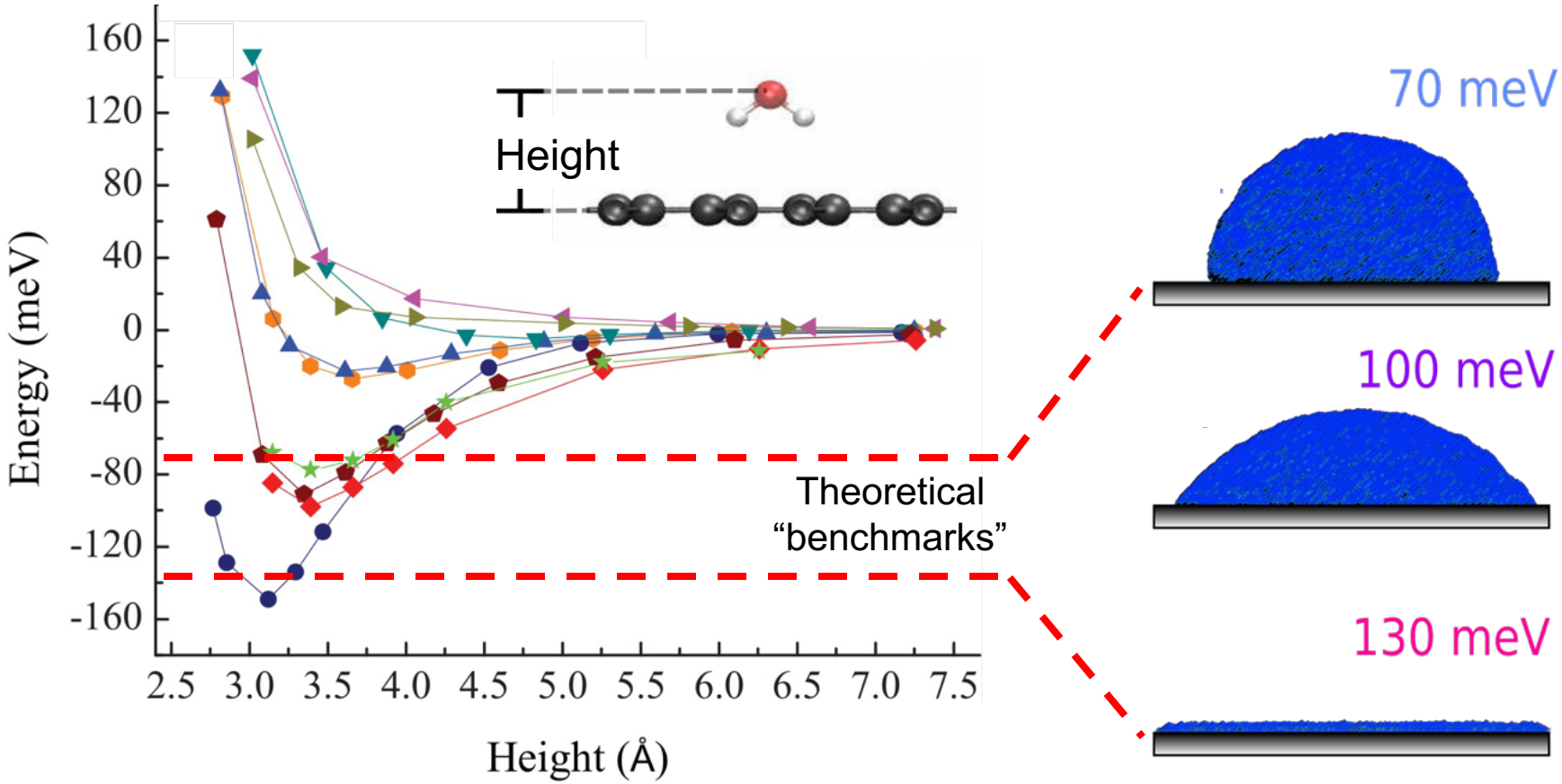
UNIVERSITÀ DEGLI STUDI DI NAPOLI
FEDERICO II

Bridging Quantum Monte Carlo and High-Performance Simulations
Luxembourg

#TREXSymposium
February 8th , 2024

Accuracy is a challenge: water at graphene

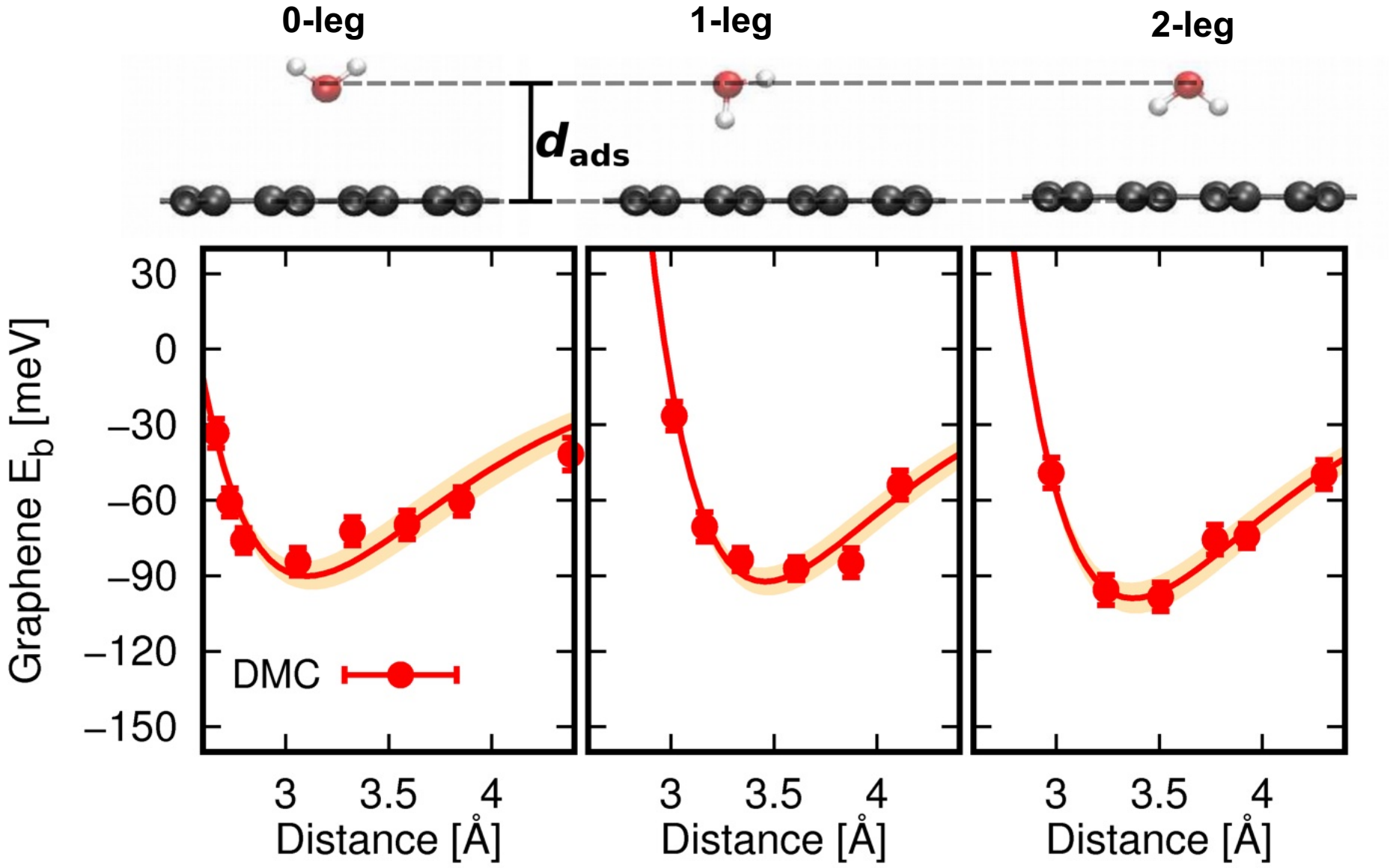
Binding curve from density functional theory [PRB, 84, 033402 (2011)].



Lots of computational interest: Bludsky, Hamada, Jordan, Kresse, Paesani, Paulus, Peeters, Silvestrelli…

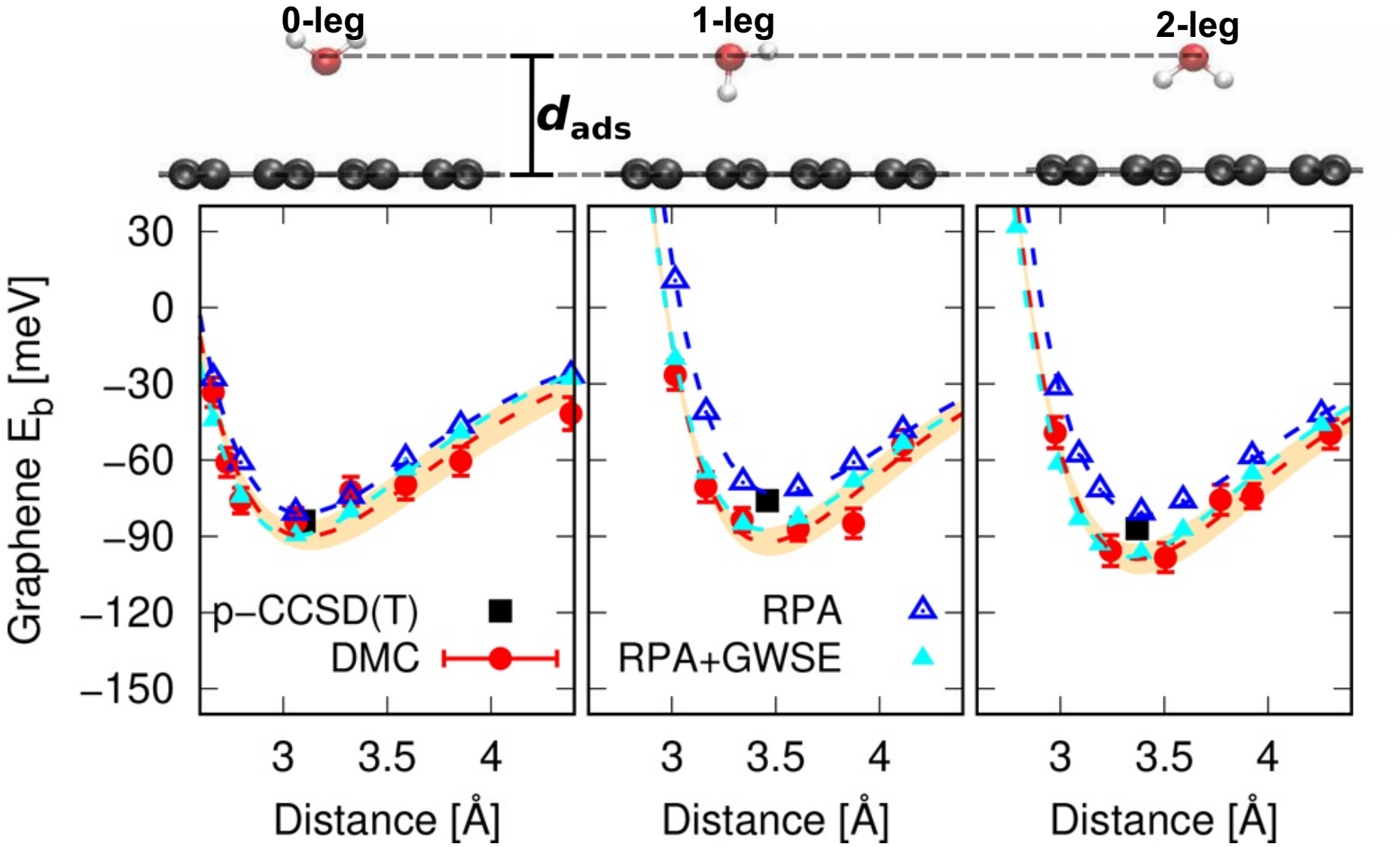
More accurate approaches are needed!

Adsorption of water on graphene with DMC



J.G. Brandenburg, A. Zen, M. Fitzner, B. Ramberger, G. Kresse, T. Tsatsoulis, A. Grüneis, A. Michaelides, D Alfè, *JPCL* **2019,** 10, 358

… and with periodic CCSD(T), and RPA+GWSE



- Consensus among DMC, CCSD(T), RPA+GWSE

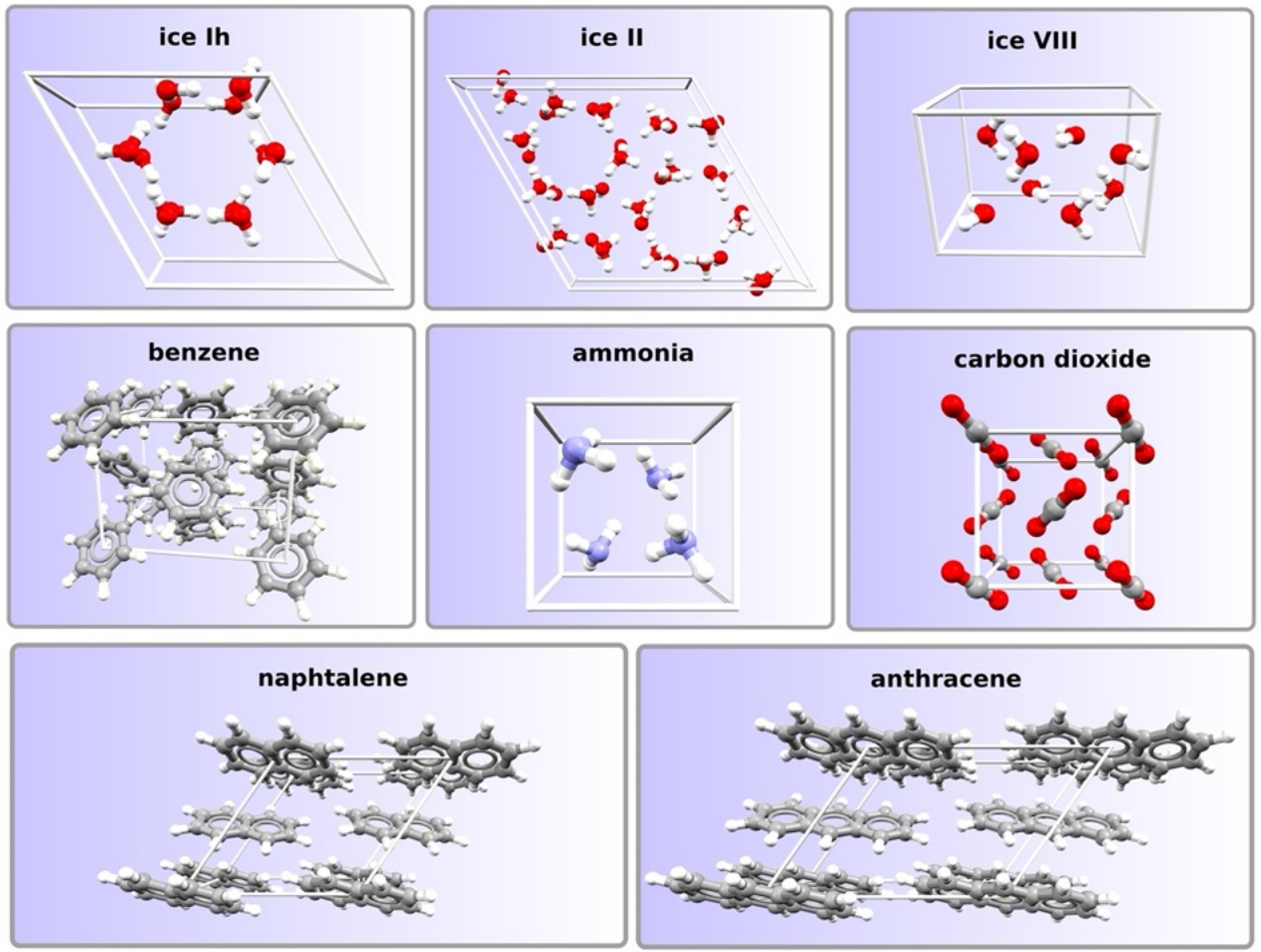
J.G. Brandenburg, A. Zen, M. Fitzner, B. Ramberger, G. Kresse, T. Tsatsoulis, A. Grüneis, A. Michaelides, D Alfè, *JPCL* **2019,** 10, 358

Discrepancies resolved

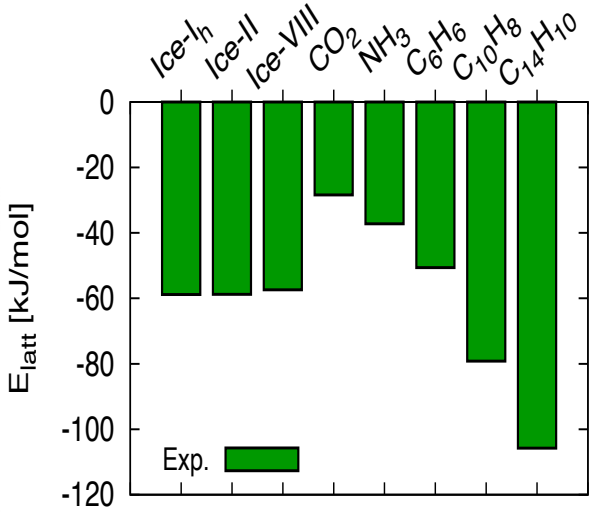
Ref.	E_b	Method	PRO / CON
1	**-130 meV**	**DFT/CC**	Unreliable extrapolation
2	**-130 meV**	**DFT-SAPT**	Unreliable extrapolation, SAPT is not a reference method
3	**-70 ± 10 meV**	**DMC**	Large stochastic error, finite-size effects are neglected
4	**-135 meV**	**i-CCSD(T)**	Single particle basis set too small
5	**-99 ± 6 meV** **-87 meV** **-98 meV**	**DMC** **p-CCSD(T)** **RPA+GWSE**	**Consensus** between independent evaluations from **UCL group** (DMC), **Grüneis' group** (pCCSD(T)), **Kresse's group** (RPA+GWSE)

1. Miroslav Rubes *et al., JPC C* **2009**, 113, 8412
2. G.R. Jenness, O. Karalti and K.D. Jordan, *PCCP* **2010**, 12, 6375
3. J. Ma, A. Michaelides, D. Alfè, L. Schimka, G. Kresse, and E. Wang, *Phys. Rev. B* **2011**, 84, 033402
4. E. Voloshina, D. Usvyat, M. Schutz, Y. Dedkov and B. Paulus *PCCP* **2011**, 13, 12041
5. J.G. Brandenburg, A. Zen, M. Fitzner, B. Ramberger, G. Kresse, T. Tsatsoulis, A. Grüneis, A. Michaelides, D Alfè, *JPCL* **2019,** 10, 358

Accuracy in molecular crystals: Lattice energy from theory and experiments



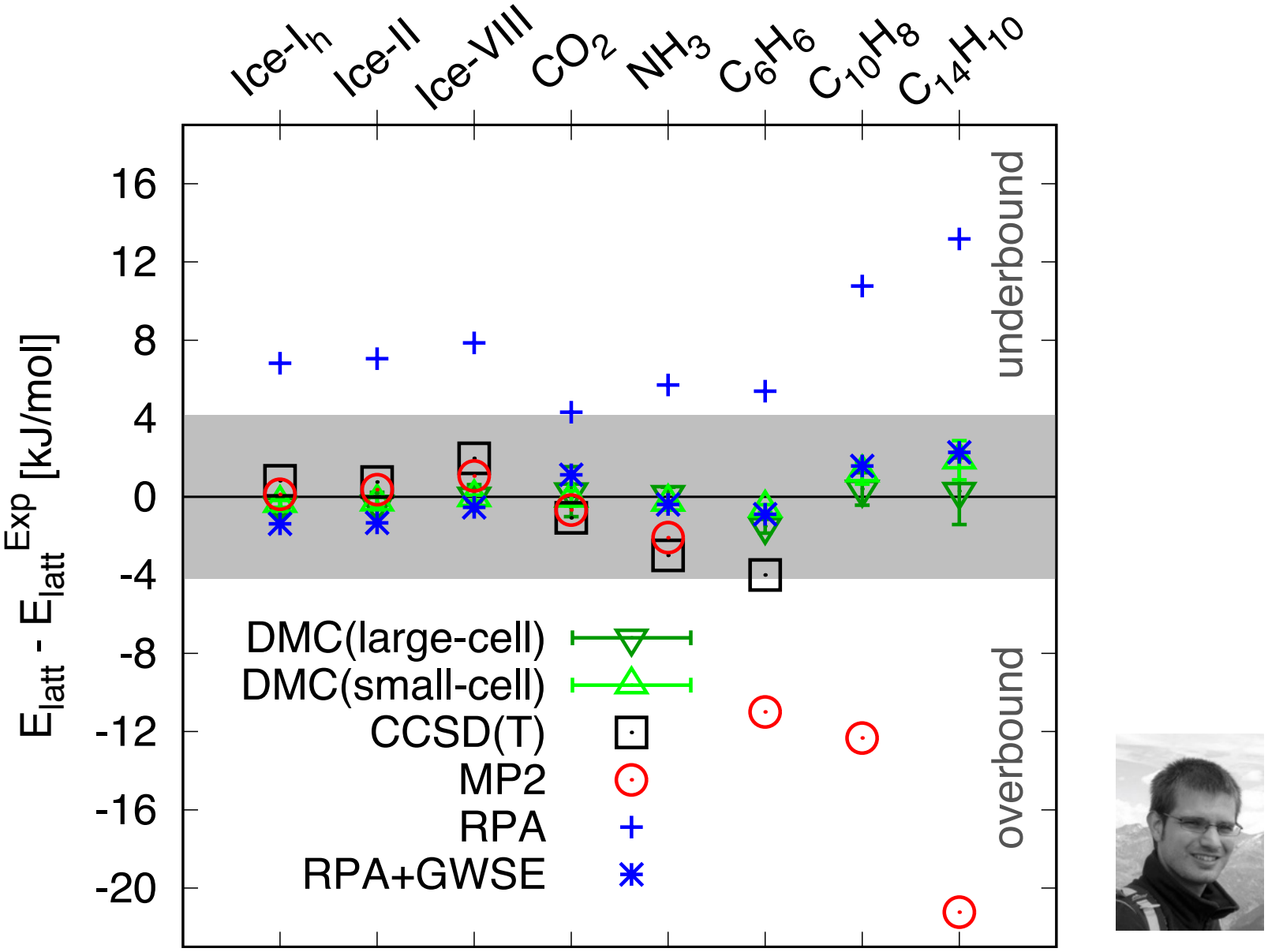
A representative set of 8 molecular crystals, comprising a diversity in intermolecular interactions, from strong **hydrogen bonds** to **London dispersion**



Three polymorphs of ice and five crystals from the C21 test set of Otero-de-la-Roza and Johnson [JCP137:054103]

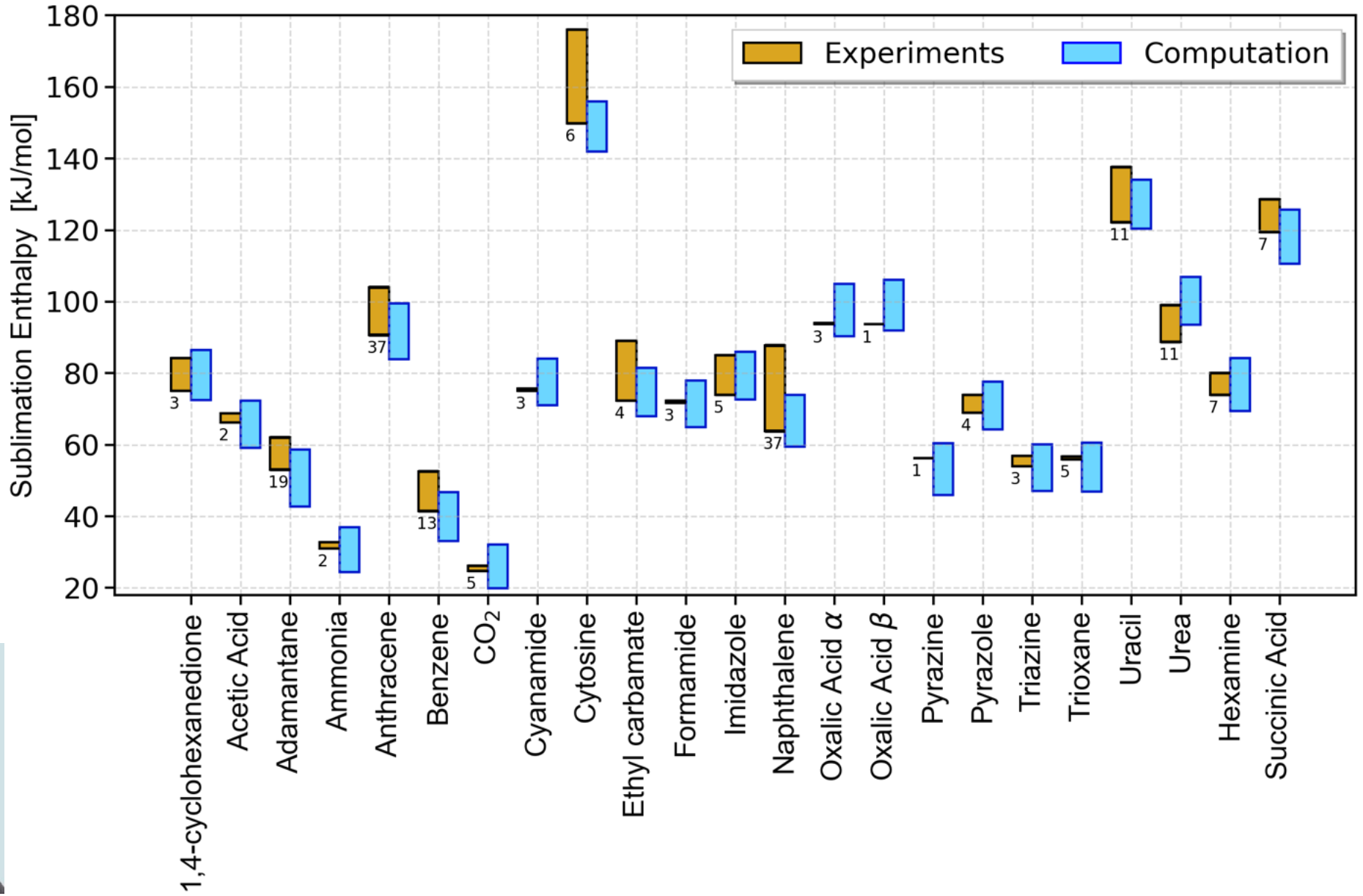
A. Zen, J.G. Brandenburg, J. Klimeš, A. Tkatchenko, D. Alfè, A. Michaelides, PNAS,115, 1724 (2018)

Agreement between DMC and experiments



A. Zen, J.G. Brandenburg, J. Klimeš, A. Tkatchenko, D. Alfè, A. Michaelides, PNAS,115, 1724 (2018)

… but comparison might be challenging



Flaviano Della Pia, Andrea Zen, Dario Alfè, Angelos Michaelides, *in preparation*

DMC in a nutshell

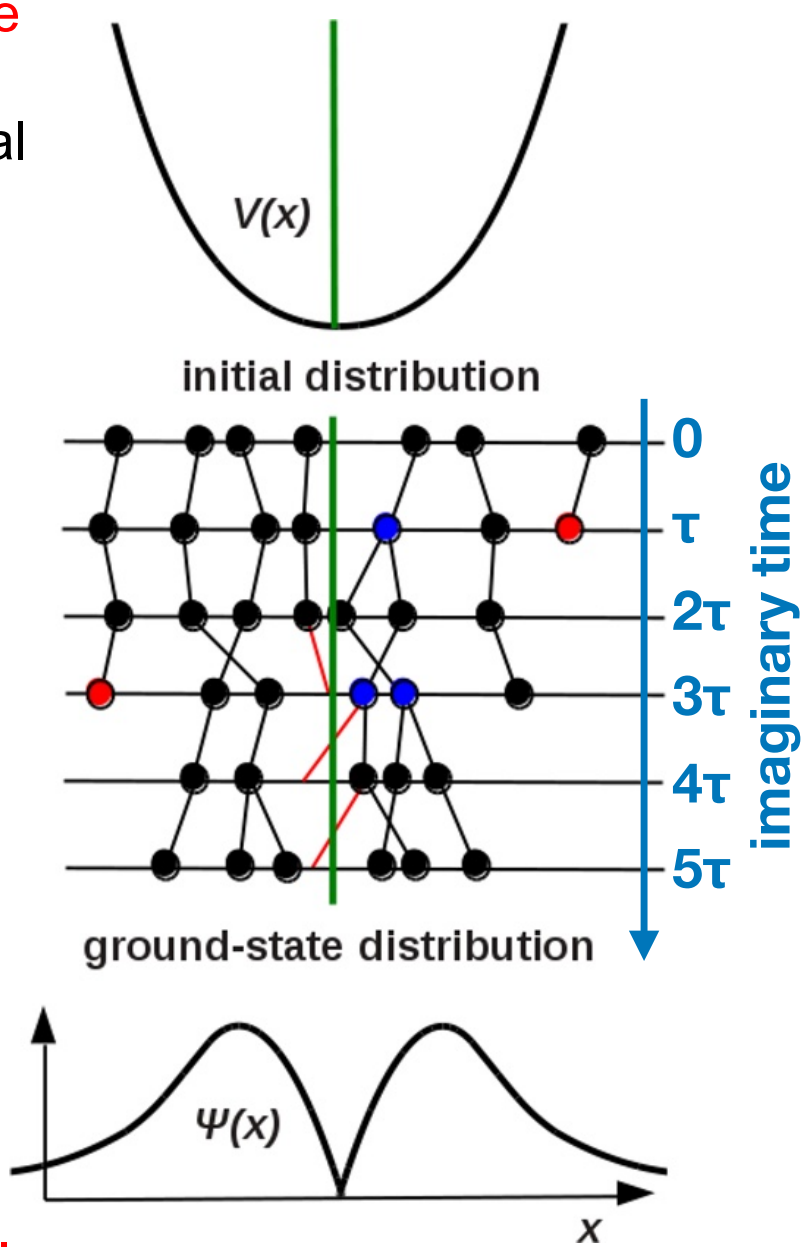
A **propagation** according to the imaginary time Schrodinger equation is performed to **project out the "exact" ground state** $\Phi(\mathbf{R})$ from a trial wave function $\Psi_T(\mathbf{R})$.

- Generate a set of configurations (**walkers**) according to a trial wave function $\Psi_T(\mathbf{R})$
- Propagate in time, with finite **time-step** τ, according to the Green's function (*branching-drift-diffusion process*)
- The set of walkers determines $f(\mathbf{R},t)$, converging to $\Phi(\mathbf{R})\Psi_T(\mathbf{R})$ for large time.
- Fixed node constraint

Two phases in DMC simulation:

- **Equilibration** (project out the exact G.S.)
- **Statistical sampling** (stochastic method, autocorrelation time)

DMC computational **cost** is proportional to **1/τ**.



Chem. Rev., 116, 5188, (2016)

Feature and Approximations in FN-DMC

In traditional FN-DMC the guiding function $\Psi_T(\mathbf{R})$ is a Slater-Jastrow wave function.

Features:
- **Accuracy** (improving DFT, typically comparable to CCSD(T), reference method)
- **Size scaling** (typically N^3, same as DFT but with large prefactor: between 10^3 and 10^4)
- **Ideal for HCP facilities** (Parallel algorithms and limited memory requirements, also GPUs)

Approximations involved:
- **Fixed-node/phase** (fermionic systems)
- **Pseudo-potentials** (non-local terms)
- **Finite time-step τ** (Green's function is known exactly for infinitesimal **τ**)
- **Modified Green's function** (stability)
- **Finite size errors** in periodic systems

Approximations in FN-DMC

How bad are the **approximations** involved?

- **Fixed-node/phase** (fermionic systems)
- **Pseudo-potentials** (non-local terms)

First two are usually not an issue in non-covalent interactions (almost perfect error cancellation).

- **Finite time-step τ** (Green's function is known exactly for infinitesimal **τ**)

Value of time-step τ is crucial:
Trade-off between accuracy and efficiency

- **Modified Green's function** (stability)

If a walker goes close to the nodal surface, its branching weight can diverge. Avoid that!

- **Finite size errors**

Either do an extrapolation to the thermodynamic limit via expensive supercell simulations or rely on correction schemes.

Methodological features for binding energies

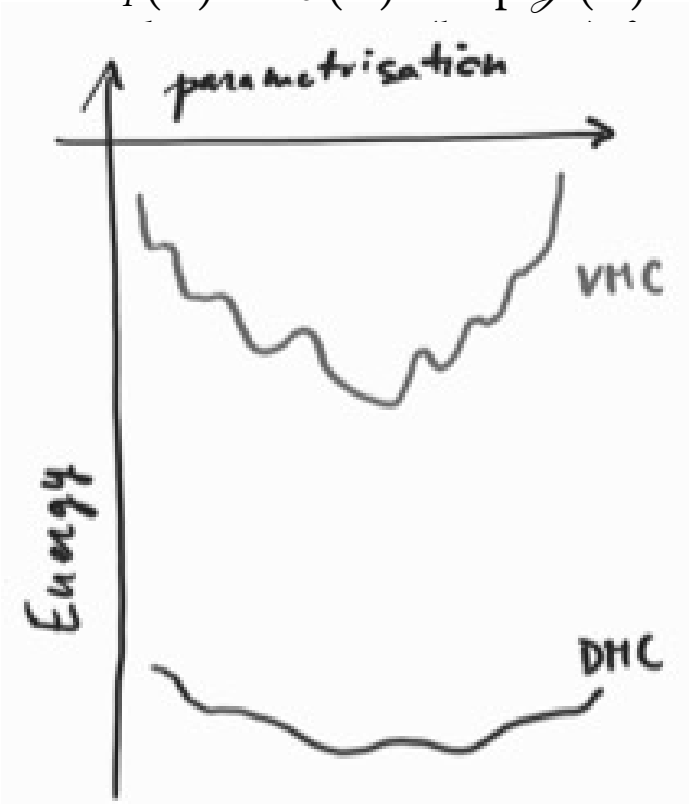
Keep uncertainty small
(both for sampling and for optimisation)
Promote reproducibility & accuracy

Methodological contributions

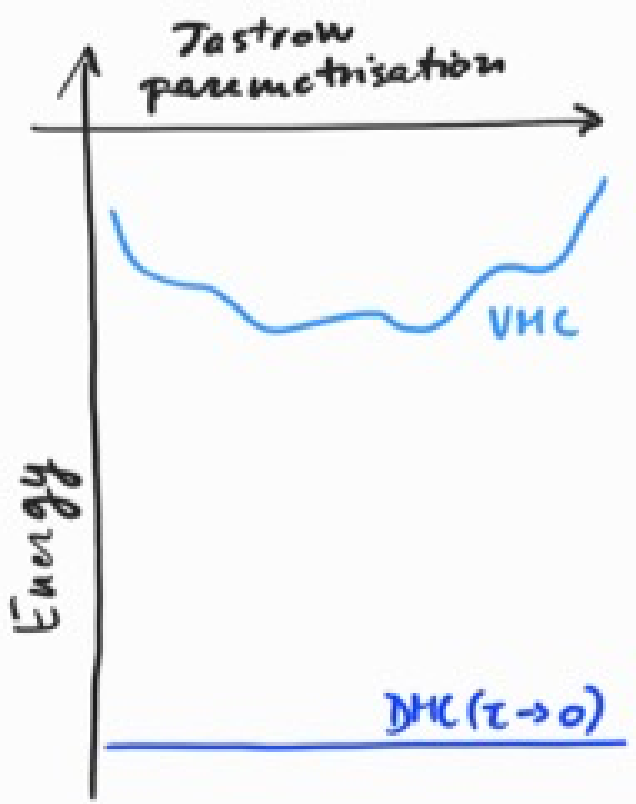
- Enforce **size-consistency** and reduce time-step bias
  A Zen, et al., Phys. Rev. B, **93**, 241118(R) (2016)
- Accurate and fast in **periodic systems**
  A Zen et al., Proc. Natl. Acad. Sci. U.S.A.,**115**, 1724 (2018)
- Improving **reproducibility** and reducing the optimization bias
  A Zen et al., J. Chem. Phys. **151**, 134105 (2019)

Uncertainty in FN-DMC, case AE

$$\Psi_T(\mathbf{R}) = \mathscr{D}(\mathbf{R}) * \exp \mathscr{J}(\mathbf{R})$$



Get the determinant from a deterministic method (HF, DFT, CASSCF, …)



- DMC dependece on parameters is smoother than in VMC
- Many local minima in VMC energy
- No general correspondence between VMC and DMC minima
- Optimese parameters minimising the VMC energy or the VMC variance
- DMC energy satisfies variational principle

- DMC energy, in the limit of infinitisemal timestep, is independent on the Jastrow parametrisation
- DMC energy depends on determinant
- Jastrow parametrisation affects efficiency and timestep dependency

DMC w/o pseudopotentials

	w/o PPs All electrons (AE)	
Hamiltonian	$\hat{H} = \hat{K} + \hat{V}_L$	
Time evolution $f(\mathbf{R}, t) = \Psi_T(\mathbf{R})\psi(\mathbf{R}, t)$	$\frac{\partial}{\partial t} f(\mathbf{R}, t) = \frac{1}{2}\nabla^2 f(\mathbf{R}, t) - \nabla \cdot (\mathbf{V}(\mathbf{R}) f(\mathbf{R}, t)) - [E_L(\mathbf{R}) - E_T] f(\mathbf{R}, t),$	
Green's function $G(\mathbf{R}' \leftarrow \mathbf{R}, \tau) \equiv \frac{\Psi_T(\mathbf{R}')}{\Psi_T(\mathbf{R})} \langle \mathbf{R}' \vert \exp(-\tau \hat{H}) \vert \mathbf{R} \rangle$	Branching drift diffusion $G_{BDD}(\mathbf{R}' \leftarrow \mathbf{R}, \tau) \approx G_B(\mathbf{R}' \leftarrow \mathbf{R}, \tau) G_{DD}(\mathbf{R}' \leftarrow \mathbf{R}, \tau)$	
Needed	1) Fixed node approximation	

DMC w/o and w/ pseudopotentials

	w/o PPs All electrons (AE)	w/ PPs Effective core potentials (ECP)
Hamiltonian	$\hat{H} = \hat{K} + \hat{V}_L$	$\hat{H} = \hat{K} + \hat{V}_L + \hat{V}_{NL}$
Time evolution $f(\mathbf{R}, t) = \Psi_T(\mathbf{R})\psi(\mathbf{R}, t)$	$\frac{\partial}{\partial t} f(\mathbf{R}, t) = \frac{1}{2}\nabla^2 f(\mathbf{R}, t) - \nabla \cdot (\mathbf{V}(\mathbf{R}) f(\mathbf{R}, t)) - [E_L(\mathbf{R}) - E_T] f(\mathbf{R}, t),$	$\frac{\partial}{\partial t} f(\mathbf{R}, t) = \frac{1}{2}\nabla^2 f(\mathbf{R}, t) - \nabla \cdot (\mathbf{V}(\mathbf{R}) f(\mathbf{R}, t)) - \left[\frac{(\hat{K} + \hat{V}_L)\Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})} + \frac{\hat{V}_{NL}\psi(\mathbf{R}, t)}{\psi(\mathbf{R}, t)} - E_T \right] f(\mathbf{R}, t)$
Green's function $G(\mathbf{R}' \leftarrow \mathbf{R}, \tau) \equiv \frac{\Psi_T(\mathbf{R}')}{\Psi_T(\mathbf{R})} \langle \mathbf{R}' \vert \exp(-\tau \hat{H}) \vert \mathbf{R} \rangle$	Branching drift diffusion $G_{BDD}(\mathbf{R}' \leftarrow \mathbf{R}, \tau) \approx G_B(\mathbf{R}' \leftarrow \mathbf{R}, \tau) G_{DD}(\mathbf{R}' \leftarrow \mathbf{R}, \tau)$	$G(\mathbf{R}' \leftarrow \mathbf{R}, \tau) \sim \int T_{NL}(\mathbf{R}' \leftarrow \tilde{\mathbf{R}}, \tau) * G_L(\tilde{\mathbf{R}} \leftarrow \mathbf{R}, \tau) d\tilde{\mathbf{R}}$ $G_L(\mathbf{R}' \leftarrow \mathbf{R}, \tau) \equiv \frac{\Psi_T(\mathbf{R}')}{\Psi_T(\mathbf{R})} \langle \mathbf{R}' \vert e^{-\tau(\hat{K} + \hat{V}_L)} \vert \mathbf{R} \rangle$ $T_{NL}(\mathbf{R}' \leftarrow \mathbf{R}, \tau) \equiv \frac{\Psi_T(\mathbf{R}')}{\Psi_T(\mathbf{R})} \langle \mathbf{R}' \vert e^{-\tau \hat{V}_{NL}} \vert \mathbf{R} \rangle$
Needed	1) Fixed node approximation	1) Fixed node approximation 2) Either the locality approximation (LA), the T-move (TM), or the determinant locality approximation (DLA), to project the NL terms

[LA] L. Mitas, E. L. Shirley, and D. M. Ceperley, J. Chem. Phys. 95, 3467 (1991).

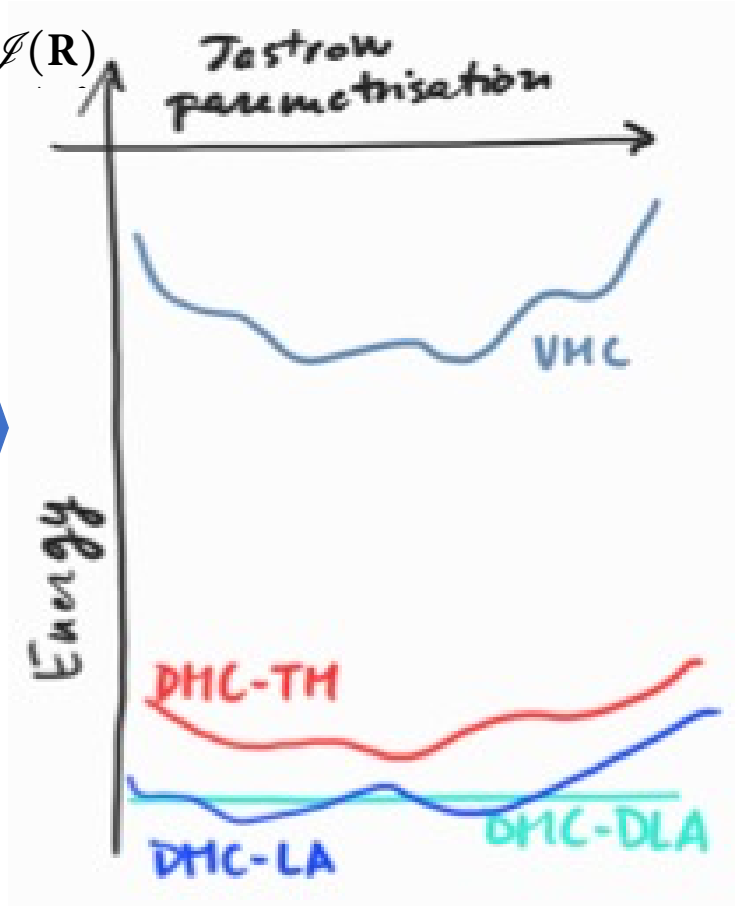
[TM] M. Casula, Phys. Rev. B 74, 161102 (2006); M. Casula, S. Moroni, S. Sorella, and C. Filippi, J. Chem. Phys. 132, 154113 (2010).

[DLA] B. L. Hammond, P. J. Reynolds, and W. A. Lester, J. Chem. Phys. 87, 1130 (1987);
A. Zen, JG Brandenburg, A. Michaelides, D. Alfè, J. Chem. Phys. 151, 134105 (2019)

Uncertainty in FN-DMC, case PPs

$$\Psi_T(\mathbf{R}) = \mathscr{D}(\mathbf{R}) * \exp \mathscr{J}(\mathbf{R})$$

Get the determinant from a deterministic method (HF, DFT, CASSCF, …)

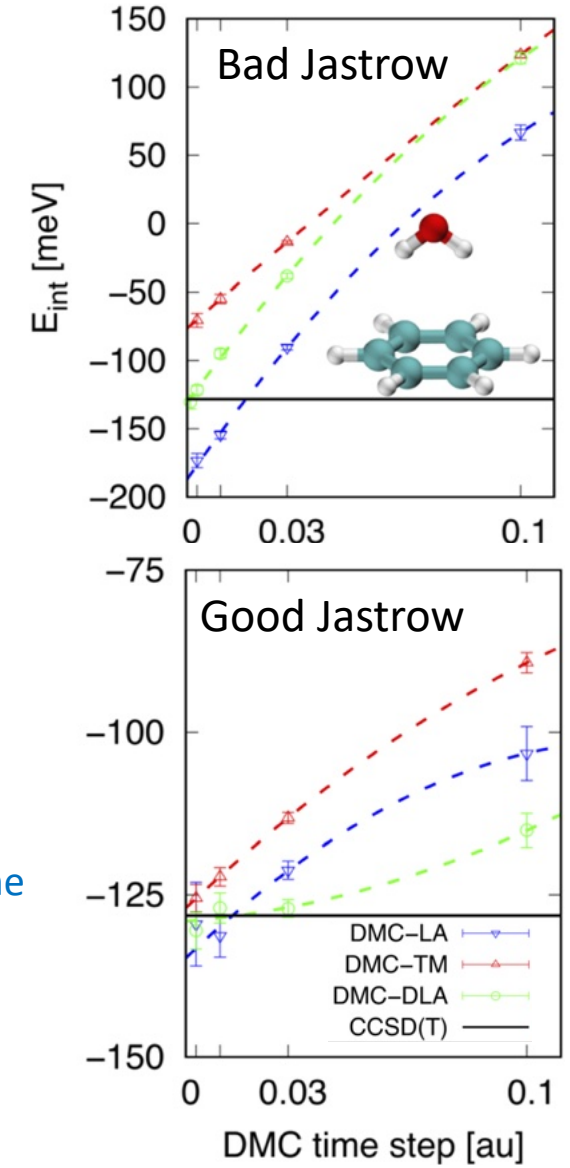


$$\hat{H}_{FN}^{LA} \equiv \hat{K} + \hat{V}_L + \frac{\hat{V}_{NL}\Psi_T}{\Psi_T} + \hat{V}_{FN}$$

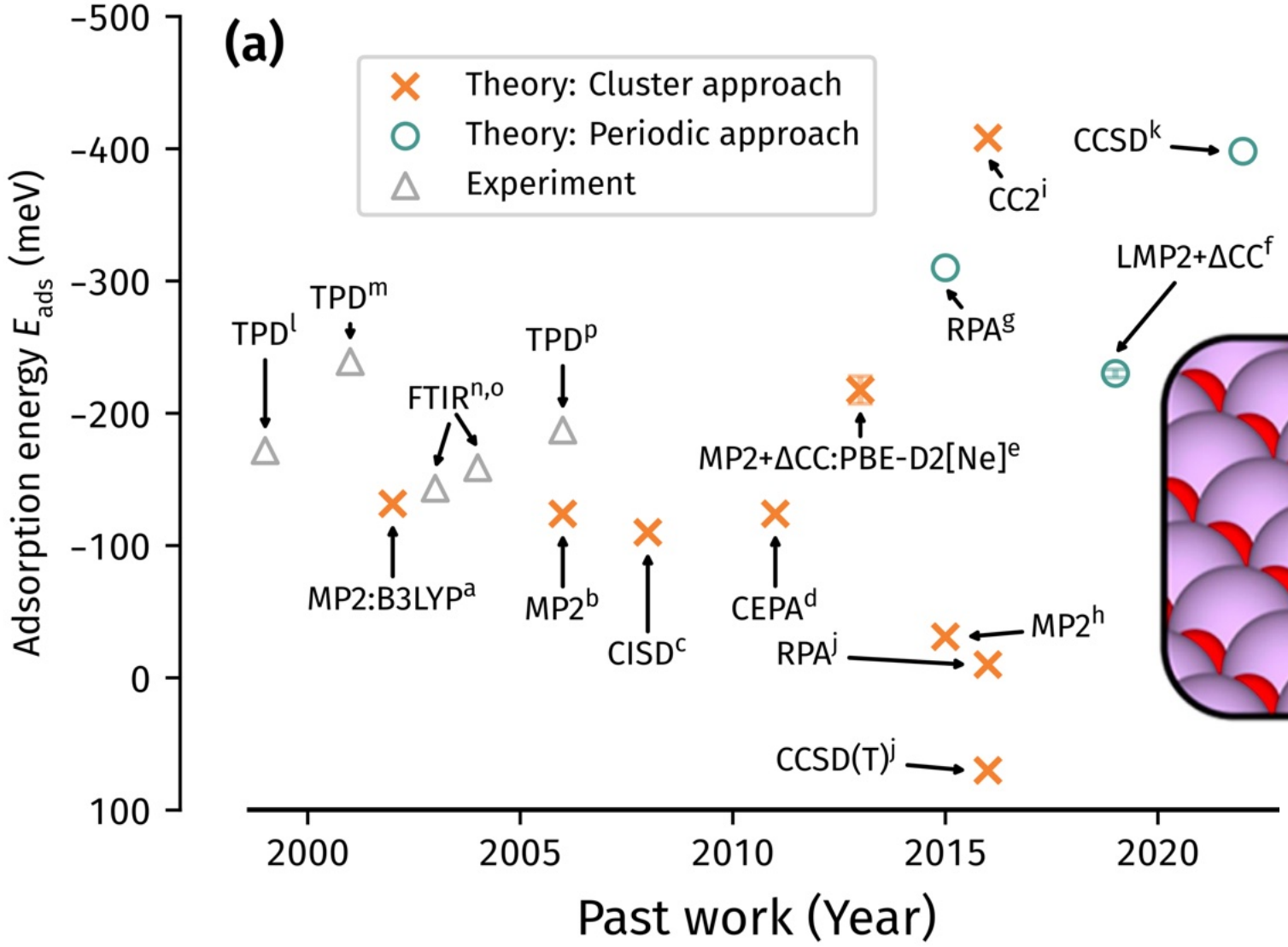
$$\hat{H}_{FN}^{TM} \equiv \hat{K} + \hat{V}_L + \hat{V}_{NL}^{-} + \frac{\hat{V}_{NL}^{+}\Psi_T}{\Psi_T} + \hat{V}_{FN}$$

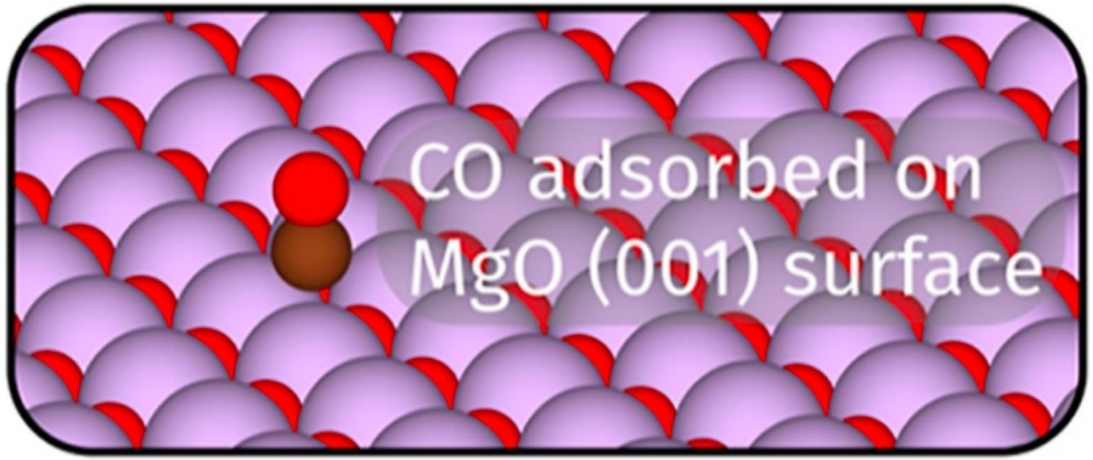
$$\hat{H}_{FN}^{DLA} \equiv \hat{K} + \hat{V}_L + \frac{\hat{V}_{NL}\mathscr{D}}{\mathscr{D}} + \hat{V}_{FN}$$

- Both DMC-LA and DMC-TM depends on the parametrisation, thus on optimisation
- Reproducibility issue in TM and LA among codes implementing different Jastrow
- DMC-DLA is flat, as in AE case: no dependence on optimisation or parametrization; reproducibility



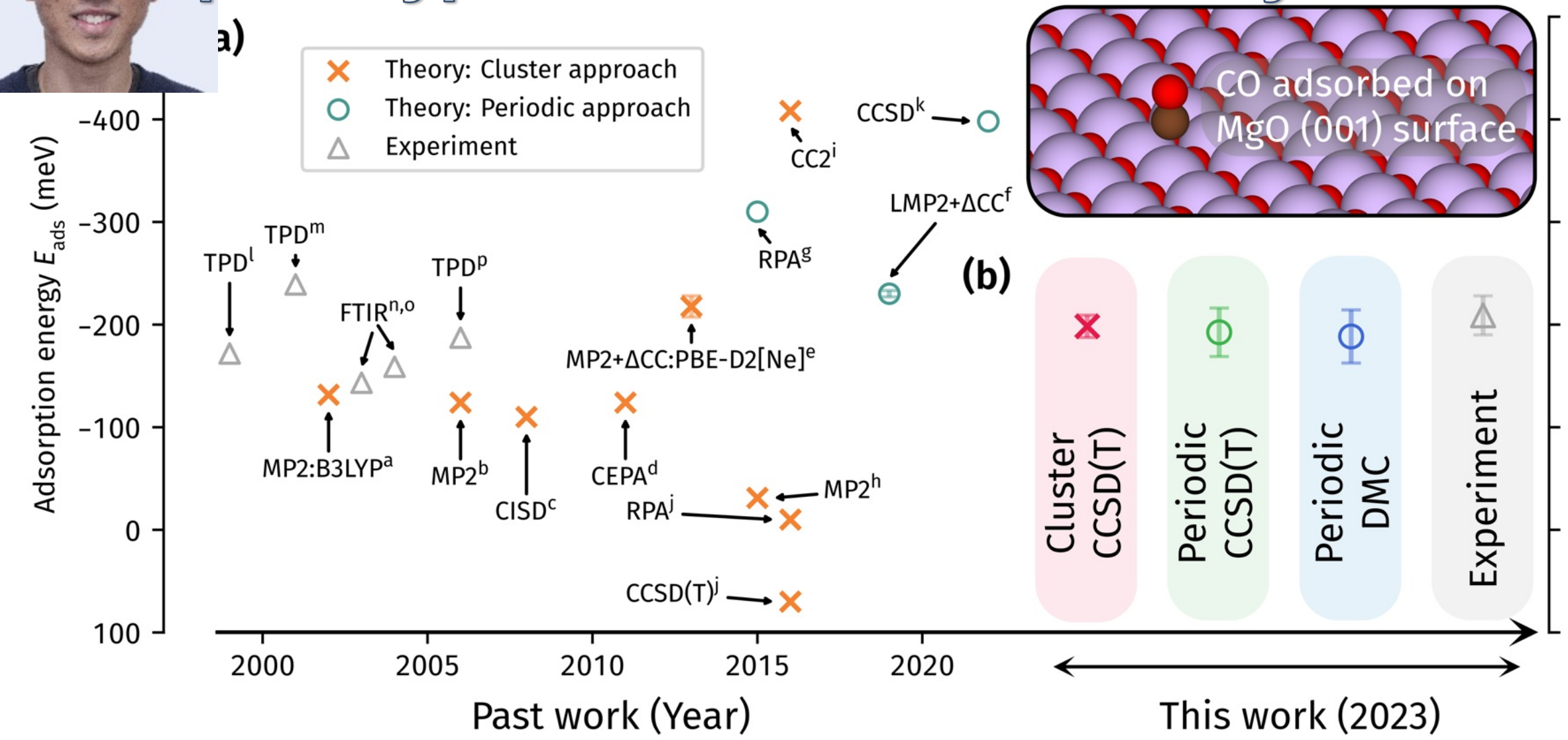
Accuracy in surface chemistry: a prototypical molecule-surface system





Benjamin Shi, A. Zen, V. Kapil, P.R. Nagy, A. Grüneis, A. Michaelides, JACS **145**, 25372–25381 (2023).

Accuracy in surface chemistry: a prototypical molecule-surface system



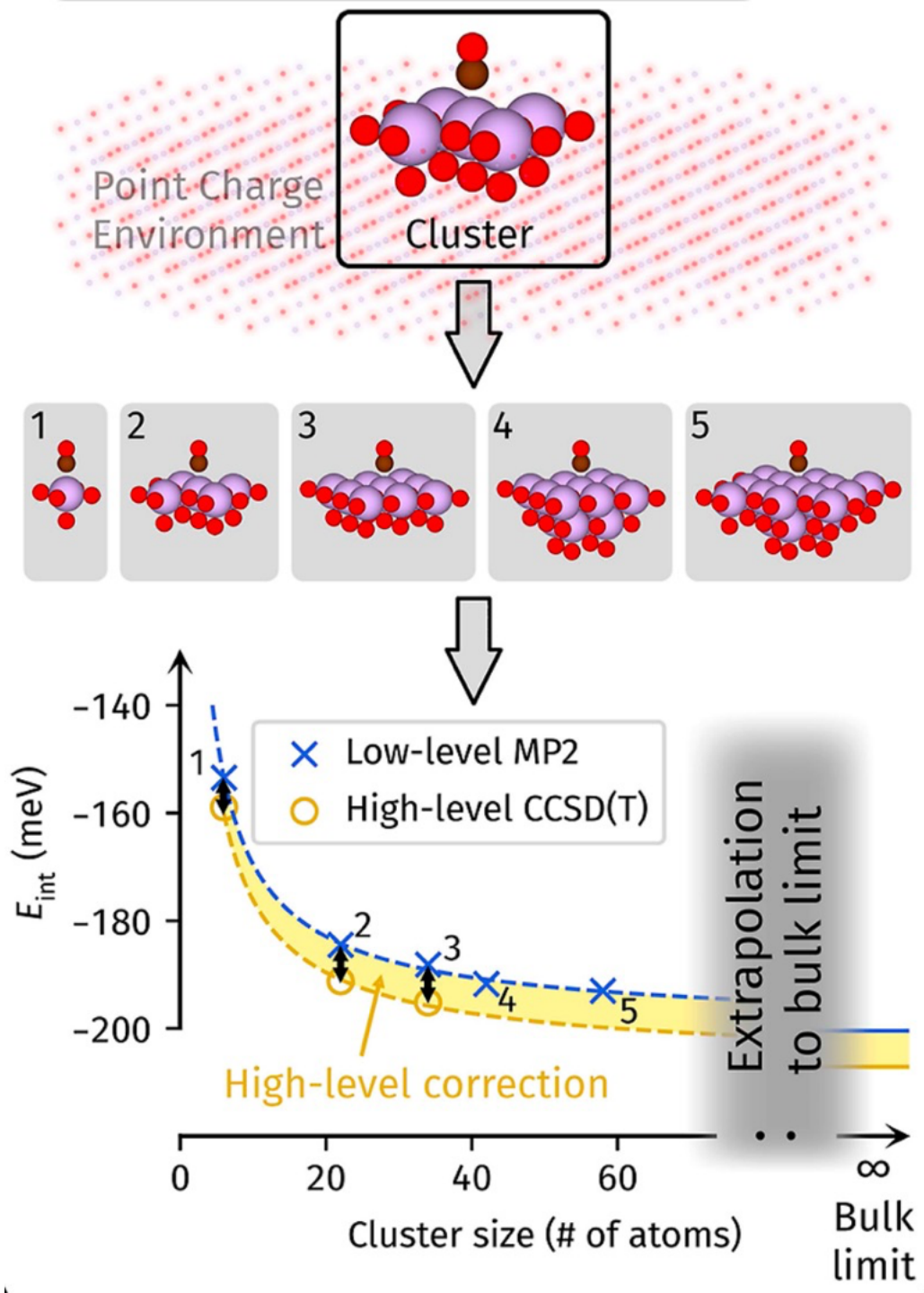
Benjamin Shi, A. Zen, V. Kapil, P.R. Nagy, A. Grüneis, A. Michaelides, JACS **145**, 25372–25381 (2023).

Cluster CCSD(T): SKZCAM approach

General embedded cluster protocol for accurate modeling of oxygen vacancies in metal-oxides

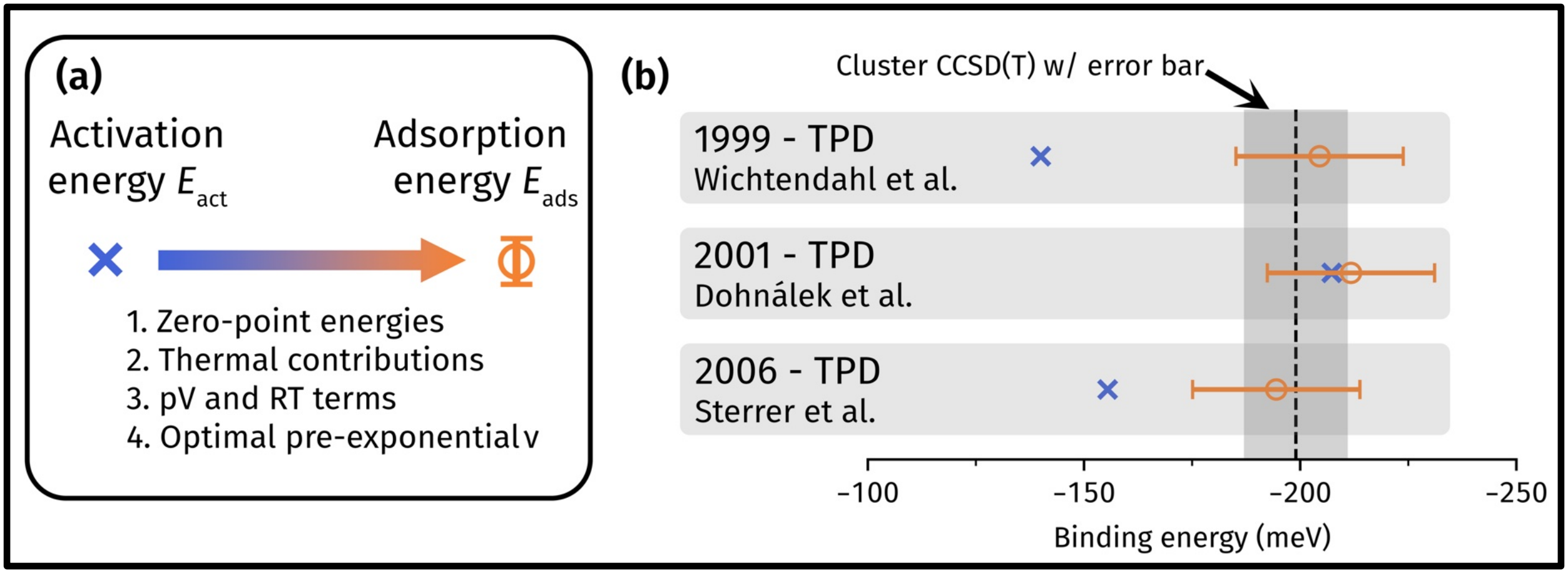
Benjamin X. Shi, Venkat Kapil, Andrea Zen, Ji Chen, Ali Alavi and Angelos Michaelides

J. Chem. Phys. **156**, 124704 (2022);
https://doi.org/10.1063/5.0087031 Open URL



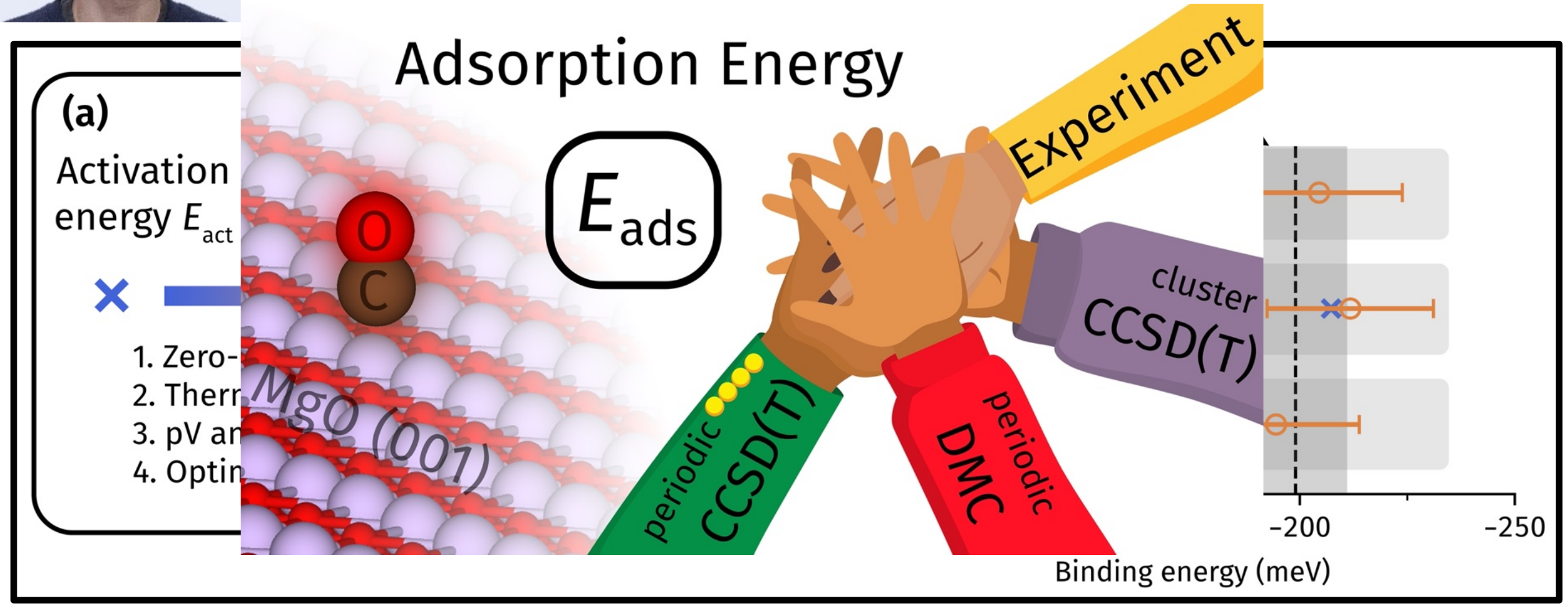
Benjamin Shi, A. Zen, V. Kapil, P.R. Nagy, A. Grüneis, A. Michaelides, JACS **145**, 25372–25381 (2023).

Accuracy in surface chemistry: a prototypical molecule-surface system



Benjamin Shi, A. Zen, V. Kapil, P.R. Nagy, A. Grüneis, A. Michaelides, JACS **145**, 25372–25381 (2023).

Accuracy in surface chemistry: a prototypical molecule-surface system

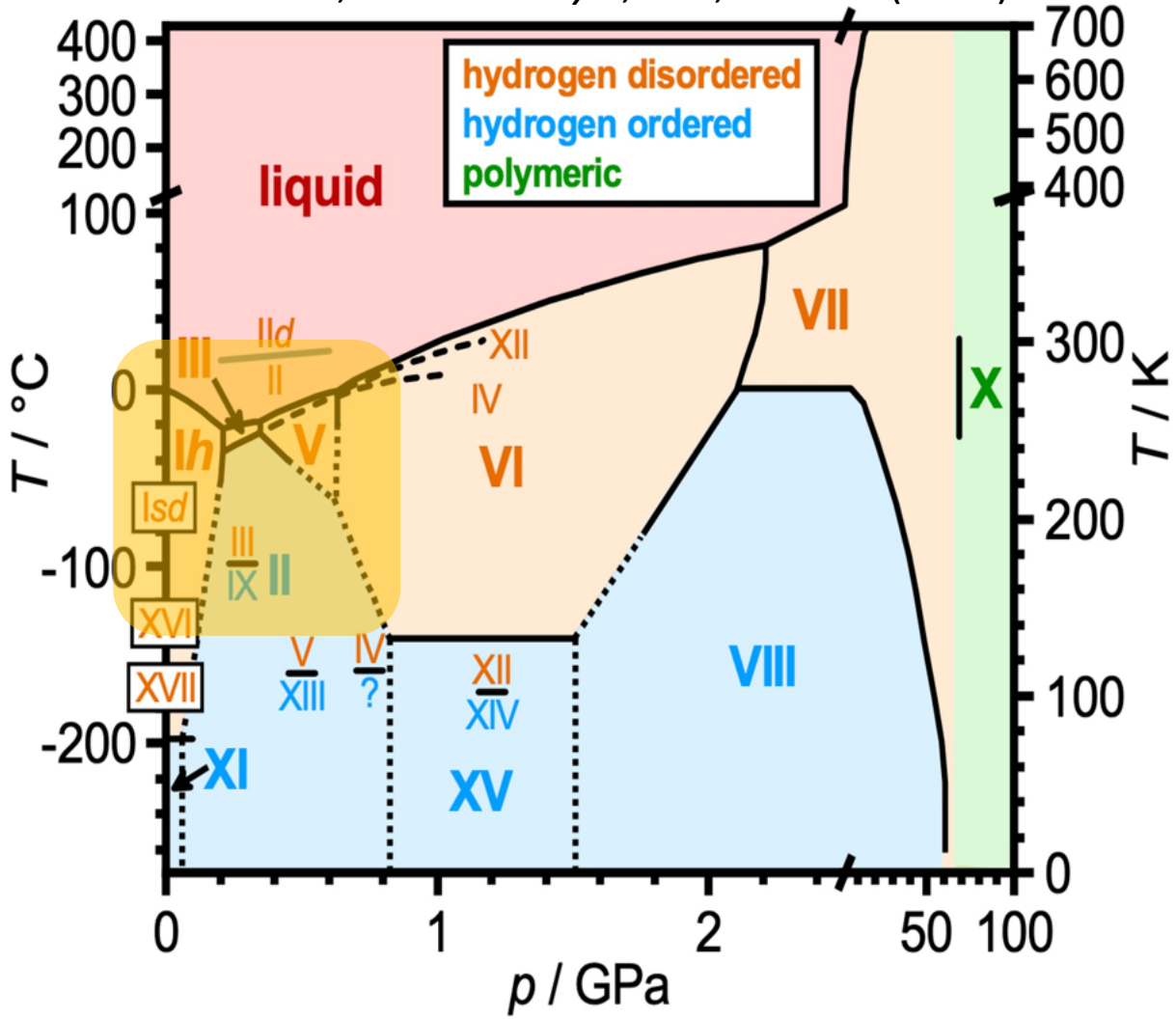


Benjamin Shi, A. Zen, V. Kapil, P.R. Nagy, A. Grüneis, A. Michaelides, JACS **145**, 25372–25381 (2023).

Accuracy is a challenge: ice phase diagram

Experiment

C. Salzmann, *J. Chem. Phys.*, **150**, 060901 (2019)



Theory

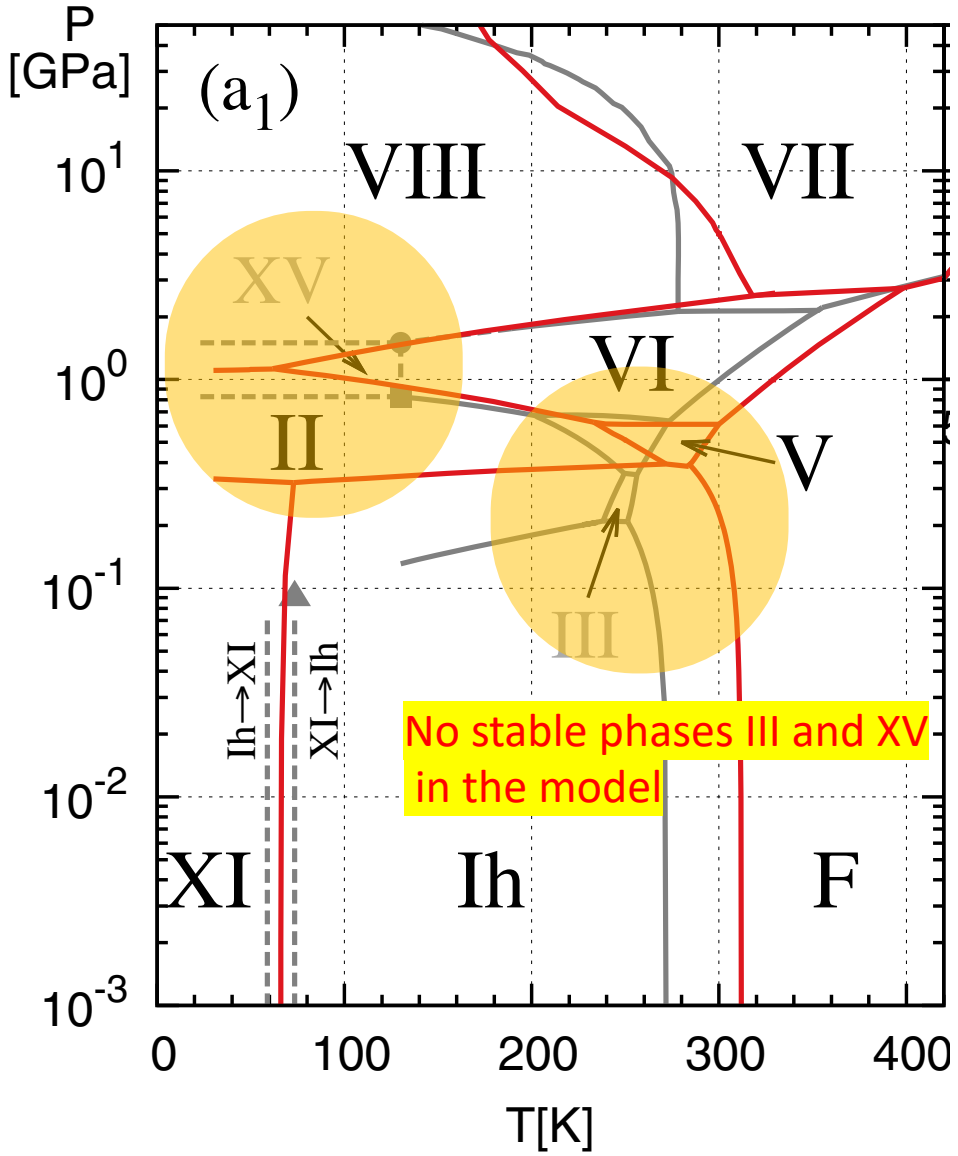
DFT + machine learning

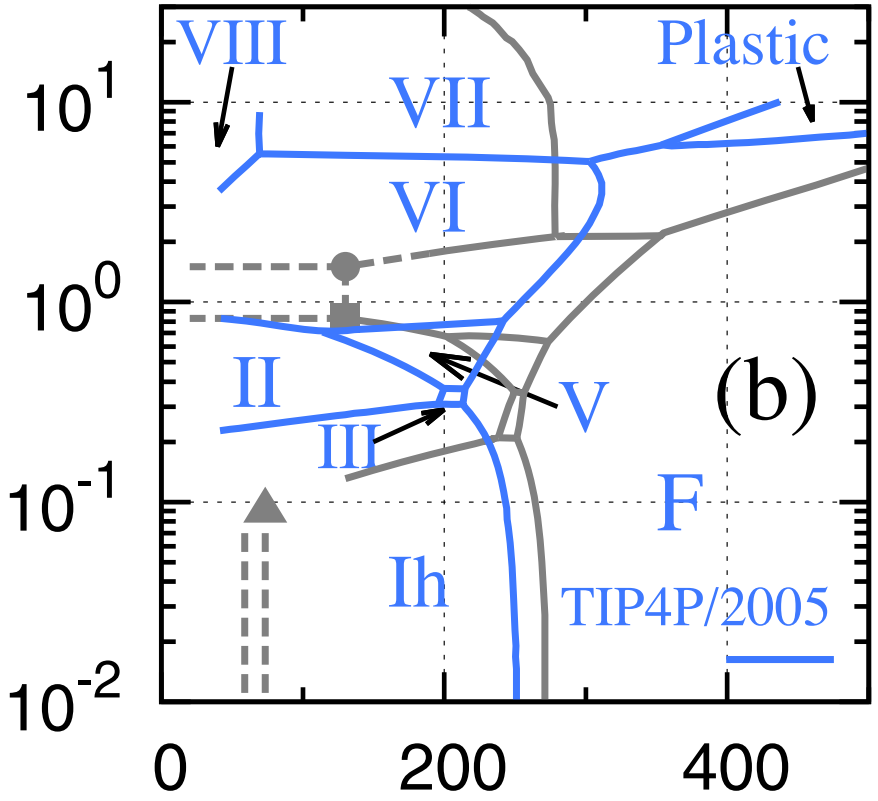
L. Zhang *et al., Phys. Rev. Lett.* **126**, 236001 (2021)

A. Reinhardt and B. Cheng, *Nat. Communication,* **12**, 588 (2021)

Ice phase diagram from DFT (1/2)

L. Zhang *et al., Phys. Rev. Lett.* **126**, 236001 (2021)





Gray: Experiments

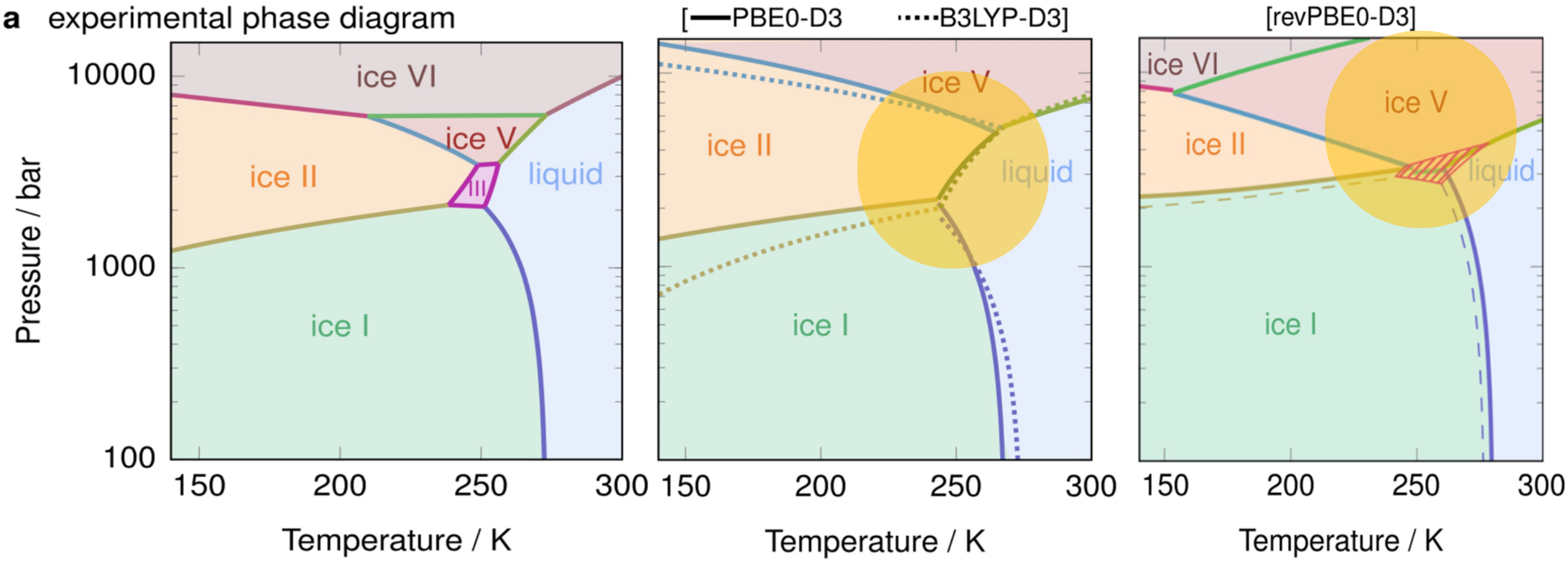
Red: ab-initio DFT (SCAN) + machine learning

Blue: semiempirical force field (TIP4P/2005)

Ice phase diagram from DFT (2/2)

A. Reinhardt and B. Cheng, *Nat. Communication,* **12**, 588 (2021)

No stable phases III in the model



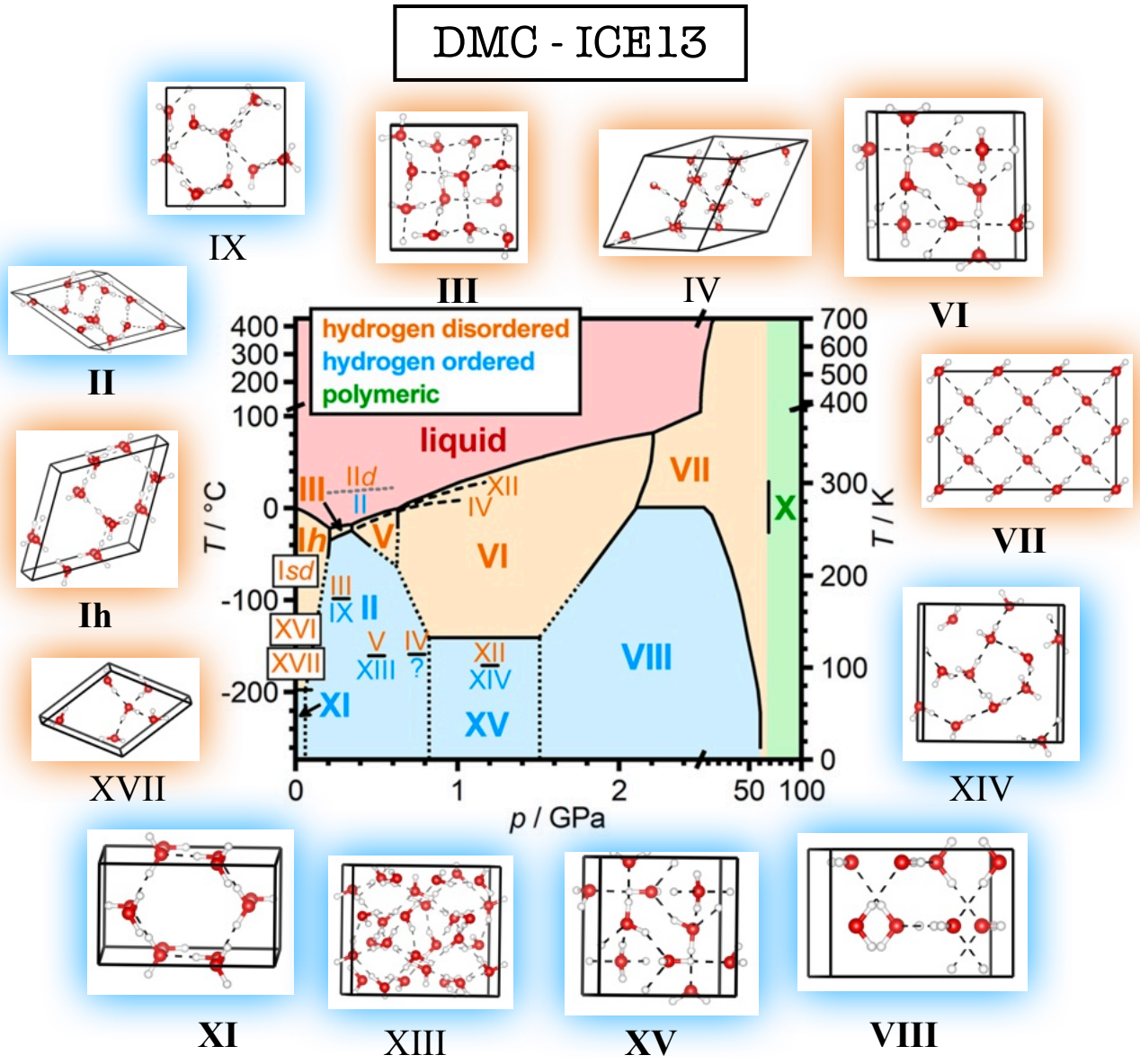
Stability of ice polymorphs from DMC

Flaviano Della Pia, Andrea Zen, Dario Alfè, Angelos Michaelides, J. Chem. Phys. 157, 134701 (2022)

Stability of ice polymorphs from DMC

1) Constructing the dataset (**DMC-ICE13**)

- 7 hydrogen-ordered phases
- 6 hydrogen-disordered phases
- 13 polymorphs covering a broad part of the phase diagram



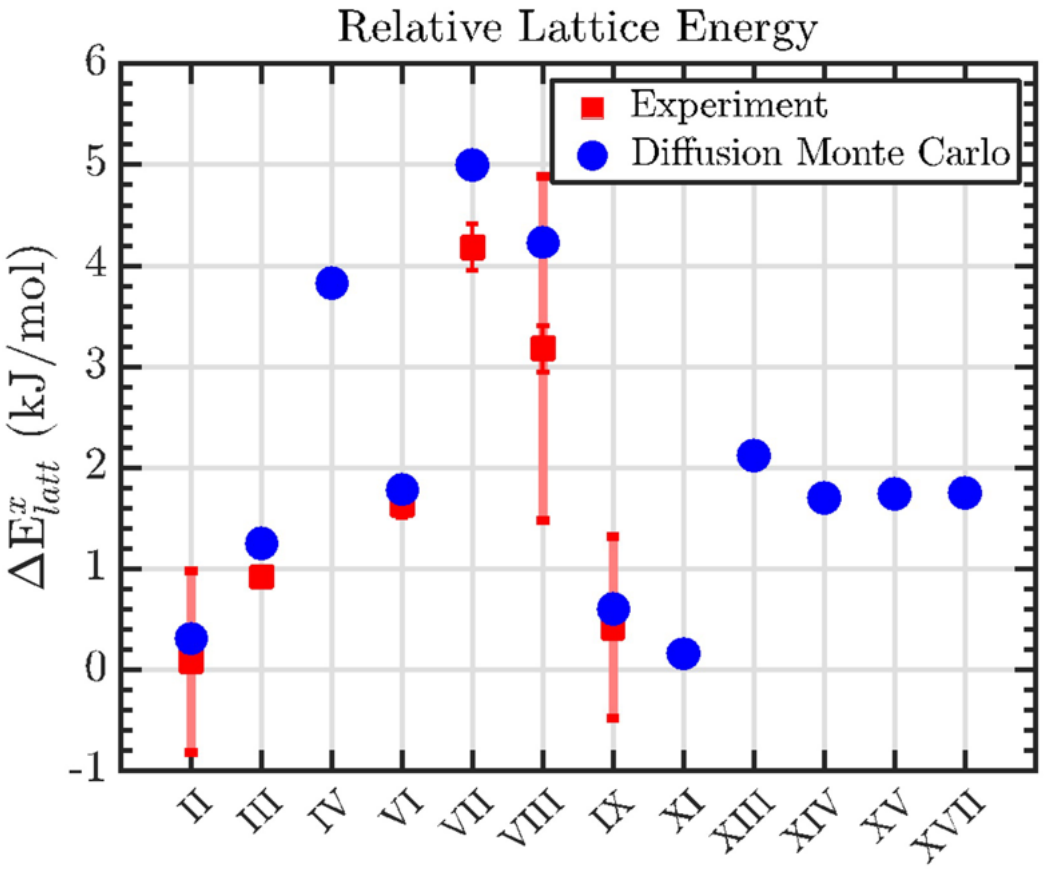
Flaviano Della Pia, Andrea Zen, Dario Alfè, Angelos Michaelides, J. Chem. Phys. 157, 134701 (2022)

Stability of ice polymorphs from DMC

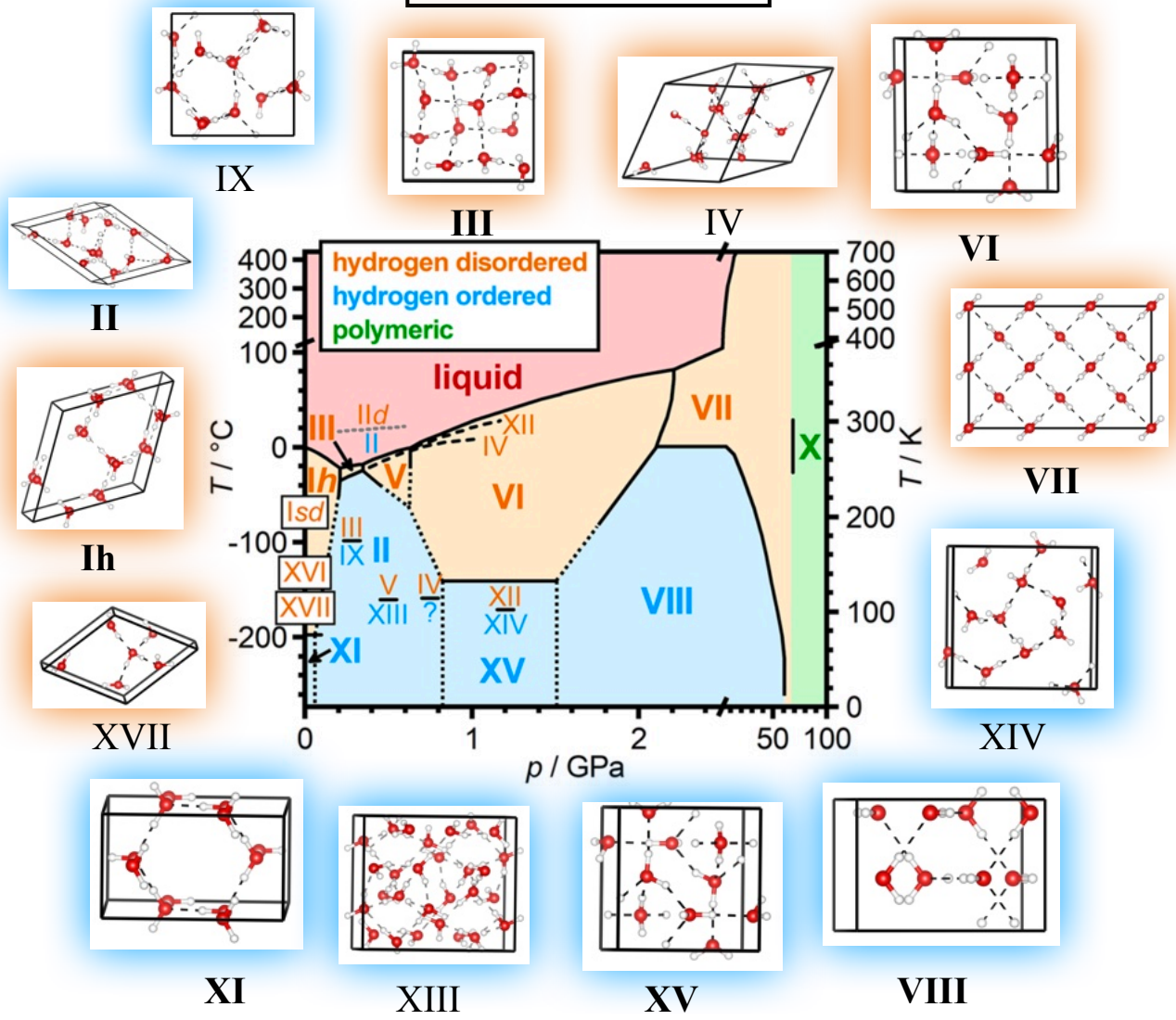
1) Constructing the dataset (**DMC-ICE13**)

2) Compare with experiments (when available)

1 E. Whalley, *J. Chem. Phys*. **81**, 4087 (1984);



DMC - ICE13



Flaviano Della Pia, Andrea Zen, Dario Alfè, Angelos Michaelides, J. Chem. Phys. 157, 134701 (2022)

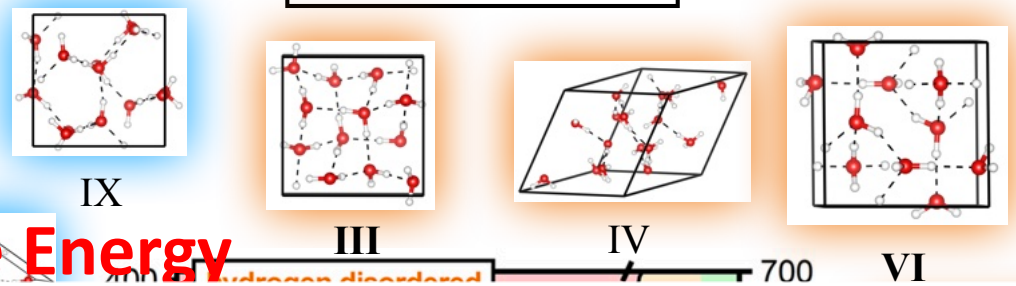
Stability of ice polymorphs from DMC

1) Constructing the dataset (**DMC-ICE13**)

2) Compare with experiments (when available)
   1 E. Whalley, *J. Chem. Phys.* **81**, 4087 (1984);

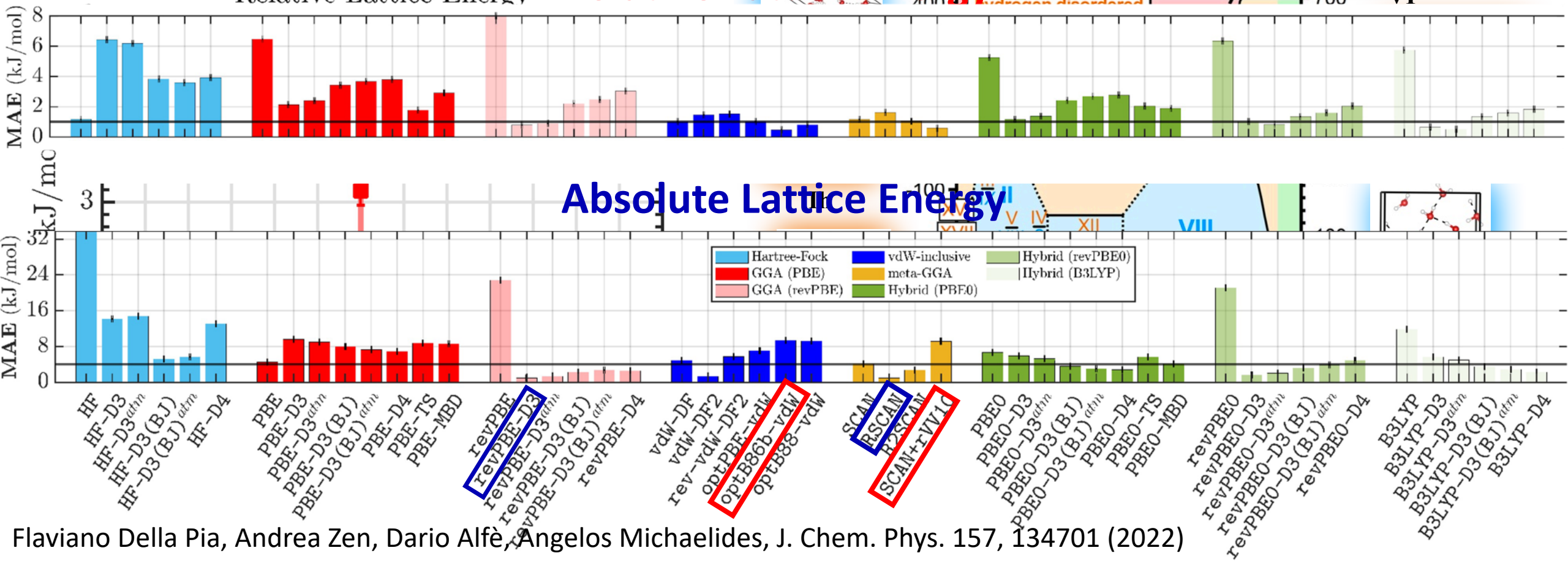
3) Benchmark DFT functionals and vdW corrections

DMC - ICE13



Relative Lattice Energy

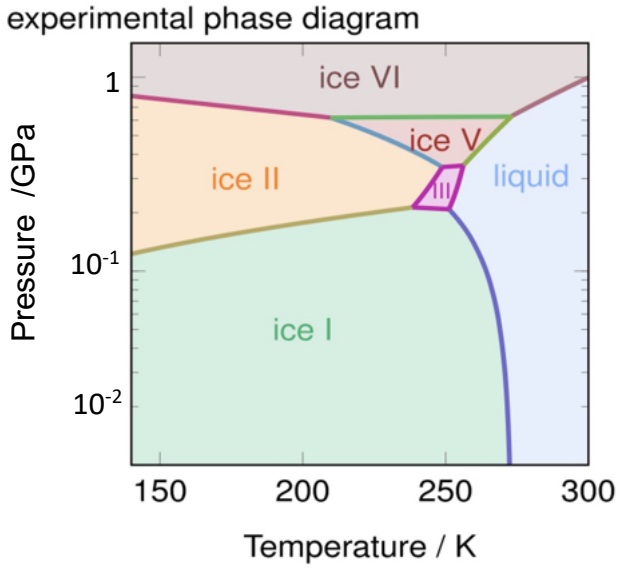
Absolute Lattice Energy



Flaviano Della Pia, Andrea Zen, Dario Alfè, Angelos Michaelides, J. Chem. Phys. 157, 134701 (2022)

Stability of ice polymorphs from DMC

DMC-ICE13 allows for a qualitative prediction of the V-III transition pressure



Approximation

- Ice XIII instead of disordered ice V;
- Neglecting temperature contribution, ZPE, and considering the zero pressure volumes;

$$p_{tr}^{DMC} = -\frac{\Delta E_{XIII-III}^{DMC}}{\Delta V_{XIII-III}^{0}} \sim 0.48 \text{ GPa}$$

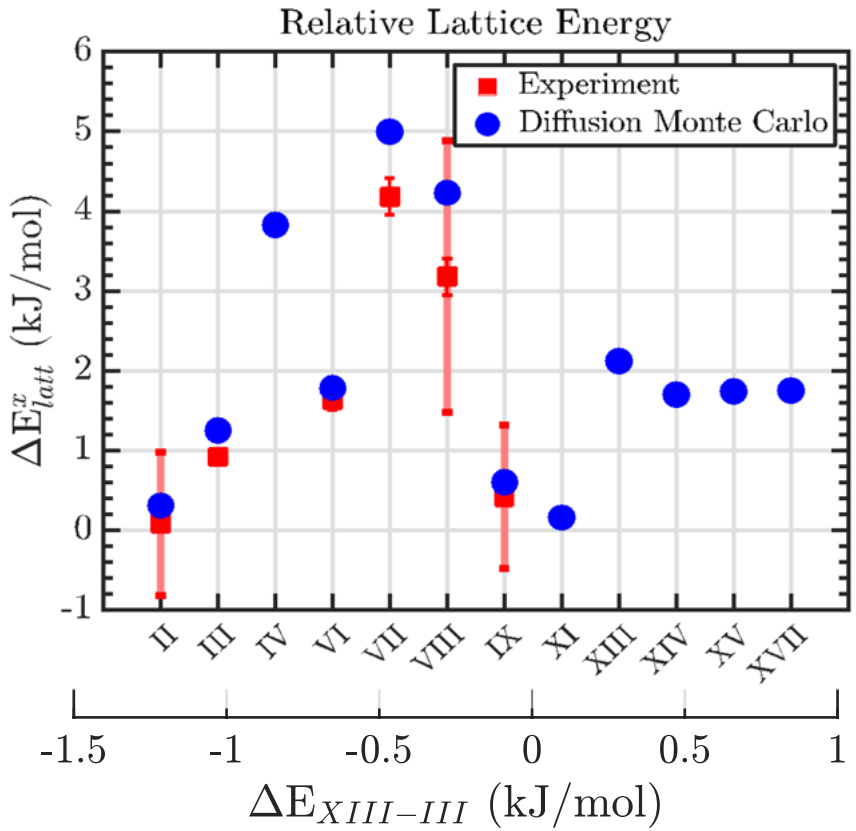
- The condition that a functional needs to satisfy to predict the stability of ice III is

$$\Delta E_{XIII-III}^{XC} \sim \Delta E_{XIII-III}^{DMC}$$

B3L

revPBE

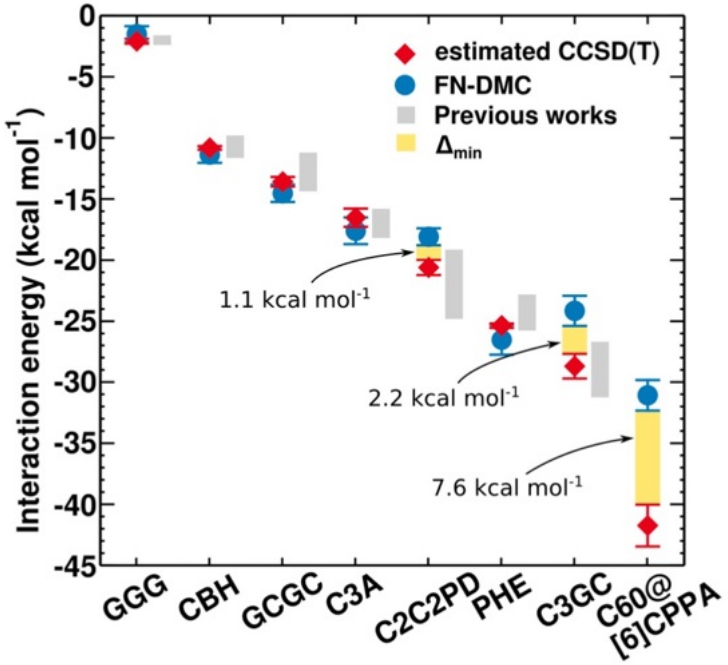
PBE



Flaviano Della Pia, Andrea Zen, Dario Alfè, Angelos Michaelides, J. Chem. Phys. 157, 134701 (2022)

Comparing FN-DMC & CCSD(T)

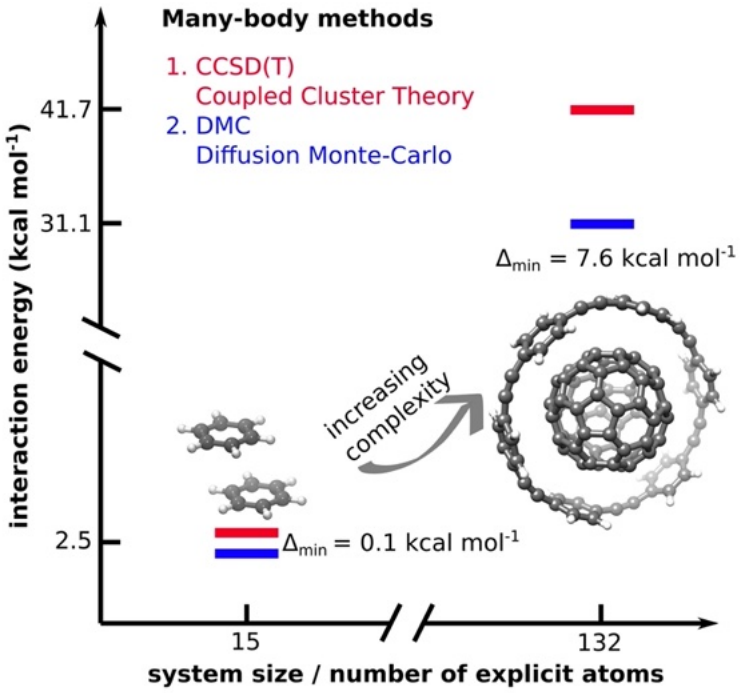
- Generally observed a **good agreement between** the **CCSD(T)** and the **FNDMC** (with a **Slater-Jastrow** guide function) evaluation of **non-covalent interactions.**
- Disagreement, when observed, could be explained from know issues, such as small basis in CCSD(T) or optimization / timestep / size-consistency issues in FN-DMC.
- Recently observed a **disagreement in large complexes** not coming from the above issues






Y.S. Al-Hamdani, P.R. Nagy, A. Zen, D. Barton, M. Kállay, J.G. Brandenburg, A. Tkatchenko, *Interactions between Large Molecules: Puzzle for Reference Quantum-Mechanical Methods*, Nature Communications 12, 3927 (2021).

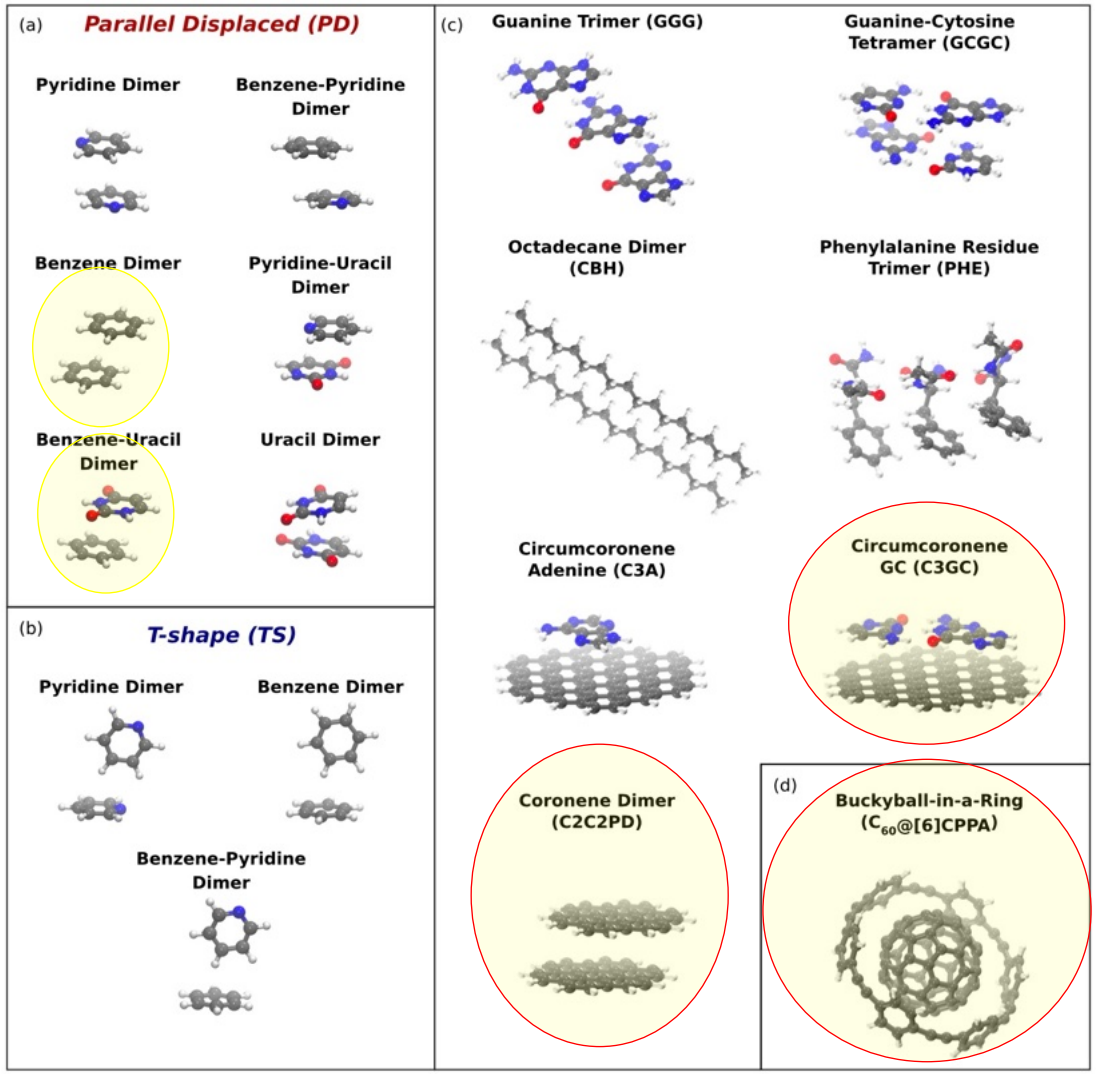
Discrepancies in large molecules between CCSD(T) and FN-DMC

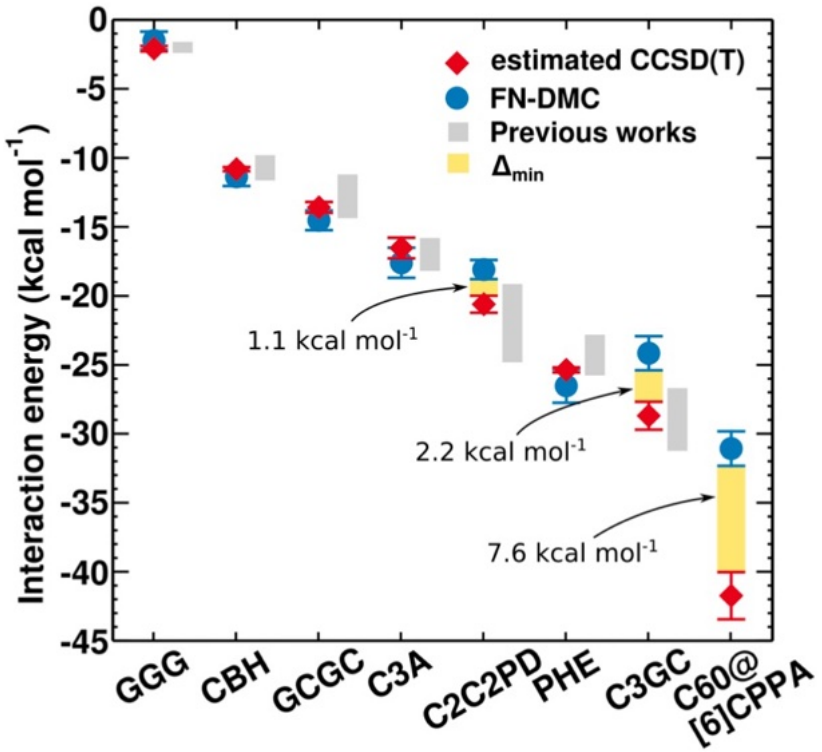


Complex	No. of atoms	CCSD(T)	FN-DMC	Δ_{min}	
pyridine-pyridine PD (No. 25)	22	-3.70 ± 0.08	-3.51 ± 0.20	0.0	indistinguishable
pyridine-pyridine TS (No. 48)	22	-3.48 ± 0.06	-3.44 ± 0.20	0.0	indistinguishable
benzene-pyridine PD (No. 27)	23	-3.28 ± 0.07	-3.03 ± 0.16	0.0	indistinguishable
benzene-pyridine TS (No. 49)	23	-3.24 ± 0.05	-3.08 ± 0.16	0.0	indistinguishable
pyridine-uracil PD (No. 29)	23	-6.61 ± 0.09	-6.38 ± 0.18	0.0	indistinguishable
benzene-benzene PD (No. 24)	24	-2.67 ± 0.07	-2.38 ± 0.12	0.1	consistent
benzene-benzene TS (No. 47)	24	-2.81 ± 0.06	-2.71 ± 0.12	0.0	indistinguishable
uracil-uracil PD (No. 26)	24	-9.61 ± 0.10	-9.40 ± 0.16	0.0	indistinguishable
benzene-uracil PD (No. 28)	24	-5.48 ± 0.11	-5.11 ± 0.18	0.1	consistent
GGG	48	-2.1 ± 0.2	-1.5 ± 0.6	0.0	indistinguishable
CBH	112	-11.0 ± 0.2	-11.4 ± 0.8	0.0	indistinguishable
GCGC	58	-13.6 ± 0.4	-12.4 ± 0.8	0.1	consistent
C3A	87	-16.5 ± 0.8	-15.0 ± 1.0	0.0	indistinguishable
C2C2PD	72	-20.6 ± 0.6	-18.1 ± 0.8	1.1	inconsistent
PHE	87	-25.4 ± 0.2	-26.5 ± 1.3	0.0	indistinguishable
C3GC	101	-28.7 ± 1.0	-24.2 ± 1.3	2.2	inconsistent
C_{60}@[6]CPPA	132	-41.7 ± 1.7	-31.1 ± 1.4	7.6	inconsistent

Y.S. Al-Hamdani, P.R. Nagy, A. Zen, D. Barton, M. Kállay, J.G. Brandenburg, A. Tkatchenko, *Interactions between Large Molecules: Puzzle for Reference Quantum-Mechanical Methods*, Nature Communications 12, 3927 (2021).

Issue with pi-pi interactions?



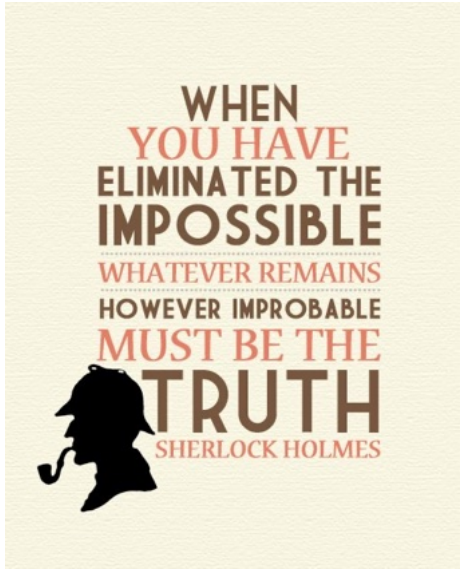


There is an inconsistency between CCSD(T) and FN-DMC in some systems with pi-pi interactions

Is FN-DMC or CCSD(T) having accuracy issues?

Y.S. Al-Hamdani, P.R. Nagy, A. Zen, D. Barton, M. Kállay, J.G. Brandenburg, A. Tkatchenko, *Interactions between Large Molecules: Puzzle for Reference Quantum-Mechanical Methods*, Nature Communications 12, 3927 (2021).

Inspecting FN-DMC weaknesses



Weaknesses in FN-DMC:

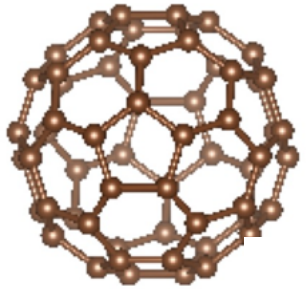
- ~~Bugs in the code~~ *No, 2 codes agree* [1,2]
- ~~Pseudopotentials~~ *No, AE and PP agree* [1,2]
- ~~Optimization of Jastrow~~ *No: tested LA, TM & DLA* [1]
- ~~Determinant initialization~~ *No: tested LDA, PBE, PBE0* [1]
- FN beyond single Slater (?) ***Work in progress***

[1] Y.S. Al-Hamdani, P.R. Nagy, A. Zen, D. Barton, M. Kállay, J.G. Brandenburg, A. Tkatchenko, *Interactions between Large Molecules: Puzzle for Reference Quantum-Mechanical Methods*, Nature Communications 12, 3927 (2021).
[2] A. Benali, H. Shin and O. Heinonen, Quantum Monte Carlo benchmarking of large noncovalent complexes in the L7 benchmark set, JCP 153, 194113 (2020).

[1] uses CASINO, DMC with PPs testing LA/TM/DLA; [2] uses QMCPACK, DMC with all-electrons

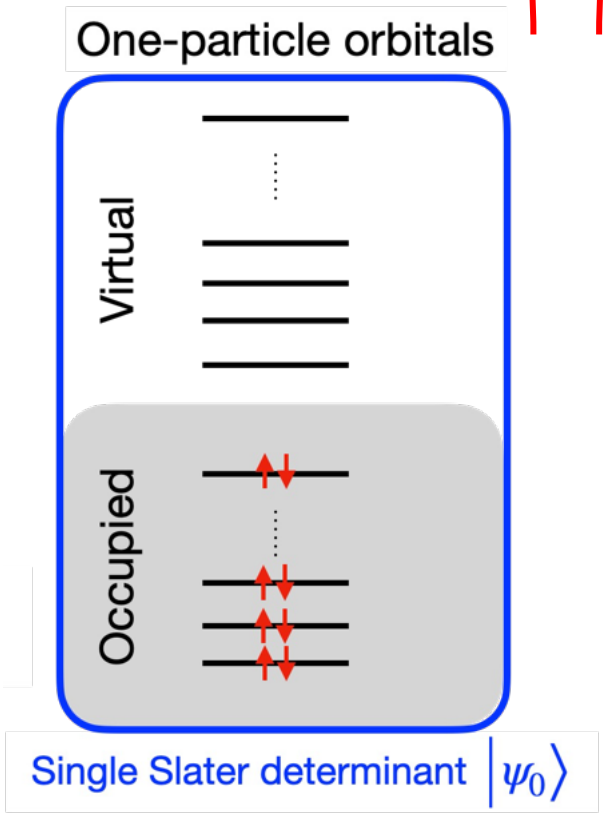
Slater-Jastrow wave function

$$\Psi_T(\mathbf{R}) = \mathscr{D}(\mathbf{R}) * \exp \mathscr{J}(\mathbf{R})$$



Schrödinger Equation
$\hat{H}\Psi = E\Psi$

approximate solution
using Hartree-Fock (HF),
density functional theory (DFT),



Jastrow factor
electron-electron and
electron-nucleus terms

captures dynamical correlation

Parameters are optimised
to minimise the
energy (variational principle)
or the **variance**

The Jastrow does **not** change
the nodal surface

FN-DMC beyond the single Slater-Jastrow w.f.
How to make it possible with TurboRVB!

Wave functions implemented in TurboRVB

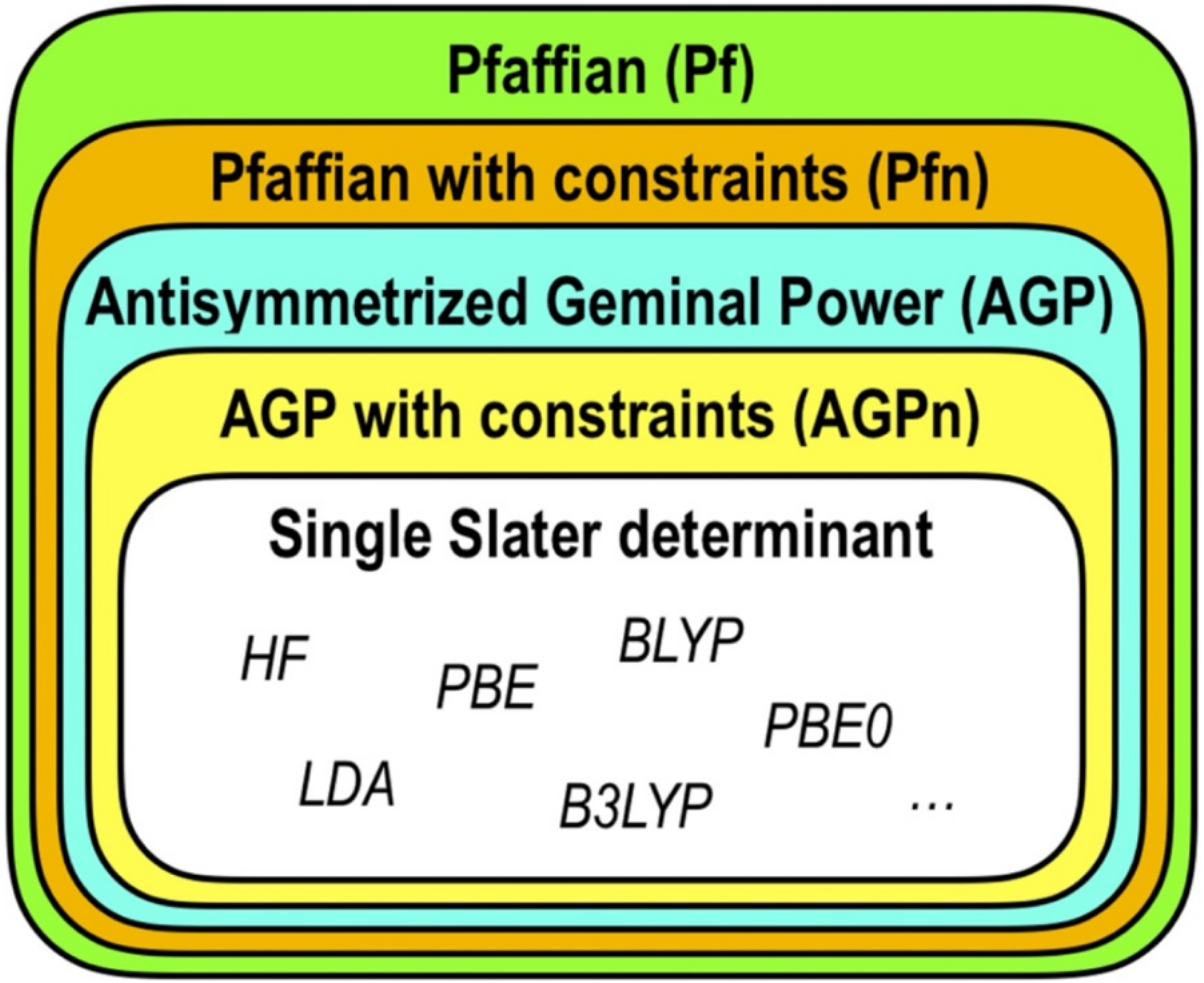


FIG. 3. *Ansatz* hierarchy. The output of Hartree–Fock (HF) or DFT simulations with different exchange-correlation functionals are special instances of the SD *Ansatz*.

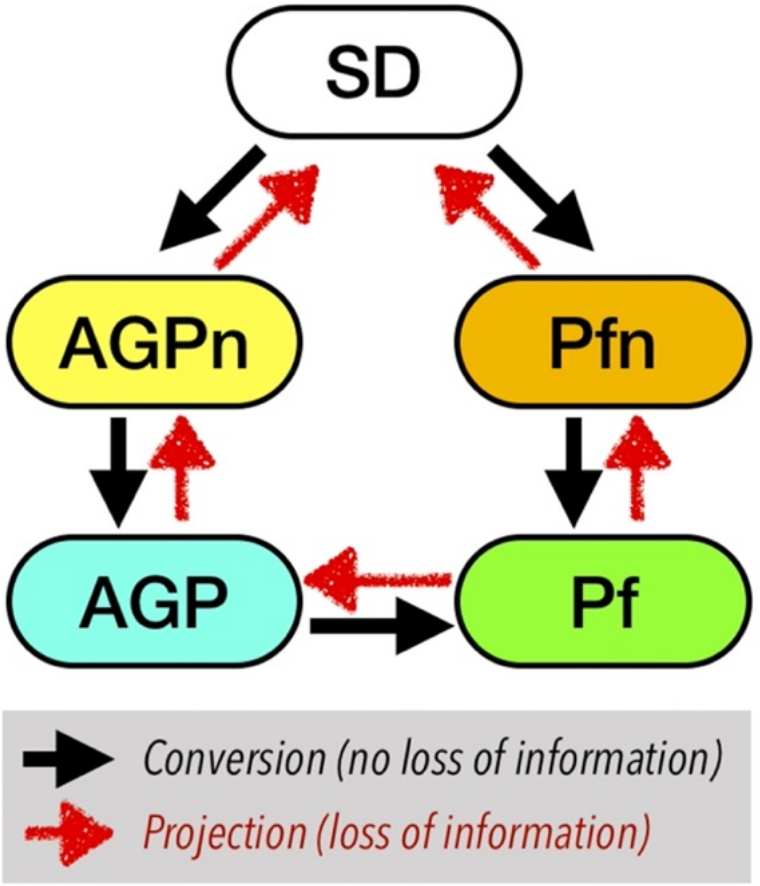
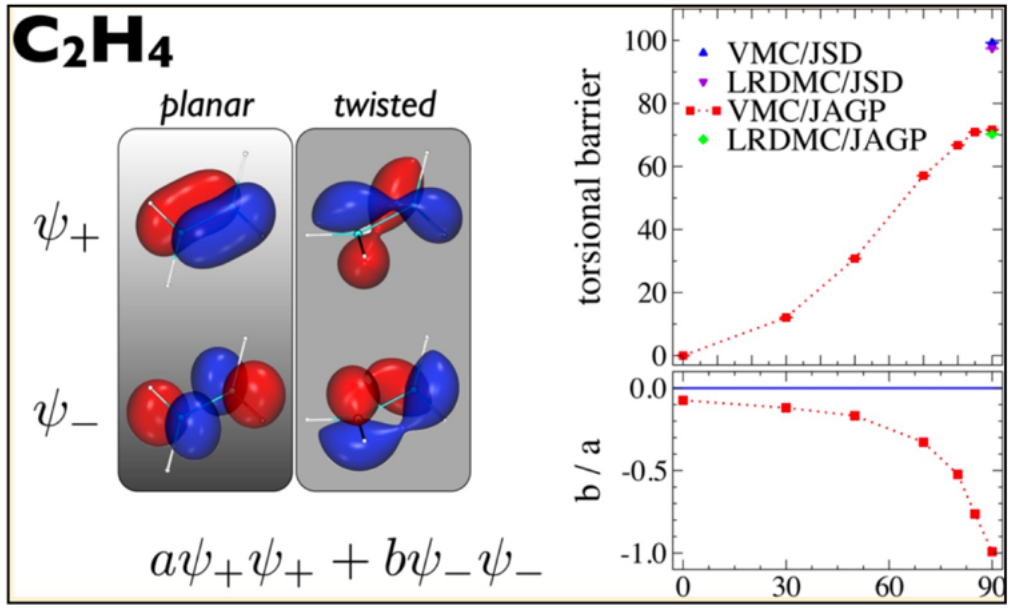


FIG. 4. *Ansatz* conversion.

K. Nakano, et al. J. Chem. Phys. 152, 204121 (2020) TurboRVB: A many-body toolkit for ab initio electronic simulations by quantum Monte Carlo

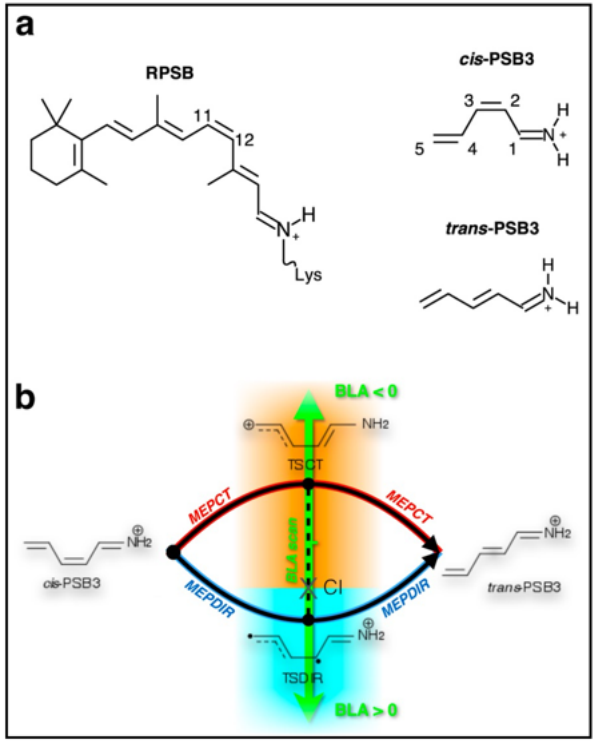
When do we need a multideterminant Slater-Jastrow w.f.?

Breaking bonds



J. Chem. Theory Comput. 2014, 10, 1048−1061
Static and Dynamical Correlation in Diradical Molecules by Quantum Monte Carlo Using the Jastrow Antisymmetrized Geminal Power Ansatz

Transition states



J. Chem. Theory Comput. 2015, 11, 992−1005
Quantum Monte Carlo Treatment of the Charge Transfer and Diradical Electronic Character in a Retinal Chromophore Minimal Model

JAGP ansatz

unpolarized system
(zero total spin S)

N electrons
$N_p = N/2$ singlet pairs

AGP

$$\Psi_{AGP}(\bar{\mathbf{x}}) = \hat{\mathcal{A}}\left[\prod_i^{N_p} G\left(\mathbf{x}_i; \mathbf{x}_{N_p+i}\right)\right]$$

$$G(\mathbf{x}_i; \mathbf{x}_j) = \mathcal{G}\left(\mathbf{r}_i, \mathbf{r}_j\right) \frac{\alpha(i)\beta(j) - \beta(i)\alpha(j)}{\sqrt{2}}$$

$$\mathcal{G}\left(\mathbf{r}_i, \mathbf{r}_j\right) = \sum_\mu^L \sum_\nu^L g_{\mu\nu} \phi_\mu\left(\mathbf{r}_i\right) \phi_\nu\left(\mathbf{r}_j\right)$$

L(L+1)/2 determinantal parameters

polarized system
(S not zero)

N electrons
$N_u = 2*S$ unpaired electrons
$N = N_u + 2*N_p$

GAGP

$$\Psi_{GAGP}(\bar{\mathbf{x}}) = \hat{\mathcal{A}}\left\{\left[\prod_i^{N_p} G\left(\mathbf{x}_i; \mathbf{x}_{N_p+i}\right)\right]\left[\prod_j^{N_u} \chi_j\left(\mathbf{x}_{2N_p+j}\right)\right]\right\}$$

$$\chi_j(\mathbf{x}_i) = \left[\sum_\mu^L f_{j,\mu} \phi_\mu(\mathbf{x}_i)\right] \alpha(i)$$

atomic orbitals (uncontracted/contracted/hybrid)

L(L+1)/2+2SL determinantal parameters

JAGP ansatz

unpolarized system
(zero total spin S)

N electrons
$N_p = N/2$ singlet pairs

AGP

$$\Psi_{AGP}(\bar{\mathbf{x}}) = \hat{\mathcal{A}}\left[\prod_i^{N_p} G\left(\mathbf{x}_i; \mathbf{x}_{N_p+i}\right)\right]$$

$$G(\mathbf{x}_i; \mathbf{x}_j) = \mathcal{G}\left(\mathbf{r}_i, \mathbf{r}_j\right) \frac{\alpha(i)\beta(j) - \beta(i)\alpha(j)}{\sqrt{2}}$$

$$\mathcal{G}\left(\mathbf{r}_i, \mathbf{r}_j\right) = \sum_\mu^L \sum_\nu^L g_{\mu\nu} \phi_\mu\left(\mathbf{r}_i\right) \phi_\nu\left(\mathbf{r}_j\right)$$

atomic orbitals (uncontracte

L(L+1)/2 determinantal parameters

General diagonalization
of the symmetric matrix G
(S is the overlap matrix)

$$S^{1/2} G S^{1/2} = U^\dagger \Lambda U$$

yields

$$G(\mathbf{r}_i, \mathbf{r}_j) = \sum_\mu^L \lambda_\mu \psi_\mu(\mathbf{r}_i) \psi_\mu(\mathbf{r}_j)$$

where we can assume

$$|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_L|$$

AGP equivalently rewritten as a multi-determinant wave function - molecular (natural) orbital expansion -

case of unpolarized systems

$$\Psi_{AGP} = c_0 \left|\Psi_0\right\rangle + \sum_{i=1}^{N_p} \sum_{a=N_p+1}^{L} c_{ii}^{aa} \left|\Psi_{ii}^{aa}\right\rangle + \sum_{\substack{i,j=1 \\ i\neq j}}^{N_p} \sum_{\substack{a,b=N_p+1 \\ a\neq b}}^{L} c_{iijj}^{aabb} \left|\Psi_{iijj}^{aabb}\right\rangle + \dots$$

(constrained) coefficients

$$c_0 = \prod_i^{N_p} \lambda_i \,; \qquad c_{ii}^{aa} = c_0 \frac{\lambda_a}{\lambda_i} \,; \qquad c_{iijj}^{aabb} = c_0 \frac{\lambda_a \lambda_b}{\lambda_i \lambda_j} \,; \qquad \dots$$

leading closed-shell Slater determinant

$$\left|\Psi_0\right\rangle = \hat{\mathcal{A}} \left\{ \left[\prod_i^{N_p} \psi_i(\mathbf{r}_i)\alpha(i) \right] \left[\prod_j^{N_p} \psi_j(\mathbf{r}_{N_p+j})\beta(j) \right] \right\}$$

excitation Slater determinants

$$\left|\Psi_{ii}^{aa}\right\rangle \quad , \ \dots$$

zero seniority subset of a FCI expansion
(with constrained coefficients)

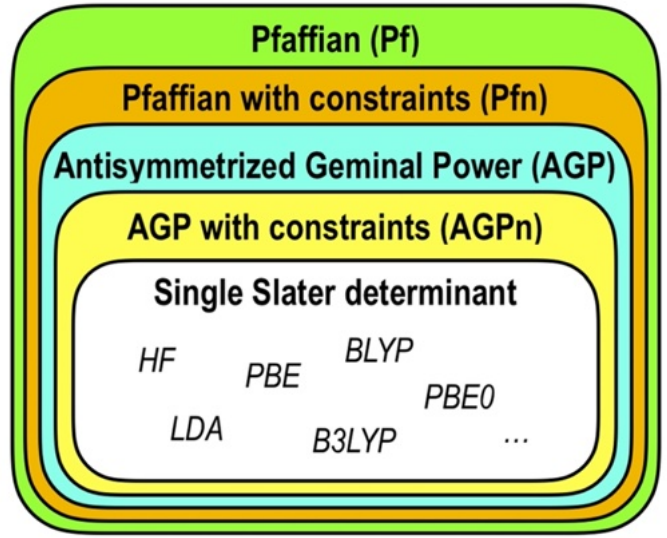
AGP*n*

truncated AGP function

truncate the pairing function using ***n*** molecular orbitals

$$\mathcal{G}_n(\mathbf{r}_i,\mathbf{r}_j)=\sum_{k=1}^{n}\lambda_k\psi_k(\mathbf{r}_i)\,\psi_k(\mathbf{r}_j) \qquad N_p \le n \ll L$$

$$\Psi_{AGP} = c_0\left|\Psi_0\right\rangle + \sum_{i=1}^{N_p}\sum_{a=N_p+1}^{n} c_{ii}^{aa}\left|\Psi_{ii}^{aa}\right\rangle + \sum_{\substack{i,j=1\\ i\neq j}}^{N_p}\sum_{\substack{a,b=N_p+1\\ a\neq b}}^{n} c_{iijj}^{aabb}\left|\Psi_{iijj}^{aabb}\right\rangle + \dots$$

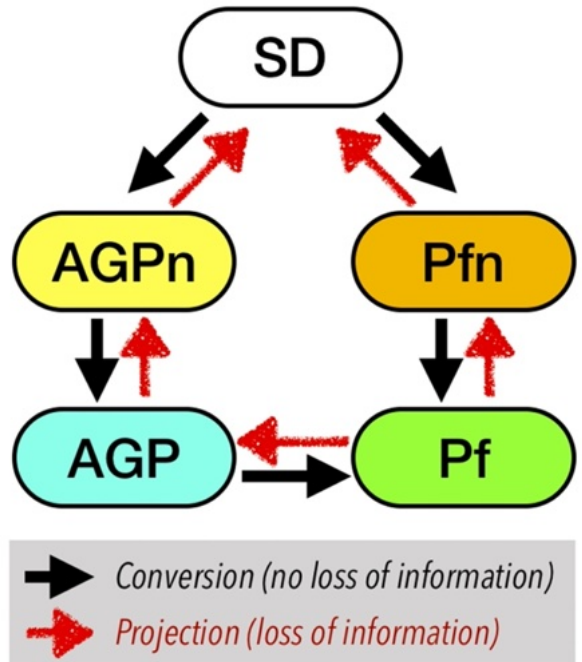


single Slater determinant (special case of AGP***n*** function with ***n*** = N_p)

RHF *or* SD

AGP***n****

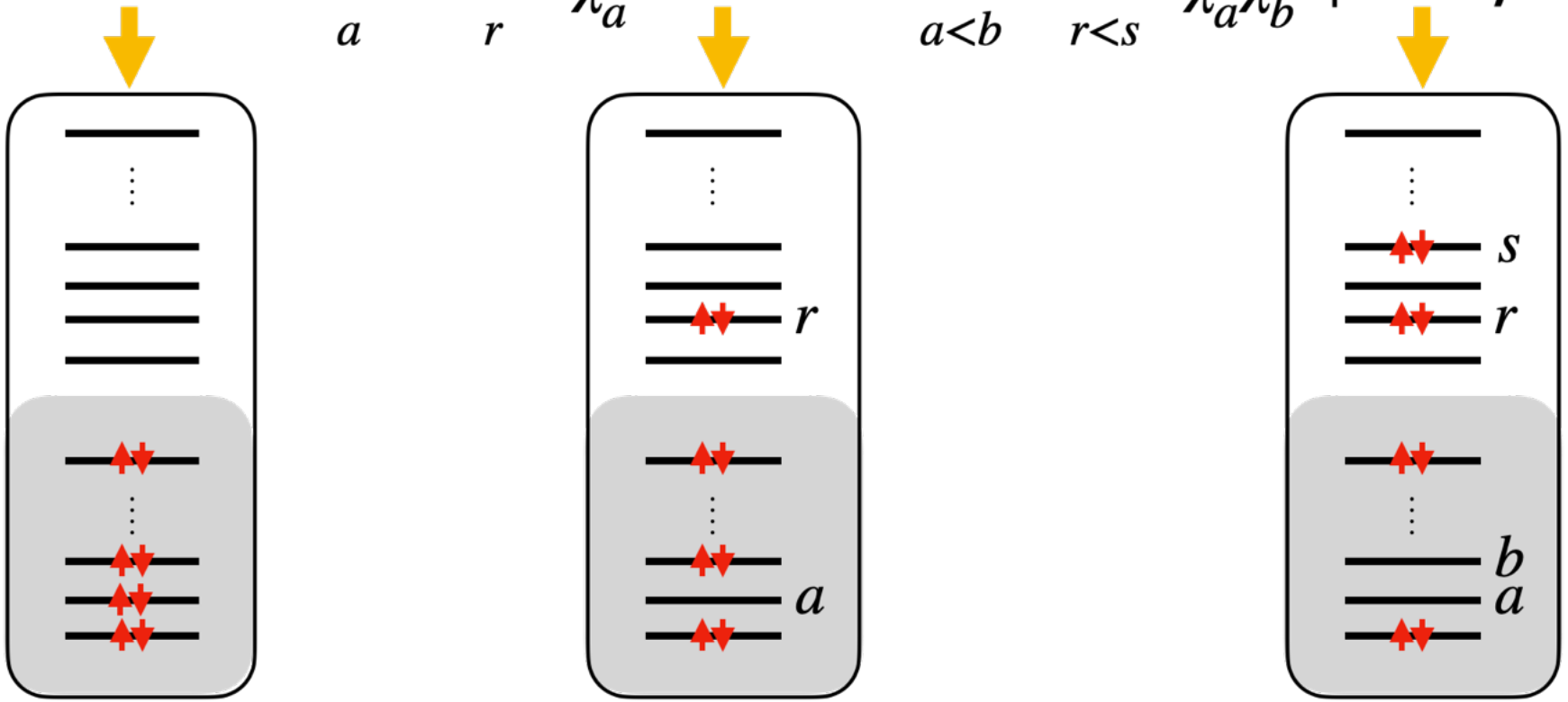
special case of AGP***n***, with ***n*** = ***n**** such that atoms are SD



Marchi, Mariapia, Sam Azadi, Michele Casula, and Sandro Sorella. JCP 2009.

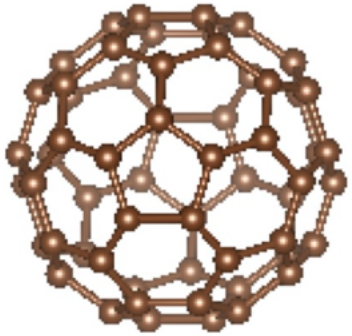
From one Slater determinant to the AGPn

$$\left|\psi_{\mathbf{AGPn}}\right\rangle = \left|\psi_0\right\rangle + \sum_{a}^{\text{occ.}} \sum_{r}^{\text{virt.}} \frac{\lambda_r}{\lambda_a} \left|\psi_{a\bar{a}}^{r\bar{r}}\right\rangle + \sum_{a<b}^{\text{occ.}} \sum_{r<s}^{\text{virt.}} \frac{\lambda_r \lambda_s}{\lambda_a \lambda_b} \left|\psi_{a\bar{a}b\bar{b}}^{r\bar{r}s\bar{s}}\right\rangle + \cdots$$



Lambda's of virtual orbitals are zero in the single Slater determinant.
If we allow them to be optimized we have a multideterminant wave function which could improve the nodal surface!

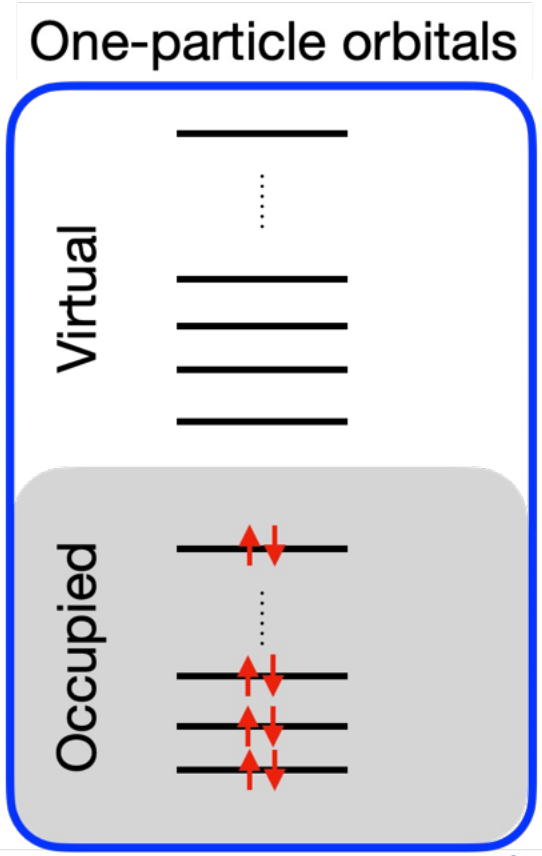
Idea: get NOs from a deterministic method



Schrödinger Equation

$$\hat{H}\Psi = E\Psi$$

approximate solution
using Hartree-Fock (HF),
density functional theory (DFT),
or 2nd order Møller–Plesset
perturbation theory (MP2)



HF or DFT $\Longrightarrow$ molecular orbitals (MOs)

MP2, … $\Longrightarrow$ natural orbitals (NOs)

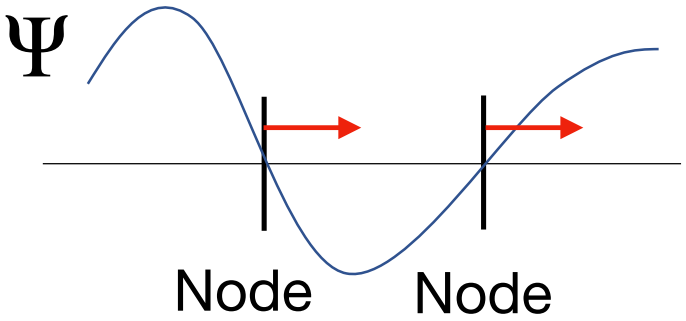
Single Slater determinant $|\psi_0\rangle$

pyscf

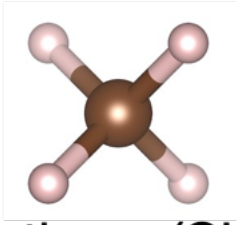
TREX
Targeting Real chemical accuracy at the EXascale

TREX-I/O

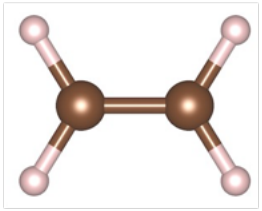
Optimize only a few coefficients lambda at the FN-DMC level



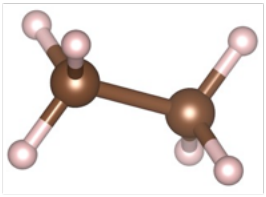
Test on systems of growing size



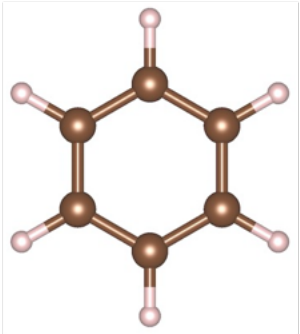
Methane (CH_4)



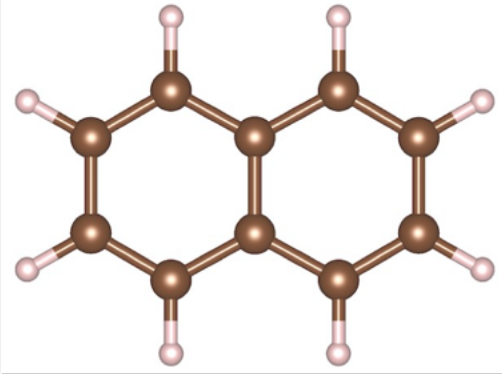
Ethylene (C_2H_4)



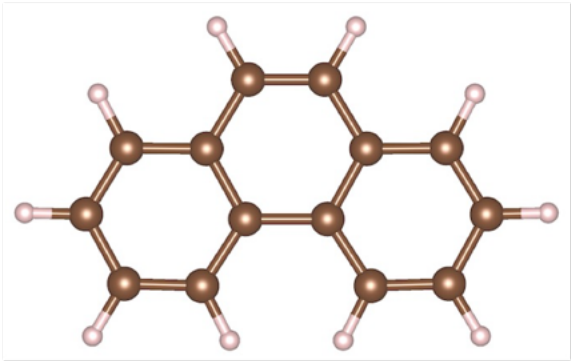
Ethane (C_2H_6)



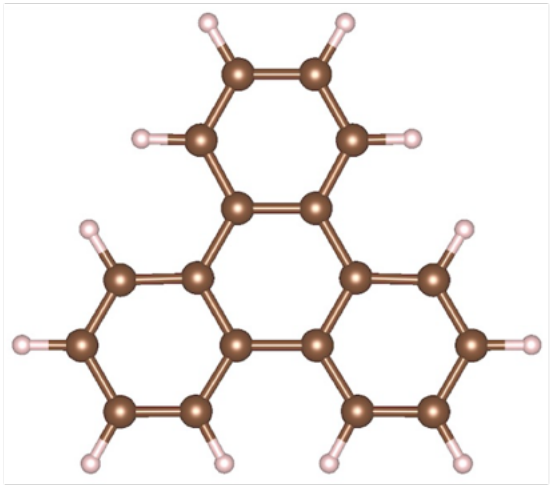
Benzene (C_6H_6)



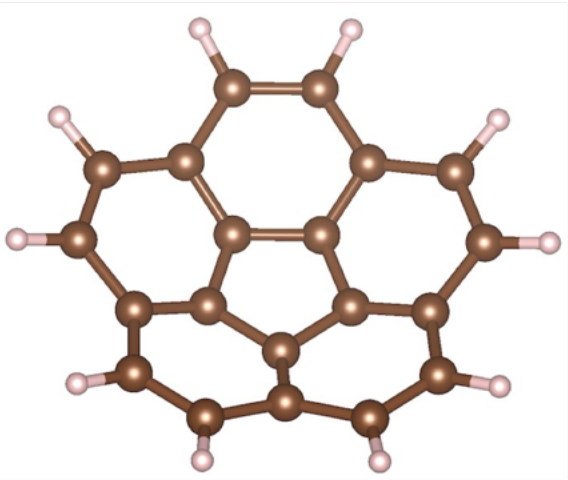
Naphthalene ($C_{10}H_8$)



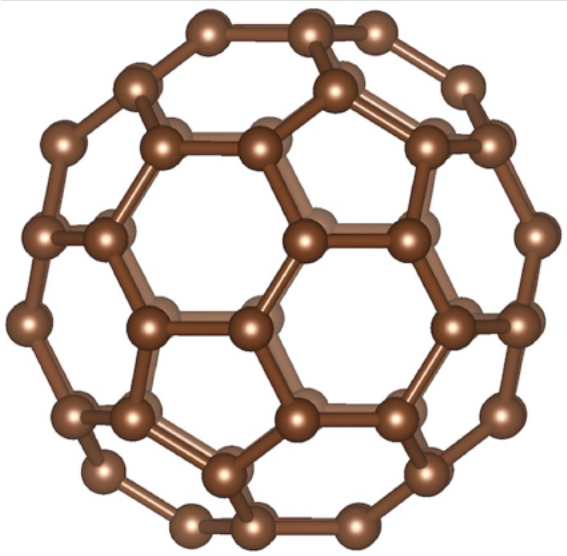
Phenanthrene ($C_{14}H_{10}$)



Triphenylene ($C_{18}H_{12}$)

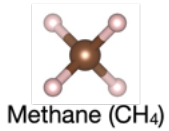


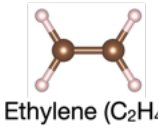
Corannulene ($C_{20}H_{10}$)

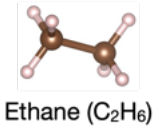


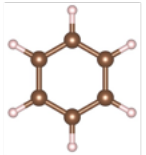
Fullerene (C_{60})

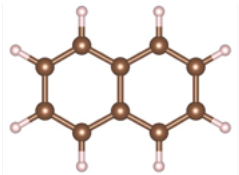
Parameters in JSD, JAGPn, JAGP

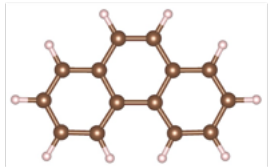

Methane (CH_4)

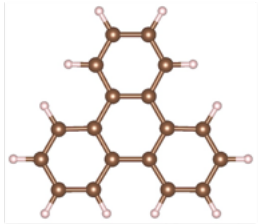

Ethylene (C_2H_4)

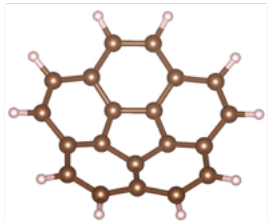

Ethane (C_2H_6)

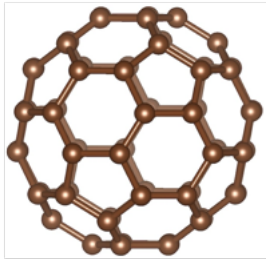

Benzene (C_6H_6)


Naphthalene ($C_{10}H_8$)

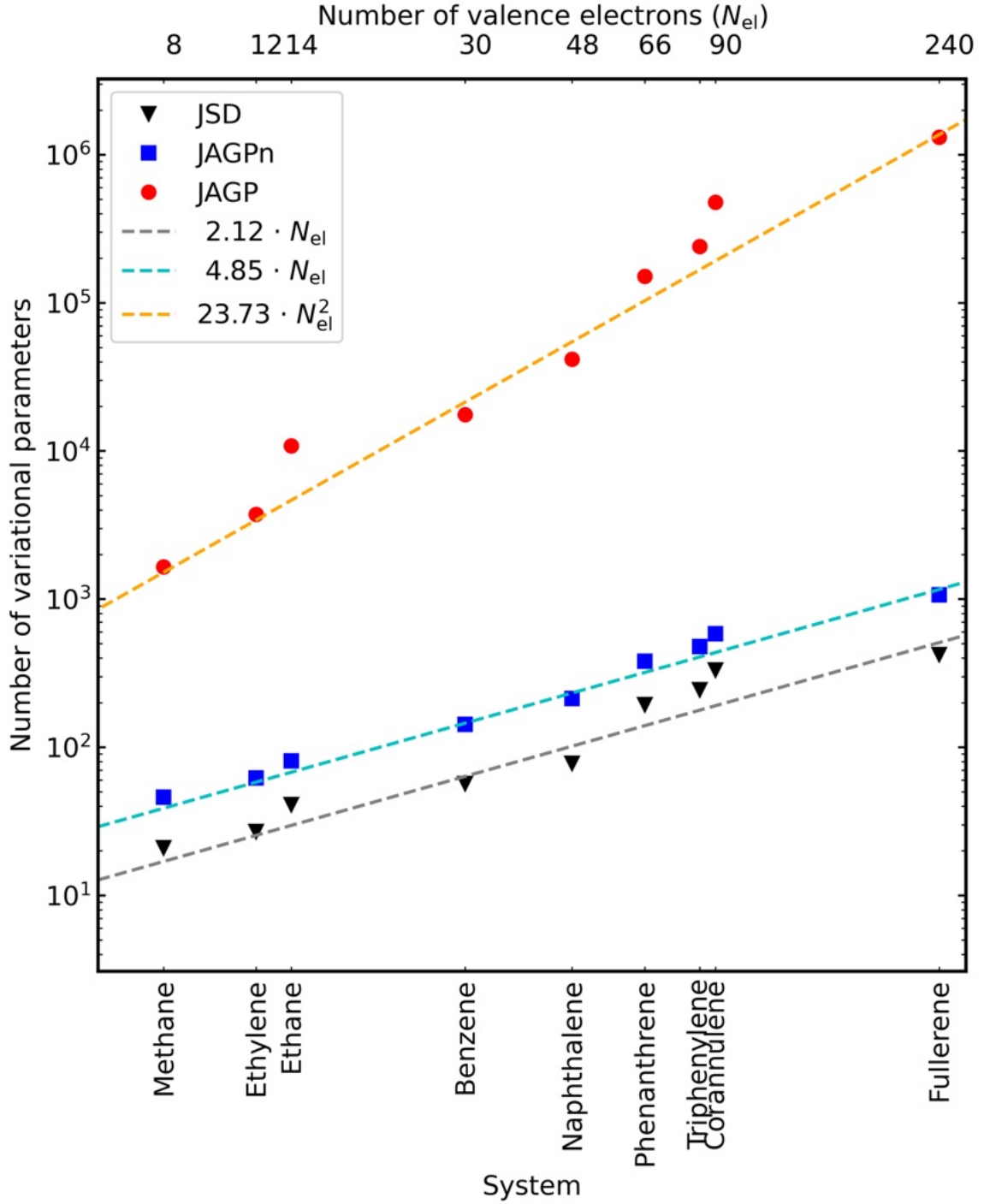

Phenanthrene ($C_{14}H_{10}$)


Triphenylene ($C_{18}H_{12}$)

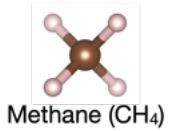

Corannulene ($C_{20}H_{10}$)

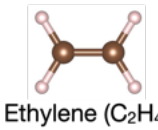

Fullerene (C_{60})

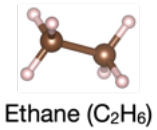
K. Nakano, S. Sorella, D. Alfè, A. Zen, *Beyond single-reference fixed-node approximation in ab initio Diffusion Monte Carlo*, arXiv:2402.01458

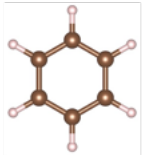


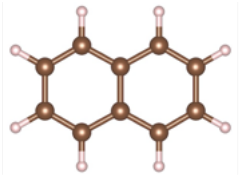
Correlation energy recovered by JAGPn and JAGP over JSD at the FN-DMC level

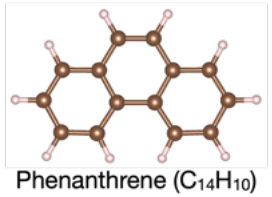

Methane (CH_4)

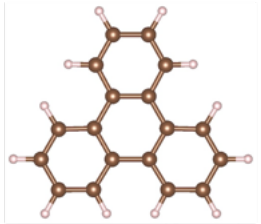

Ethylene (C_2H_4)

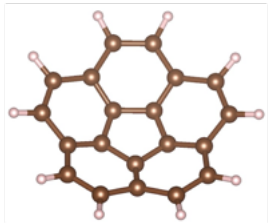

Ethane (C_2H_6)

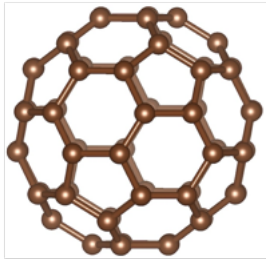

Benzene (C_6H_6)


Naphthalene ($C_{10}H_8$)

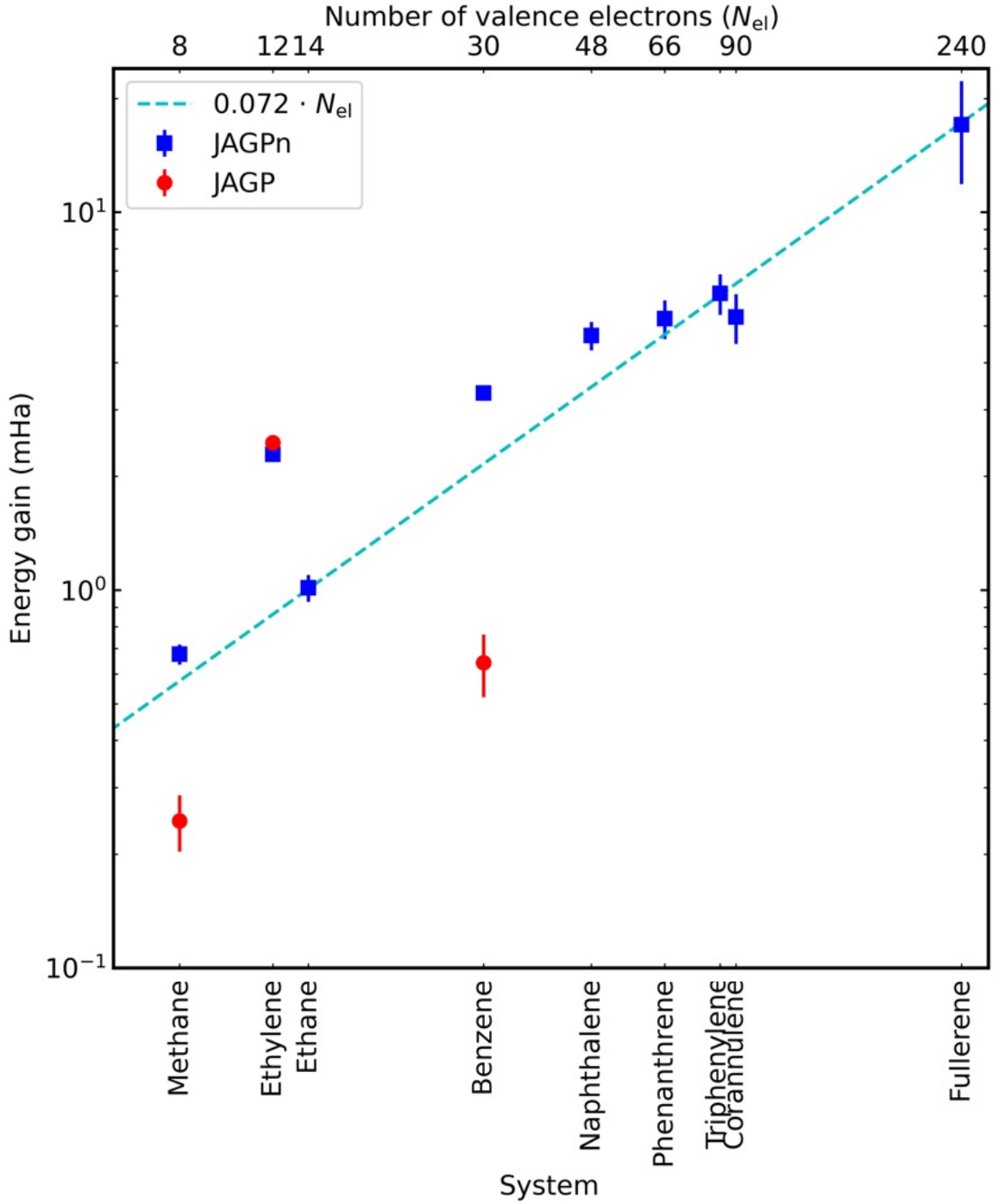

Phenanthrene ($C_{14}H_{10}$)


Triphenylene ($C_{18}H_{12}$)

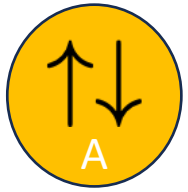

Corannulene ($C_{20}H_{10}$)


Fullerene (C_{60})

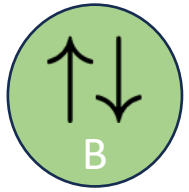
K. Nakano, S. Sorella, D. Alfè, A. Zen, *Beyond single-reference fixed-node approximation in ab initio Diffusion Monte Carlo*, arXiv:2402.01458



Size-consistency: AGP is not size-consistent!



$$\hat{G}_A|0\rangle$$



$$\hat{G}_B|0\rangle$$

$$(\hat{G}_A+\hat{G}_B)^2|0\rangle = \hat{G}_A\hat{G}_B + \hat{G}_B\hat{G}_A + \underbrace{\hat{G}_A\hat{G}_A + \hat{G}_B\hat{G}_B}_{\text{unphysical charge fluctuations}}|0\rangle$$

JAGP is size-consistent if Jastrow is “good”!

S. Sorella, M. Casula, D. Rocca. J. Chem. Phys. 127, 014105 (2007).
E. Neuscamman. Phys. Rev. Lett. 109, 203001 (2012).

… at VMC level. What about the FN level?

FN-DMC of JAGP is size-consistent if parameters are optimized at the FN level.

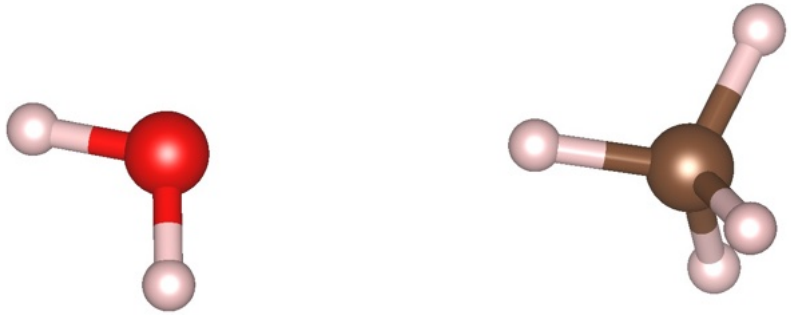
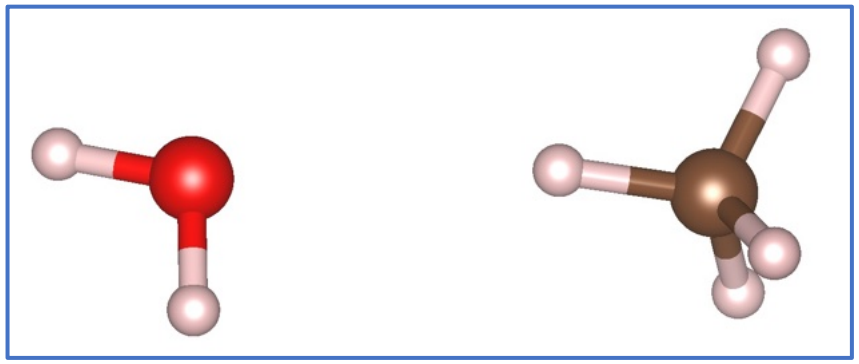


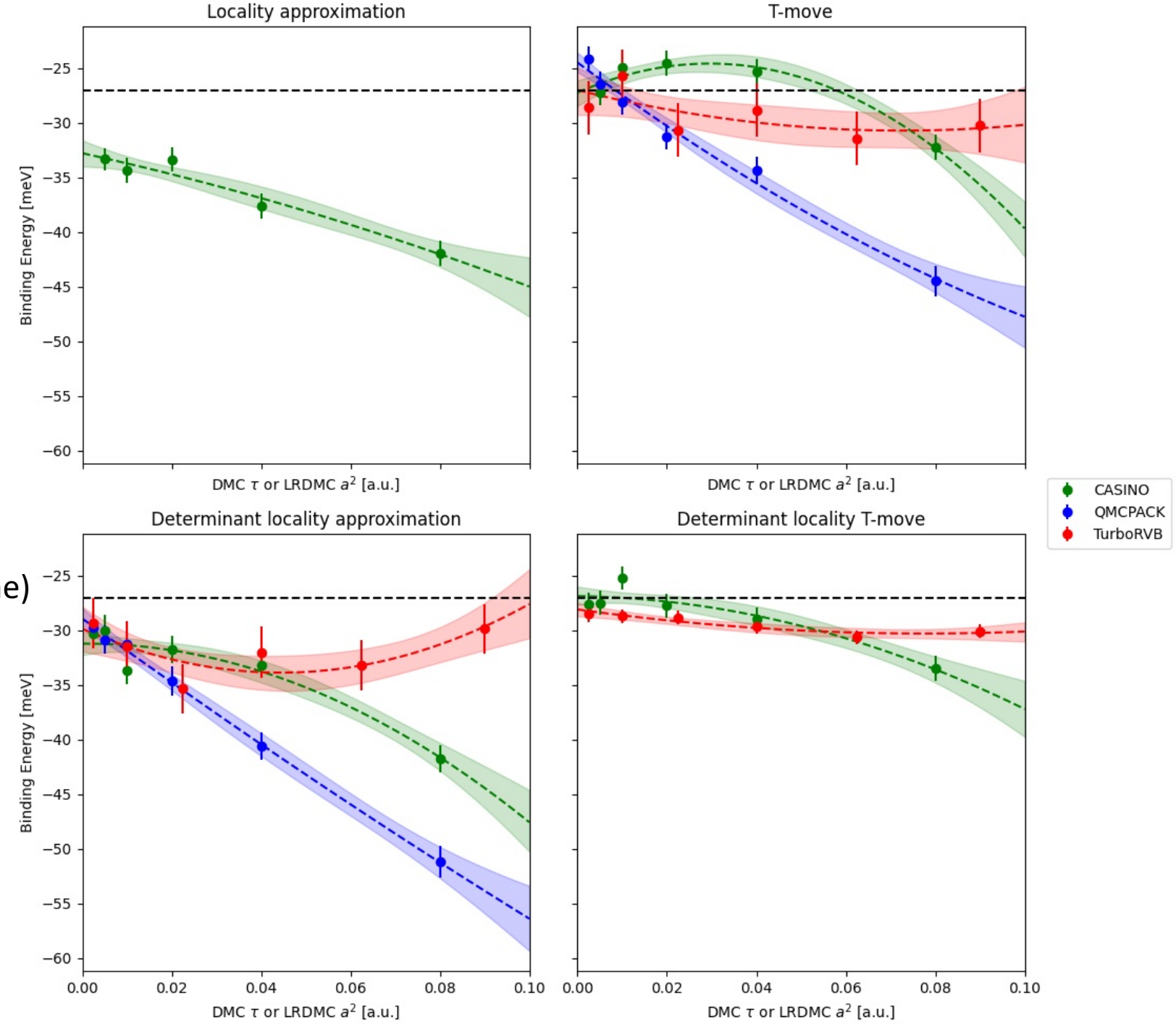
Table 1: FN binding energy E_b and size consistency energy error E_{SCE}, computed with LRDMC $a \to 0$, as obtained with the JSD, JAGPn and JAGP wave functions. For JAGPn we consider both the case of using VMC and FN gradients to optimize the nodal surface. The latter is the scheme dubbed FNAGPAS in this work.

Ansatz	Nodes Opt.	E_b (meV)	E_{SCE} (meV)
JSD	-	-27(2)	-1(1)
JAGPn	VMCopt	-46(2)	10(2)
JAGPn	FNopt	-29(2)	-2(2)
JAGP	VMCopt	-41(3)	11(3)
CCSD(T)	-	-27	0

Good agreement between QMCPACK, CASINO, TurboRVB!

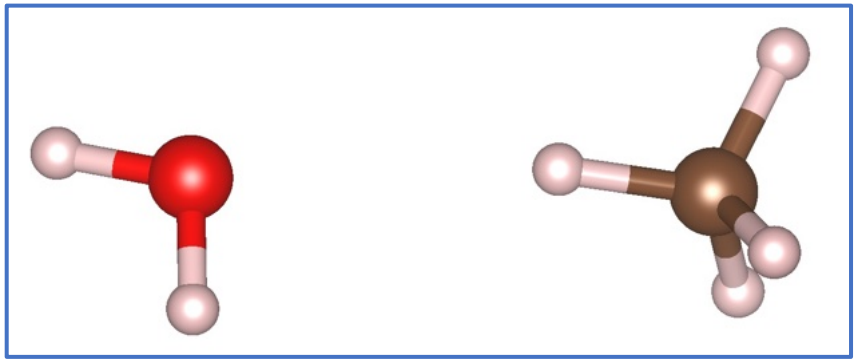
Binding energy
Eb = E(water+methane) - E(water) - E(methane)





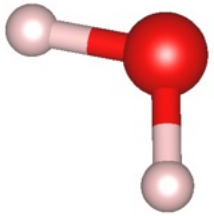
DMC reproducibility across 11 QMC codes?

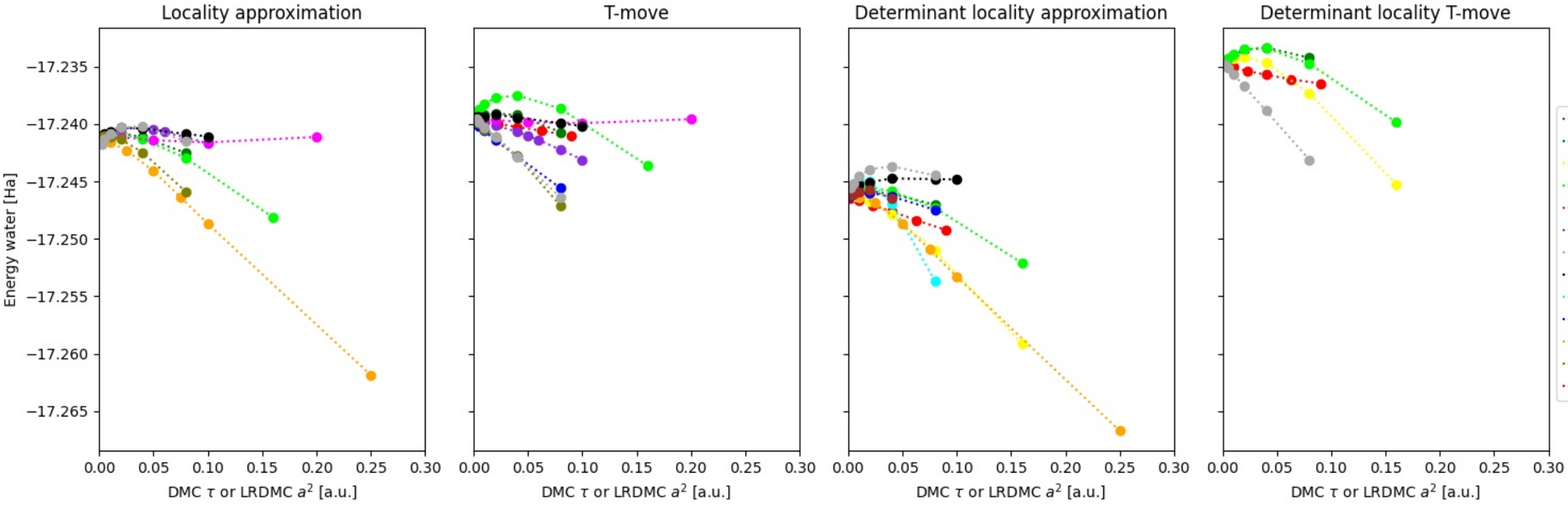
- CASINO
- QMCPACK
- TurboRVB
- QMC=Chem
- QMeCha
- PyQMC
- Amolqc
- QWalk
- CMQMC
- CHAMP Cyrus Umrigar's version
- CHAMP Claudia Filippi's version



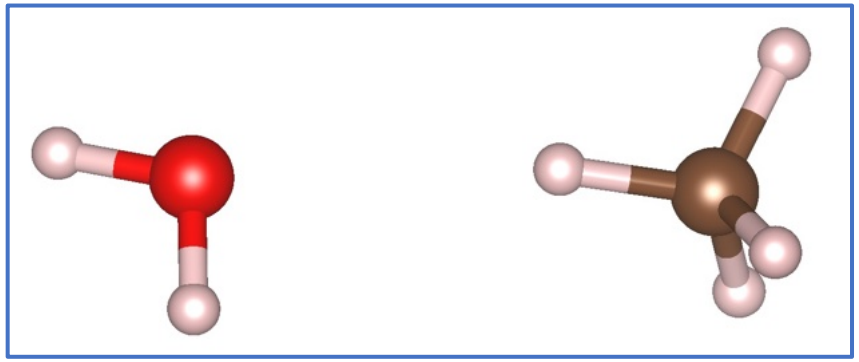
Use **Slater-Jastrow ansatz** and
same **geometry**, **pseudo-potential** (ccECP), **basis set** (ccECP-ccpVTZ), **determinant** (from Perdew-Zunger LDA),
different implementations of **Jastrow** factors and FN-DMC **algorithms**.

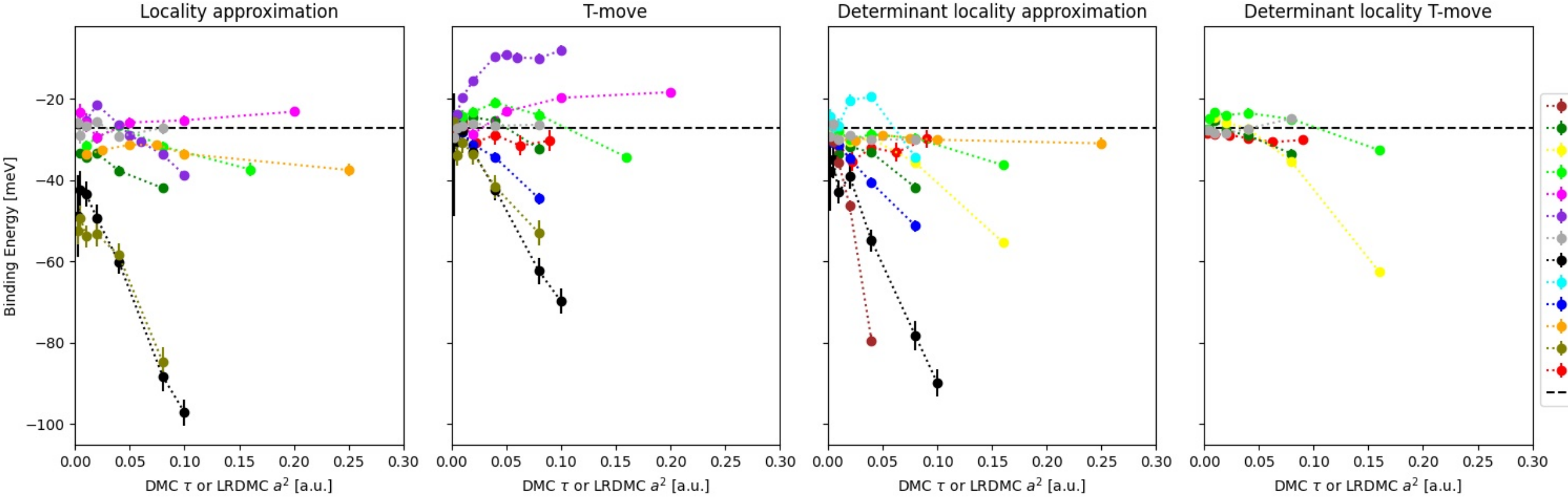
Water molecule - DMC total energy



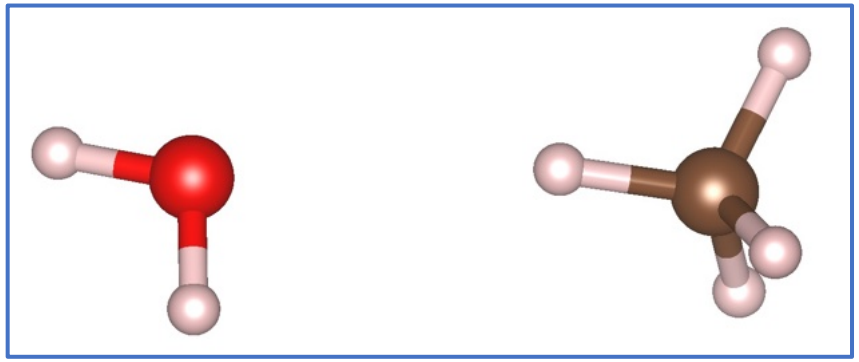


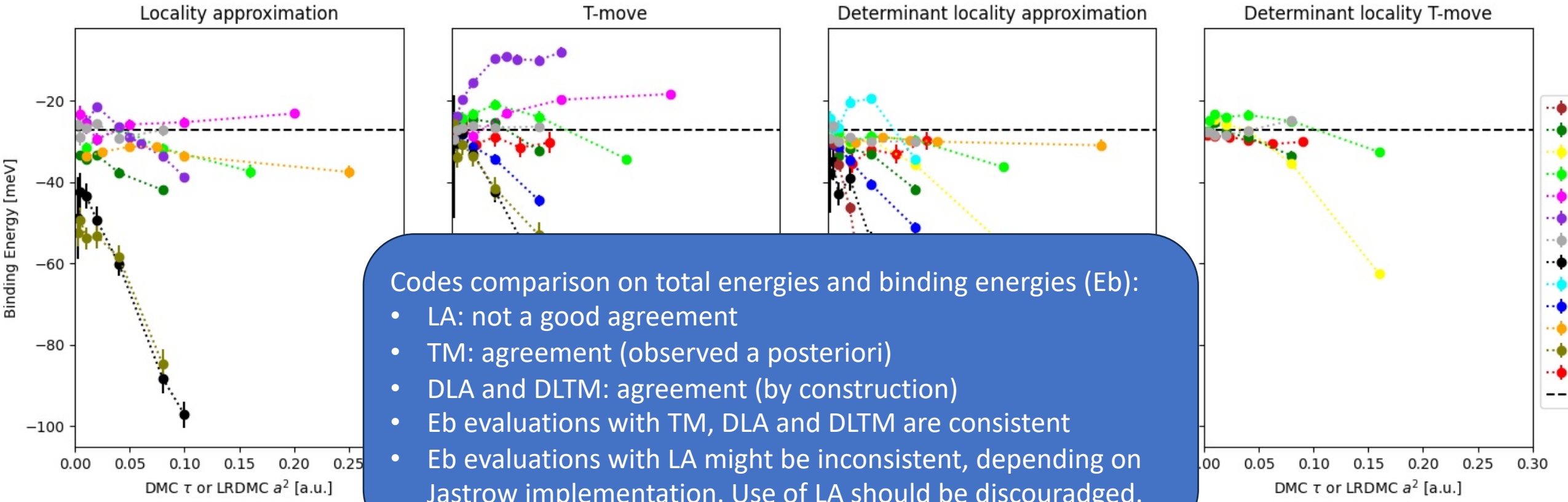
Water-Methane - DMC binding energy





Water-Methane - DMC binding energy





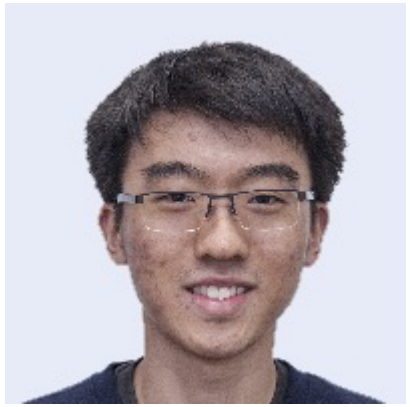
Codes comparison on total energies and binding energies (Eb):

- LA: not a good agreement
- TM: agreement (observed a posteriori)
- DLA and DLTM: agreement (by construction)
- Eb evaluations with TM, DLA and DLTM are consistent
- Eb evaluations with LA might be inconsistent, depending on Jastrow implementation. Use of LA should be discouradged.

Thanks












Università degli Studi di Napoli Federico II

Ministero dell'Università e della Ricerca

Finanziato dall'Unione europea NextGenerationEU

MATERIALS AND MOLECULAR MODELLING HUB

INTERFACES CATALYTIC ENVIRONMENTAL

THE THOMAS YOUNG CENTRE TyC

archer