Atomistic simulations of materials

Dave Drabold, Ohio University

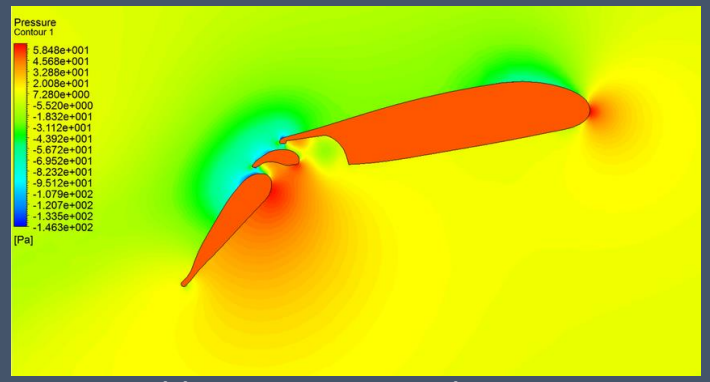
Lehigh Grand Rounds, 2021

Roadmap

- Introduction
- Simulation: some history
- Example: amorphous graphene
- Machine Learning: a new frontier
- Prospects and outlook

Some important applications/frontiers

- Aid to design (ex: air foils)



https://cfd2012.com/wings.html

- Drug discovery

Traditional drug discovery: costs (on average) about $3 billion and takes 12 years. Computational science is having helpful impact to improve these numbers, esp. for cancer drugs (Cui *et al*, Front. Phar, 2020)

- MGI (Pres. Obama, ca. 2011)

https://obamawhitehouse.archives.gov/mgi

Atomistic simulation of materials

- The goal: prediction, understanding, optimization
- Such simulations are based on classical dynamics: **F**=m**a**
  - Connected to Newtonian determinism: give me the forces and the initial conditions and I will predict the future!
  - *nb*1: Jurassic Park
  - *nb*2: Quantum mechanics is probabilistic not deterministic
- Big picture
  - Input: force laws and initial conditions
  - Output: structure, vibrations, electronic, optical, magnetic, transport

History: realistic atomistic modeling of materials

History I

- 17th century: Newton The universe is a big coupled systems of second order nonlinear differential equations!
- Newton was initially concerned about gravity and dynamics of the solar system. Easy to write down the equations, impossible to solve analytically for more than two gravitating bodies.
- The first simulation (without a computer!): Galactic dynamics, Eric Holmberg and his photocells.

History II: Statistical Mechanics (good news)

- If we have a macroscopic number of particles (say 10^{23}) in a gas, solid or liquid, can learn much about the properties of the collective:
  - Microscopic understanding of Thermodynamics
  - Critical phenomena/phase transitions
  - Near equilibrium response (such as fluctuation-dissipation theorem)
  - Transport theory
  - Classical theory of liquids
- Summary (E. T. Jaynes): Equations of motion are incredibly complex and impossible to simulate or even store. And we usually don't care about irrelevant details (the velocity of atom number $10^{17)}$. S.M. elegantly "marginalizes" this cluttering complexity and focuses on the observables that matter.

History II: Stat Mech: what it *can't* do

- Processes very far from equilibrium (for example: growth)
- Local quantities
  - What is the vibrational frequency of an atom in a given configuration
  - What is the charge on a defect site
  - Where in my material does my electronic defect state come from
- Complicated interaction (I dare you to compute the partition function of amorphous silicon with energies from quantum chemistry!)
- Not an effective way to *predict* undiscovered materials, drugs…

History III: Atomistic simulations: heroic age

- First atomistic simulations in 50's and 60's: Guess an energy function (given the positions of *N* atoms, what is the potential energy of the configuration) – and from this compute forces and carry out a MD simulation (integrate **F**=m**a**). Rahman, Stillinger,…
- Early calculations for example on liquid water.
- Some qualitative resemblance to experiments. Harbinger of better things to come.
- Consequences:
  - Good: broke open a new field of science
  - Bad: interactions too crude for most problems of interest or predictive accuracy

History III: Toward realism and prediction

- Interatomic interactions
  - Q: Where do interatomic forces “come from”?
  - A: Chemistry and chemical bonds.
- The way out
  - Q. How do we compute generally applicable and accurate interactions?
  - A: Admit the root of your woes: grapple with the quantum mechanics!
  - Memorable cocky quote of science: Dirac statement:

  The fundamental laws necessary for the mathematical treatment of a large part of physics and the whole of chemistry are thus completely known, and the difficulty lies only in the fact that application of these laws leads to equations that are too complex to be solved.

History IV: Quantum Mechanics of Solids

- Looks hopeless at first glance:

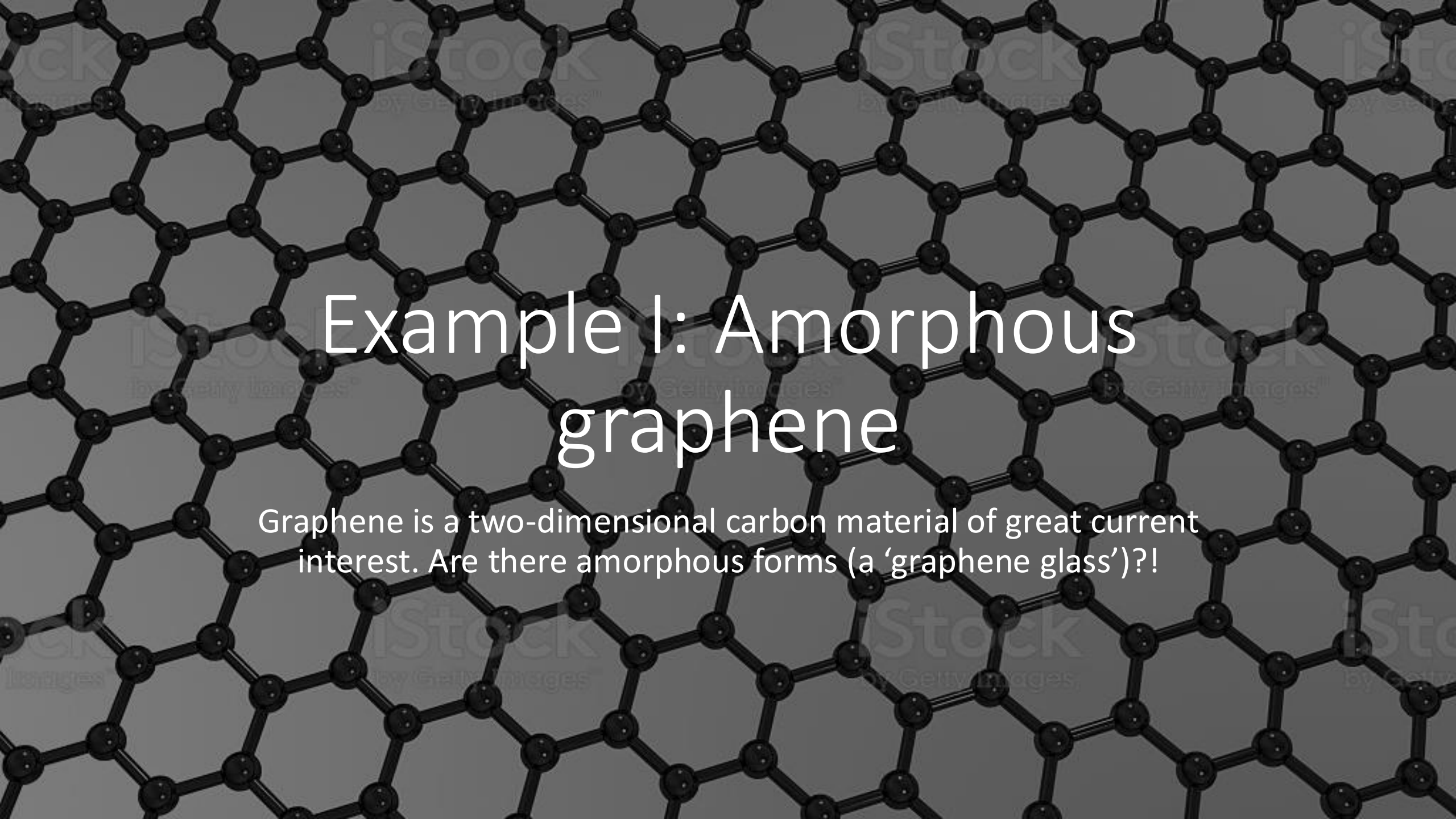
$$\hat{H}(\{\vec{r}_i\},\{\vec{R}_\nu\},t) = \hat{T}_e + \hat{V}_{e-e} + \hat{T}_N + \hat{V}_{N-N} + \hat{V}_{e-N} + \hat{H}_{ext}$$

$$\hat{T}_e = \sum_{i=1}^{NZ_a} \frac{p_i^2}{2m}, \quad \hat{T}_N = \sum_{\nu=1}^{N} \frac{P_\nu^2}{2M} \quad \ldots \text{kinetic energies}$$

$$\hat{V}_{e-e} = \frac{1}{2}\sum_{i\neq j}^{NZ_a} \frac{e^2}{|\vec{r}_i - \vec{r}_j|}, \quad \hat{V}_{N-N} = \frac{1}{2}\sum_{\nu\neq\mu}^{N} \frac{Z_a^2 e^2}{|\vec{R}_\nu - \vec{R}_\mu|}, \quad \hat{V}_{e-N} = -\sum_{i=1}^{NZ_a}\sum_{\nu=1}^{N} \frac{Z_a e^2}{|\vec{r}_i - \vec{R}_\nu|}$$

$$\ldots \text{potential energies}$$

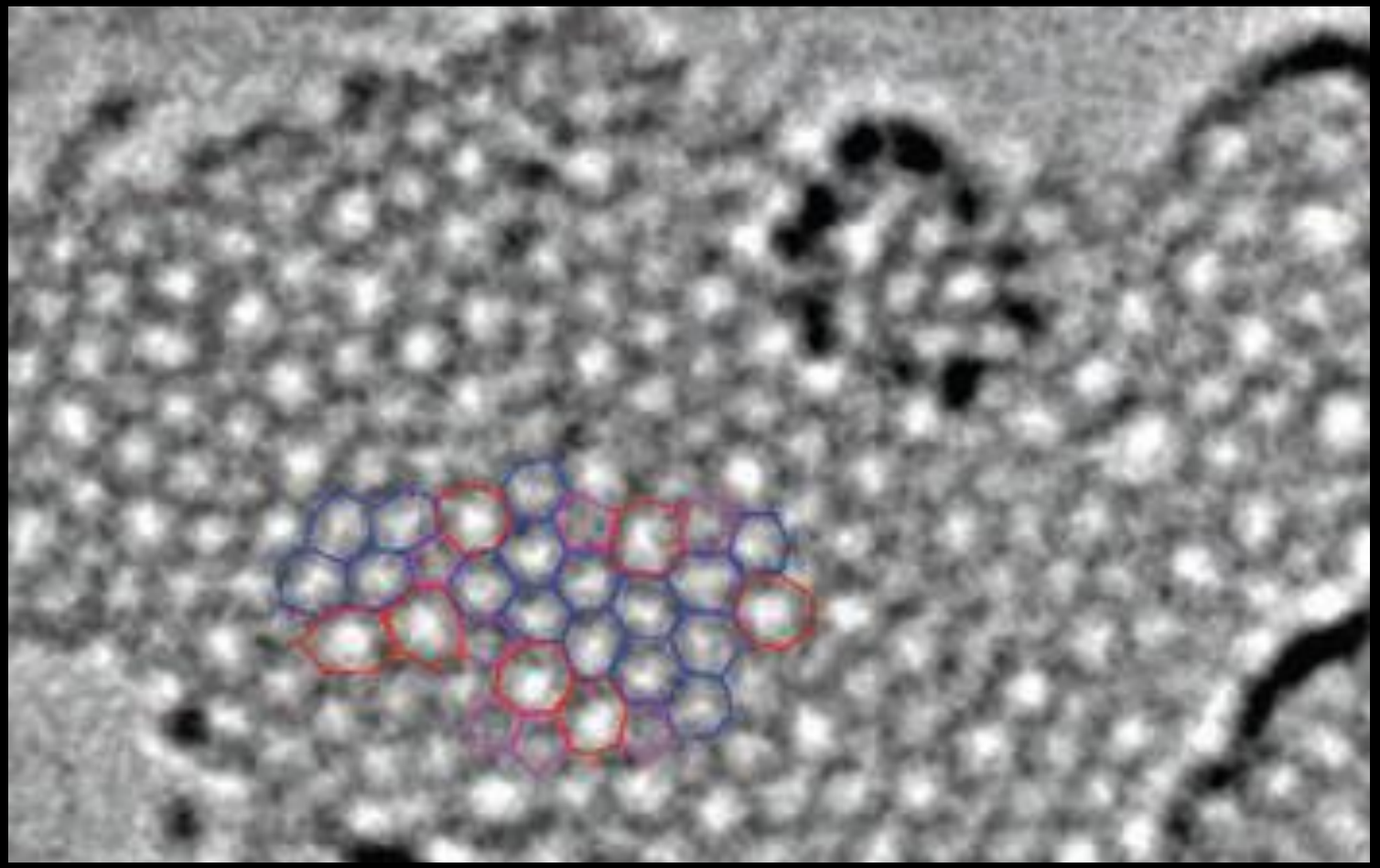
- The epiphany (Dirac, Fermi, Hartree, Slater, Kohn…). The solid/molecular binding from the “electronic glue” of solids can be estimated from the electronic charge density. This is well approximated in a complicated **single particle** theory: Density Functional Theory.
- Great recent Review, Bob Jones Rev. Mod. Phys. 87, 897 (2015).



Example I: Amorphous graphene

Graphene is a two-dimensional carbon material of great current interest. Are there amorphous forms (a ‘graphene glass’)?!

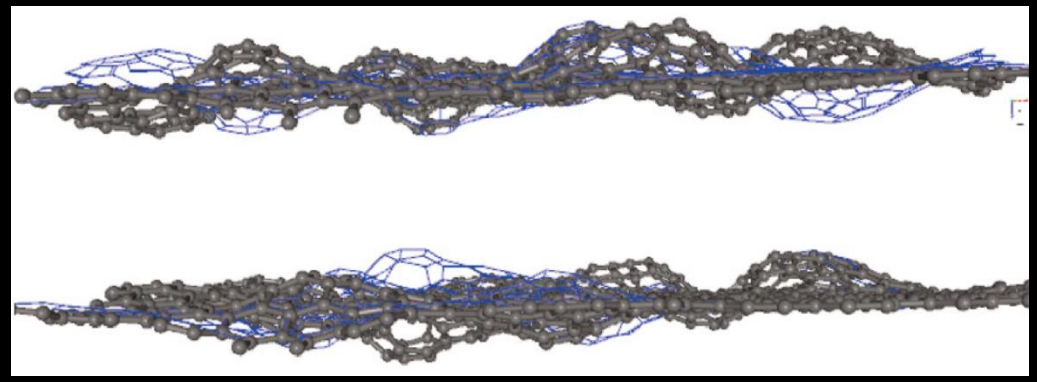
Amorphous Graphene exists. A 2D amorphous “solid”.

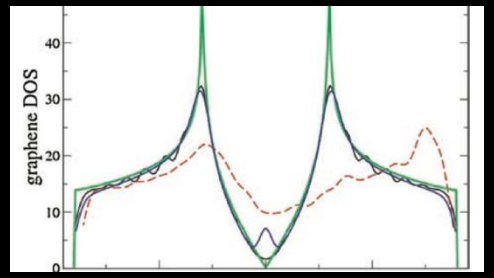


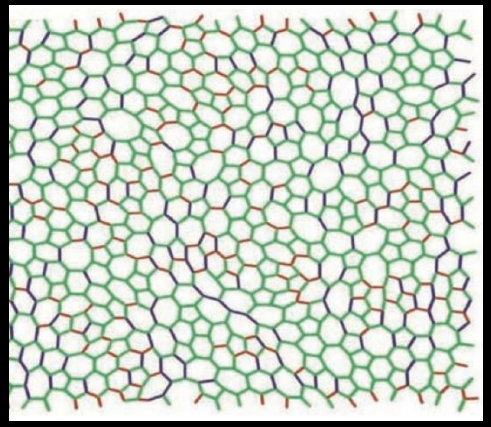
Geim *et al*,Rings 5/6/7 Purple/blue/red

We make up a model and explore it

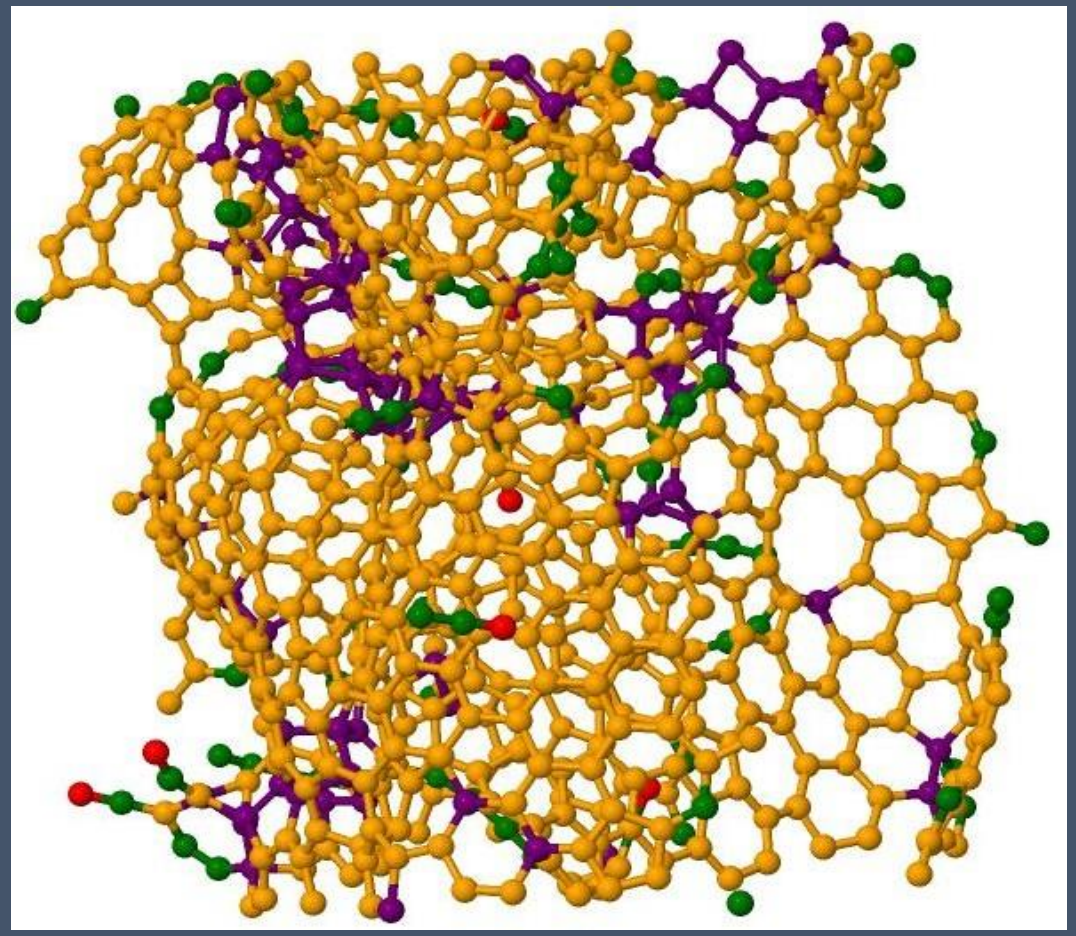
- He and Thorpe were prescient, made such models in 1980s.
- We show that: the models are realistic, that they pucker (pentagonal strain) and that electronic properties are very different in glassy graphene.

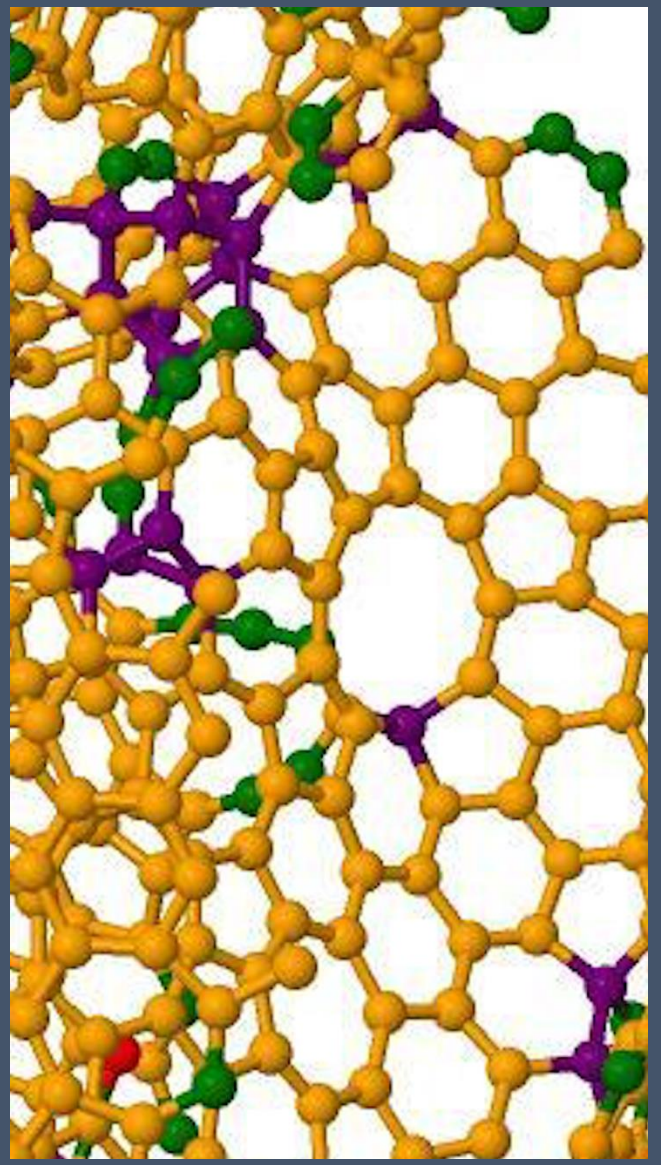






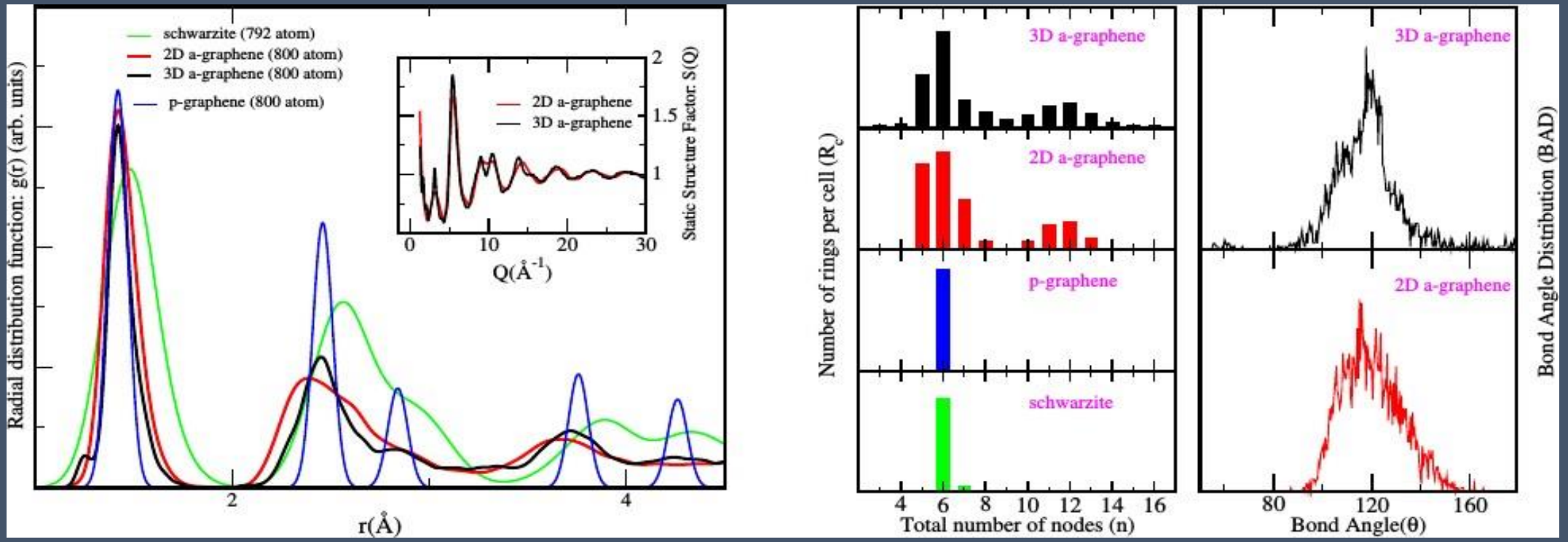
Low density 3D amorphous carbon is 3D amorphous graphene!



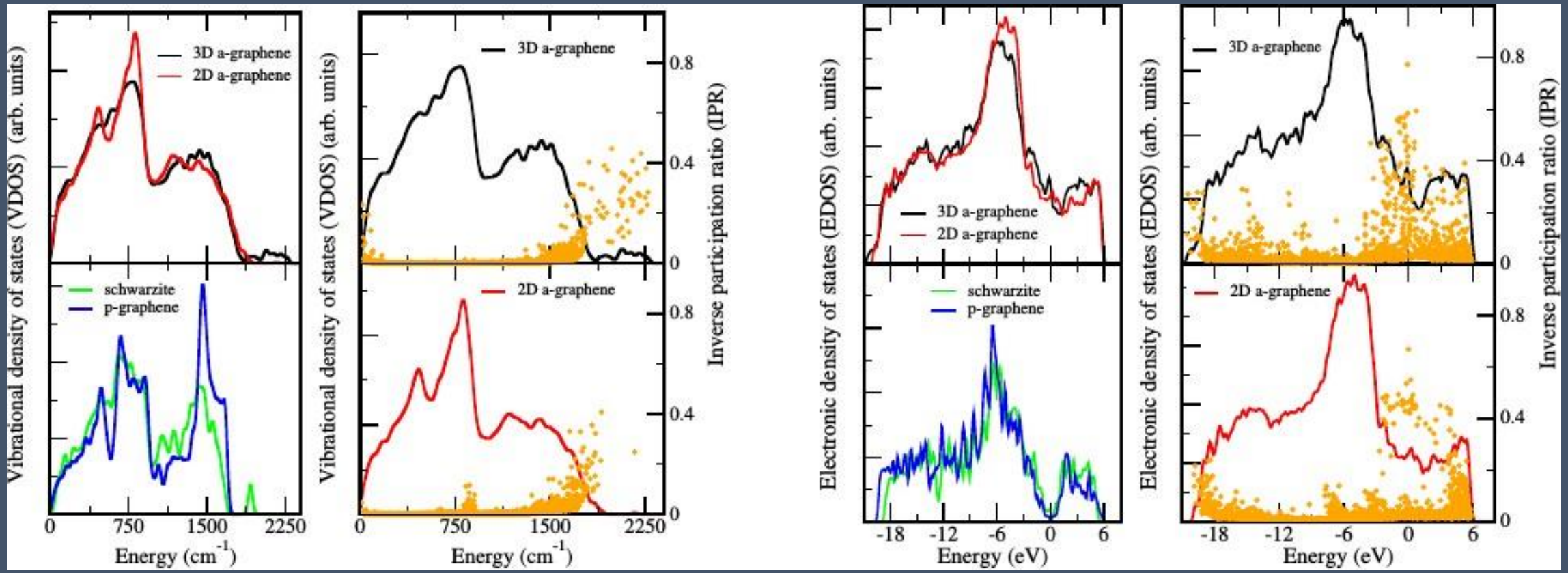


Colors: gold, sp^2;
Green sp, purple sp^3.

Experimental observables: structure



Electrons and phonons



Vibrations

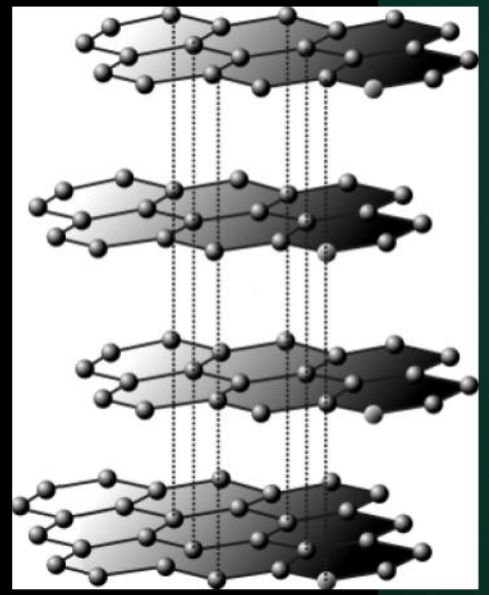
Electrons

Example II. Graphite from Coal

Graphite: a layered form of carbon (graphene hexagons separated by a large interplane distance).

Essential material for lithium-ion battery electrodes

J. Trembly (Ohio) has shown that “graphitization” occurs if coal based “foam” materials are heated to about 3000K. *What are these materials?*



wikicommons

Machine learning for potentials (the idea)

- Pick a material, say a-SiO_2. How many millions of computer core hours have been spent, a paper published and the simulation data **deleted**?
- Suppose that we try to build a database of the all the atomistic information we obtained.
  - Q: Can we use this data to predict interatomic interactions for new simulations of the same material?
  - A: Conceptually this is possible because the potential is a continuous function of the coordinates — so it is effectively a hard interpolation problem! Yes (uh, it is a little more complex than this).

Machine learning for interatomic potentials: ingredients

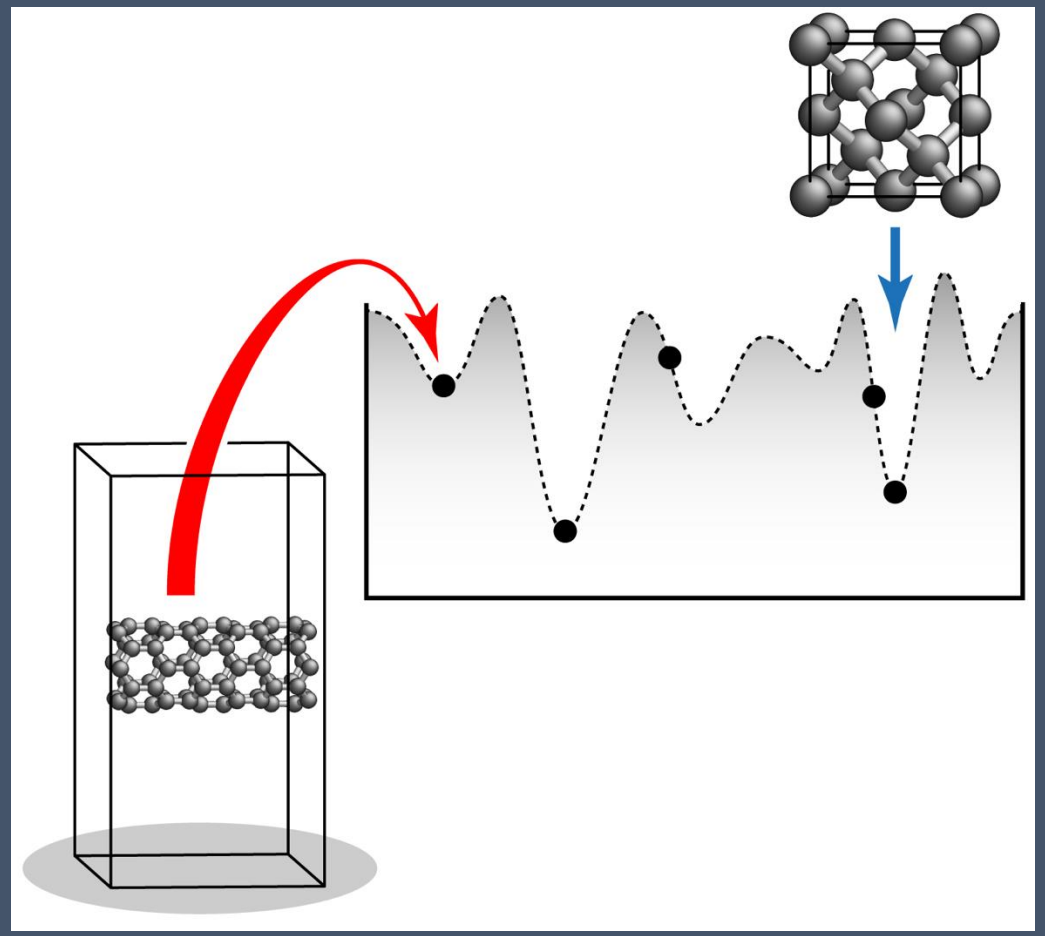
- Many different approaches, but more or less all follow a pattern like this. For a given “new” configuration encountered —-
  - 1) Identify local bonding environment in some compact form including all effects of “rotation” etc. Note: this can involve tens of atoms….
  - 2) Compare current configuration with what is in the database. Need a matrix on the space of configurations.
  - 3) Determine whether existing database is adequate to provide energy, forces etc.
    - If adequate, take a time step with the “inferred” forces
    - If not, “learn on the fly” – do a new quantum calculation on the new configuration and add it to the database. Use the computed forces to take a time step
    - If exploring a new phase, lots of learning steps are required until the database is sufficiently “complete”.

Accurate large-scale simulations of Si: representing the energy landscape

- Silicon is hard to model well. Well known that liquid and disordered phases are well modeled only with DFT.
- Furio Ercolessi had an idea in the early 90's: why not fit a parametrized functional form for an interatomic potential to *ab initio* data? "Force-Matching method". Clever, but impossible to find a good fit.
- Nowadays: non-parametric approaches and "**Machine Learning**".
- **Csanyi, Bartok and Deringer** have pioneered a successful new approach: "Gaussian Approximation Potential" (GAP).

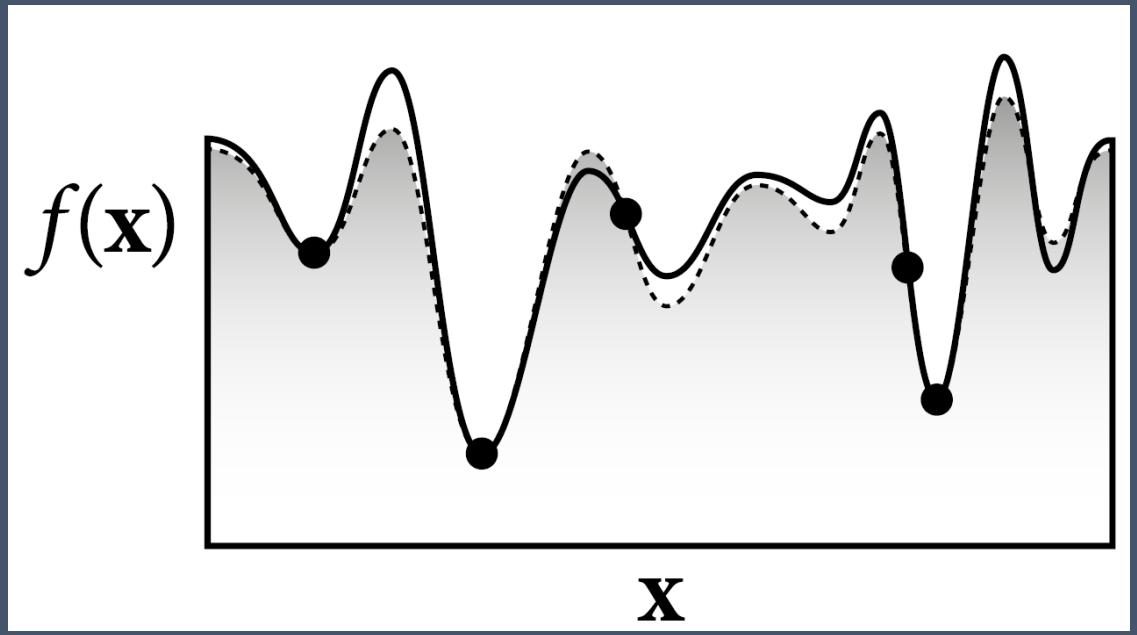
PRX 8 041048 (2018), PRL 104 136403 (2010); Adv. Mats. 31 1902765; several others.

Atomic-scale materials modelling:
Machine learning as an emerging approach



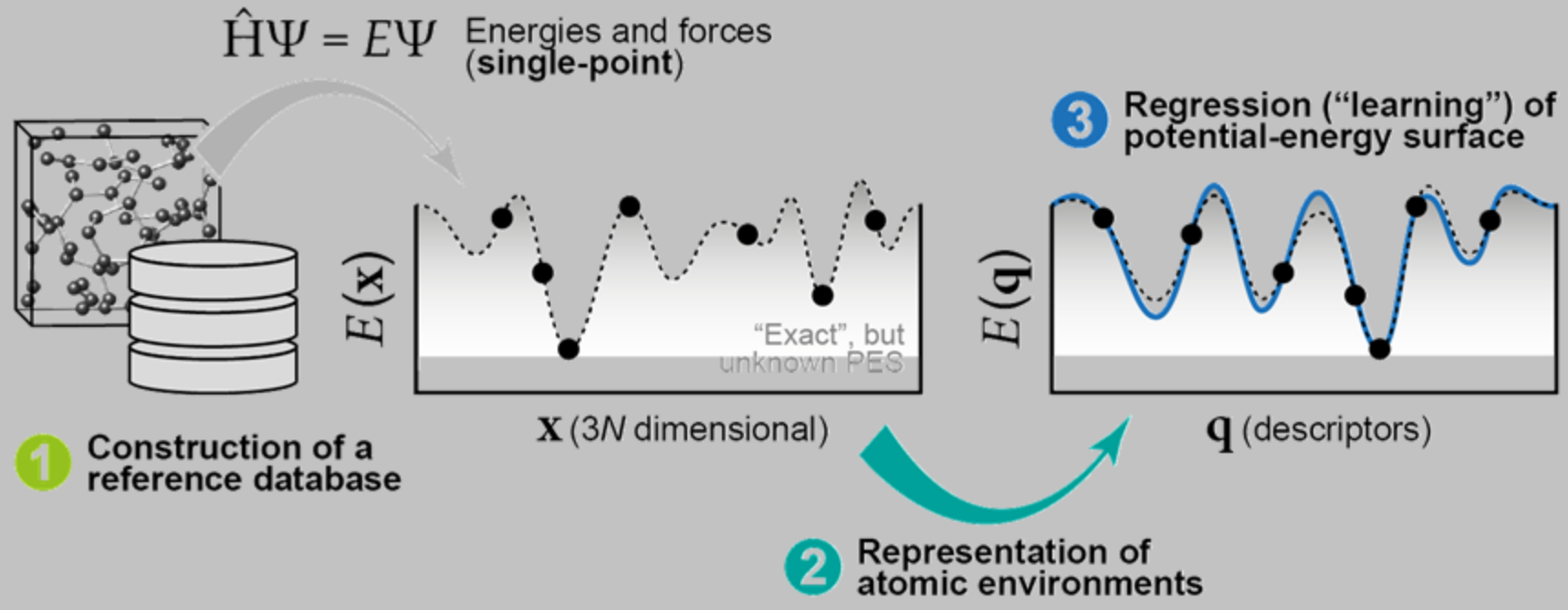
Quantum-mechanically accessible,
but only at selected points!

Atomic-scale materials modelling: Machine learning as an emerging approach



Approximate an unknown function
(*here:* the potential energy surface)
based on data alone

Atomic-scale materials modelling: Machine learning as an emerging approach

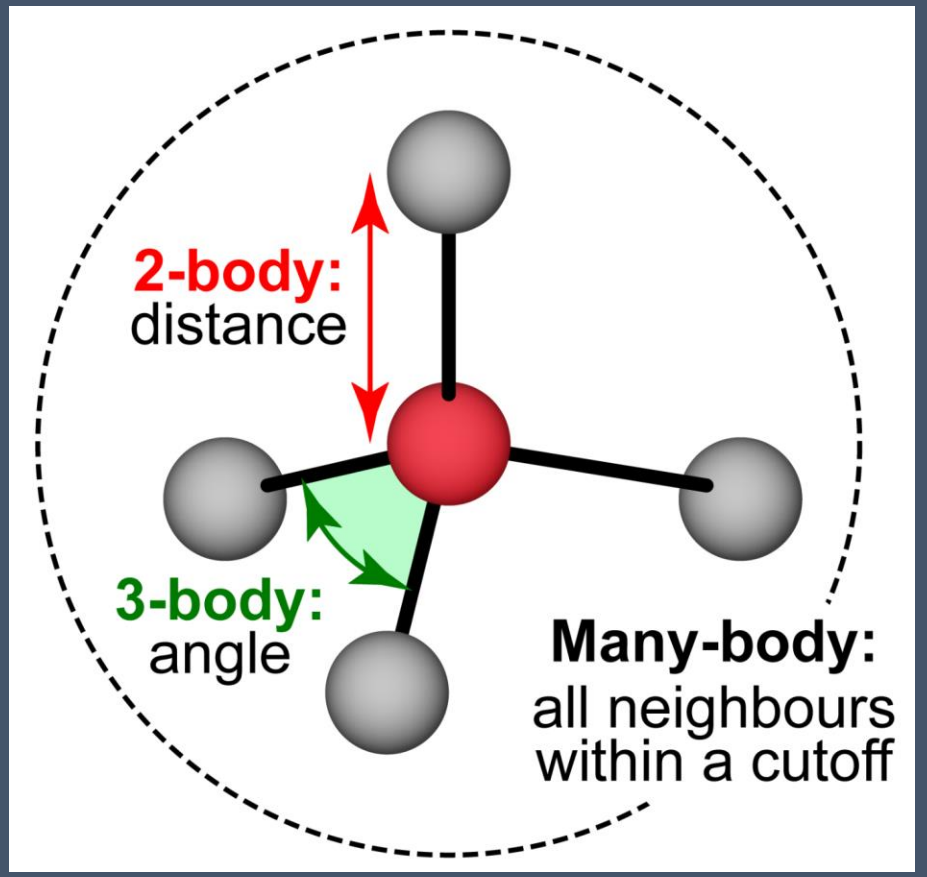


Courtesy Dr. Volker Deringer

A machine-learned potential for silicon

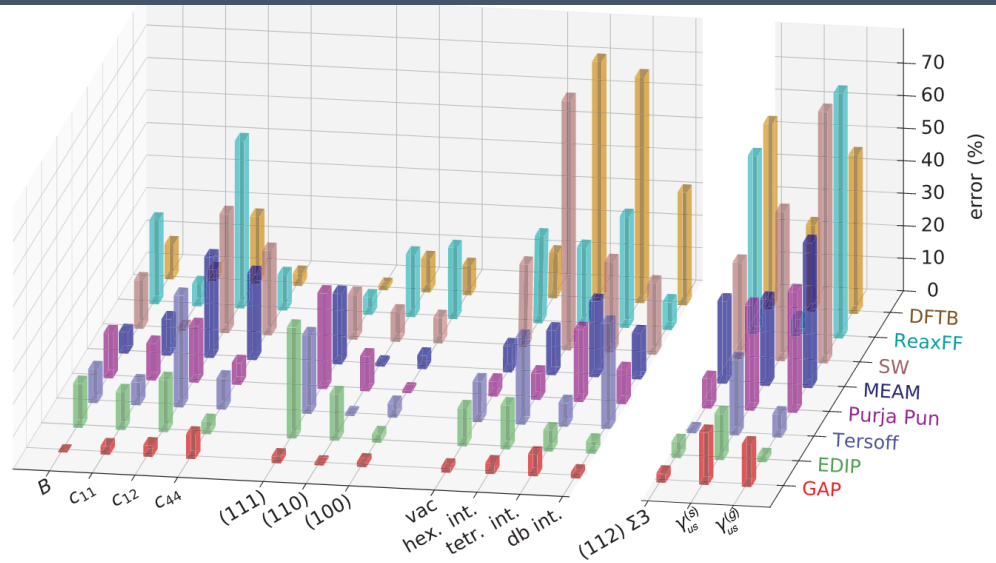
Gaussian approximation potential

- (GAP) framework: a kernel (*similarity*) based machine-learning method.
- New approach here: **combine** suitable structural descriptors.
- Provides meaningful local (site) energies.
- NB: calculations are **not** "cheap", but **are** linear scaling.



V. L. Deringer, G. Csányi, *Phys. Rev. B* **2017**, *95*, 094203

Tests (just a few of *many*)

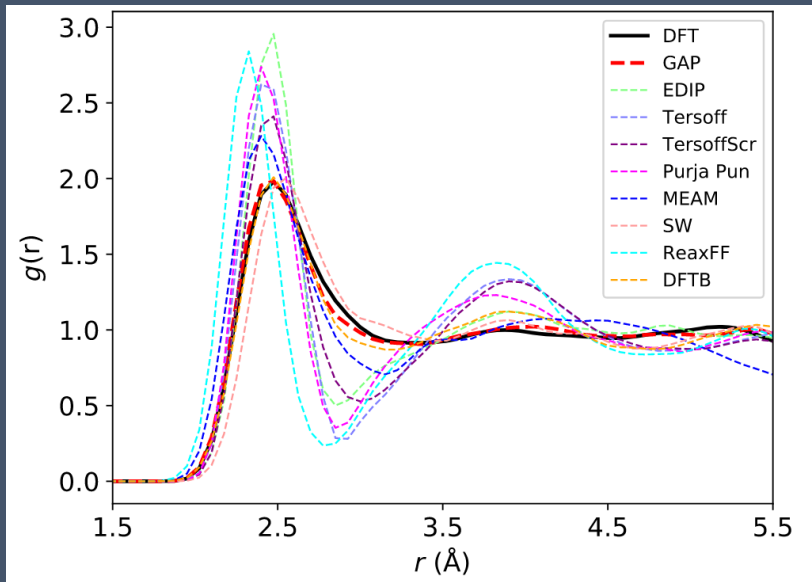


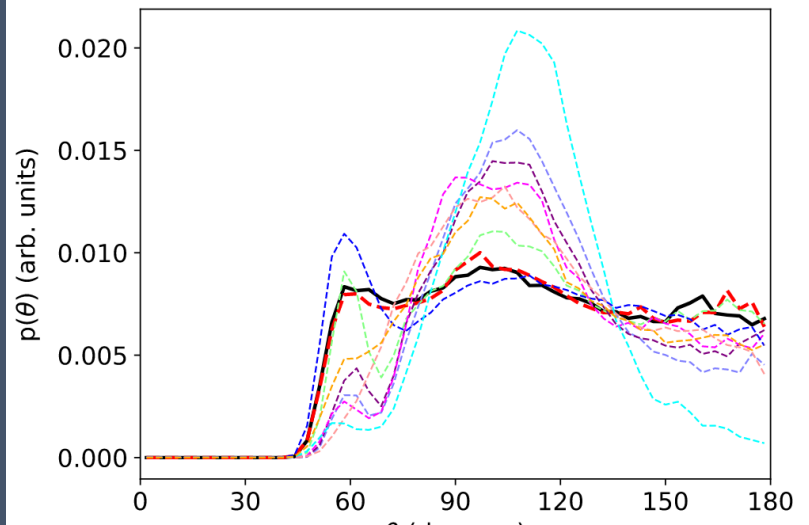
Model	Elastic props. / GPa				Surfaces / J/m^2			Point defects / eV				Planar defects / J/m^2		
	B	c_{11}	c_{12}	c_{44}	(111)	(110)	(100)	vac	hex. int.	tetr. int.	db int.	(112) Σ3	$\gamma_{us}^{(s)}$	$\gamma_{us}^{(g)}$
DFT reference	88.6	153.3	56.3	72.2	1.57	1.52	2.17	3.67	3.72	3.91	3.66	0.93	1.61	1.74
Relative error [%]														
GAP	0	-3	4	-8	-2	-1	-2	-2	-3	-7	-2	3	-16	13
EDIP	14	12	16	-4	-34	-14	-3	-12	14	6	-4	5	-14	-2
Tersoff	10	-7	34	-10	-24	-0	4	13	27	-7	32	-1	-23	10
Purja Pun	14	11	17	7	-29	-11	1	5	8	-22	-10	9	-32	37
MEAM	7	-11	31	-26	-22	-1	4	-8	-14	-23	-14	25	-26	45
SW	14	-1	36	-26	-14	9	8	-27	77	28	22	30	-46	77
ReaxFF	26	7	51	-11	-5	19	-23	28	24	34	8	55	5	75
DFTB	11	4	21	-4	1	10	10	15	74	69	35	57	27	49

FIG. 1. Comparison of percentage errors made by a range of interatomic potentials for selected properties, with respect to our DFT reference. Those on the left of the break in the axis are interpolative, i.e., well represented within a training set of the GAP model: elastic constants (bulk modulus B, stiffness tensor components C_{ij}), unreconstructed (but relaxed) surface energies [(111), (110), and (100) low-index surfaces], point-defect formation energies (vacancy and hexagonal, tetrahedral, and dumbbell interstitials); while the planar defects to the right are extrapolative: (112) Σ3 symmetric tilt grain boundary and unstable stacking-fault energies on shuffle plane $\gamma_{us}^{(s)}$ and glide plane $\gamma_{us}^{(g)}$. The first row in the corresponding table shows reference quantities computed with the DFT (units indicated in the header row).

Bartok et al.
PRX 8 041048
(2018)

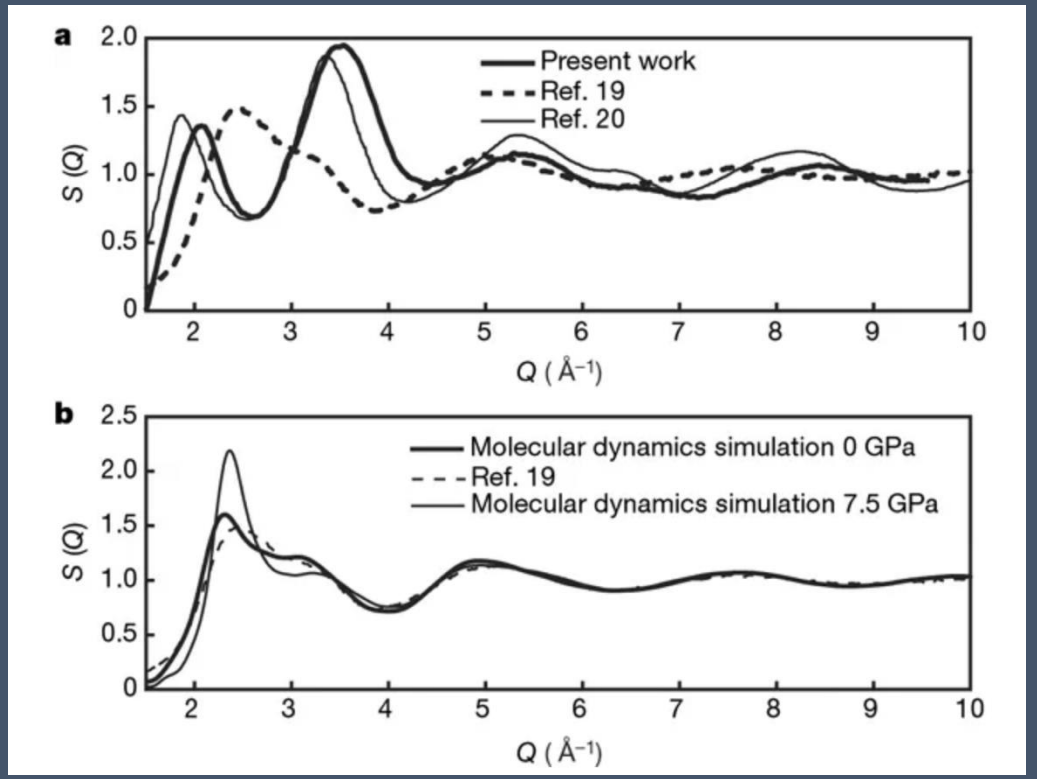
liquid





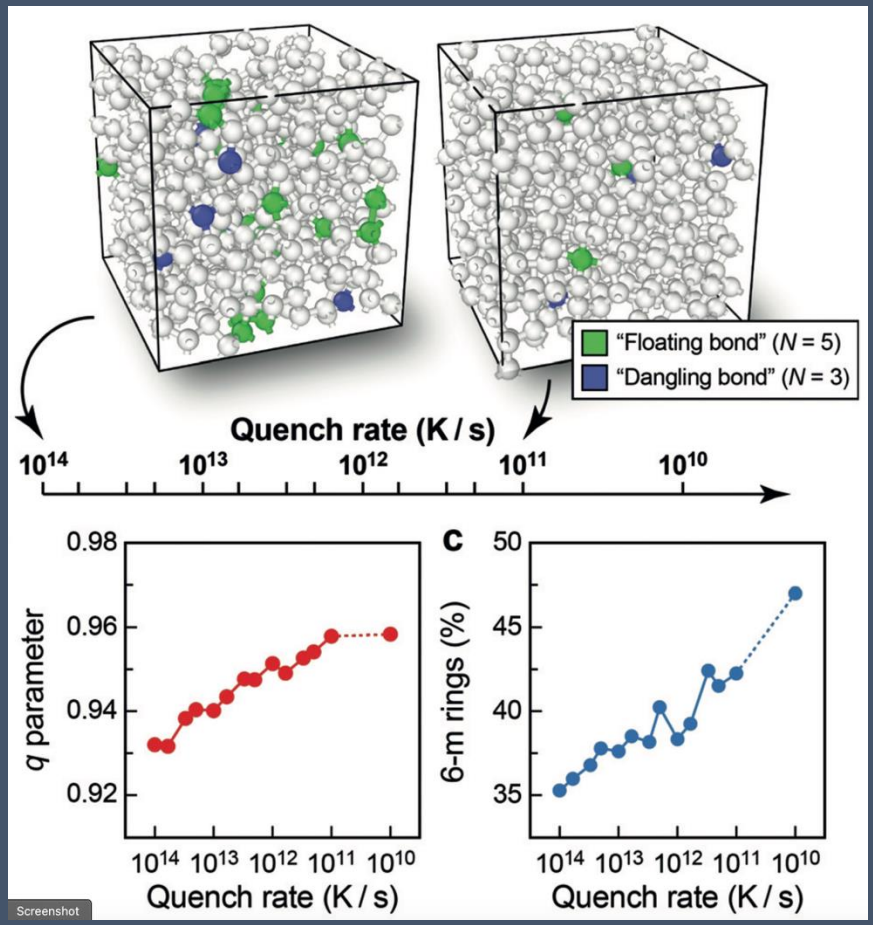
Liquid to amorphous transition

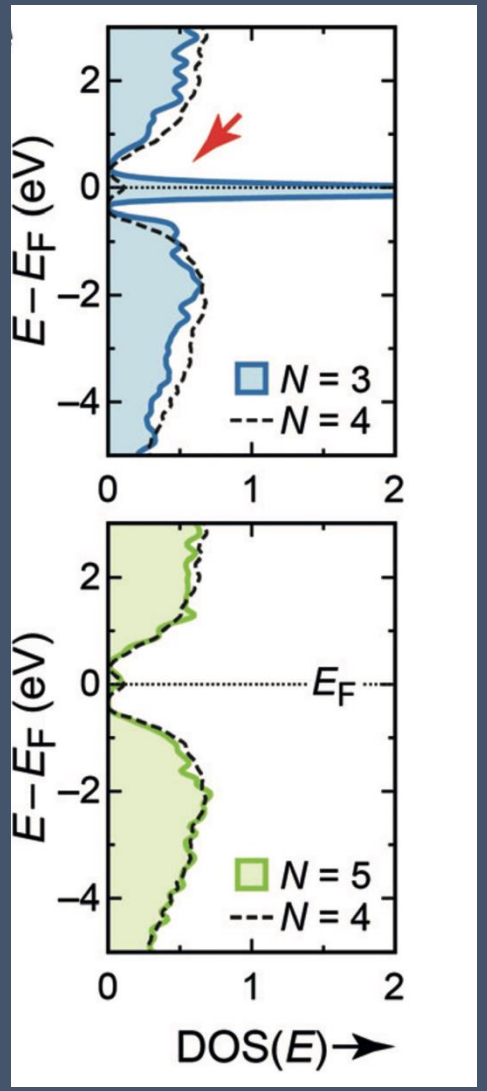
- Fact: Liquid Si is a ~6-fold coordinated metal, amorphous silicon a tetrahedral semiconductor. To my knowledge nobody has made a-Si by quench from the melt (Angell however has done it for Ge!)
- Train GAP for liquid configurations.
- We show that slow enough quenches of the liquid with GAP produces models of a-Si consistent with experiment.



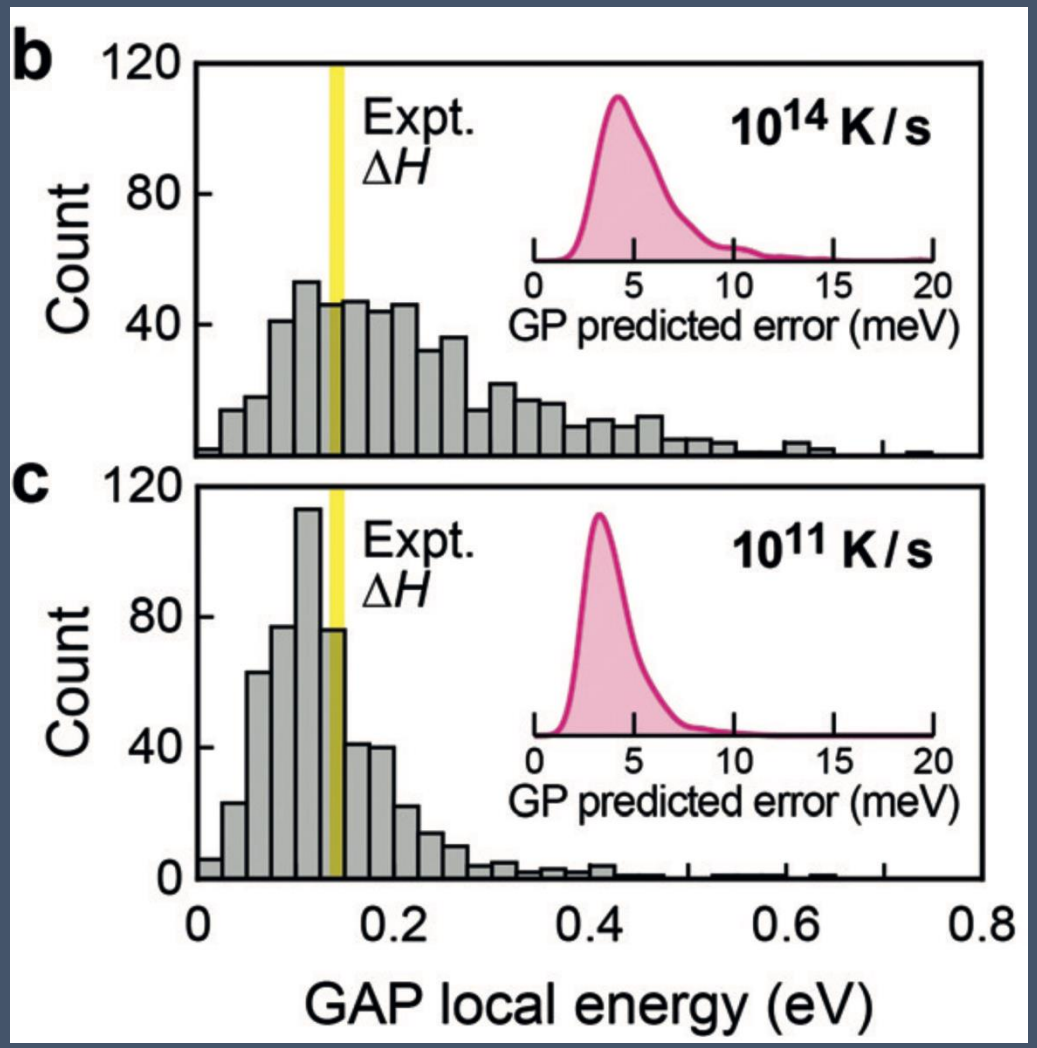
Bhatt *et al*., Nature (2007).

Take 4096 atoms, and quench the liquid…. Slooooowly…….





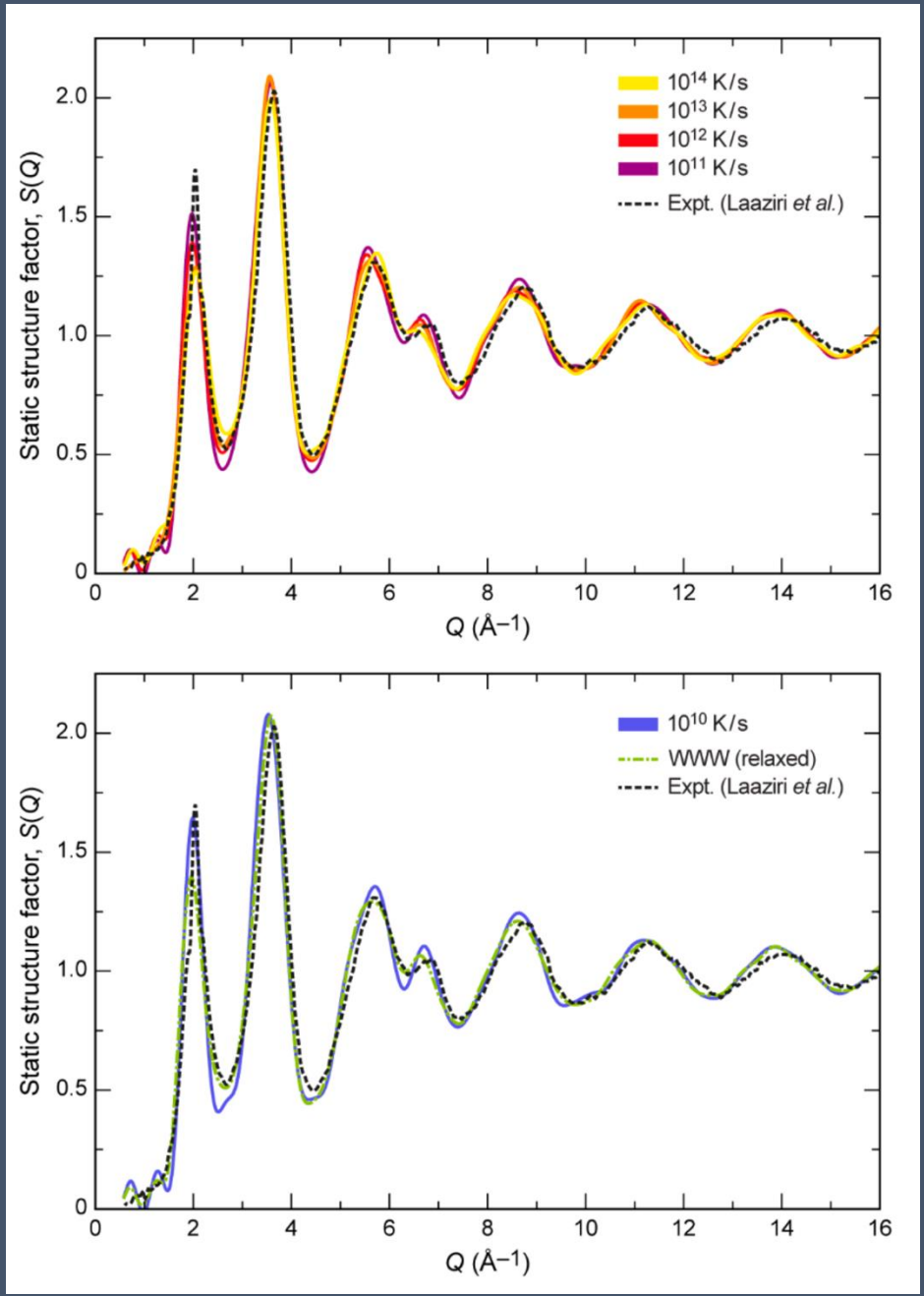
N. Bernstein, B. Bhattarai, G. Csanyi, DAD, S. R. Elliott Angew. Chem 131 7131 (2019)



Distribution of local energies (a fringe benefit of GAP)

10^{11}K/s system *slightly below* best WWW a-Si[1]

1. B. R. Djordjevic, M. F. Thorpe, F. Wooten PRB 52 5685 (1995)



Comparison to diffraction measurements on a-Si (Laaziri *et al.*)

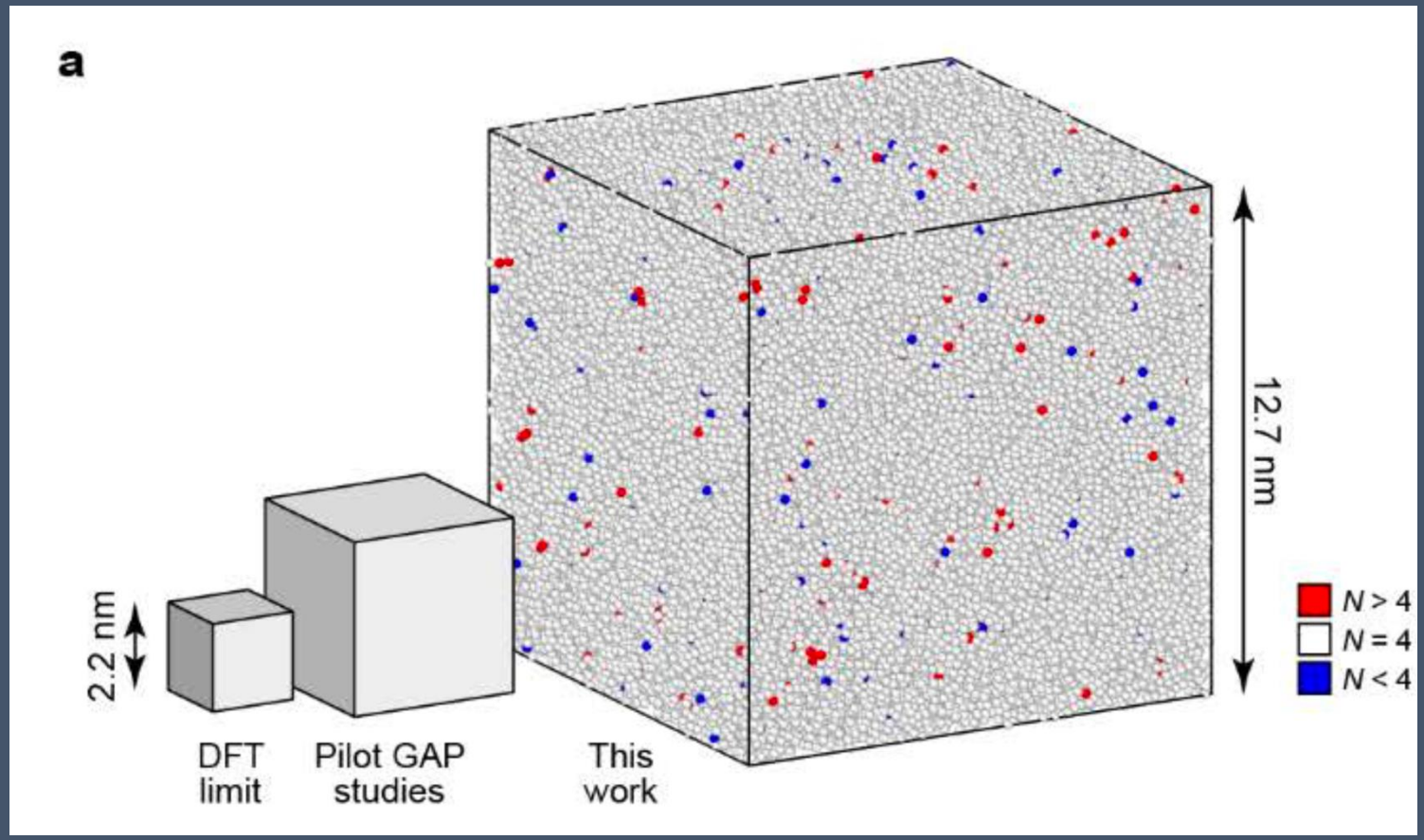
100,000 atoms with DFT-like accuracy

- GAP is linearly scaling (albeit with big prefactor). Linear scaling opens up some new realms for inquiry.

(1) Track liquid to amorphous transition

(2) Squeeze the liquid, compare to experiments

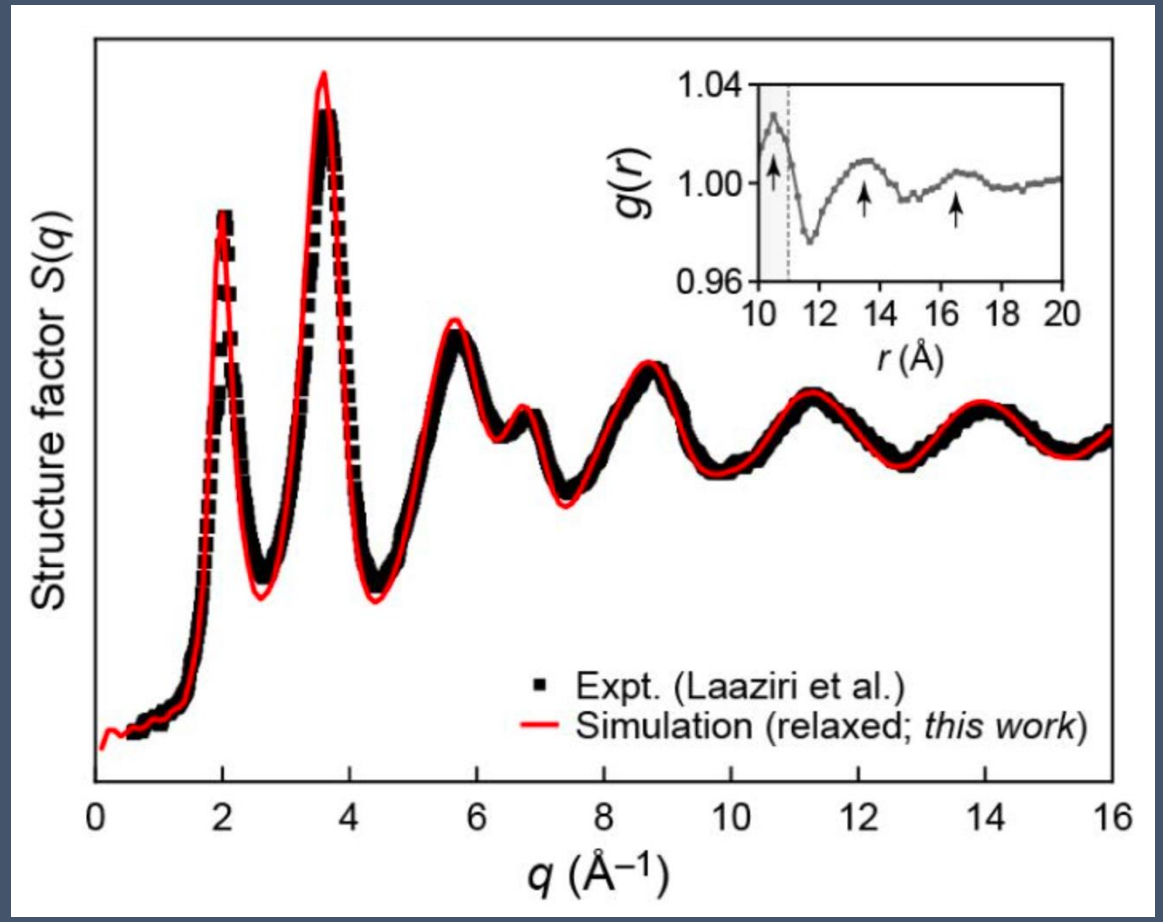
(3) Squeeze the solid, track the phase transitions. *n.b. needed to 'train' for such configurations.*

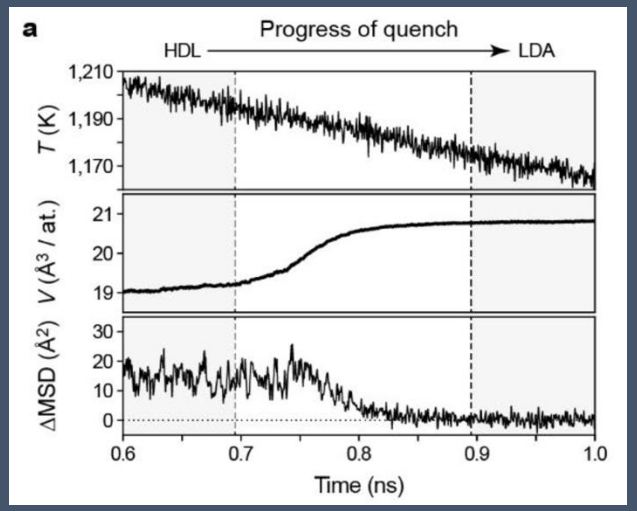


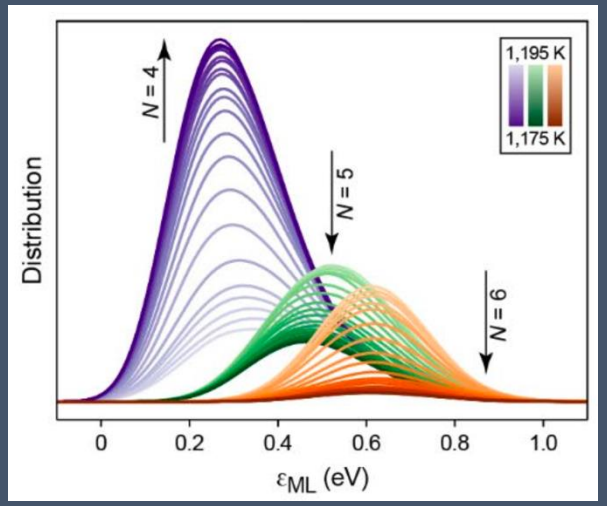
V. Deringer, N. Bernstein, G. Csanyi, M. Wilson, D. A. Drabold and S. R. Elliott, Structural transitions in dense disordered silicon from quantum-accurate ultra-large-scale simulations, Nature 589 59 (2021).

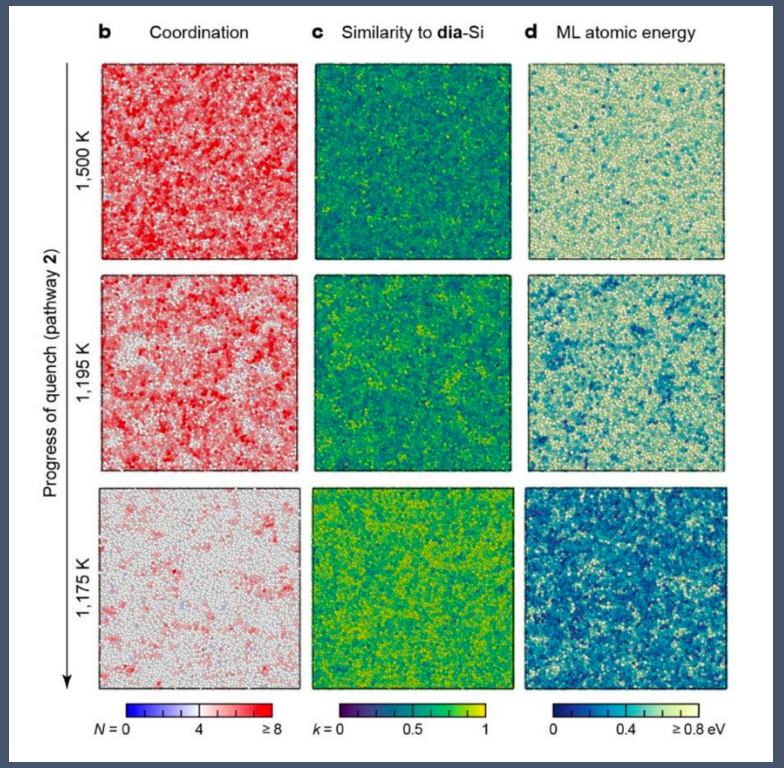
Quench the liquid to make a-Si (zero pressure)

- Similar S(q) to ideally tetrahedral model of Thorpe and coworkers.
- Statistically similar to 4096-atom mode, as expected.



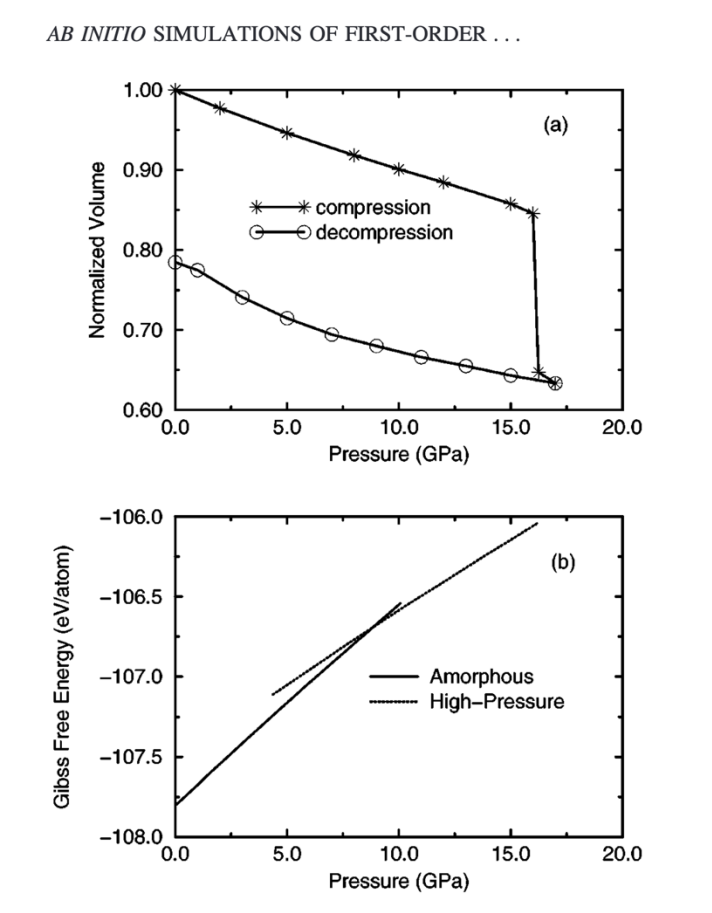


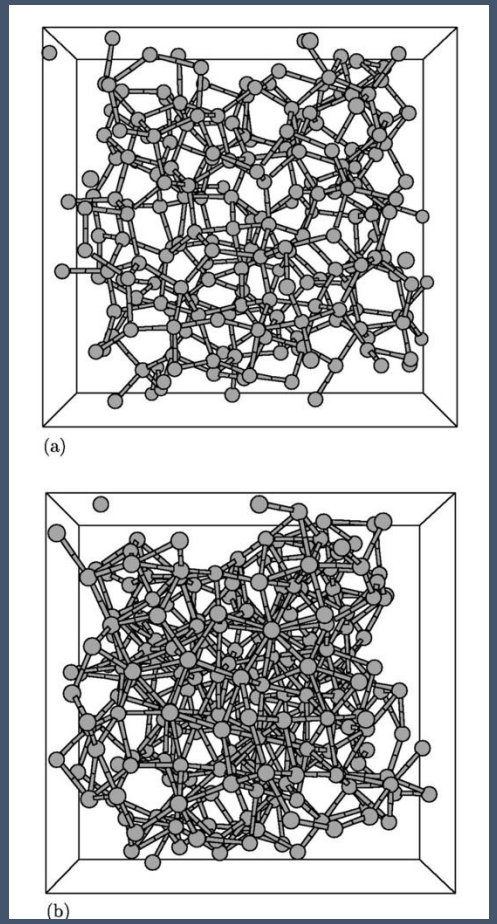




Structural evolution through the quench (500K, 10^{11} K/s)

History: pressure-induced amorphous to amorphous transition in silicon





Abrupt collapse from tetrahedral amorphous phase into dense metallic disordered phase *ca.* 16 GPa.

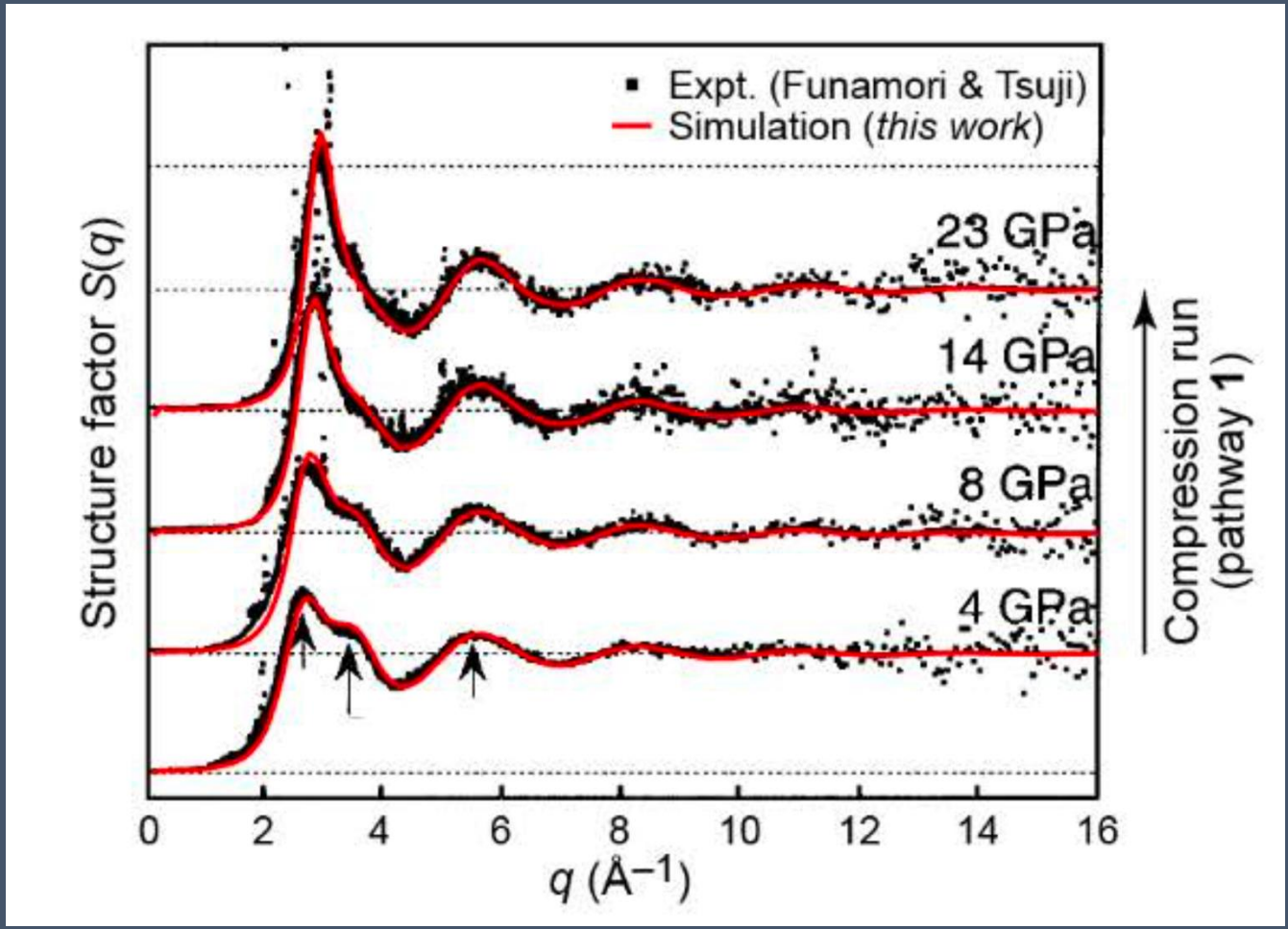
Paltry 216 atoms!

M. Durandurdu, DAD *Phys Rev B 64 014101 (2001).*

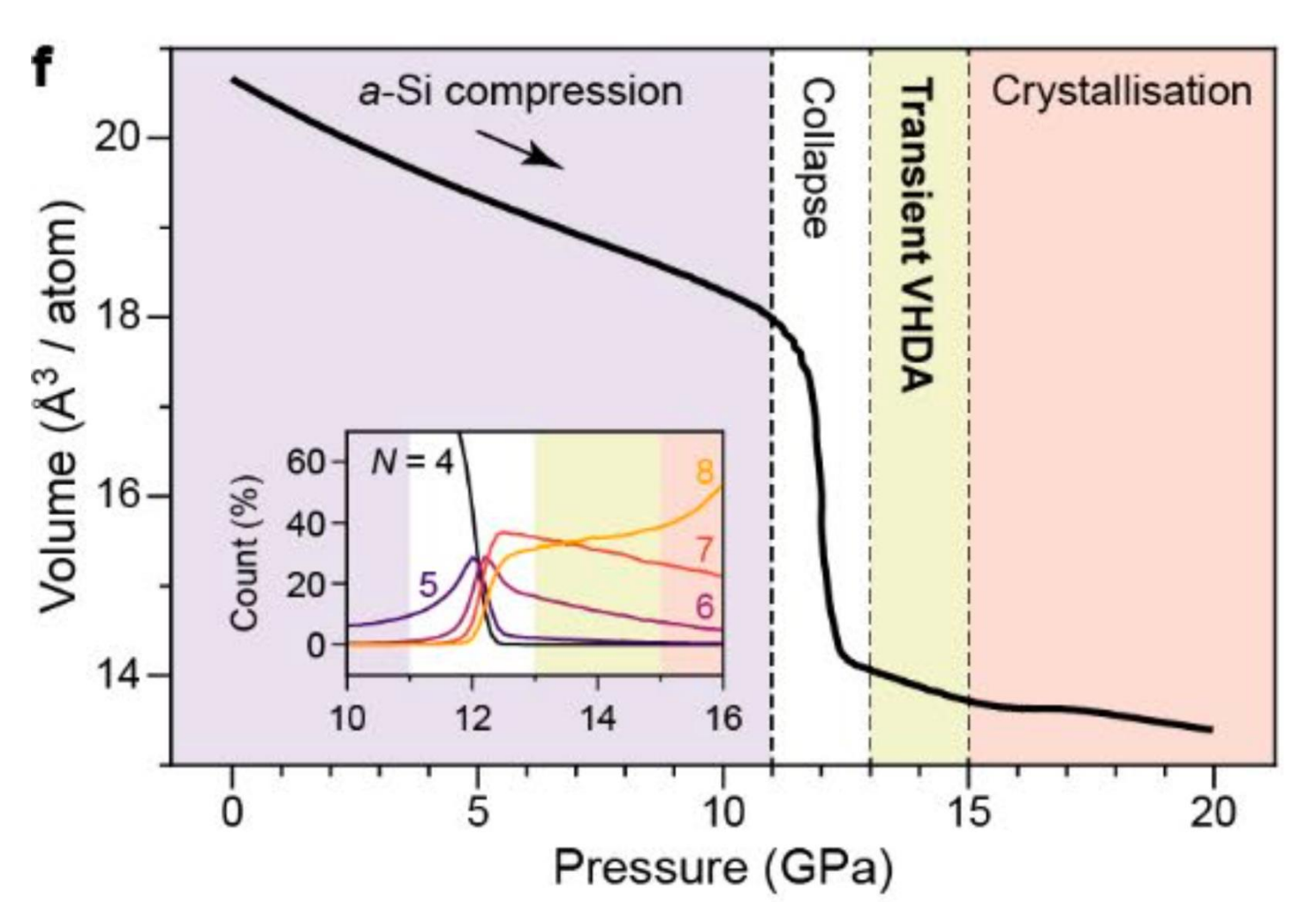
High Pressure: 100,000-atom models

- **First**, we squeeze the liquid (T=1500K). Partly to check GAP, ensure we have all the conformations required. Compare to experiments.
- **Then** we squeeze a-Si (0.1 GPa/ps and T=500K).

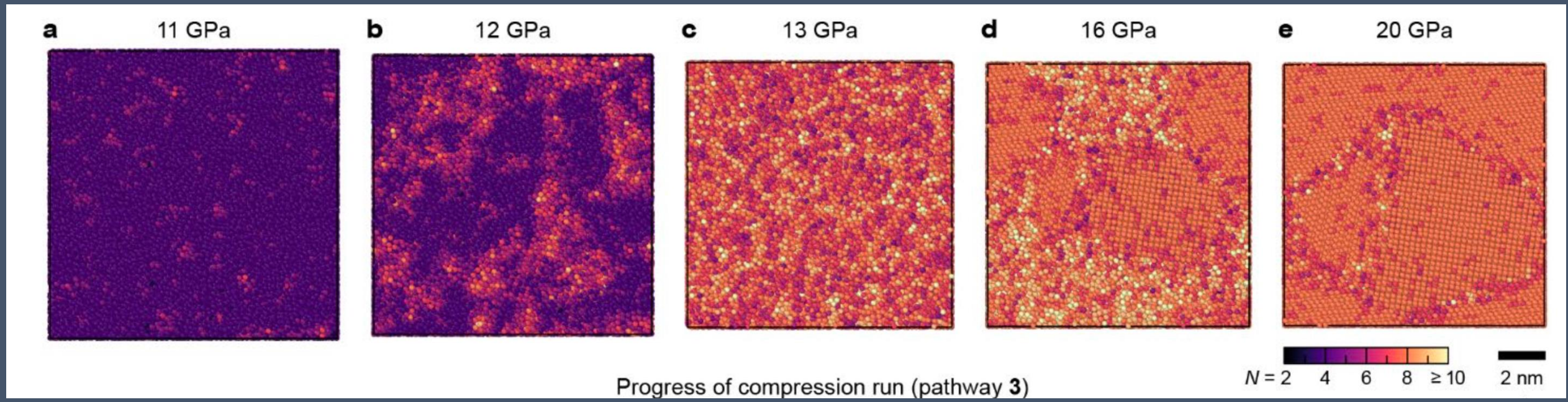
Squeeze the liquid: theory and experiment.



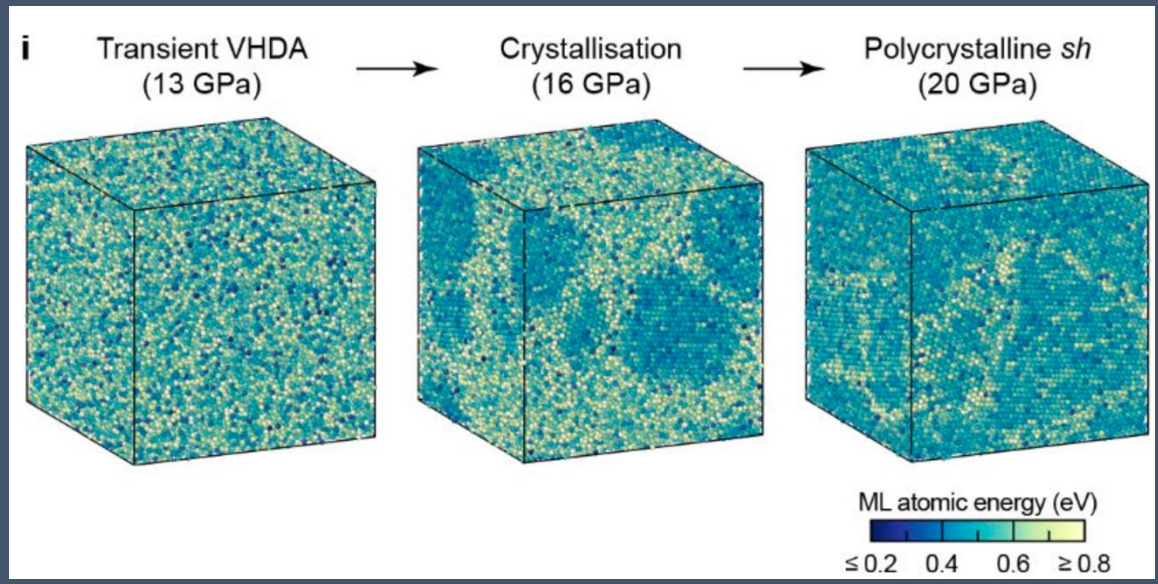
Squeeze a-Si: Results



Phase change: characterization

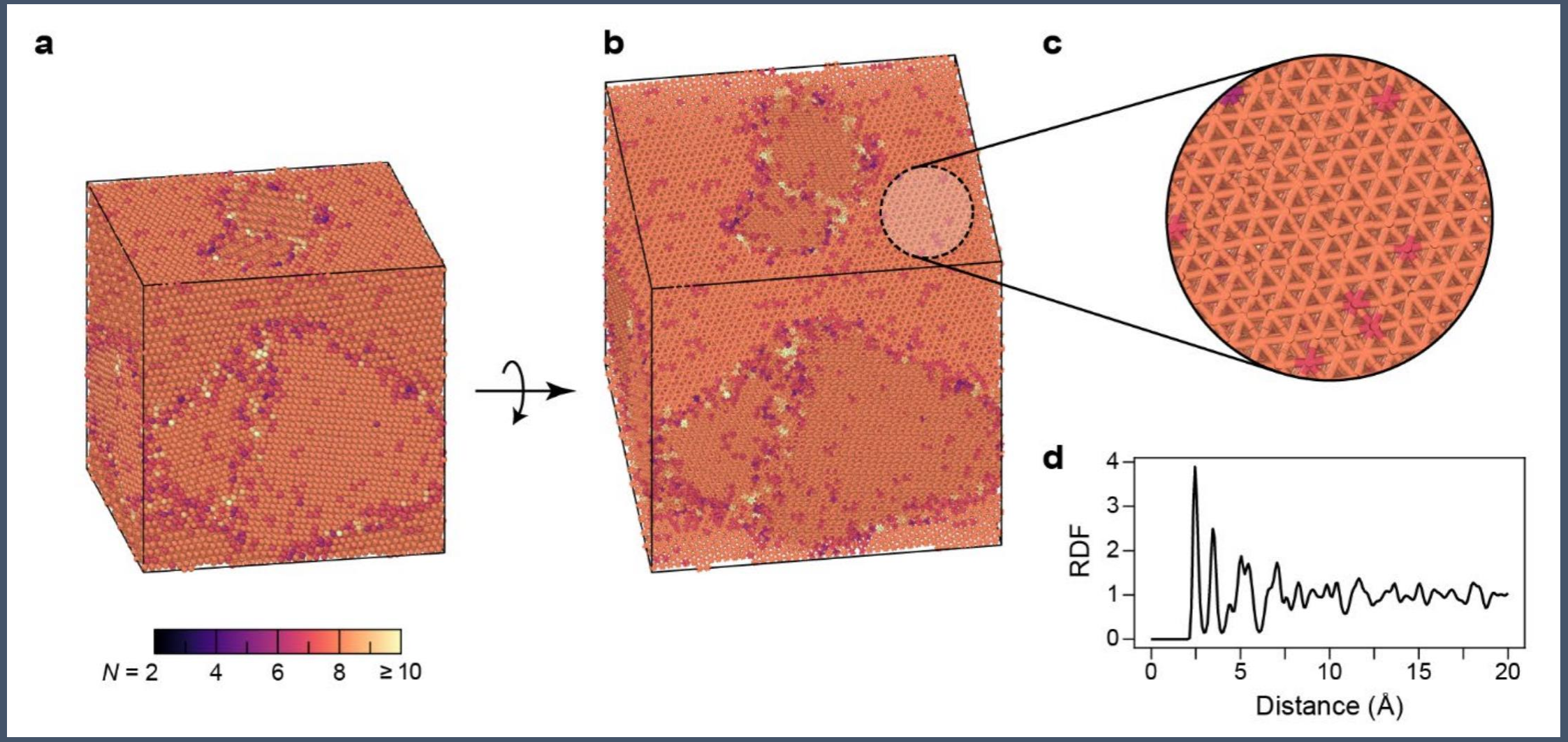


coordination



Local energy

Close up of crystallized phase



High pressure: discussion

Proceeds as follows:

(1) Some initial co-existence of High Density Amorphous (HDA) phase and low density amorphous: “Polyamorphism”.

(2) Abrupt collapse into highly disordered VHDA phase around 11 GPa. VHDA is transient, crystallization (to simple hexagonal phase) occurs at 15-16 GPa.

(3) So we have multistep crystallization originating in a precursor transient VHDA phase. **Not** direct HDA to simple hexagonal as previously believed.

(4) The crystallization does **not** occur in 1000-atom models, even up to 50 GPa. Small cell too dependent on stochastic effects?

Electronic structure

- Use orthogonal tight binding Hamiltonian (Kwon et al. PRB 1994). Four orbitals per site.
- dim(H)=400,000
- Method of DAD and Sankey (PRL 1993) to compute density of states. Ingredients:

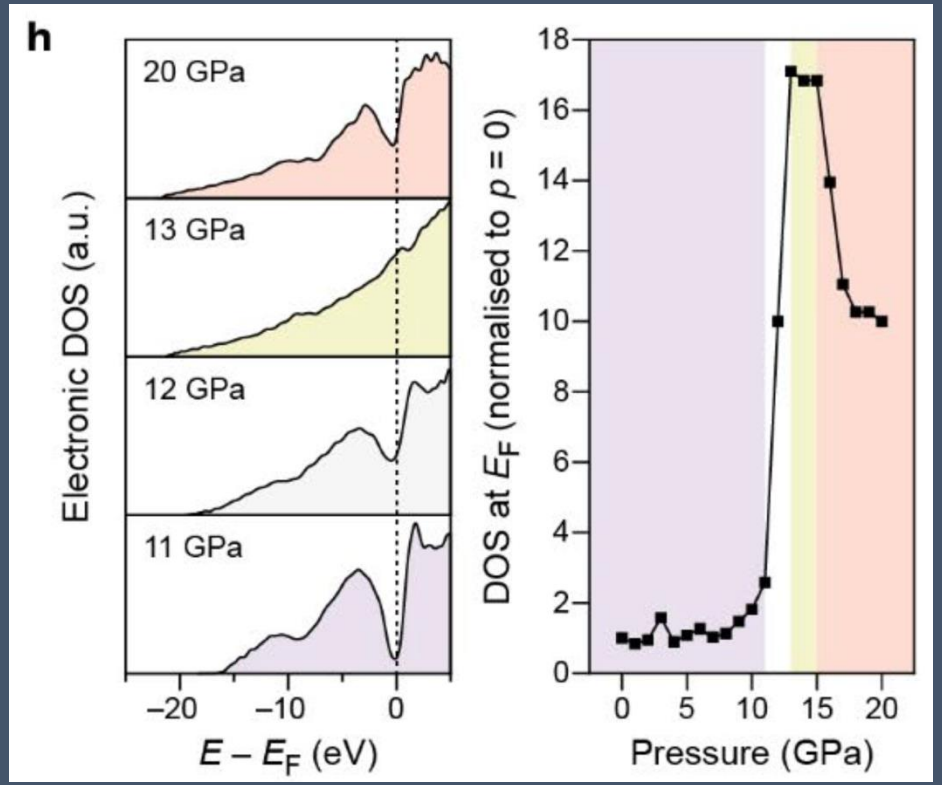
(1) Sparse matrix methods

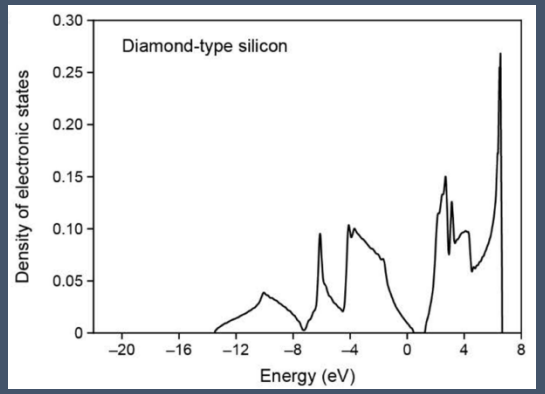
(2) Order-N computation of (many) moments of the spectral density of states

(3) Maximum-entropy reconstruction of the density of states from moments

Results: electrons

- Snapshots of the system through the pressurization run: examine the electronic density of states.
- Metallicity tracked by DOS(E_f).
- System "goes metallic" above 10 GPa, drops off some with s-h crystallization.
- Very High Density Amorphous DOS is **very** similar to 1500K liquid at similar pressure.
- *Caveat emptor*: Simple Hamiltonian, fit to some high-pressure configurations. Conduction states leave something to be desired.





2.1 million-atom fragment of diamond (same method)

Conclusion: GAP/Silicon/Pressure

- Machine Learning techniques are emerging as a meaningful tool in simulation. Opens some new doors.
- Squeeze a-Si: Abrupt collapse into a transient high density/coordination state. Then rapid crystallization to simple hexagonal phase. Does not happen in 1000-atom cell!
- Lots of new frontiers: now we are looking at surfaces. Collaborators are working on Carbon, GeSbTe (phase change memory) materials, others.

GAP: comments from a cheerleader, not an expert

- Given a very large sampling of accurate (DFT) computations of forces for an “adequately diverse and representative” set of configurations, GAP estimates the forces by fitting/interpolating from its library of configurations.
- If ever the devil is in the details, it is in building ML potentials:

1) How many configurations are enough?
2) Have we sampled all salient environments?
3) How do we represent a local environment?
4) Error estimation is built in – if there is nothing close in the database, demand a new DFT calculation.
5) When this is done properly, it is **not cheap**. For less than 200 atoms, cheaper to use planewave DFT! **But, it is order N**