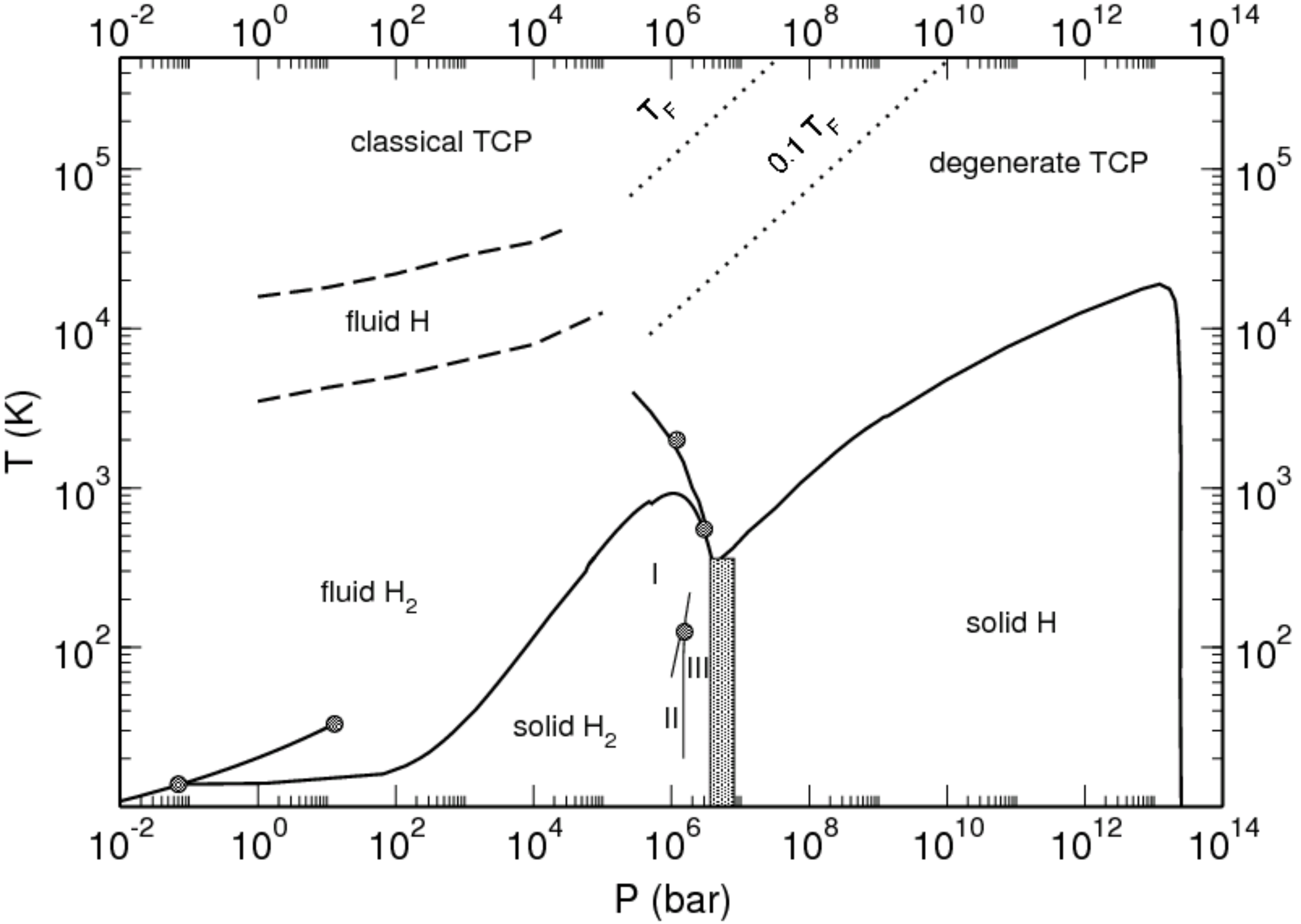
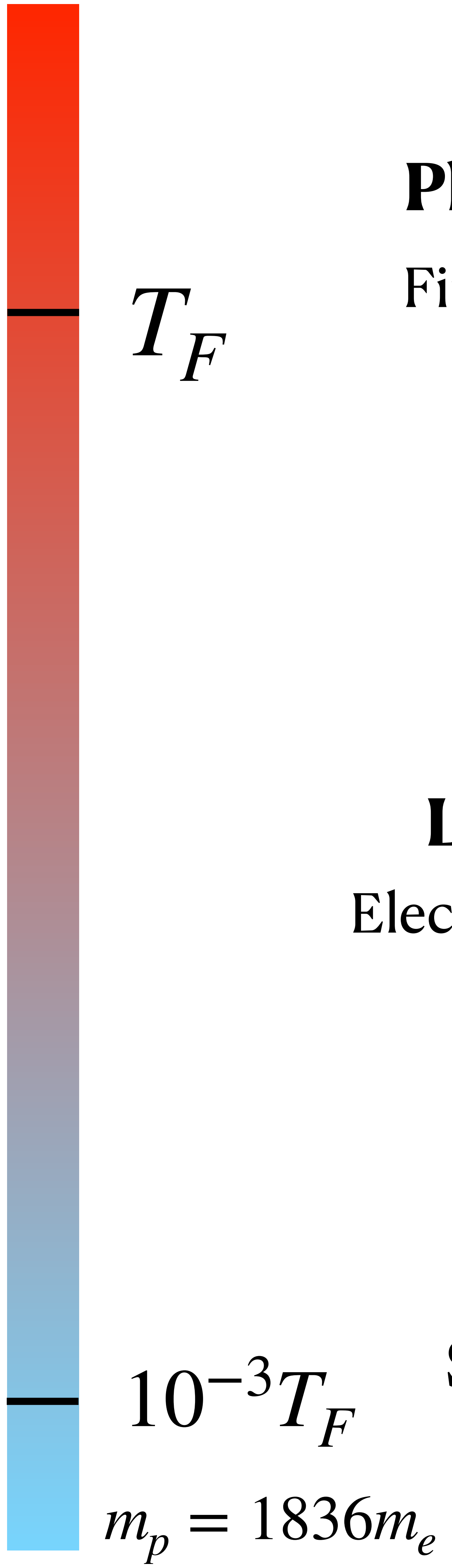
A deep variational free energy approach to dense hydrogen

Lei Wang (王磊)
Institute of Physics, CAS
https://wangleiphy.github.io


75%
in mass
H is the most abundant element in the visible Universe

McMahon et al, RMP 2012





Plasma with H^+ and e^-

Finite electron temperature

Liquid with H and H_2

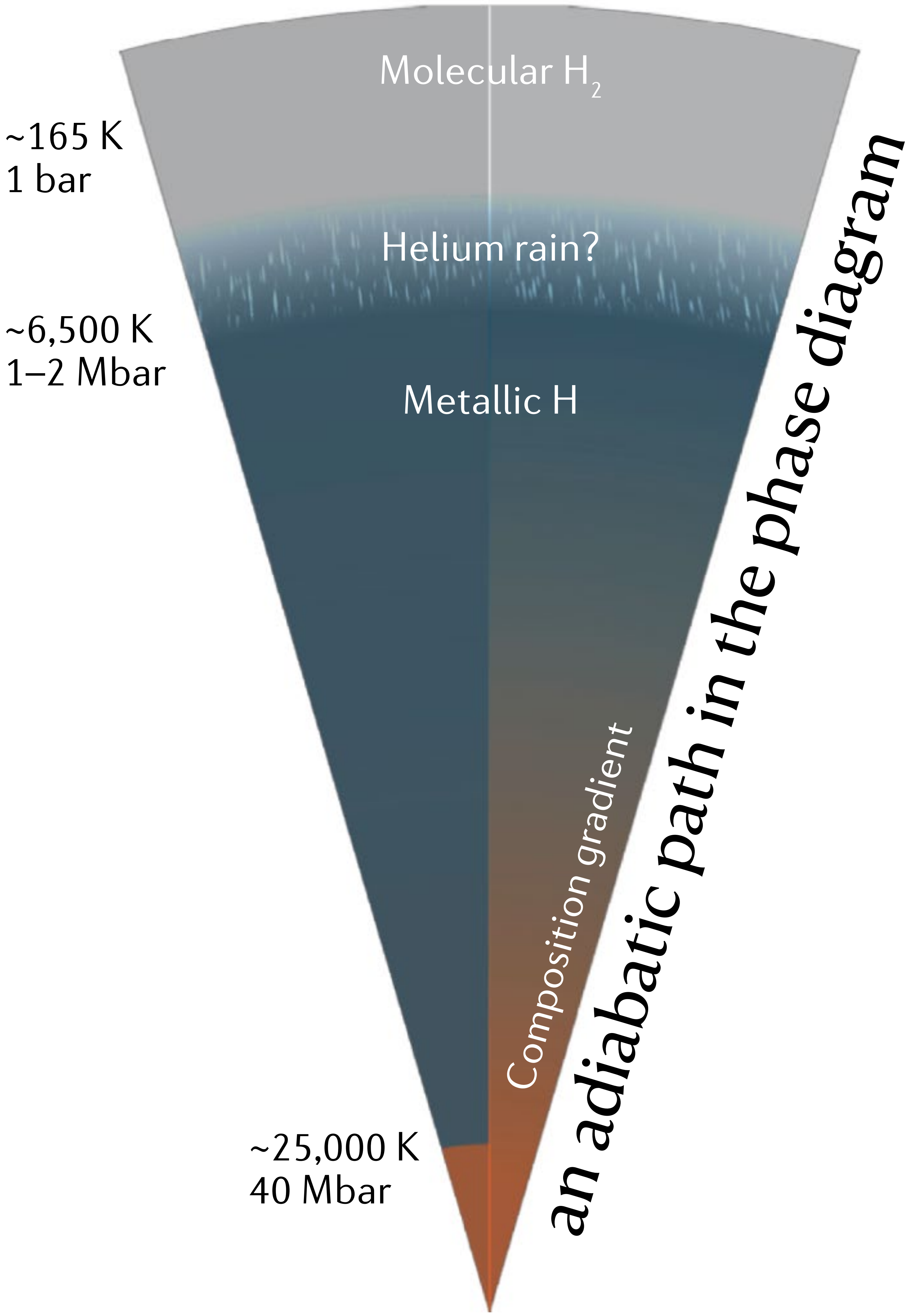
Electron stays in the ground state

Solidification
Metallization
Superconductivity...

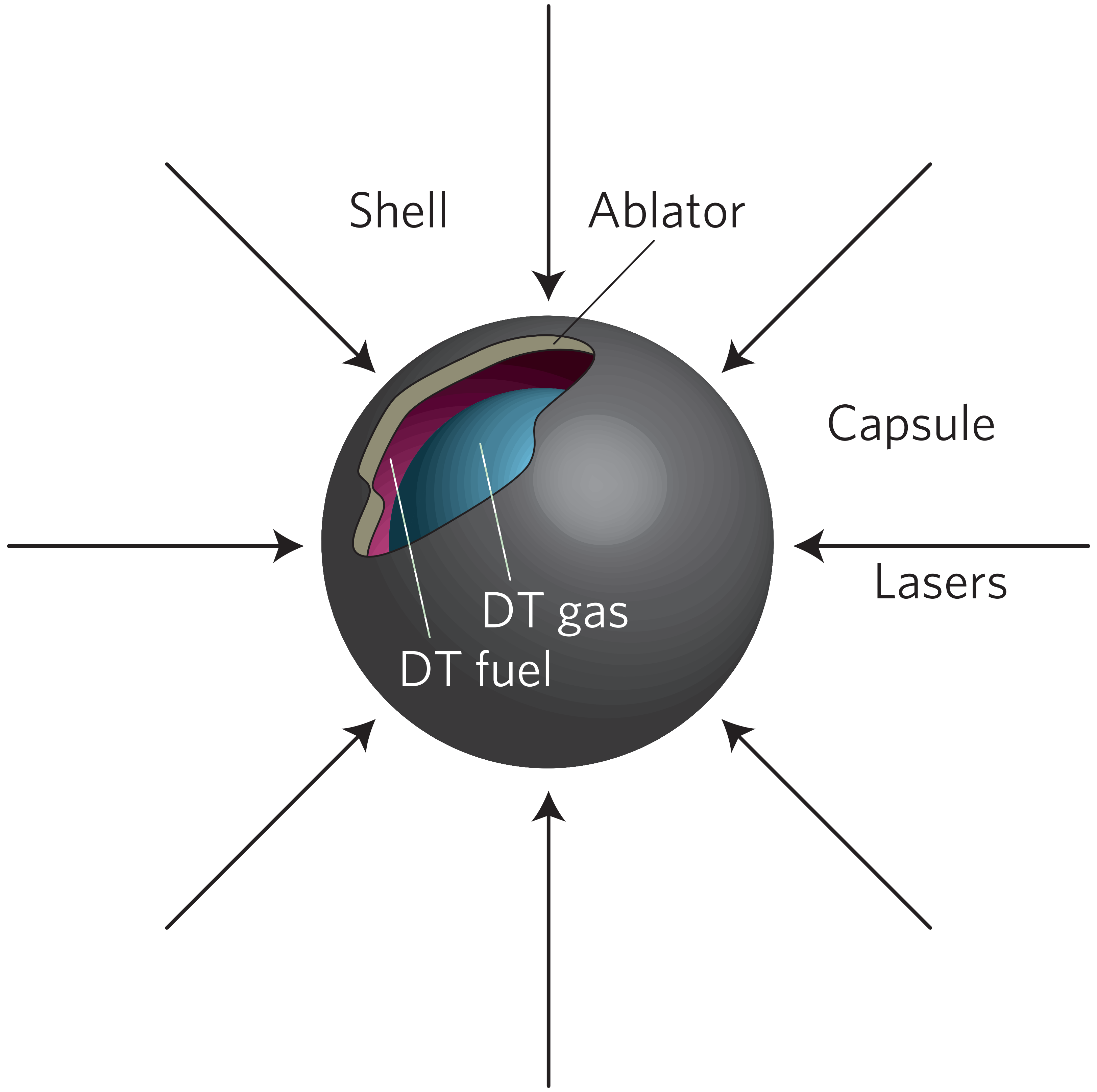
Nuclear quantum effect

Dense hydrogen in the sky and in the lab

Jupiter interior



Inertial confinement fusion



Equation-of-state is the input for hydrodynamics simulations

Superconductivity in metallic hydrogen

Wigner and Huntington 1935, Ashcroft 1968, ...

BCS theory

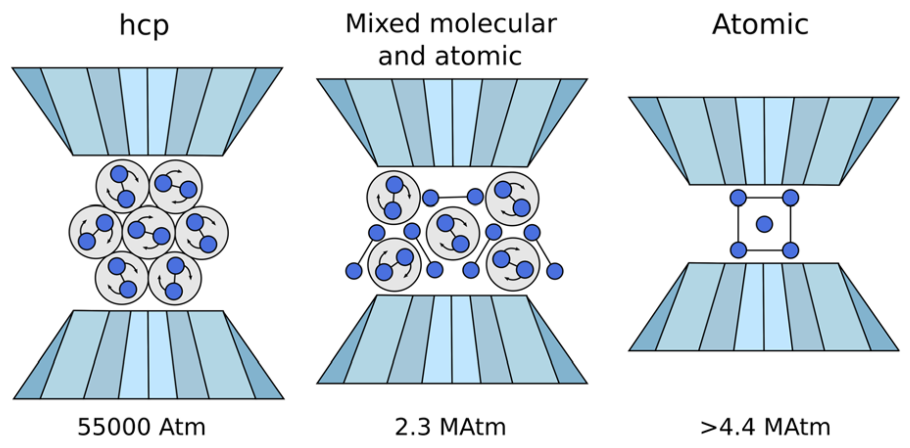
$$k_B T_c = \frac{\langle\omega\rangle\uparrow}{1.2}\exp\left[-\frac{1.04(1+\uparrow\lambda)}{\uparrow\lambda-\downarrow\mu^*(1+\uparrow 0.62\lambda)}\right],$$

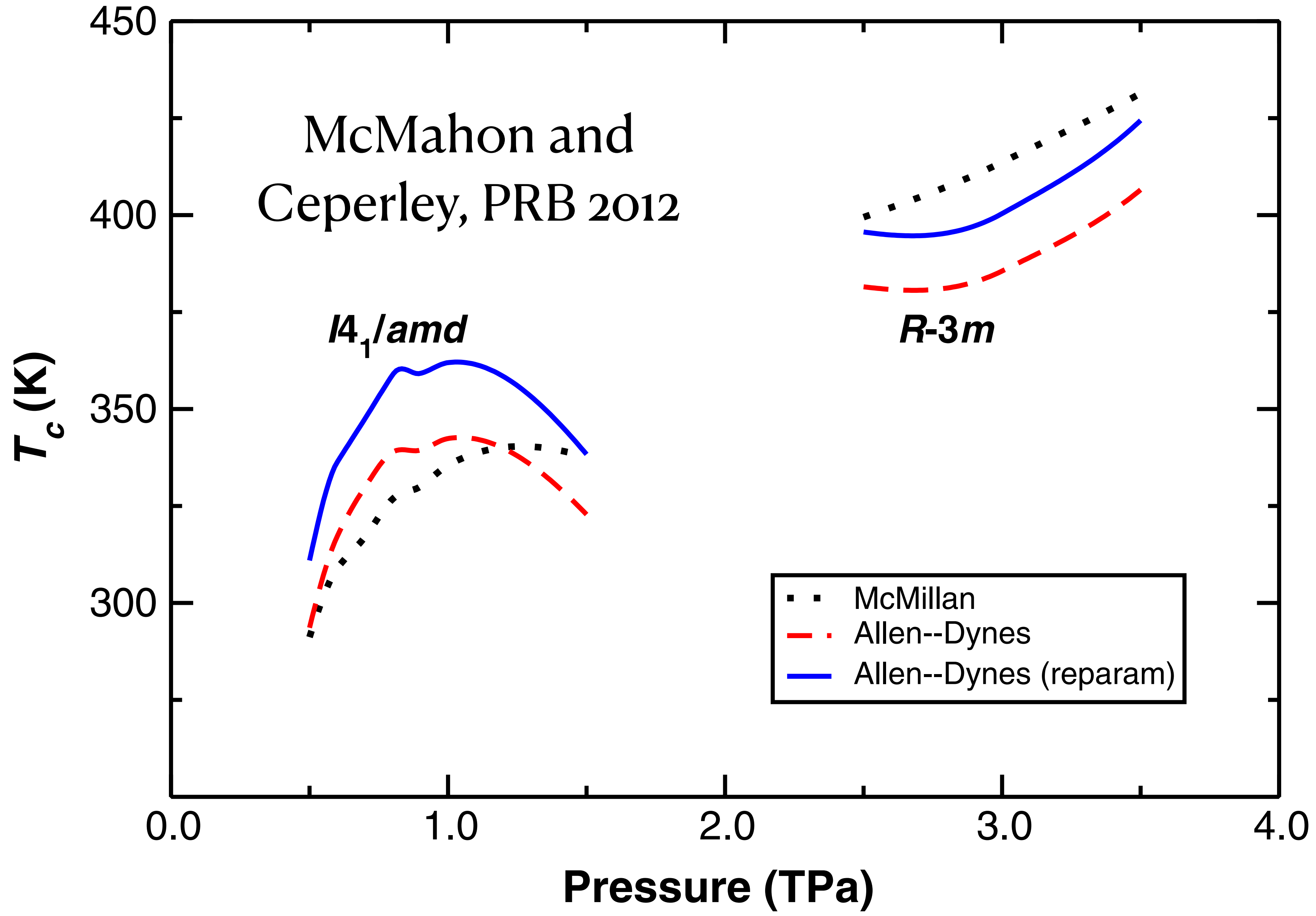
Light ion mass => higher vibrational energy scale $\langle\omega\rangle$

Bare electron-ion interaction => stronger e-p interaction λ

High density => relatively weaker e-e interaction μ^*

Higher Tc!

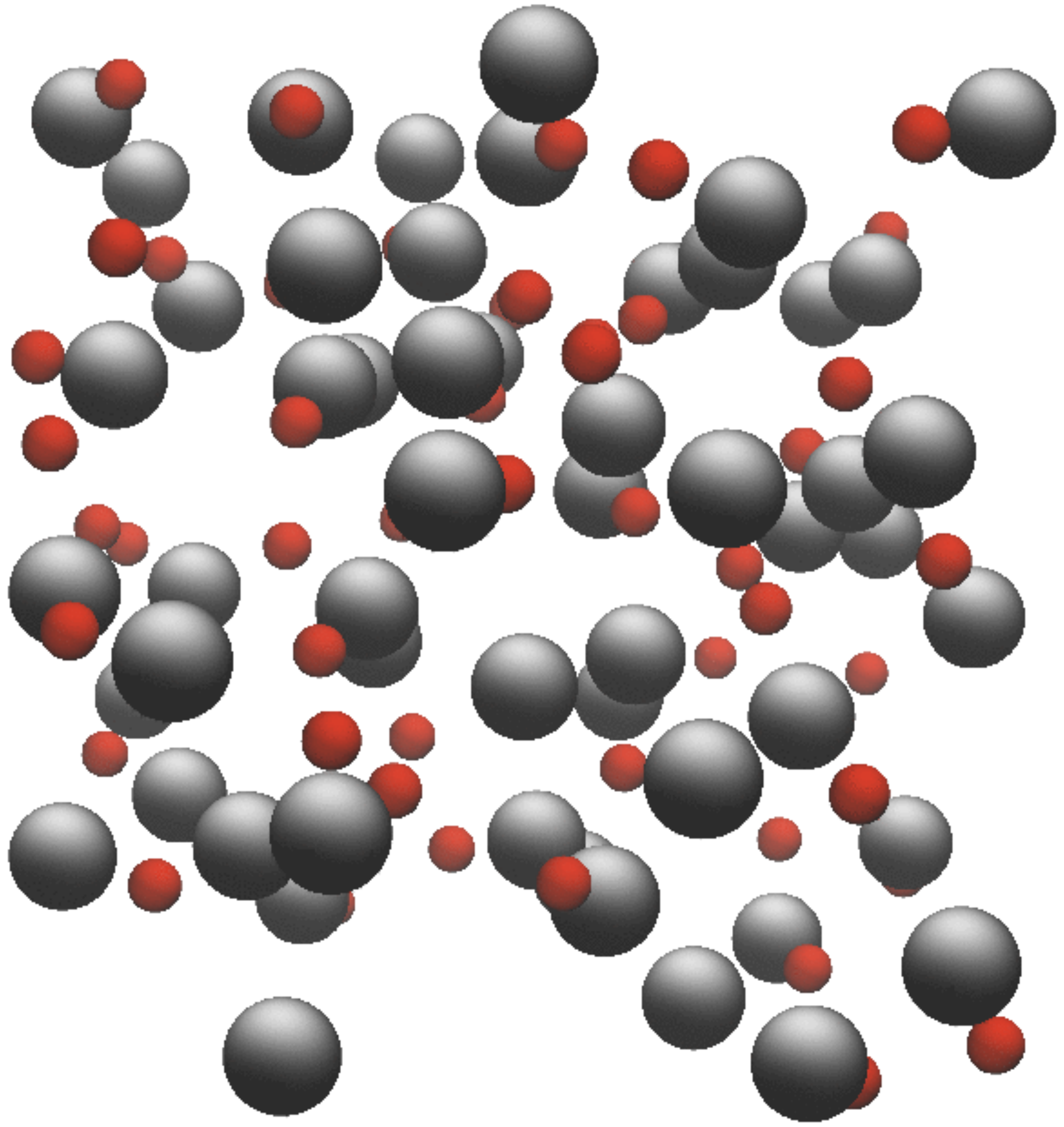




Exotic phases

Liquid superconductors: Jaffe and Aschcroft, PRB 1981, Liu et al, PRR 2020

Proton Cooper pairs: Aschcroft, JPCM 2000, Babaev et al, Nature 2004



Dense hydrogen: a simple yet fascinating quantum many-body system

Touchstone of computational methods

$T = 0$: Variational and Diffusion Monte Carlo

Mexico City, 1981

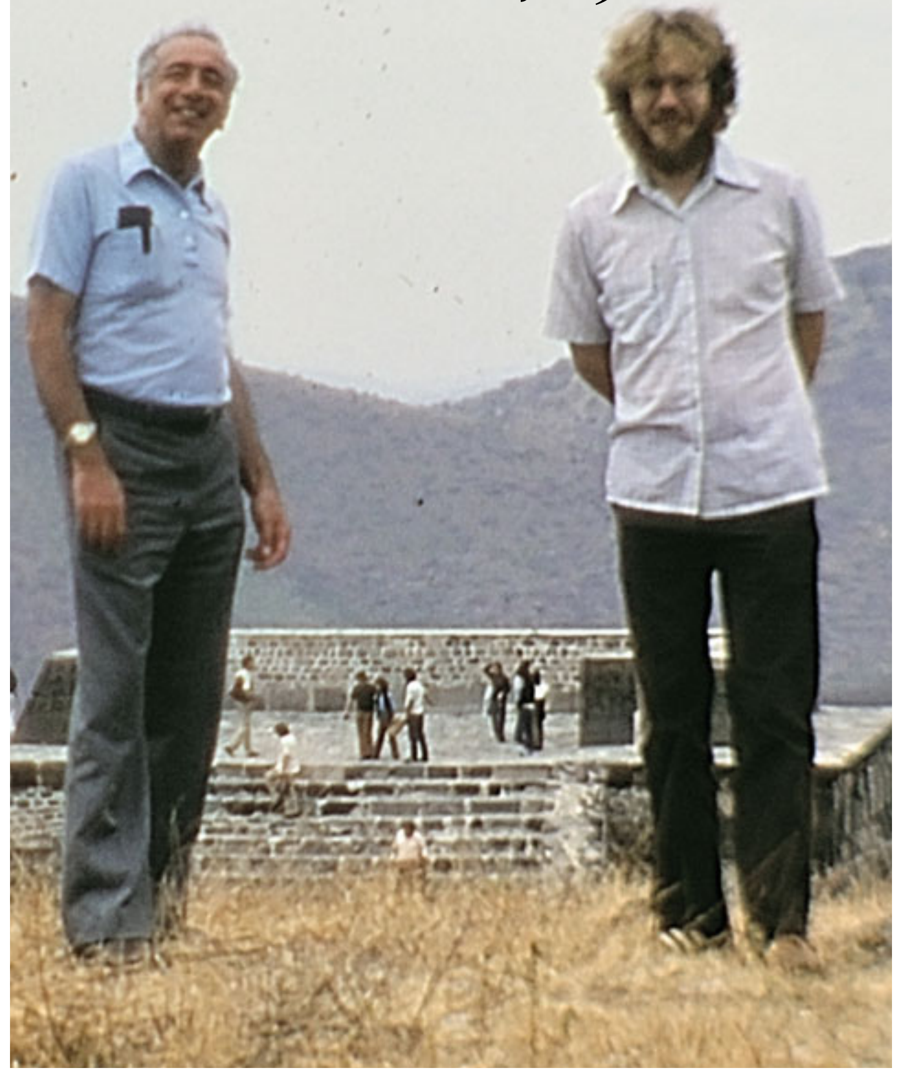


TABLE I

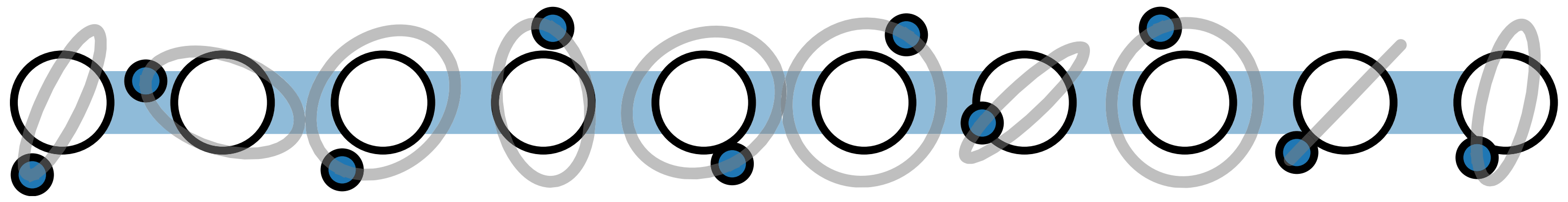
r_s	E_s	E_H	E_{CBF}	E_{PERT}	E_{LDF}
1.0	-0.726	-	-	-0.719	-
1.13	-0.892	-0.856	-0.903	-0.884	-0.906
1.31	-1.002	-0.974	-1.017	-0.996	-1.021
1.45	-1.033	-1.013	-1.054	-1.032	-1.059
1.61	-1.053	-	-1.069	-1.044	-1.074
1.77	-1.050	-1.036	-1.068	-	-1.073

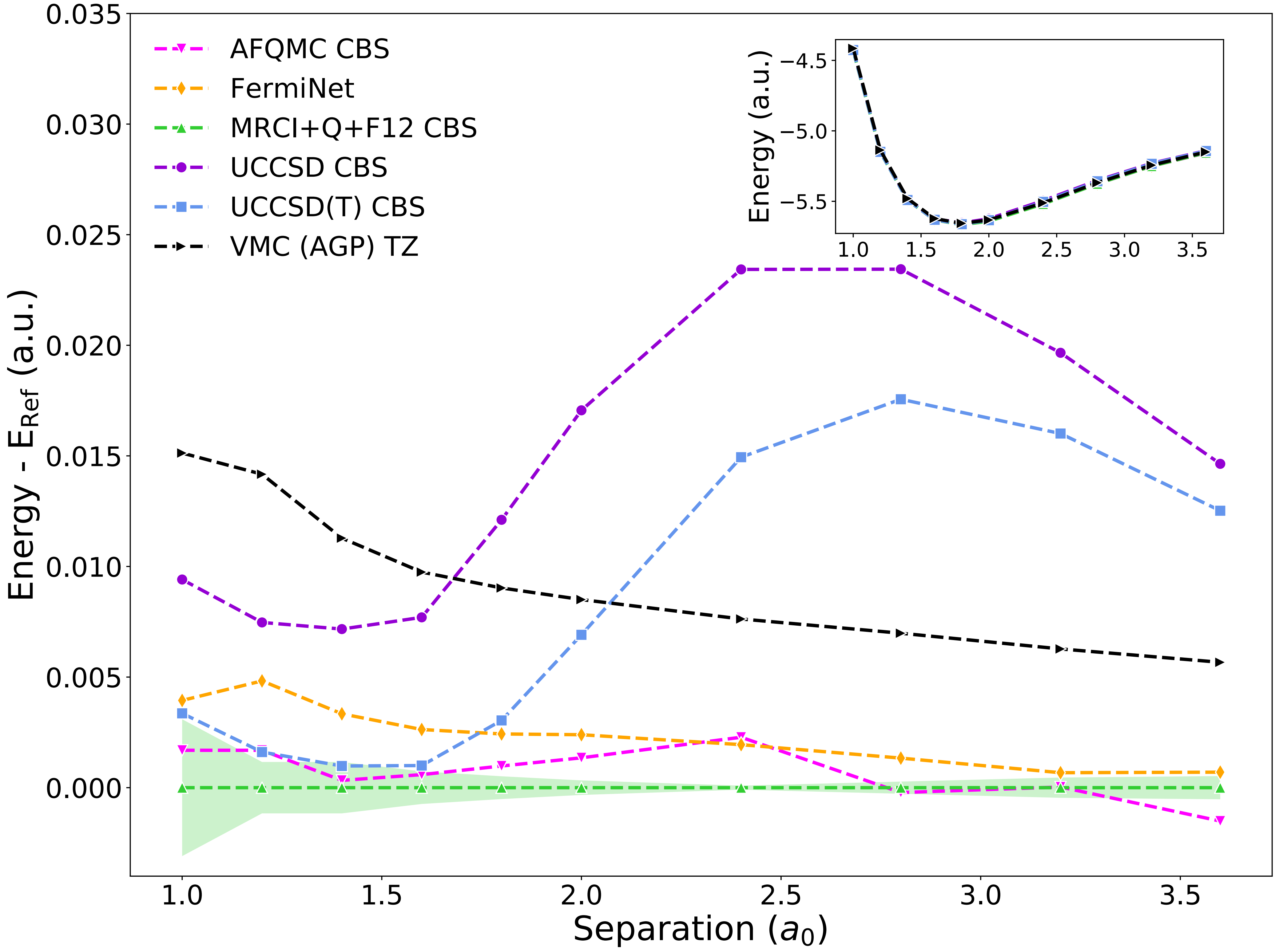
FCC lattice ground state energy
Ceperley and Alder, Physica 1981

gas model. «After I finished the electron gas calculations», Ceperley recalls, «with Berni's urging, I began to work on many-body hydrogen in 1980. An electron gas is not directly realized in any material, it's an idealized model, while hydrogen is a real material. With the hydrogen calculation we wanted to address experimental predictions, not just compare with theory. Our hydrogen calculation was the first many-electron calculation of a material to lead to important predictions».

—Computer Meets Theoretical Physics, Springer 2020

Hydrogen chain





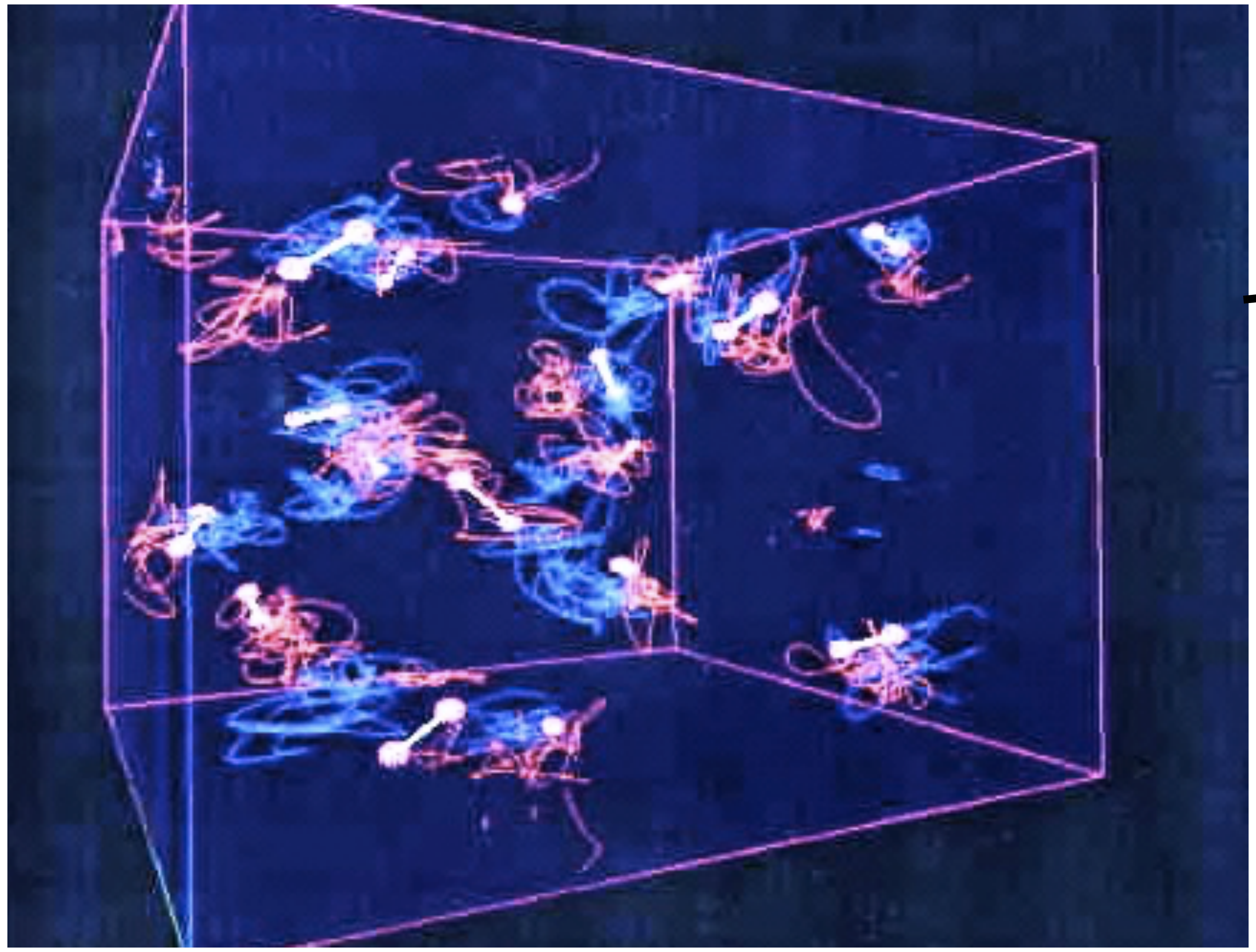
DeepMind, Pfau et al, PRR 2020
Simons collaboration, Motta el al, PRX 2017

Fixed proton configuration, no thermal effect

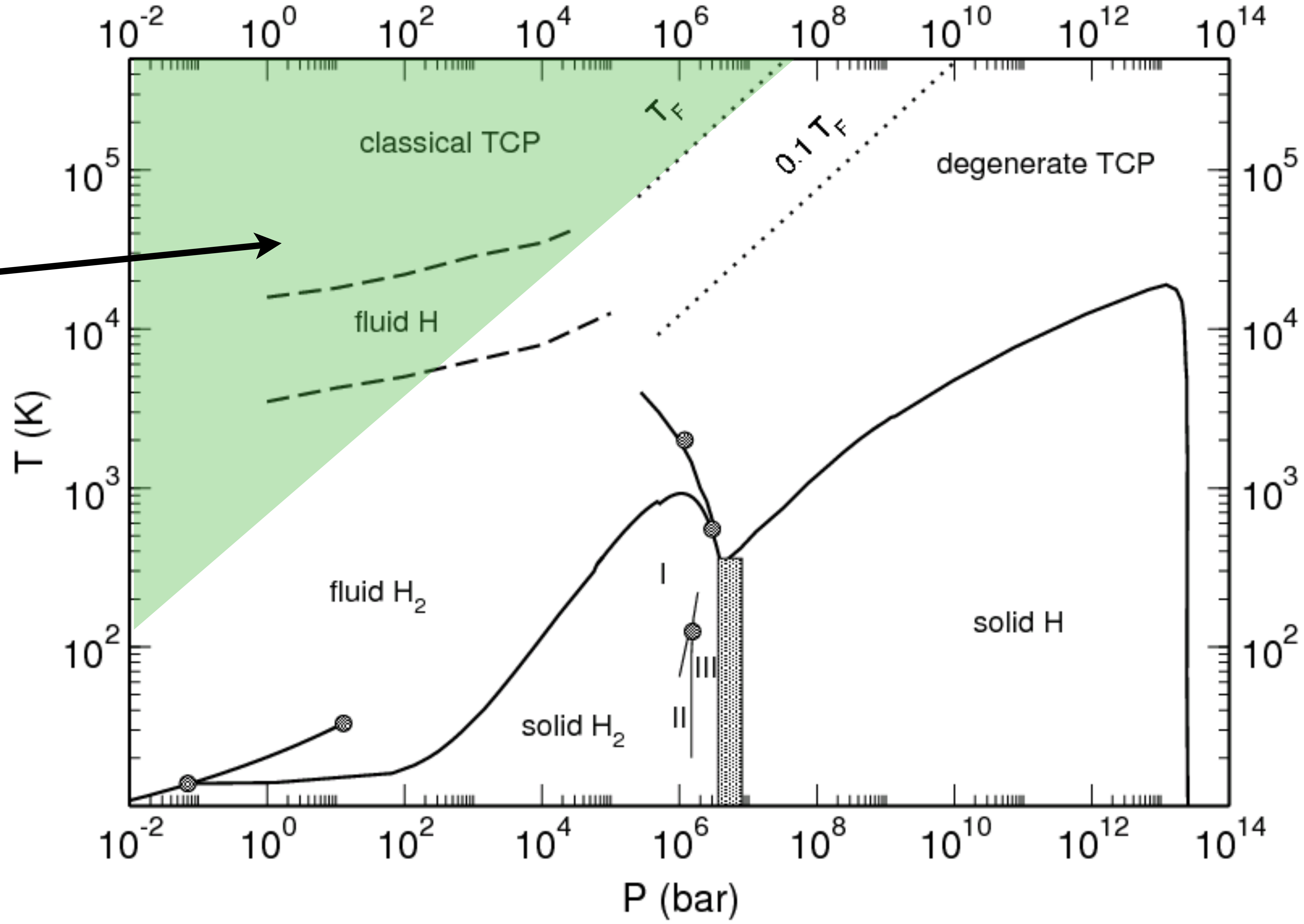
$T \gtrsim T_F$: Restricted path integral Monte Carlo

$$Z = \iint dXdR \, \langle \boldsymbol{X}, \boldsymbol{R} \, | \, e^{-\hat{H}/k_B T} \, | \, \boldsymbol{X}, \boldsymbol{R} \rangle$$

Stat-Mech problem of ring-polymers



Pierleoni et al, PRL 1994



Limited to high temperature low density region by the Fermion sign problem

$0 < T \ll T_F$: a classical-quantum coupled system

$\boldsymbol{X}$: classical proton configuration

$E(\boldsymbol{X})$: Born-Oppenheimer energy surface

Quantum

Solve $E(\boldsymbol{X})$ by DFT/VMC/QMC/...

$$E(\boldsymbol{X}) = \min_{\psi_X} \frac{\langle \psi_X | \hat{H} | \psi_X \rangle}{\langle \psi_X | \psi_X \rangle}$$

Needs a fast and accurate many-body solver as it is called repeatedly in the inner loop

Classical

Sample $\boldsymbol{X}$ with classical Monte Carlo/Molecular dynamics

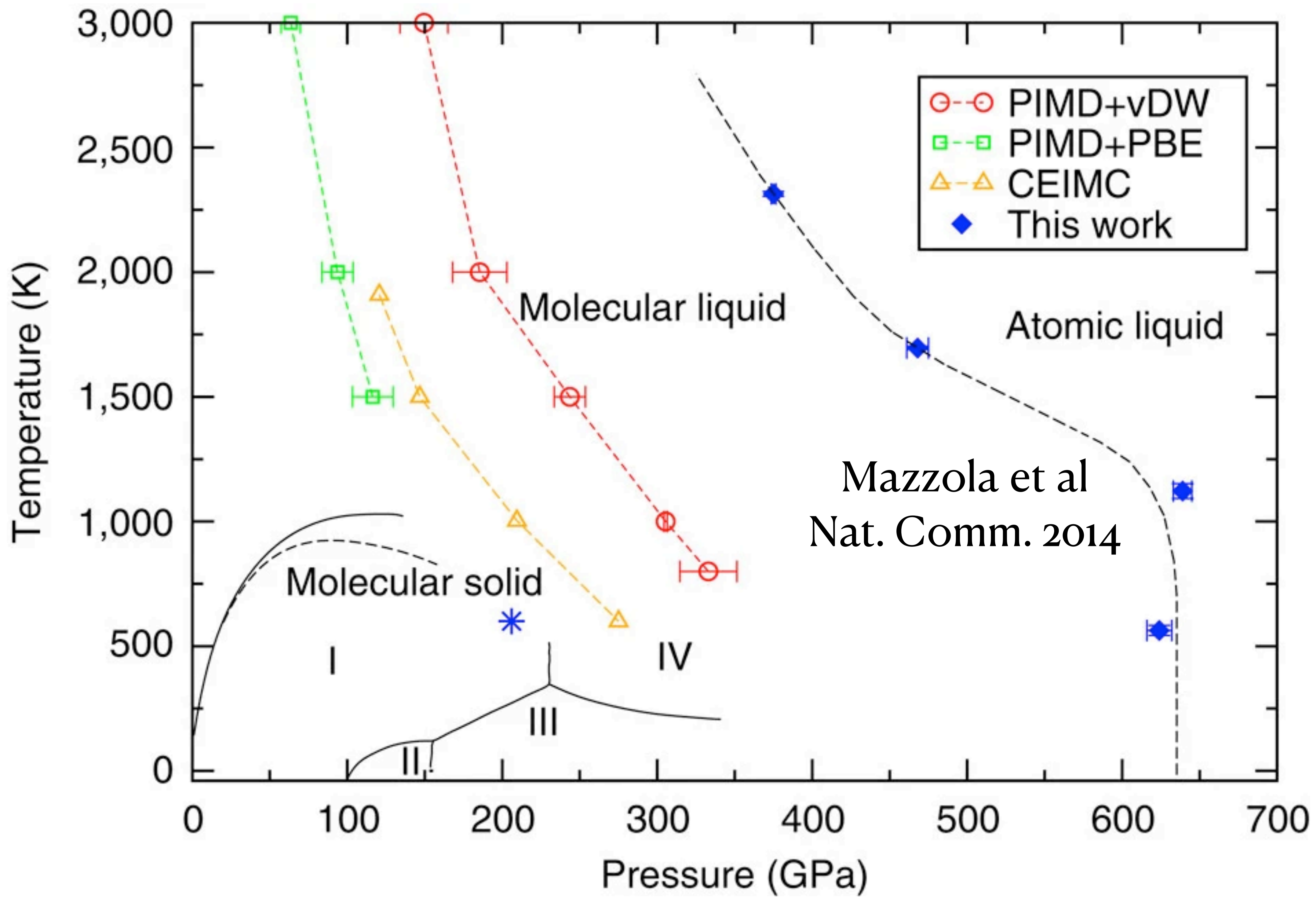
$$\min \left\{ 1, \exp \left[\frac{E(\boldsymbol{X}) - E(\boldsymbol{X}')}{k_B T} \right] \right\}$$

Tricky to sample unbiasedly with inaccurate or noisy energy estimates

Pierleoni et al, PRL 2004, Attaccalite et al, PRL 2008

$0 < T \ll T_F$: Debate on the liquid-liquid transition

Where is the transition point ?



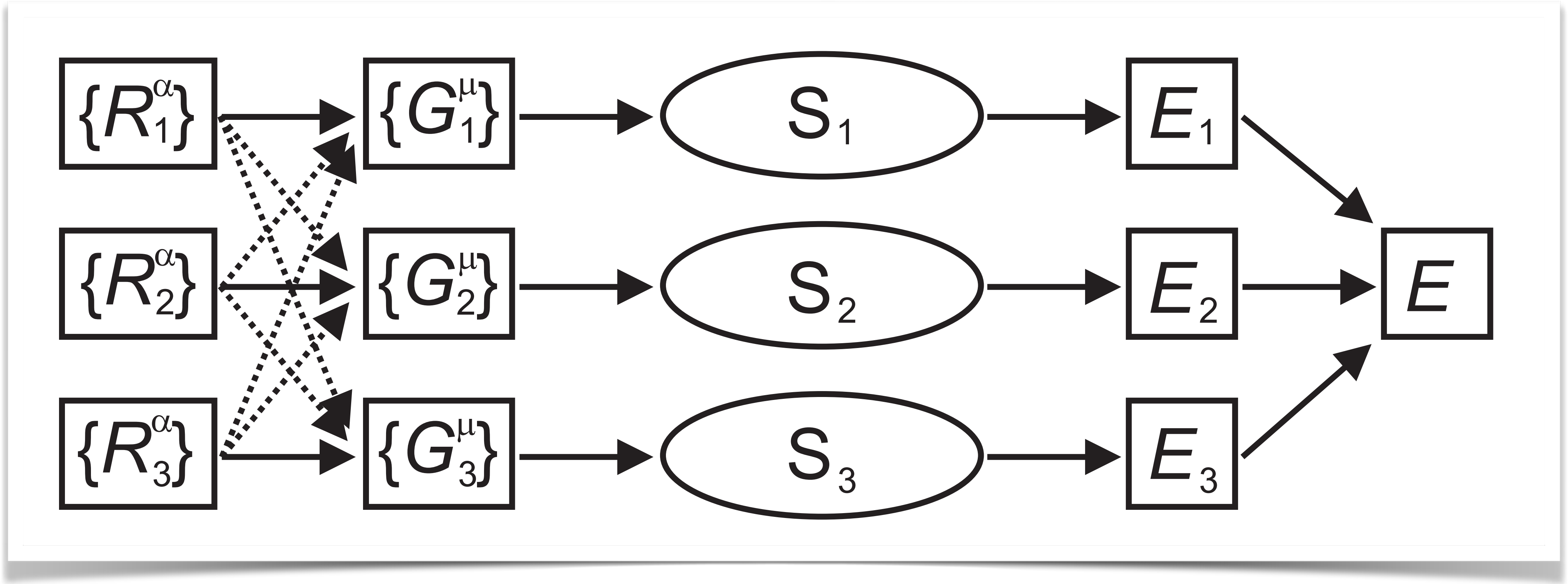
Mazzola et al
Nat. Comm. 2014

Algorithmic uncertainties coupled with finite size effect/sampling ergodicity/…

Machine learning potential

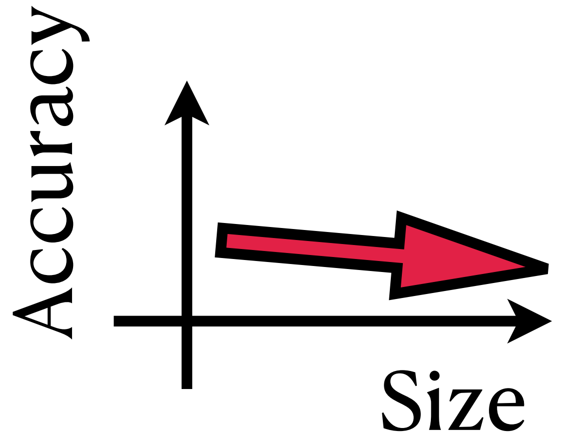
fit $E(X)$ with a ML model to DFT/VMC/QMC data

Blank, J. Chem. Phys., 1995
Behler and Parrinello, PRL 2007
...



Can reach larger system size and more samples
However, accuracy is still limited by (or worse than) DFT/VMC/QMC

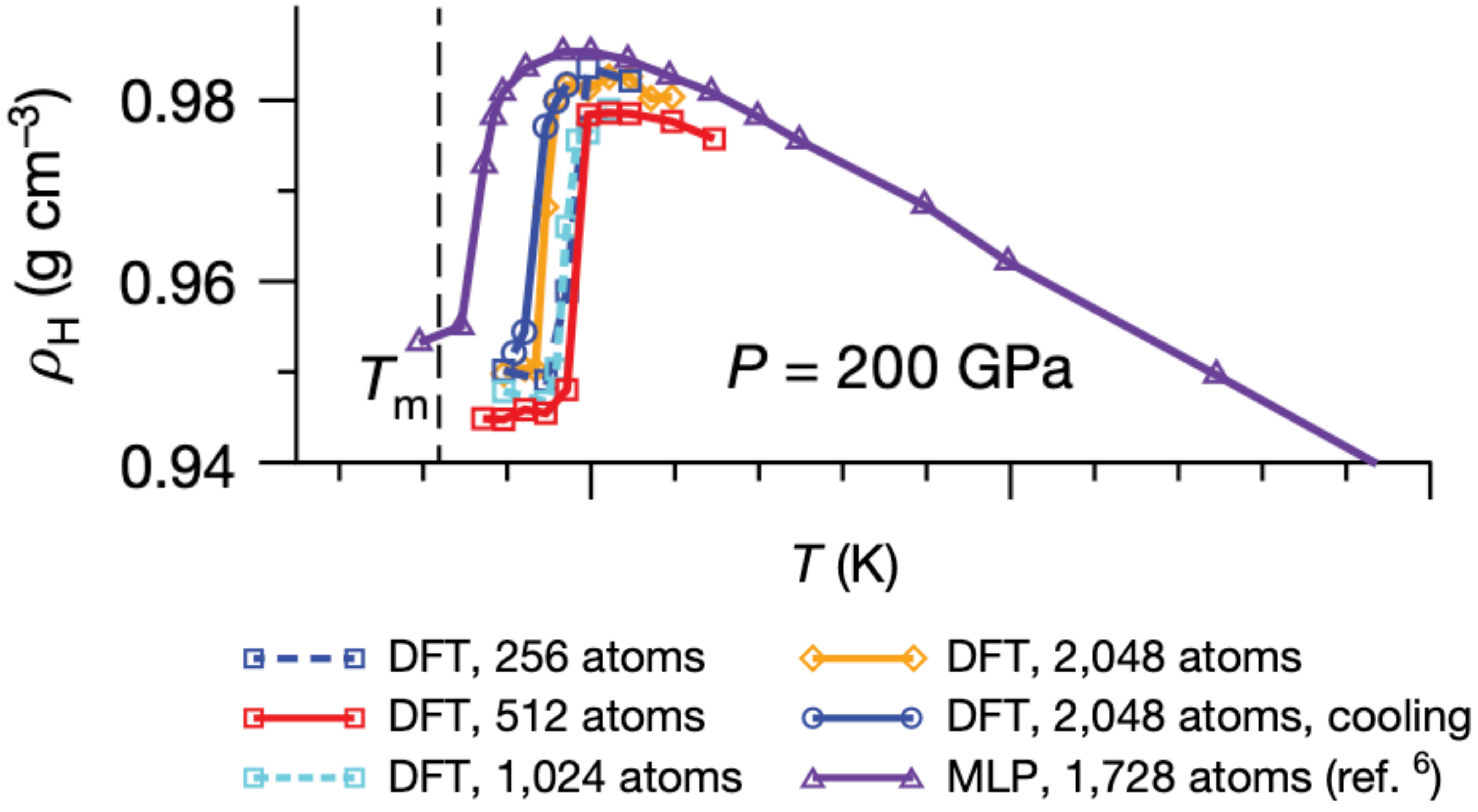
May or may not address the actual difficulty



$0 < T \ll T_F$: Debate on the liquid-liquid transition

Is it first or second order ?

Cheng et al, Nature 2020, Karasiev et al, Nature 2021



Matters arising

On the liquid–liquid phase transition of dense hydrogen

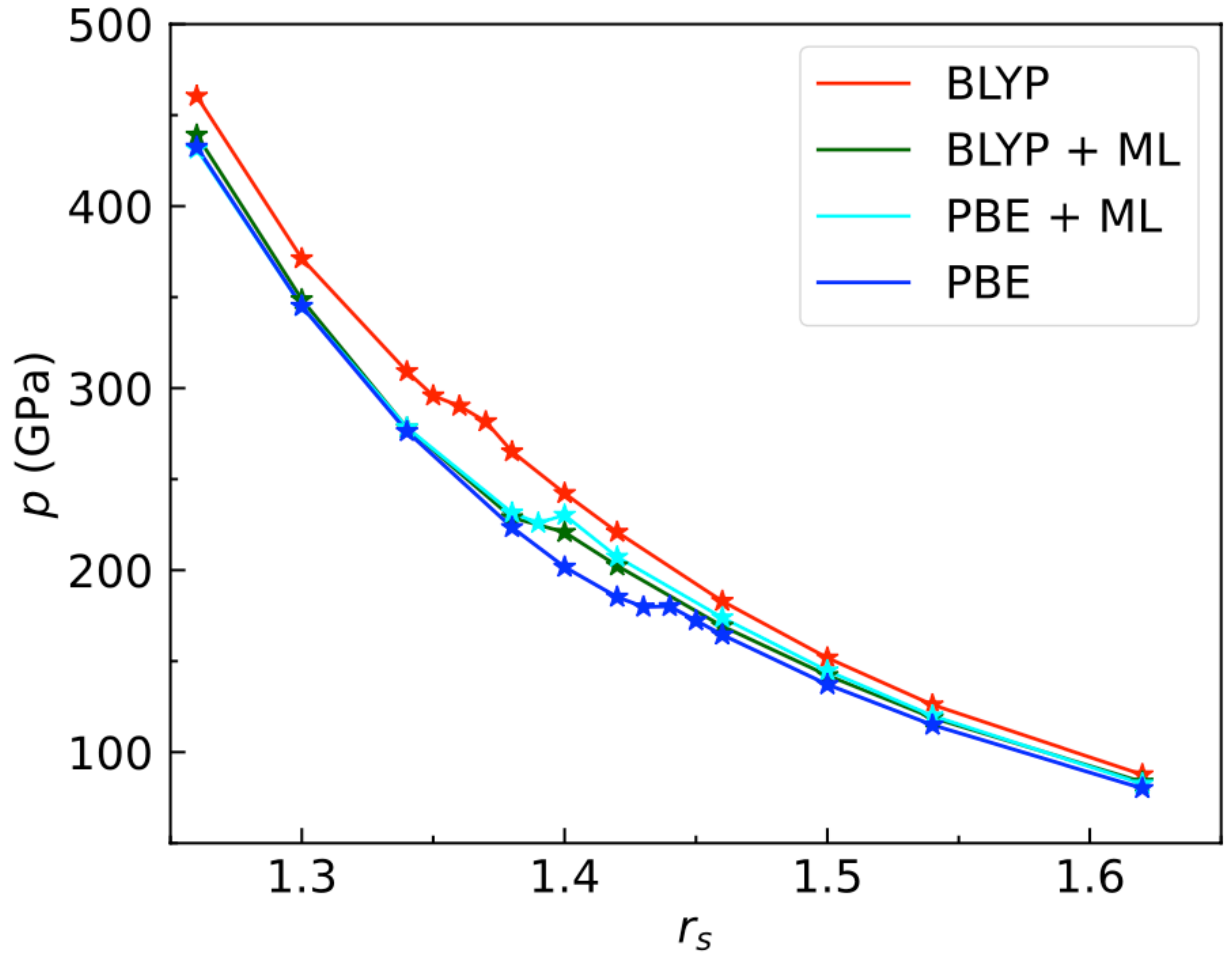
Until recently, the consensus theoretical and computational interpretation of the liquid–liquid phase transition (LLPT) of high-pressure hydrogen—which has proved challenging to determine—has been that it is first order[1–5]. Cheng et al.[6] developed a machine learning potential (MLP) that, in larger-than-previous molecular dynamics (MD) simulations, gives a continuous transition instead. We show that the MLP does not reproduce our still larger density functional theory MD (DFT-MD) calculations as it should. As the MLP is not a faithful surrogate for the DFT-MD, the prediction of a supercritical atomic liquid by Cheng et al.[6] is unfounded.

Δ-machine learning for dense hydrogen

$$E = E_{\mathrm{DFT}} + \Delta$$

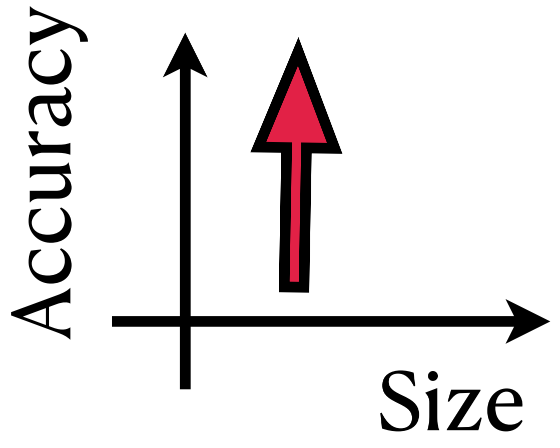
Δ is expected to be small & smooth
learn Δ from expensive & accurate QMC data

Tirelli et al, PRB 2022
Niu et al, PRL 2023



Ideally, the results will be independent of the reference

We would like to try something different



Deep variational free energy approach

Deep generative models unlocks the power of
the Gibbs–Bogolyubov-Feynman variational principle

$$F[p] = \mathop{\mathbb{E}}_{\boldsymbol{X}\sim p(\boldsymbol{X})}\Big[\underbrace{k_B T \ln p(\boldsymbol{X})}_{\text{entropy}} + \underbrace{E(\boldsymbol{X})}_{\text{energy}}\Big] \geq -k_B T \ln Z$$

Li and LW, PRL '18
Wu, LW, Zhang, PRL '19

✓ Additive statistical noises in $E(\boldsymbol{X})$ do not deteriorate stochastic optimization

✓ Turning a sampling problem to an optimization problem better leverages the deep learning engine: JAX 

Two kinds of variational Monte Carlo

Variational free energy $T > 0$

Gibbs–Bogolyubov-Feynman, Li and LW, PRL '18, Wu, LW, Zhang, PRL '19, ...

$$F[p] = \mathop{\mathbb{E}}_{\boldsymbol{X} \sim p(\boldsymbol{X})} \left[k_B T \ln p(\boldsymbol{X}) + E(\boldsymbol{X}) \right]$$

p: probabilistic models with tractable normalization

Variational ground state energy $T = 0$

McMillan 1965, Carleo & Troyer Science 2017, Pfau et al, FermiNet, ...

$$E[\psi] = \mathop{\mathbb{E}}_{\boldsymbol{R} \sim |\psi(\boldsymbol{R})|^2} \left[\frac{\hat{H}\psi(\boldsymbol{R})}{\psi(\boldsymbol{R})} \right]$$

ψ: ANY neural network that respects physical symmetries

See talks by Jannes Nys and Markus Heyl

Why does normalization matter?

Suppose $$p(\boldsymbol{X}) = \frac{e^{-E_\theta(\boldsymbol{X})/k_B T}}{Z_\theta}$$

"Boltzmann machine"
or, energy-based model

We have

$$F[p] = \mathop{\mathbb{E}}_{\boldsymbol{X} \sim p(\boldsymbol{X})} \left[E(\boldsymbol{X}) - E_\theta(\boldsymbol{X}) \right] - k_B T \ln Z_\theta \geq -k_B T \ln Z$$

Intractable!

Deep variational free energy approach

Deep generative models unlocks the power of
the Gibbs–Bogolyubov-Feynman variational principle

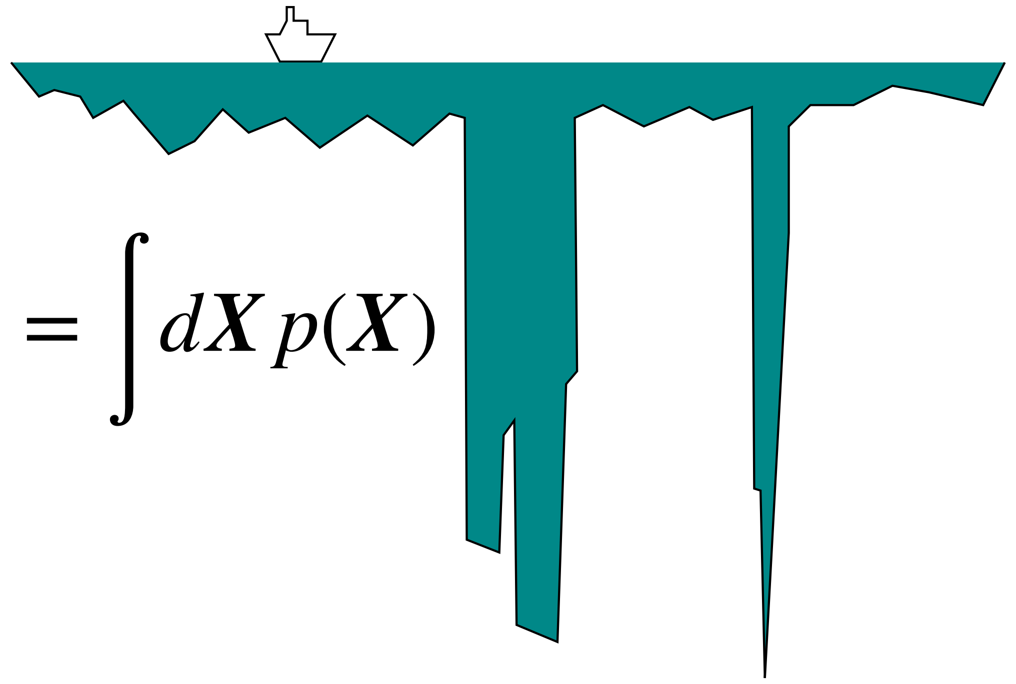
$$F[p] = \mathop{\mathbb{E}}_{\boldsymbol{X}\sim p(\boldsymbol{X})} \Big[\underbrace{k_B T \ln p(\boldsymbol{X})}_{\text{entropy}} + \underbrace{E(\boldsymbol{X})}_{\text{energy}}\Big] \geq -k_B T \ln Z$$

Li and LW, PRL '18
Wu, LW, Zhang, PRL '19

Tractable normalization

Mackay, Information Theory, Inference, and Learning Algorithms

$$1 = \int d\boldsymbol{X}\, p(\boldsymbol{X})$$



Direct sampling

$$\mathop{\mathbb{E}}_{\boldsymbol{X}\sim p(\boldsymbol{X})}$$

Krauth, Statistical Mechanics: Algorithms and Computations

Deep generative models

Autoregressive model

$$p(\boldsymbol{X}) = p(x_1)p(x_2\,|\,x_1)p(x_3\,|\,x_1, x_2)\cdots$$

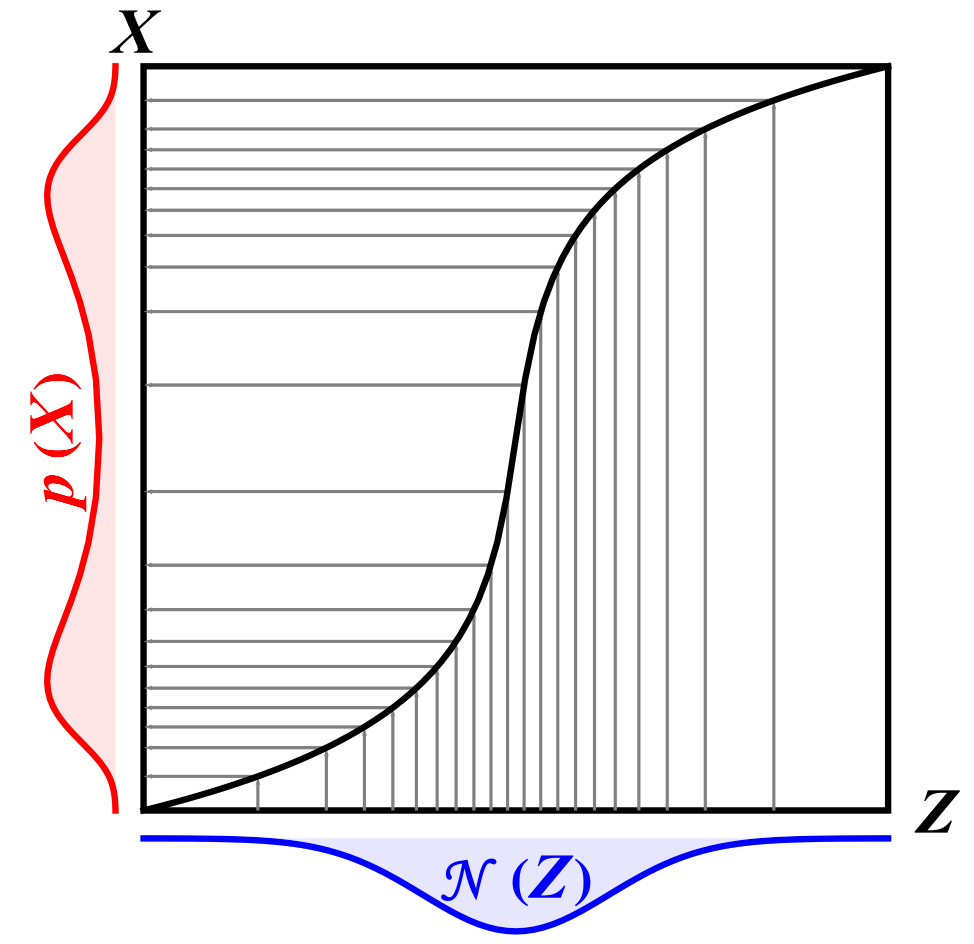
"... the murderer is ___ "

$p(\_\,|\,\ldots)$

Implementation: transformer with causal mask...

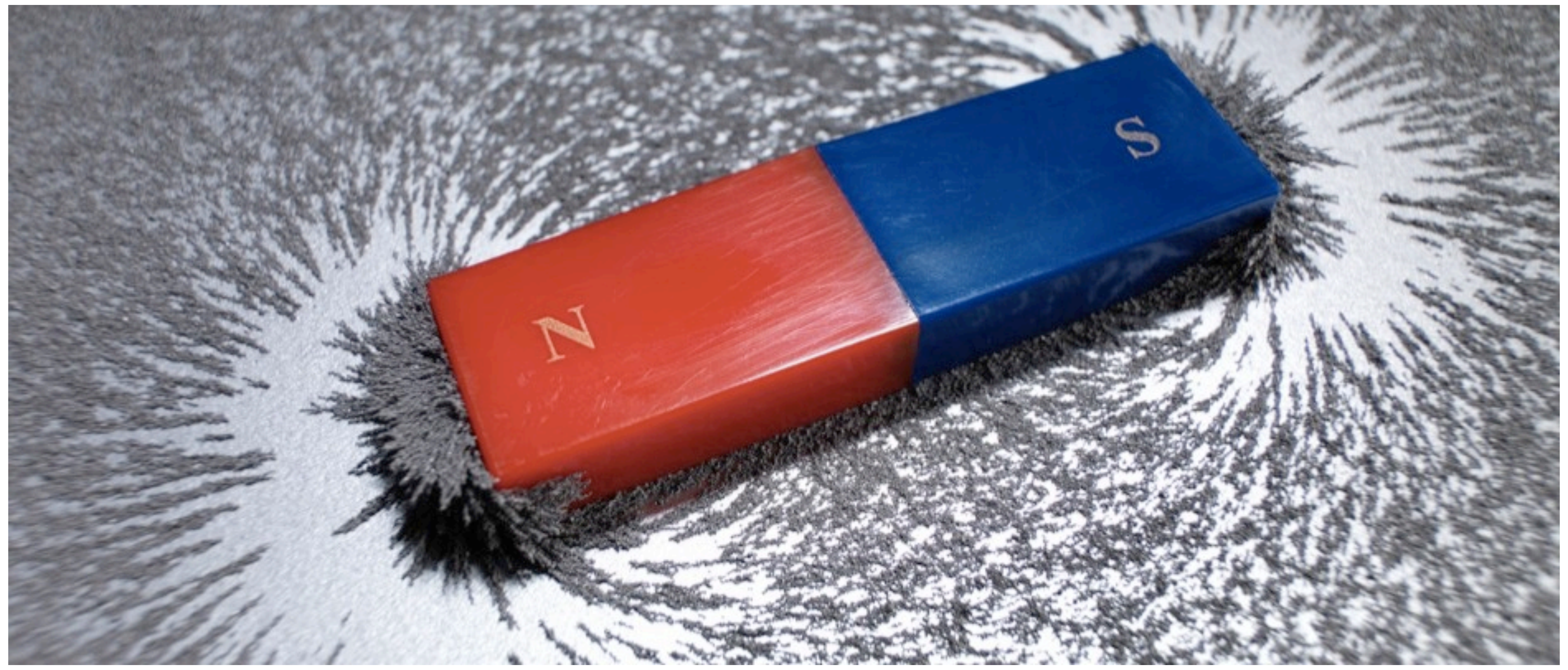
Normalizing flow

$$p(\boldsymbol{X}) = \mathcal{N}(\boldsymbol{Z})\left|\det\left(\frac{\partial \boldsymbol{Z}}{\partial \boldsymbol{X}}\right)\right|$$



Implementation: invertible Resnet (backflow)...

Variational free energy



Known: (noisy) energy function

Unknown: samples

"learn from Hamiltonian"

$$\min_{\theta} \mathbb{KL}(p_\theta \,\|\, e^{-E/k_B T})$$

Maximum likelihood estimation

Known: samples

Unknown: generating distribution

"learn from data"

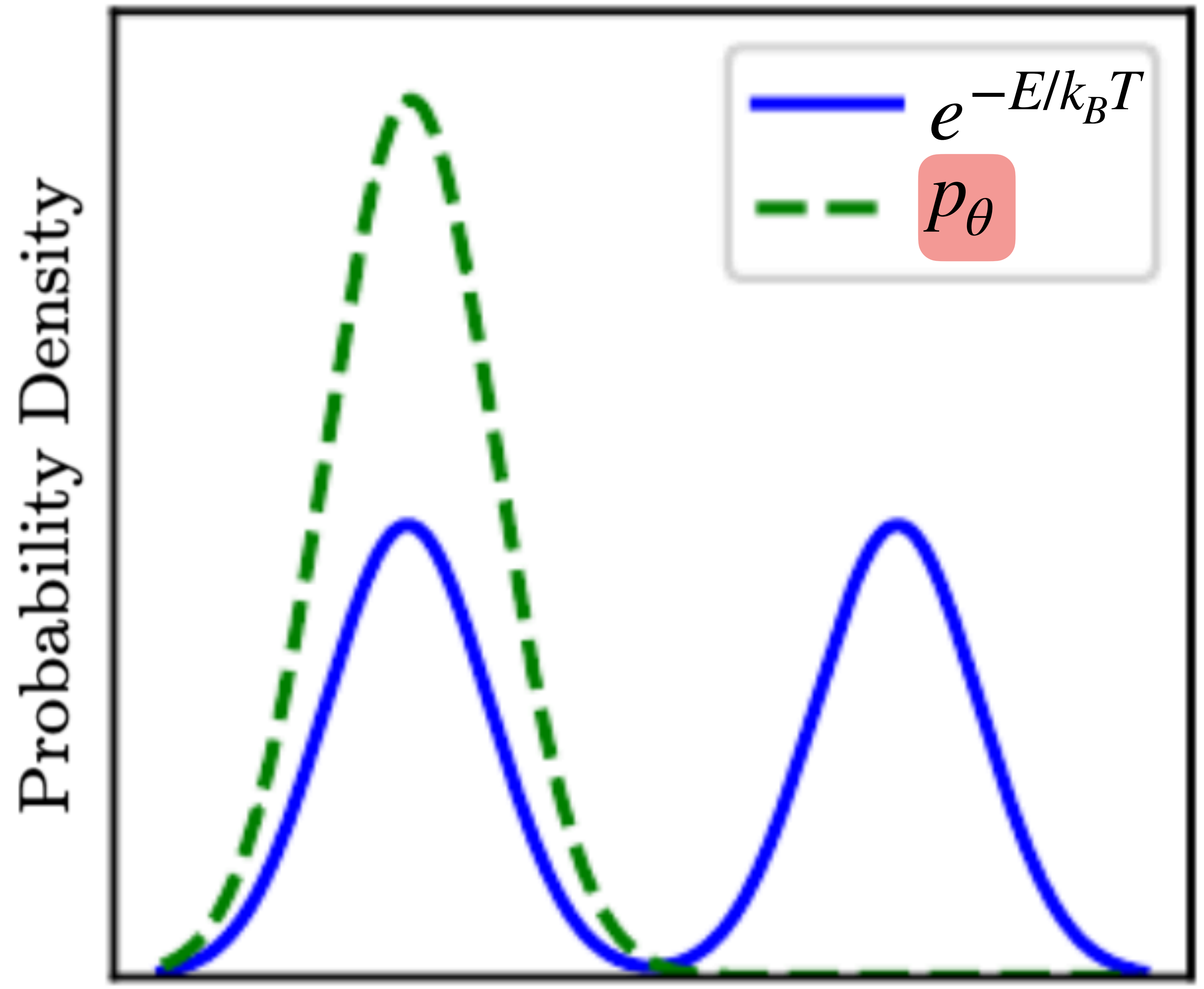
$$\min_{\theta} \mathbb{KL}(\text{data} \,\|\, p_\theta)$$

Two sides of the same coin

Pros and cons

$$\min_{\theta} \mathbb{KL}(p_\theta \,\|\, e^{-E/k_B T})$$

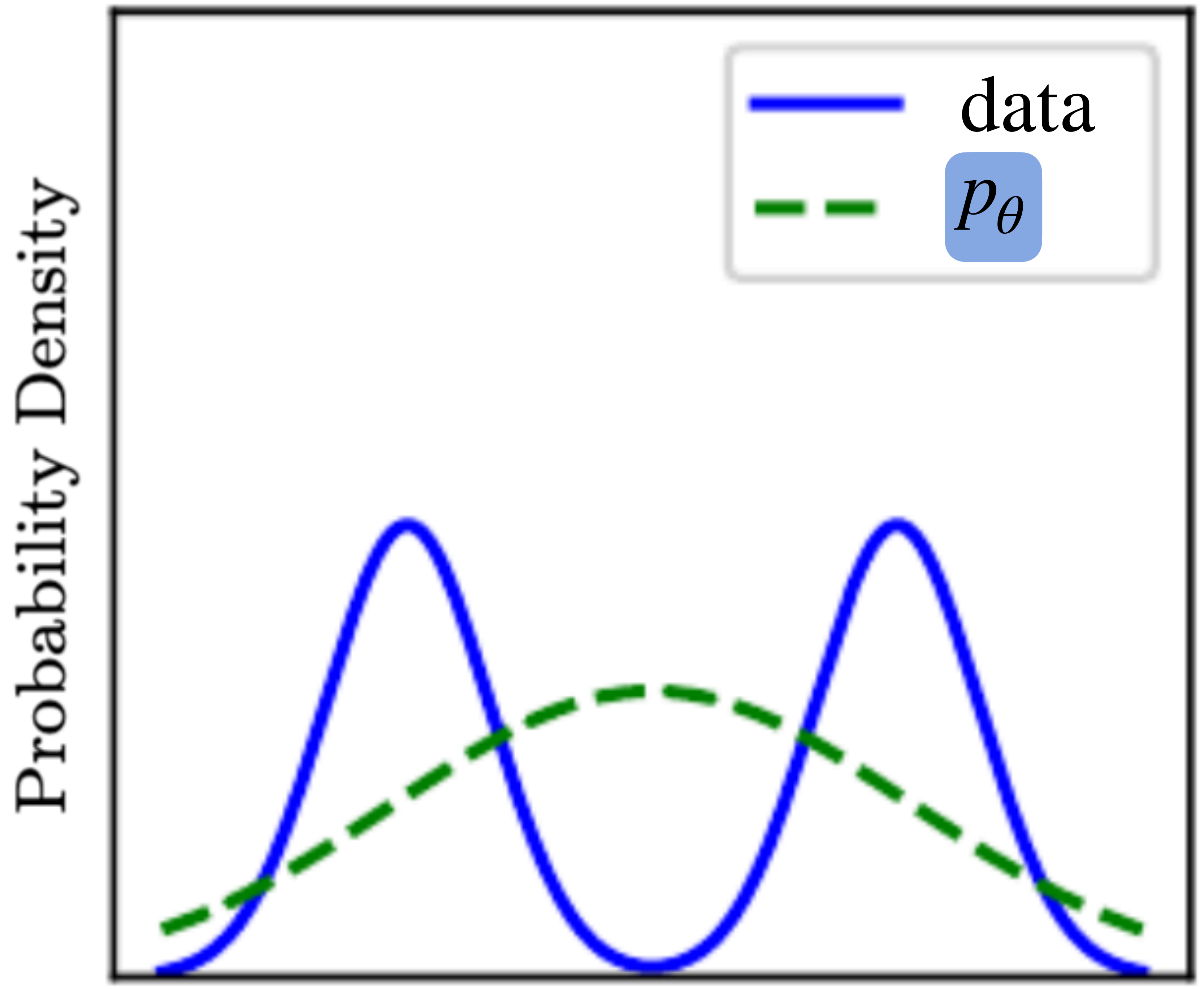
Mode seeking



Failure mode: local minima

$$\min_{\theta} \mathbb{KL}(\text{data} \,\|\, p_\theta)$$

Mode covering



Failure mode: hallucination

Goodfellow et al, Deep Learning

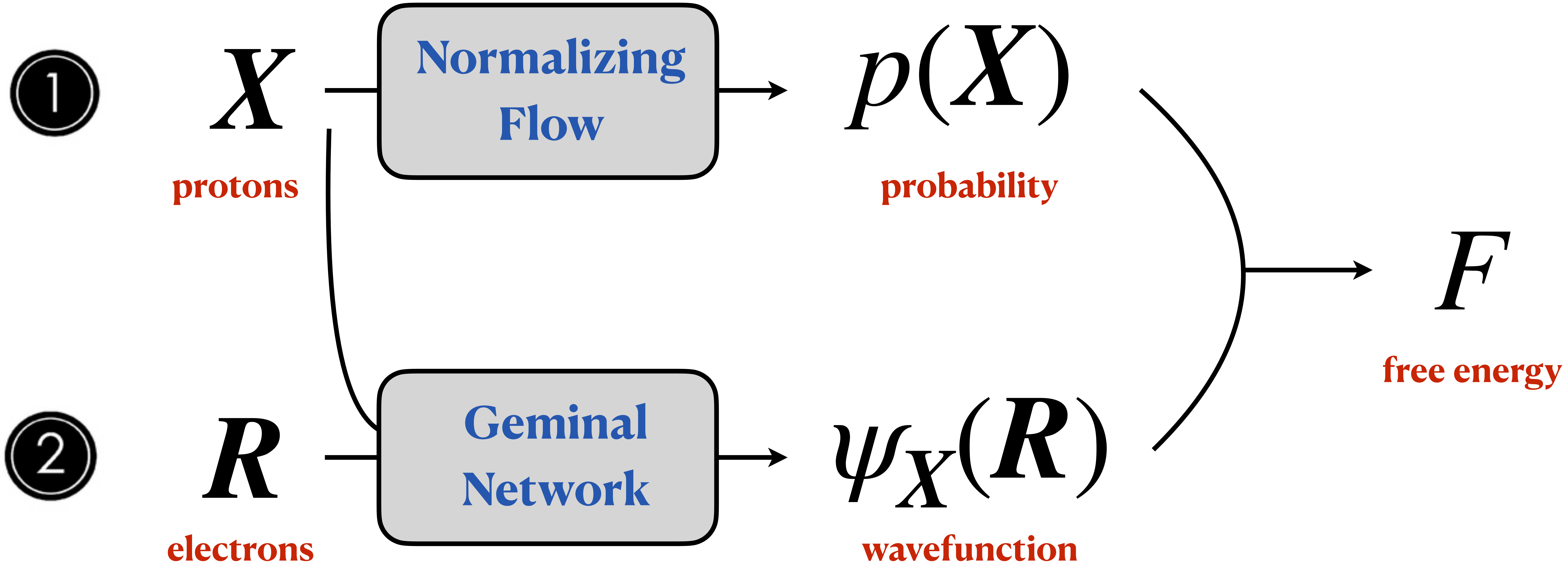
"Jack of all trades, master of none" — 2302.10724

filling the gap vs pushing the boundary of human knowledge

Deep variational free energy for dense hydrogen

Xie, Li, Wang, Zhang, LW, 2209.06095

$$F = \mathbb{E}_{\boldsymbol{X}\sim p(\boldsymbol{X})}\left[k_B T \ln p(\boldsymbol{X}) + \mathbb{E}_{\boldsymbol{R}\sim|\psi_{\boldsymbol{X}}(\boldsymbol{R})|^2}\left[\frac{\hat{H}\psi_{\boldsymbol{X}}(\boldsymbol{R})}{\psi_{\boldsymbol{X}}(\boldsymbol{R})}\right]\right]$$



❶ Normalizing flow for proton distribution

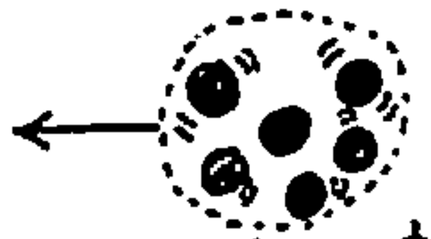
$$p(\boldsymbol{X}) = \frac{1}{L^3}\left|\det\left(\frac{\partial \boldsymbol{Z}}{\partial \boldsymbol{X}}\right)\right|$$

$\boldsymbol{X} \leftrightarrow \boldsymbol{Z}$: an invertible equivariant neural net

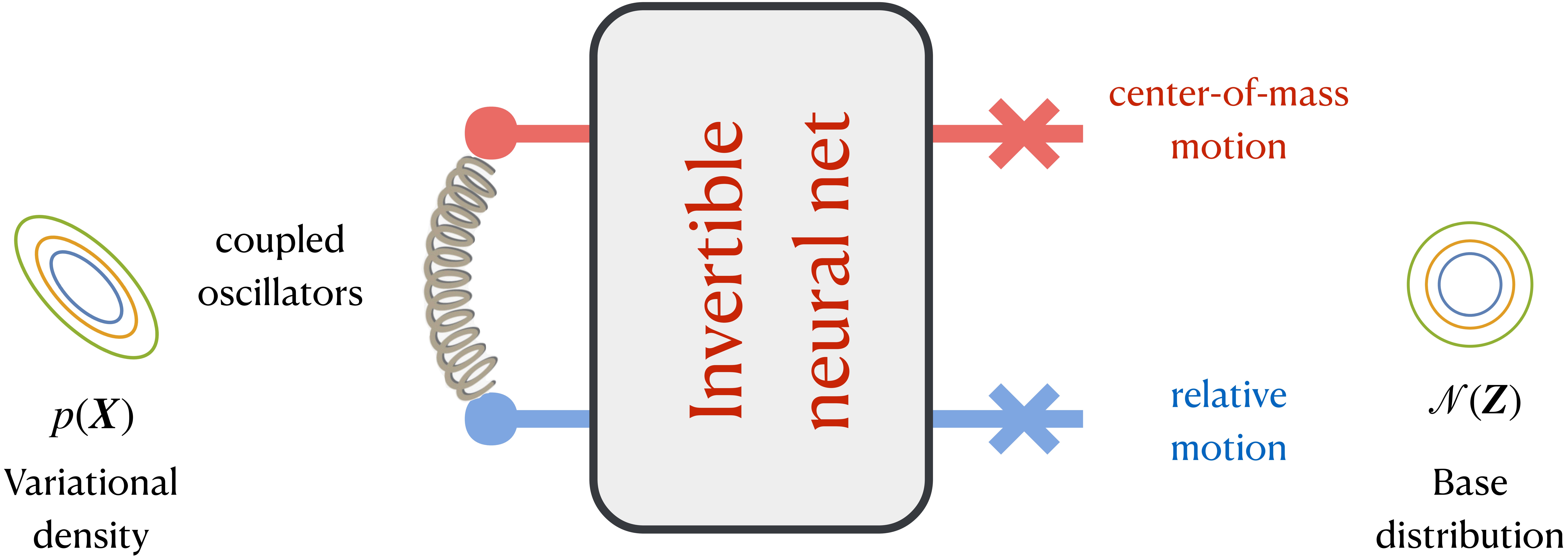
$\boldsymbol{X}$: proton coordinates $\boldsymbol{Z}$: uniform random variables


real particle

$X + \mathrm{NN}(X) = Z$


quasi particle

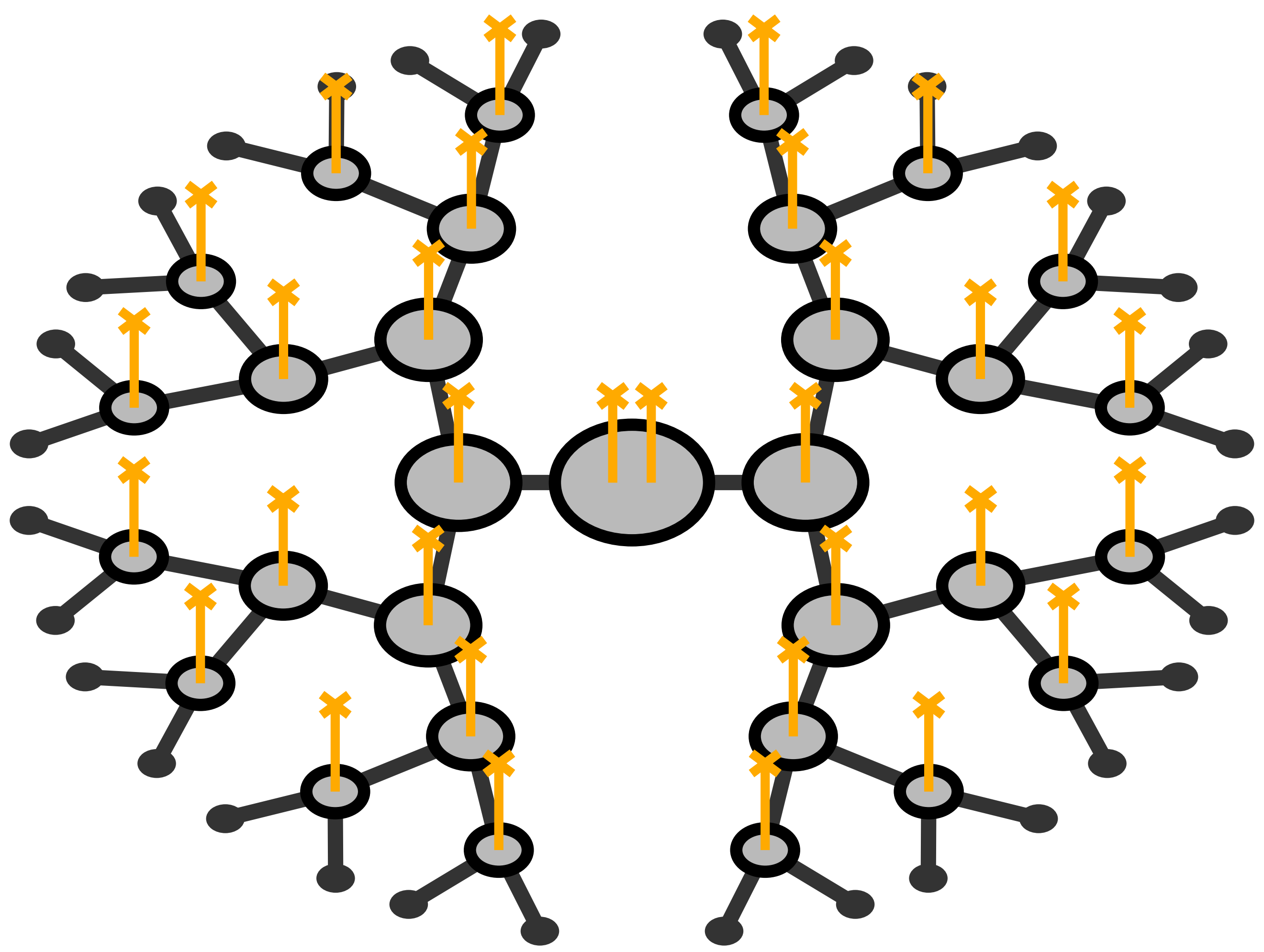
Physics intuition for normalizing flow



High-dimensional, composable, learnable, nonlinear transformations

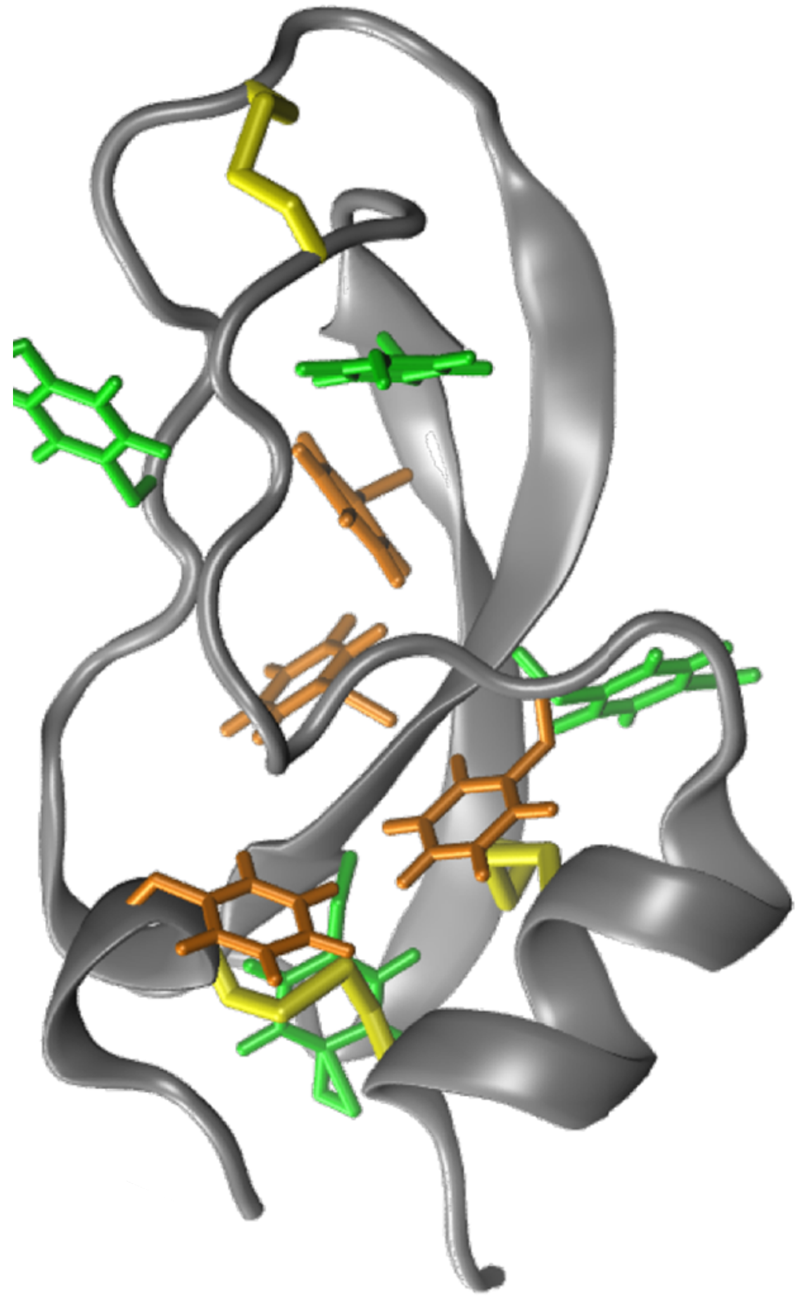
Normalizing flow in physics

Renormalization group



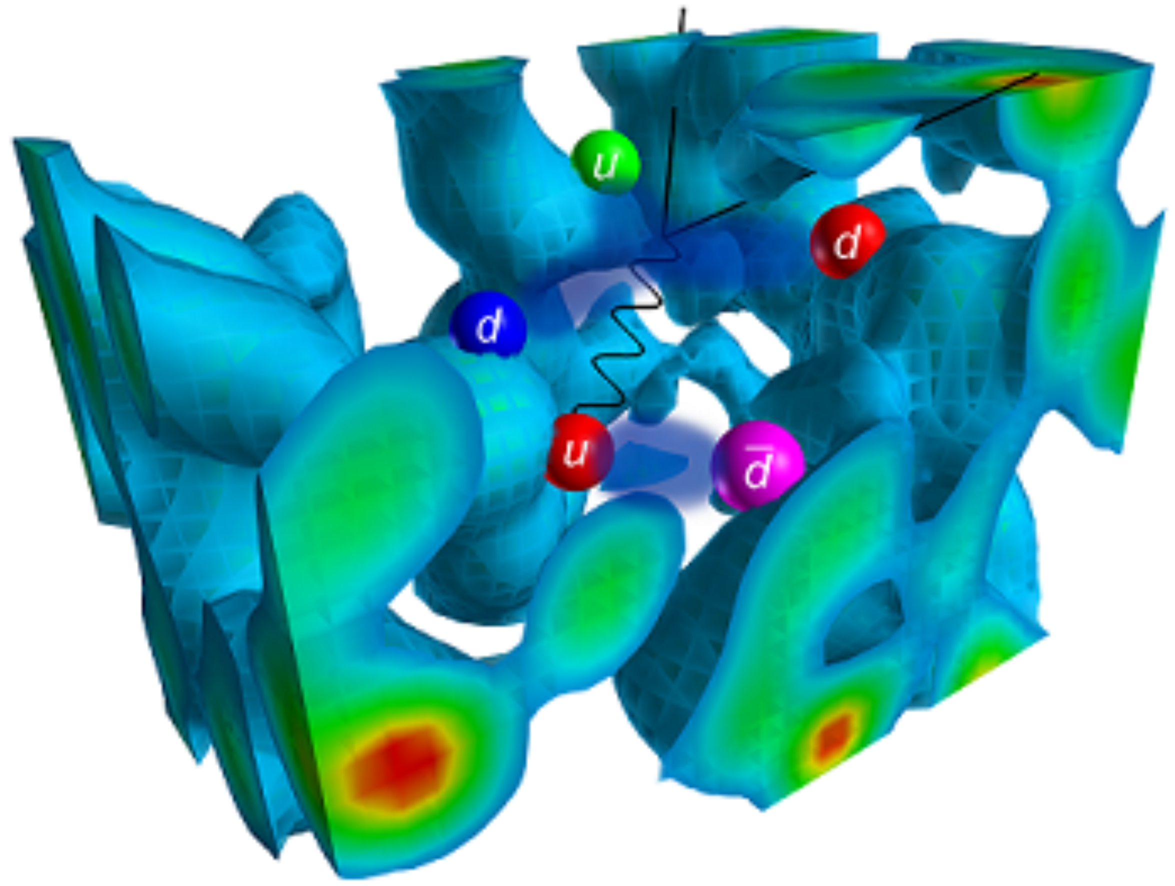
Li and LW, PRL '18
Li, Dong, Zhang, LW, PRX '20

Molecular simulation



Noe et al, Science '19
Wirnsberger et al, JCP '20

Lattice field theory



Albergo et al, PRD '19
Kanwar et al, PRL '20

② Geminal network

Xie, Li, Wang, Zhang, LW, 2209.06095

$$\psi_X(\boldsymbol{R}) = e^J \det G$$

Jastrow $J = \sum_{i,\mu} f^X_{i\mu} b_\mu$

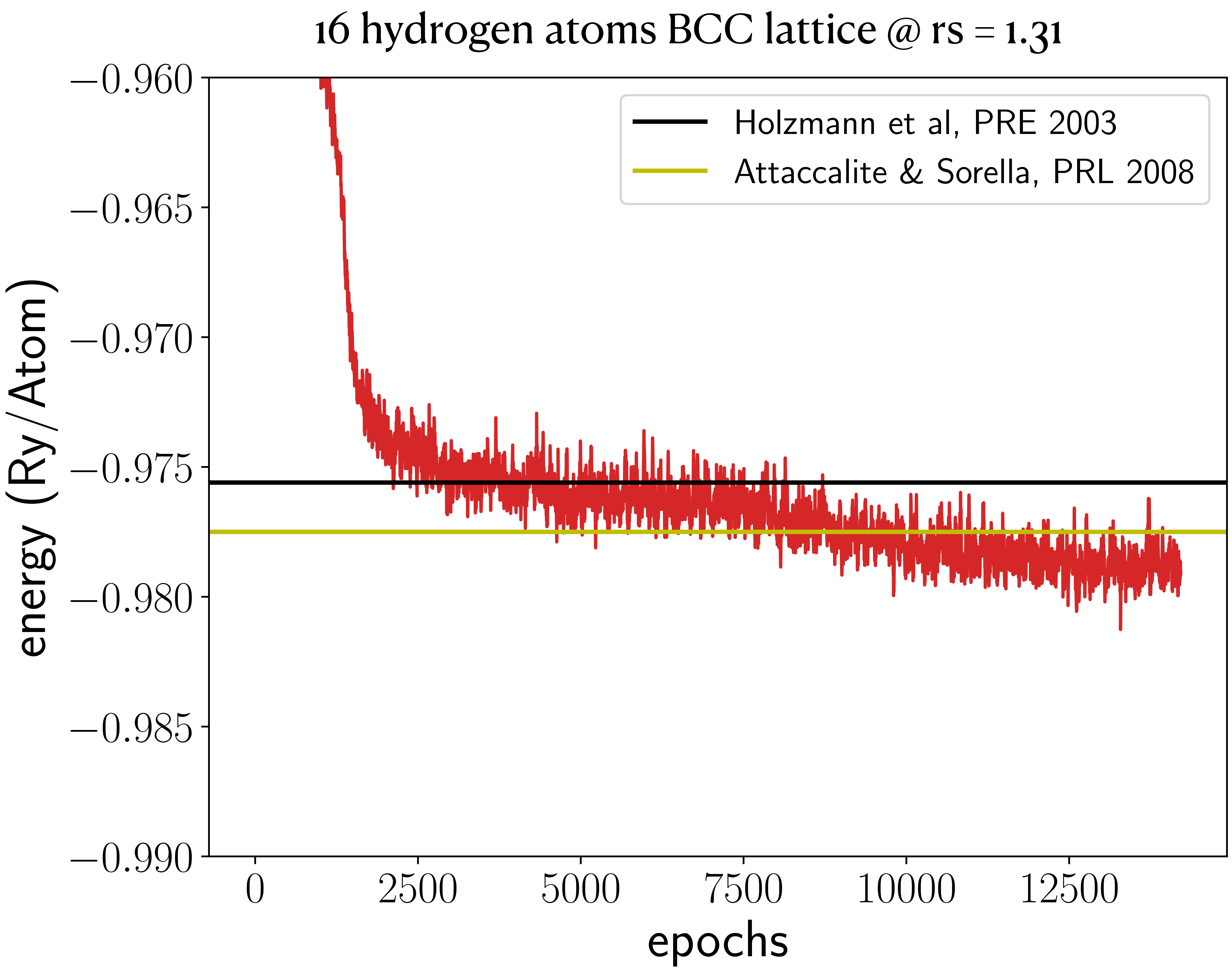
Captures atomic, molecular, and superconducting state

Bouchaud et al, '88
Casula et al, '03
Lou et al, 2305.06989

$$G_{ij} = \underbrace{\boxed{f^{\uparrow}_{i\mu} \quad \forall \mu \in 1 \cdots M}}_{\frac{N}{2} \times M} \cdot W_{\mu\nu} \cdot \underbrace{\boxed{f^{\downarrow}_{j\nu} \quad \forall \nu \in 1 \cdots M}}_{M \times \frac{N}{2}}$$

Equivariant features $\boldsymbol{f}^X, \boldsymbol{f}^\uparrow, \boldsymbol{f}^\downarrow = \mathrm{FermiNet}(\boldsymbol{X}, \boldsymbol{R}^\uparrow, \boldsymbol{R}^\downarrow)$ Pfau et al, PRR '20

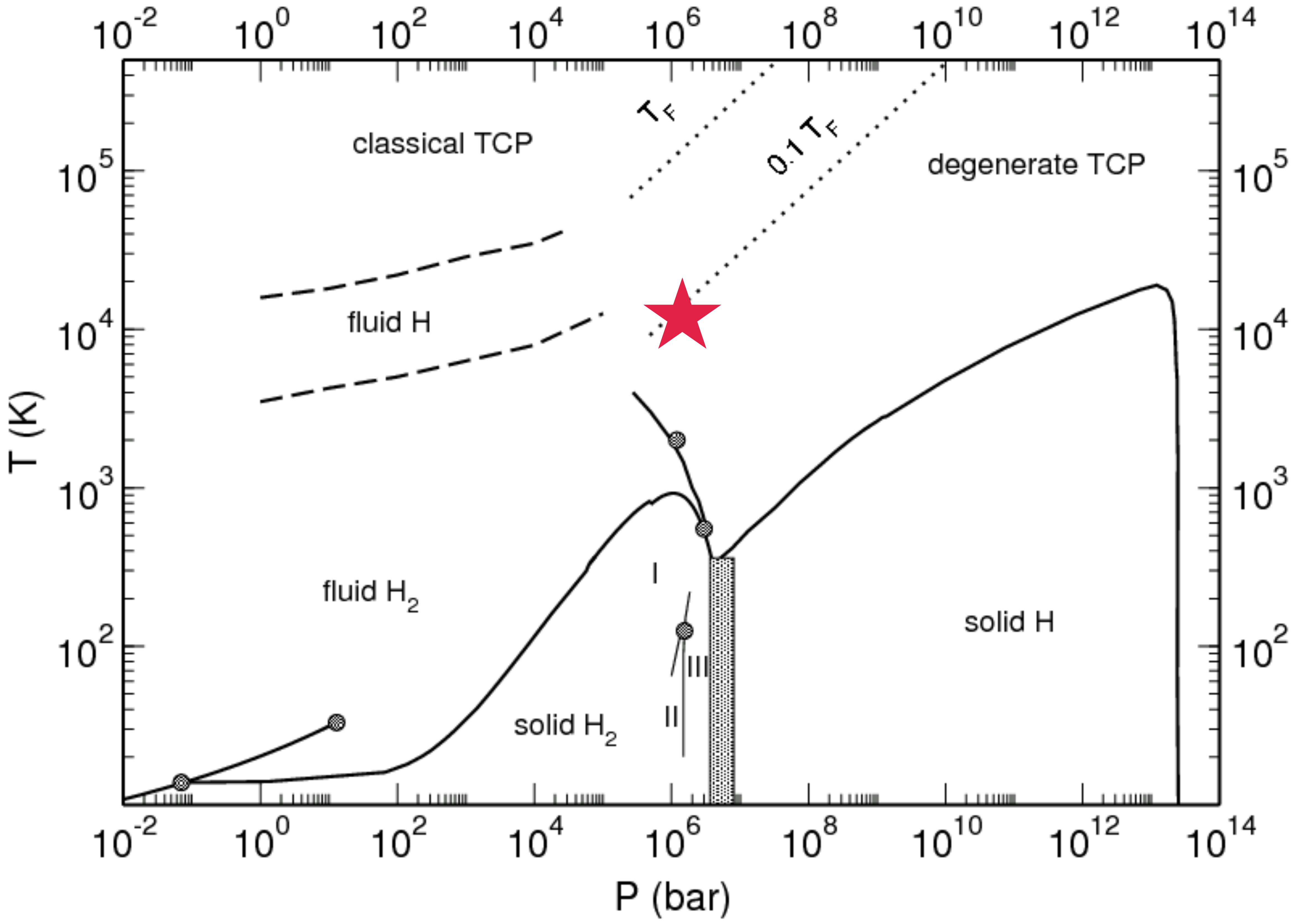
Variational ground state benchmark



This tests the quality of variational wavefunction

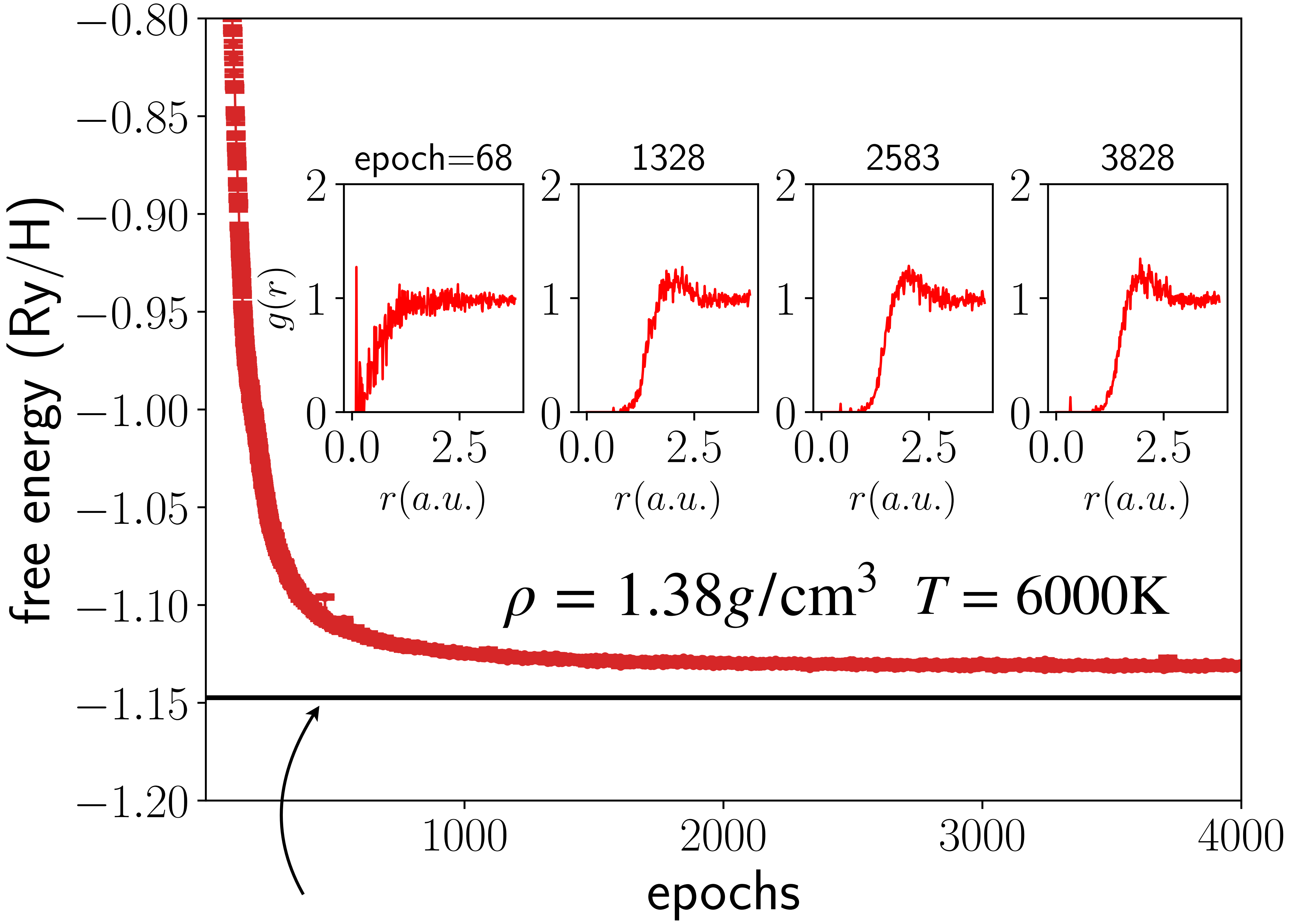
See also: Pfau et al, PRR '20, Li et al, Nat. Comm. '22

Variational free energy of dense hydrogen



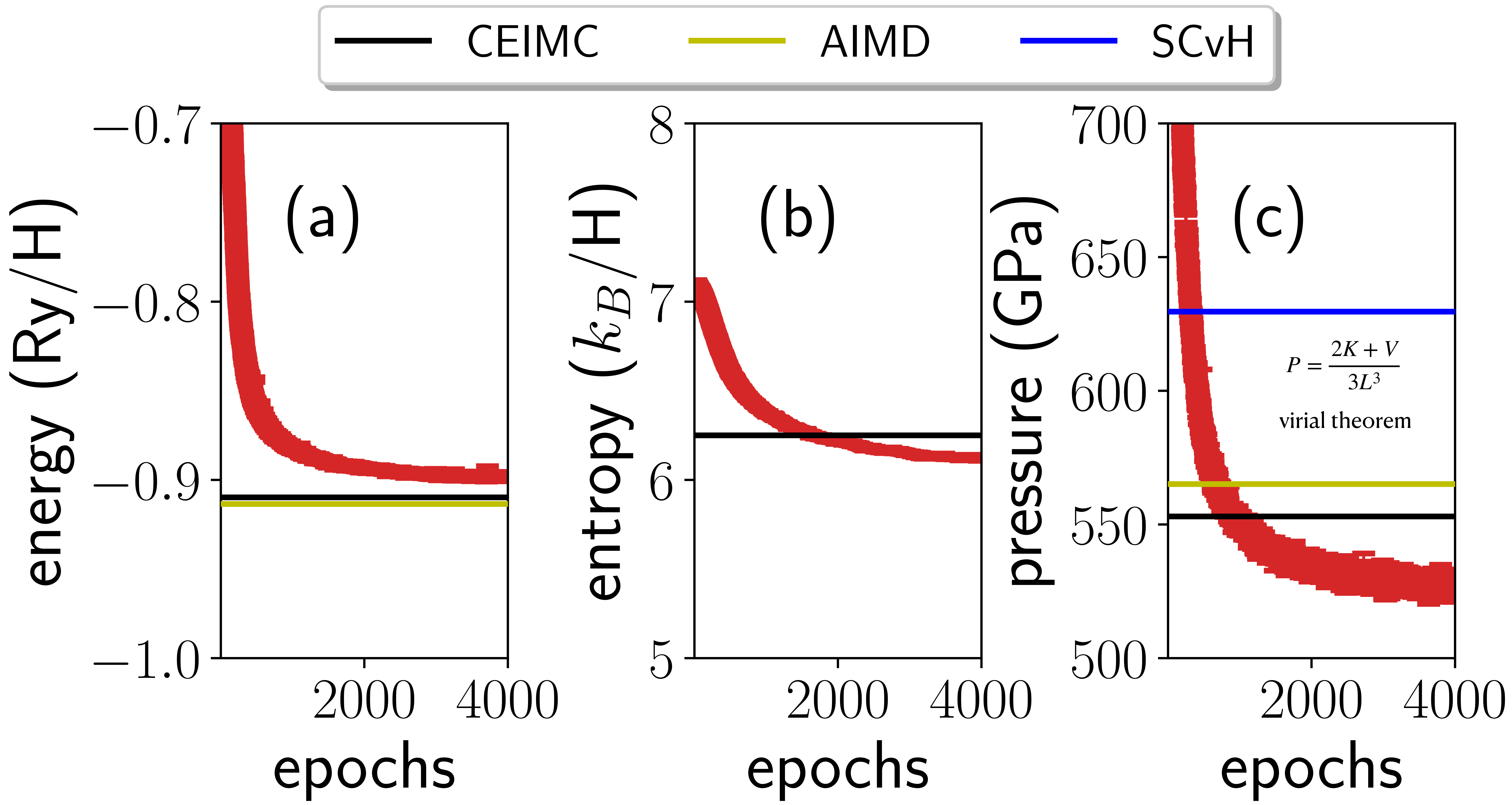
The only parameter point in the literature with published free energy value

Xie, Li, Wang, Zhang, LW, 2209.06095



Morales et al, PRE '10: two stage thermodynamic integration: ideal gas -> Yukawa gas -> Hydrogen

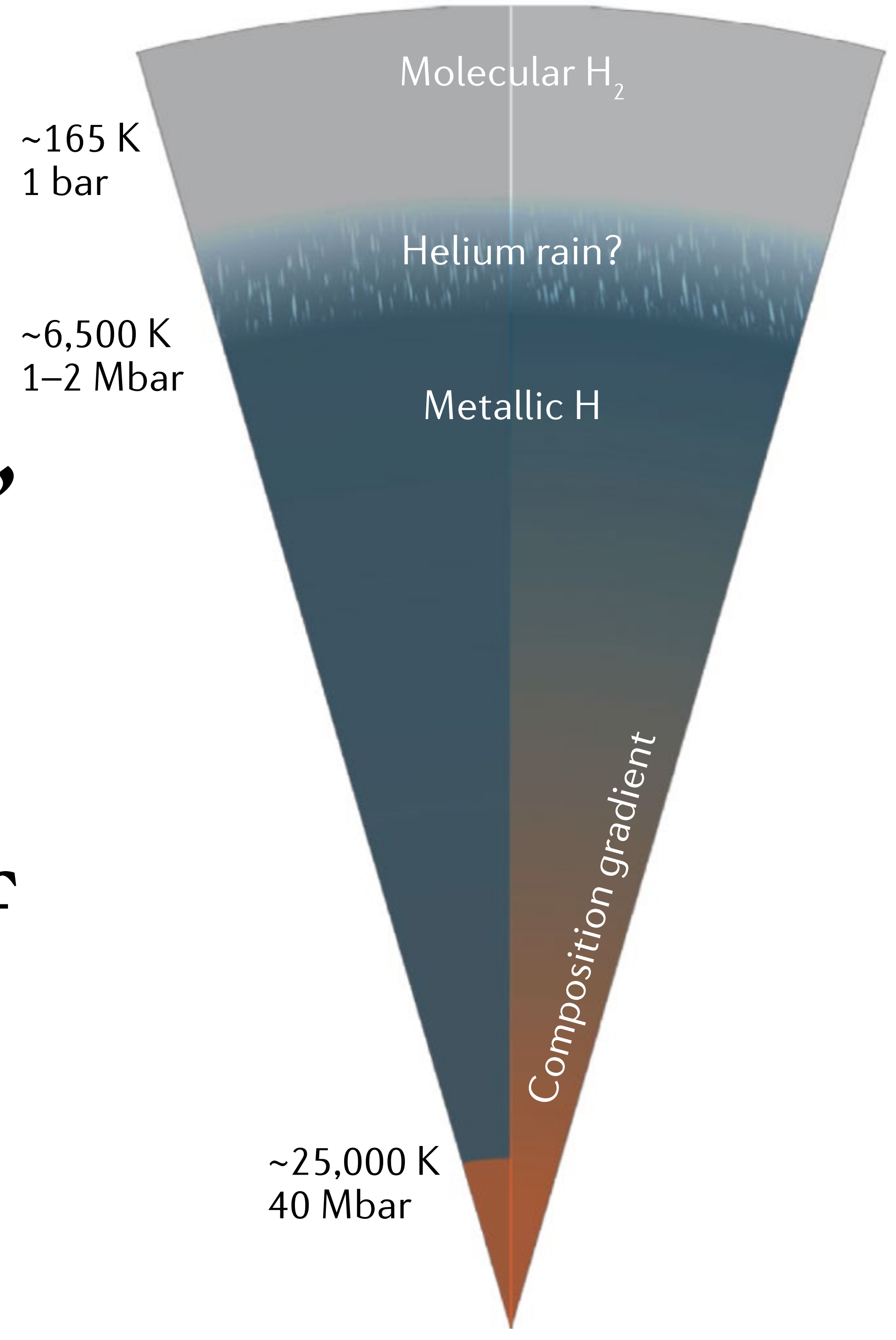
54 hydrogen atoms with twist-averaged boundary condition



CEIMC
AIMD
SCvH
(a)
(b)
(c)
energy (Ry/H)
entropy (k_B/H)
pressure (GPa)
epochs
$P = \frac{2K + V}{3L^3}$
virial theorem
−0.7
−0.8
−0.9
−1.0
8
7
6
5
700
650
600
550
500
2000
4000

Discussions

- Our calculation shows even denser equation-of-state compared to previous results. The prediction can be systematically improved with lowering the variational free energy.
- The predicted equation of state is relevant for planet modeling, where direct access to entropy is welcoming.
- This is an “uninteresting” point in the phase diagram: a soup of H^+, e^-, and H. No phase transition or other fancy physics.



Inject physics knowledge into the flow

Uninformative uniform base distribution

$$p(\boldsymbol{X}) = \boxed{\frac{1}{L^3}} \left| \det \left(\frac{\partial \boldsymbol{Z}}{\partial \boldsymbol{X}} \right) \right|$$

Absolute variational free energy for normalized variational density

$$F = \mathop{\mathbb{E}}_{\boldsymbol{X} \sim p(\boldsymbol{X})} \left[k_B T \ln p(\boldsymbol{X}) + \mathop{\mathbb{E}}_{\boldsymbol{R} \sim |\psi_{\boldsymbol{X}}(\boldsymbol{R})|^2} \left[\frac{\hat{H} \psi_{\boldsymbol{X}}(\boldsymbol{R})}{\psi_{\boldsymbol{X}}(\boldsymbol{R})} \right] \right]$$

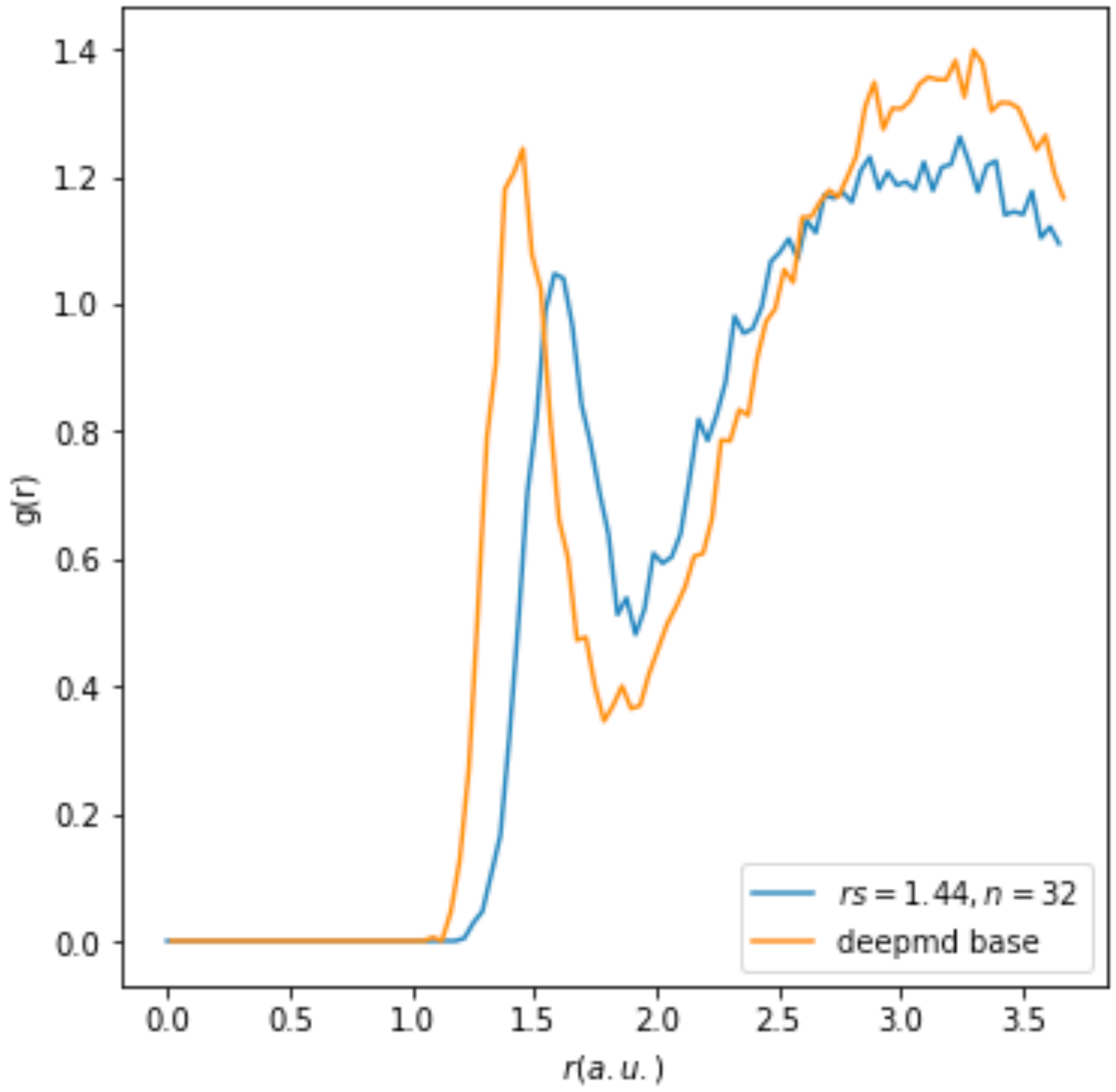
Inject physics knowledge into the flow

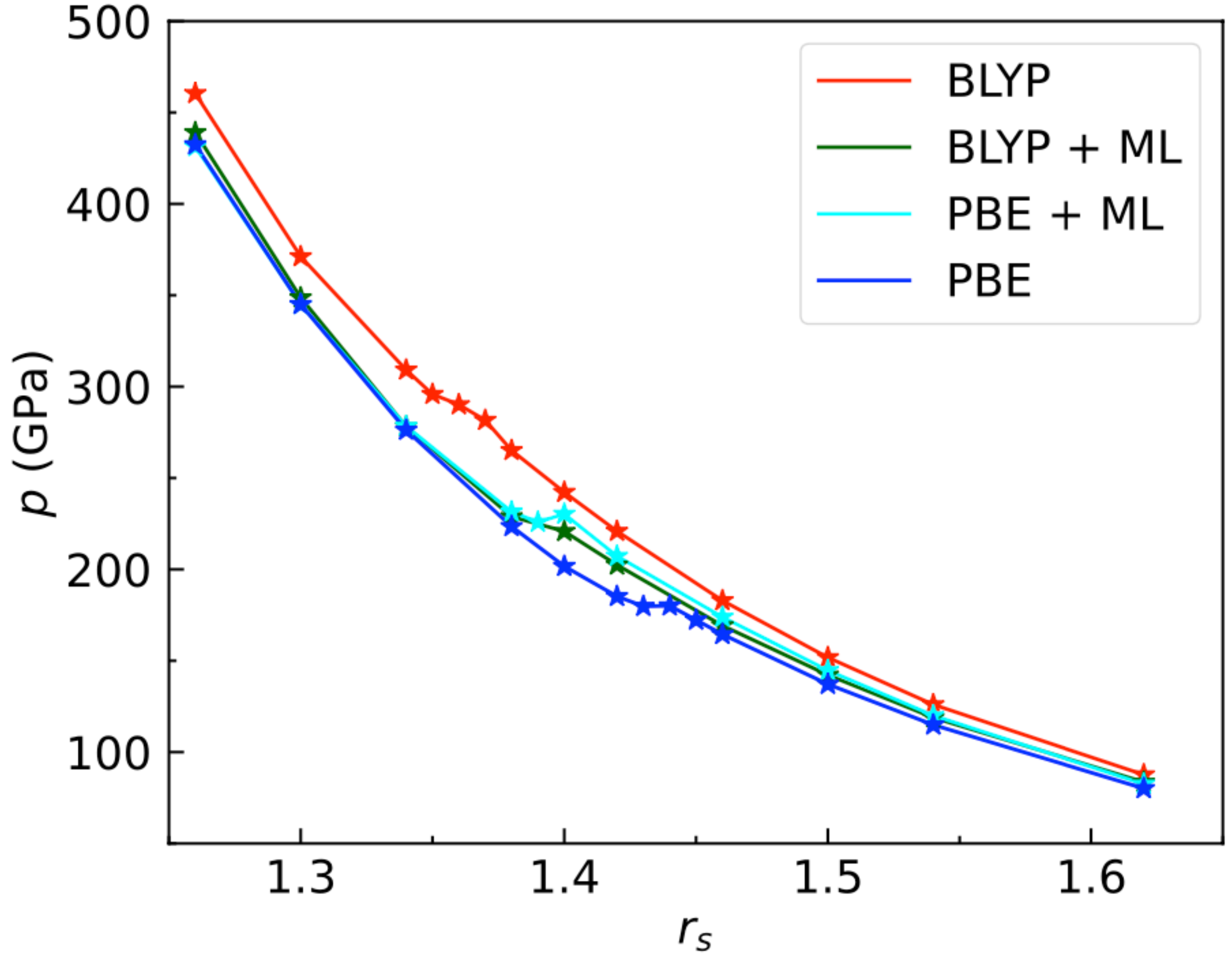
A more informative base distribution, e.g. a machine learning potential

$$p(\boldsymbol{X}) = \boxed{\frac{e^{-E_{\mathrm{ML}}(\boldsymbol{Z})/k_B T}}{\mathcal{Z}_{\mathrm{ML}}}} \left| \det \left(\frac{\partial \boldsymbol{Z}}{\partial \boldsymbol{X}} \right) \right|$$

We are optimizing free energy difference to the machine learning model

$$F = \mathop{\mathbb{E}}_{\boldsymbol{X} \sim p(\boldsymbol{X})} \left[\mathop{\mathbb{E}}_{\boldsymbol{R} \sim |\psi_{\boldsymbol{X}}(\boldsymbol{R})|^2} \left[\frac{\hat{H}\psi_{\boldsymbol{X}}(\boldsymbol{R})}{\psi_{\boldsymbol{X}}(\boldsymbol{R})} \right] - E_{\mathrm{ML}}(\boldsymbol{Z}) + k_B T \ln \left| \det \left(\frac{\partial \boldsymbol{Z}}{\partial \boldsymbol{X}} \right) \right| \right] - k_B T \ln \mathcal{Z}_{\mathrm{ML}}$$





Correcting base bias with variational optimization

Correcting baseline bias in Δ-ML

Tirelli et al, PRB 2022

Outlook: quantum protons and finite electronic temperatures

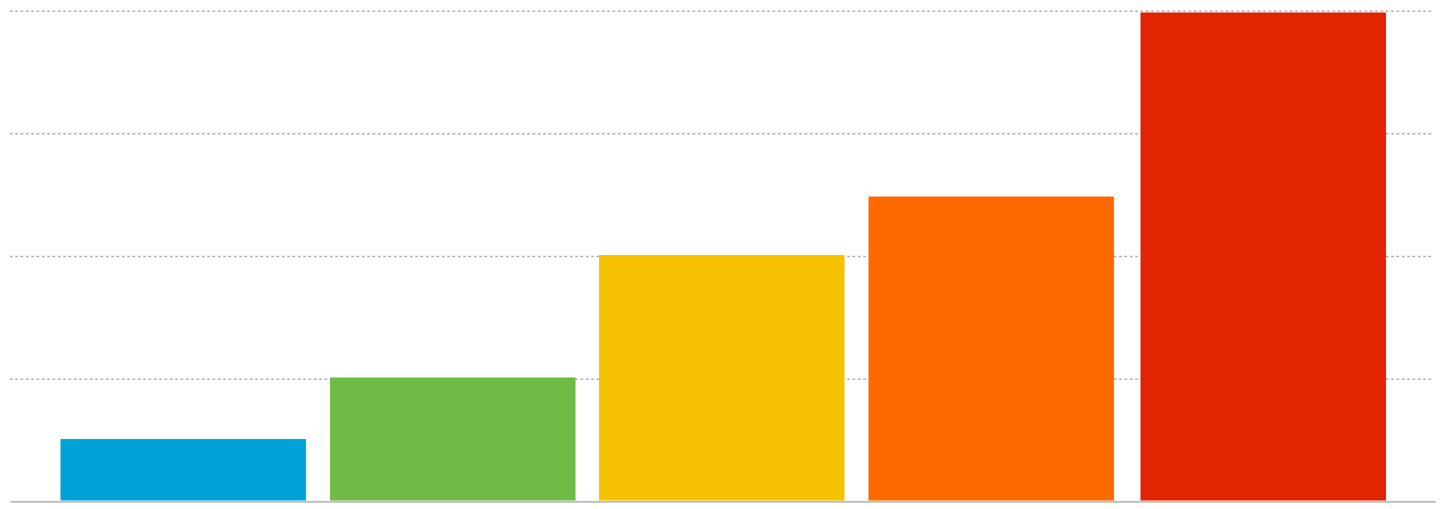
Variational density matrix with neural canonical transformations

Xie et al, 2105.08644 & 2201.03156

$$\min \; F[\rho] = k_B T \,\mathrm{Tr}(\rho \ln \rho) + \mathrm{Tr}(H\rho)$$

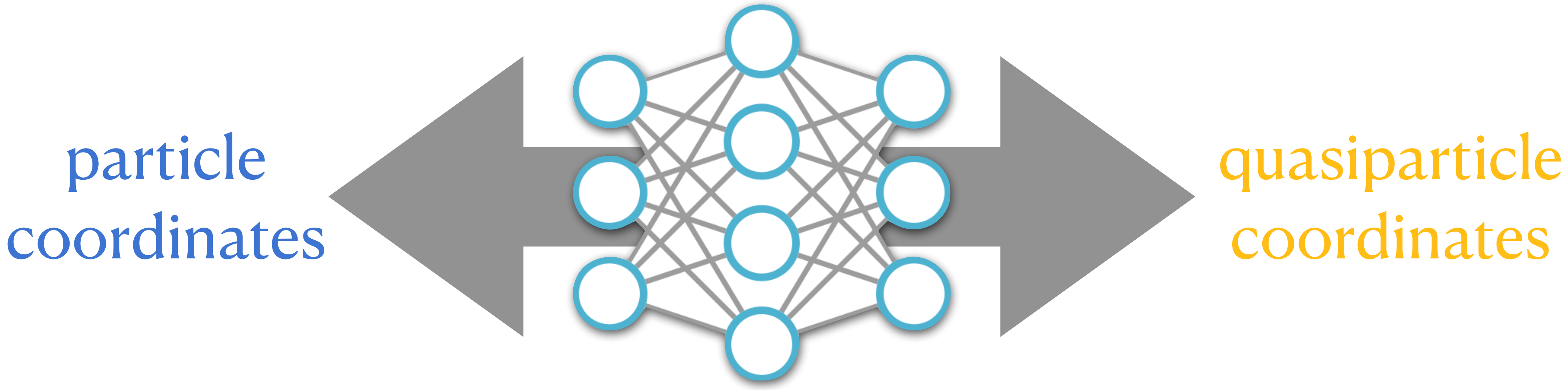
$$\rho = \sum_n p_n |\Psi_n\rangle\langle\Psi_n|$$

Classical probability p_n



masked causal transformer

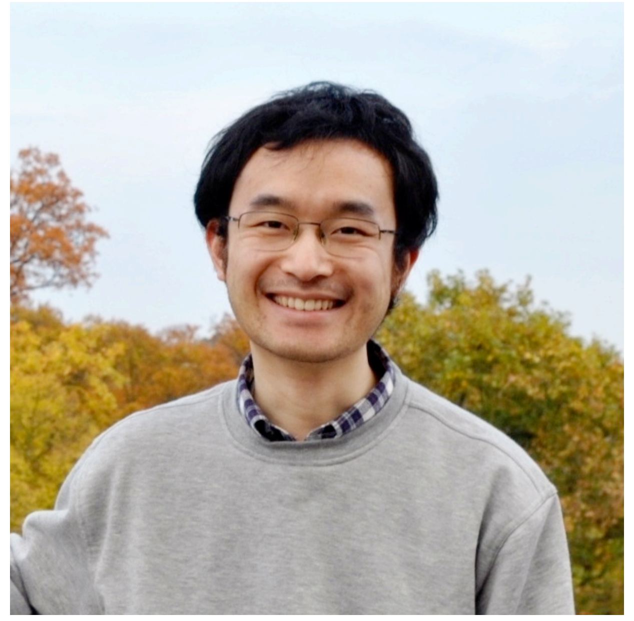
Quantum state basis $|\Psi_n\rangle$

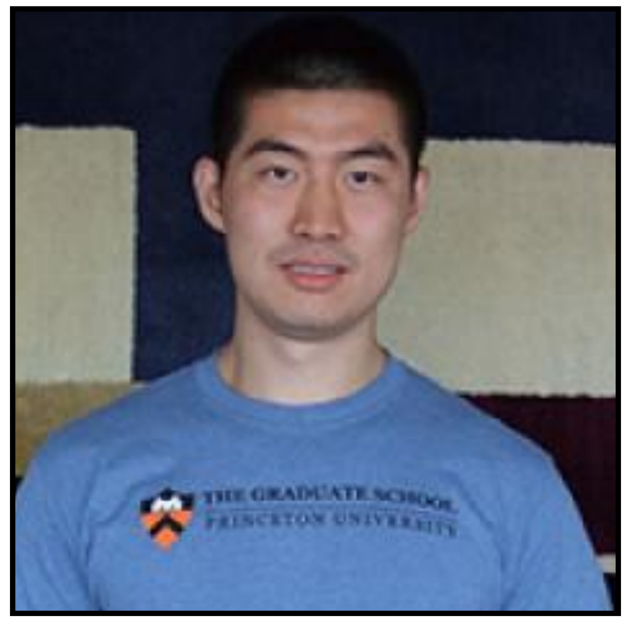


$\sqrt{\text{Normalizing flow}}$

“Using AI to accelerate scientific discovery” Demis Hassabis, co-founder and CEO of DeepMind, 2021

Thank you!





Hao Xie, Zi-Hang Li — IOP

Han Wang — IAPCM

Linfeng Zhang — DP/AISI

fermiflow theory, 2105.08644

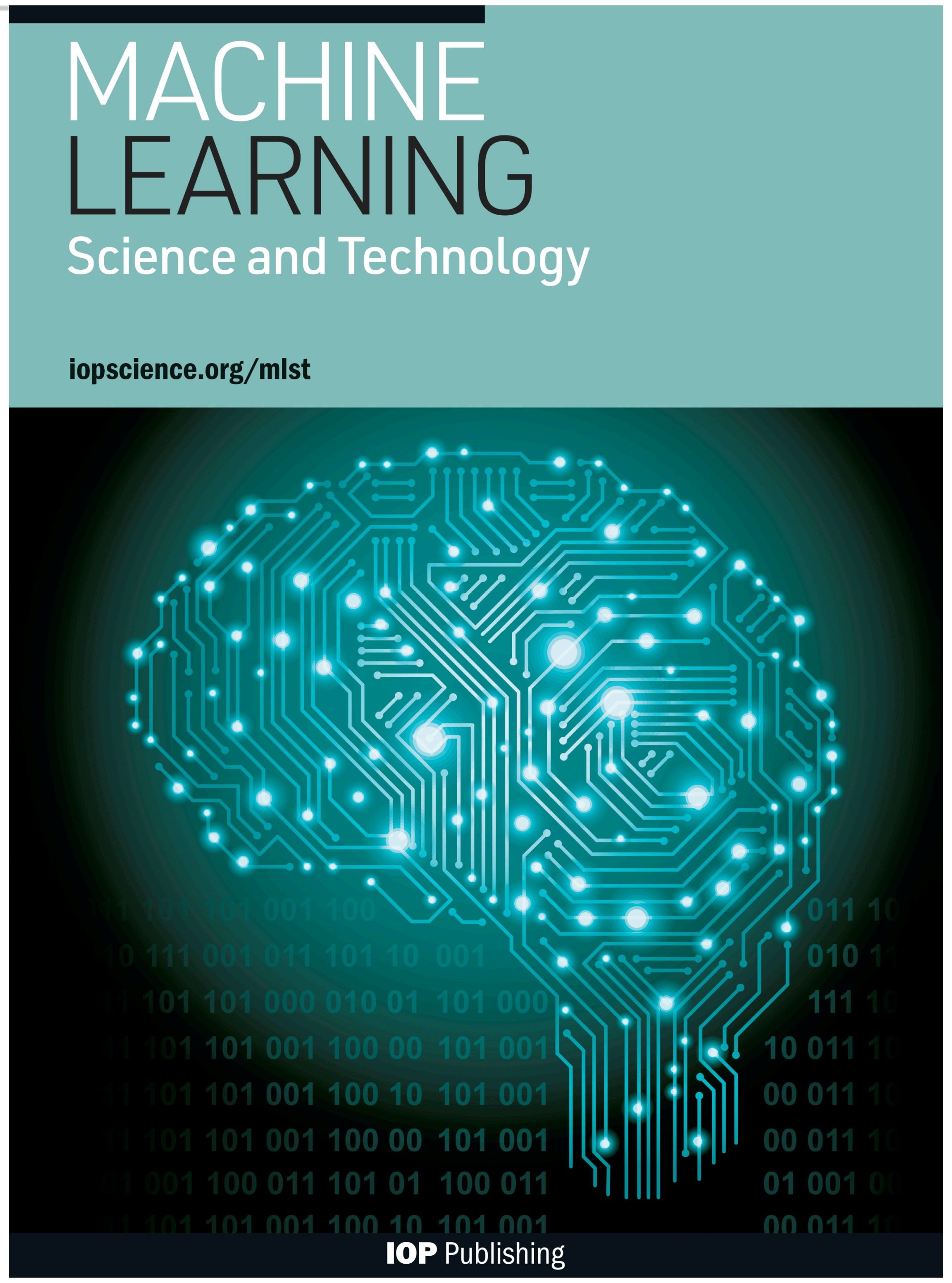
m* of electron gas, 2201.03156

dense hydrogen, 2209.06095

github.com/FermiFlow

IOPscience Journals Books Publishing Support Login

Machine Learning: Science and Technology



Focus on Generative AI in Science

Guest Editors

Juan Felipe Carrasquilla, *Vector Institute, Canada*
Stephen R. Green, *University of Nottingham, UK*
Lei Wang, *Institute of Physics, CAS, China*
Linfeng Zhang, *DP Technology/AI for Science Institute, China*
Pan Zhang, *Institute of Theoretical Physics, CAS, China*

https://iopscience.iop.org/collections/mlst-230424-207