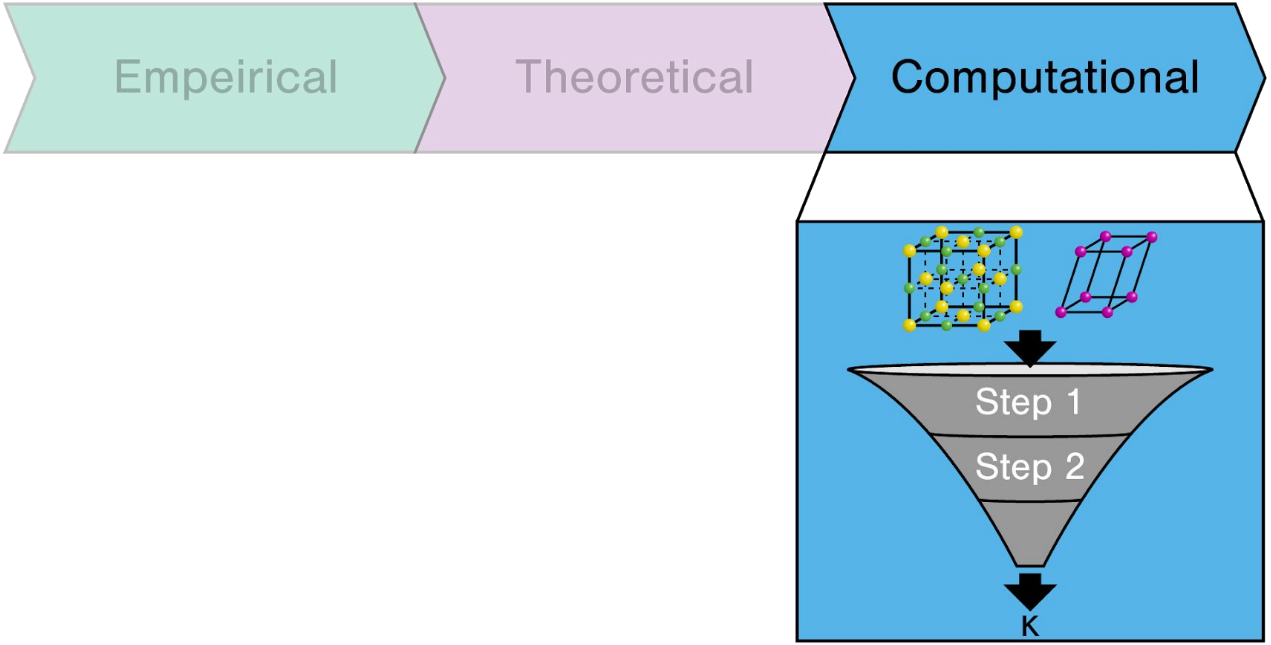
Developing High-Throughput Thermal Transport Workflows

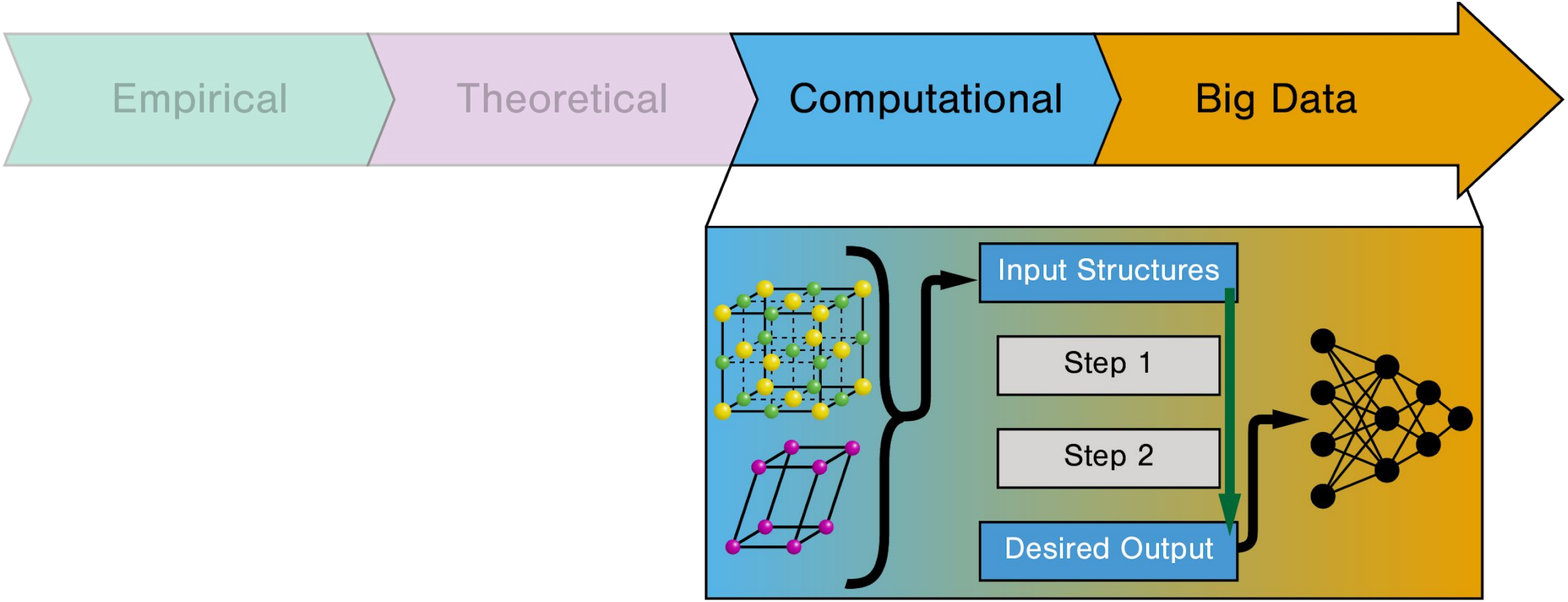
Thomas Purcell, Florian Knoop, Matthias Scheffler, Christian Carbogno

September 3, 2019
Hands on DFT 2019

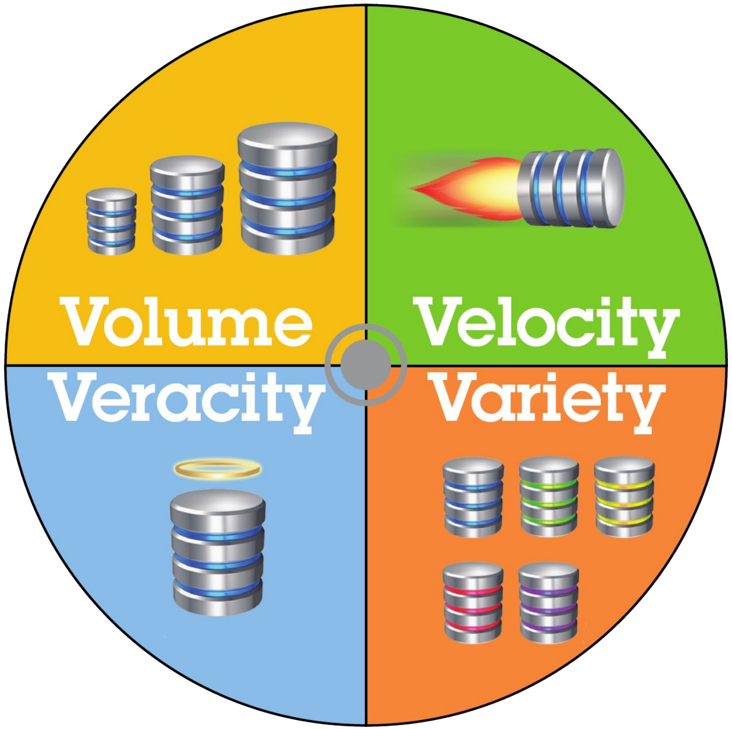
High-Throughput Computation: Faster Computational Results



High-Throughput Computation: Generate Data Sets for Big Data



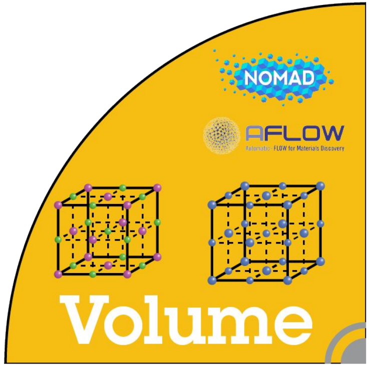
Developing Workflows with the 4 V's



- Volume
  - What is the size of the data set?
  - How much is needed to process?
- Velocity
  - How fast is data generated?
  - How fast is it processed?
- Veracity
  - Is the data accurate?
  - Can the errors be measured?
- Variety
  - How diverse sampling?
  - What does the data represent?

https://www.ibmbigdatahub.com/sites/default/files/infographic_file/4Vs_Infographic_final.pdf

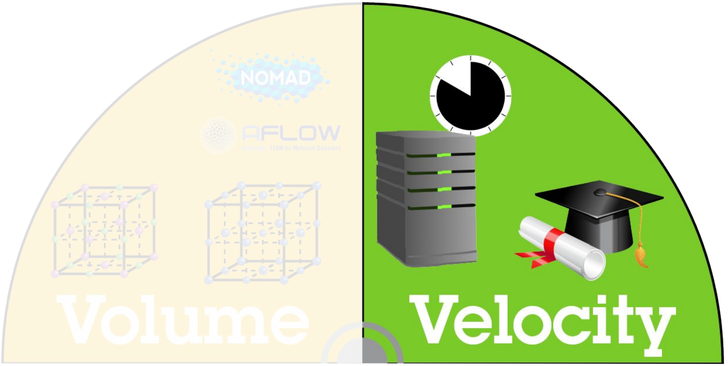
Volume: Do you actually want big data?



- Volume
    - What materials are being targeted?
    - How much data is needed to get desired results?

https://www.ibmbigdatahub.com/sites/default/files/infographic_file/4Vs_Infographic_final.pdf

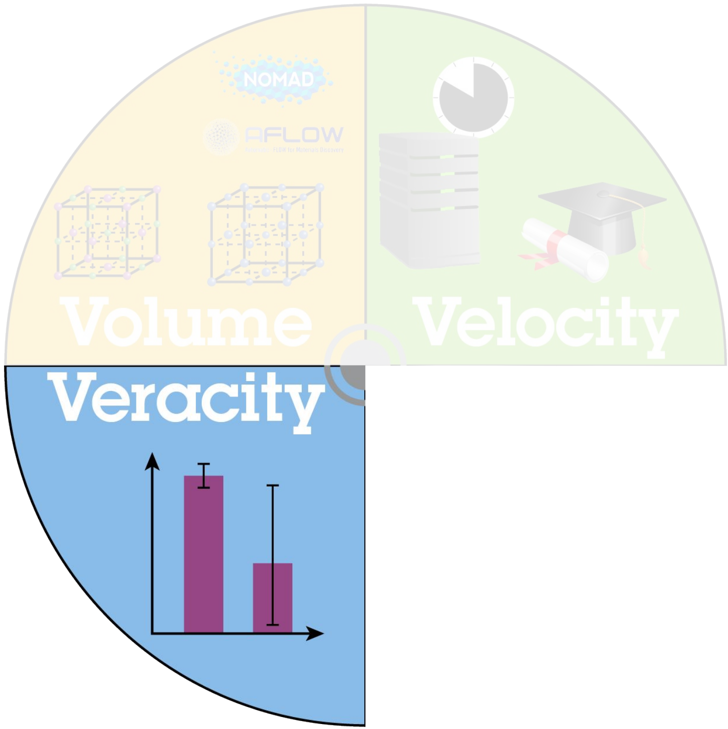
Velocity: Can you generate the data in a reasonable time-frame?



- Volume
  - What materials are being targeted?
  - How much data is needed to get desired results?
- Velocity
  - Are there enough computer resources?
  - Can this be done during a student's degree?

https://www.ibmbigdatahub.com/sites/default/files/infographic_file/4Vs_Infographic_final.pdf

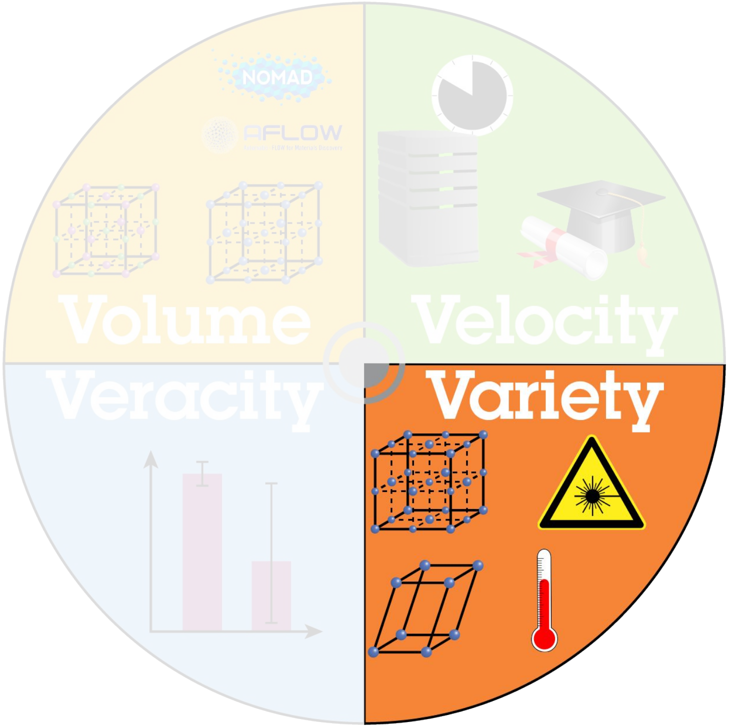
Veracity: What are the error tolerances?



- Volume
  - What materials are being targeted?
  - How much data is needed to get desired results?
- Velocity
  - Are there enough computer resources?
  - Can this be done during a student's degree?
- Veracity
  - How much error can be tolerated?
  - What trade offs between speed and accuracy can be made?

https://www.ibmbigdatahub.com/sites/default/files/infographic_file/4Vs_Infographic_final.pdf

Variety: What can the workflow be used for?



- Volume
  - What materials are being targeted?
  - How much data is needed?
- Velocity
  - Are there enough computer resources?
  - Can this be done during a student's degree?
- Veracity
  - How much error can be tolerated?
  - What trade offs between speed and accuracy can be made?
- Variety
  - How diverse is sampling?
  - How adaptable is the workflow?

https://www.ibmbigdatahub.com/sites/default/files/infographic_file/4Vs_Infographic_final.pdf

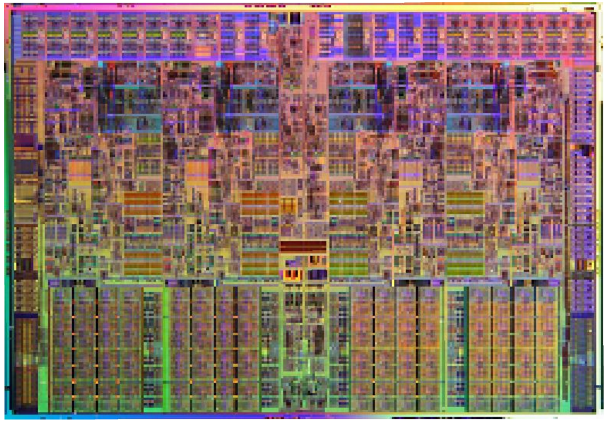
An Example Implementation

Understanding Thermal Transport Across Material Space

ZrO_2: Thermal Conductivity
Minute (3 M/mK)

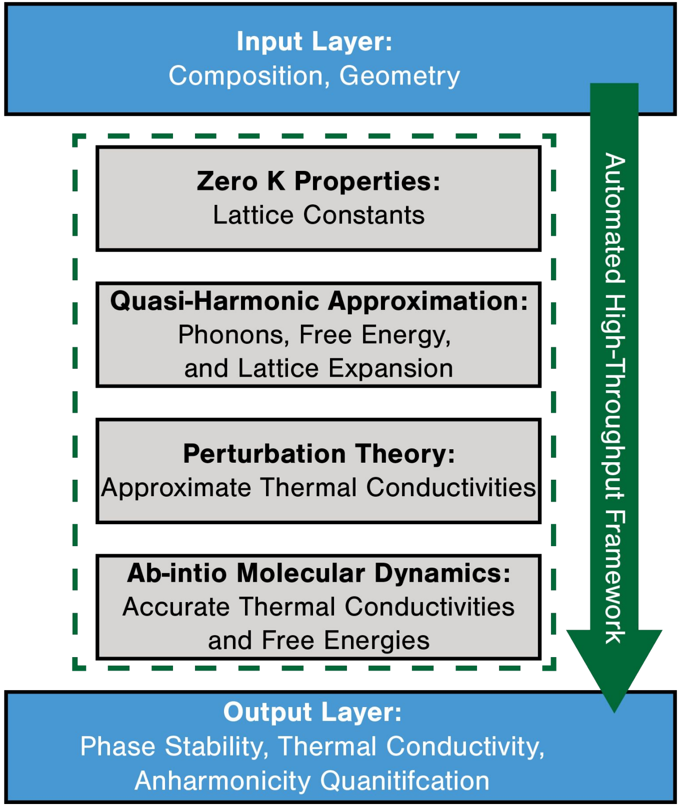
Suppress heat transfer
even further

Si: Thermal Conductivity
Huge (250 M/mK)

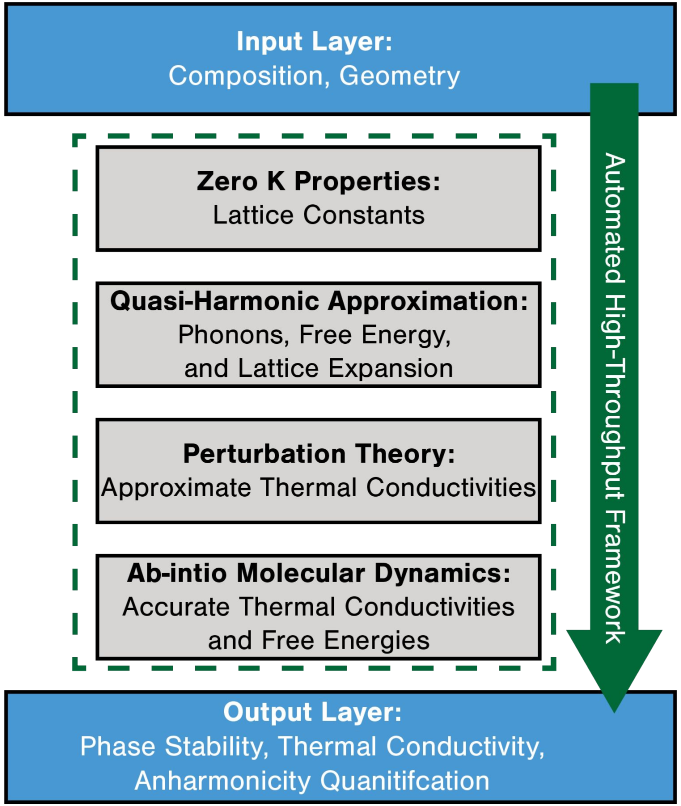


Boost heat transfer
even further

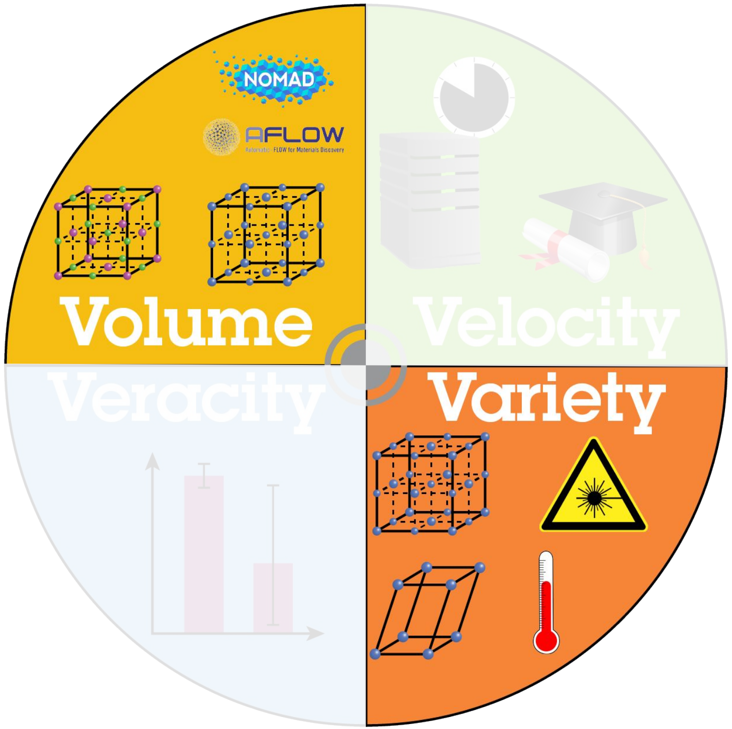
Ideal Workflow for Modeling Thermal Transport



Try to Avoid Reinventing the Wheel

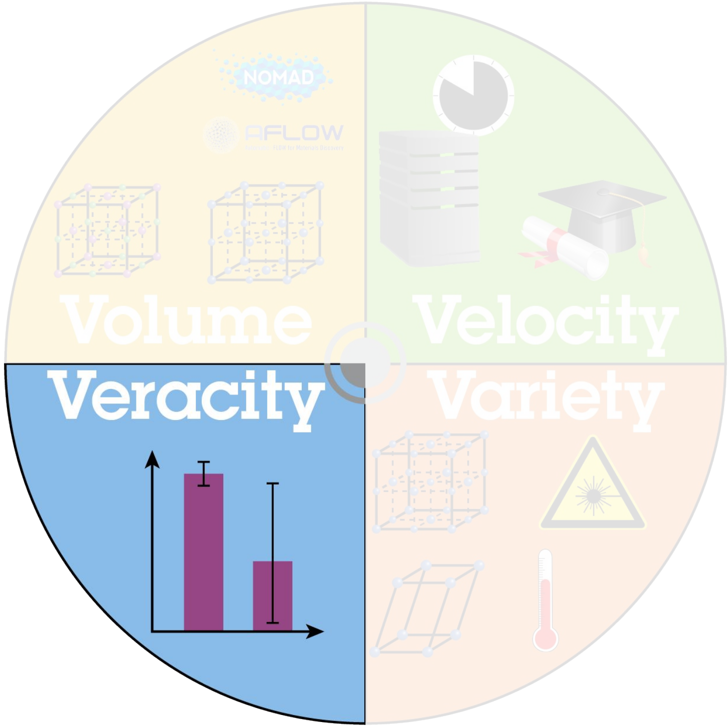


Does making a HTC workflow make sense?



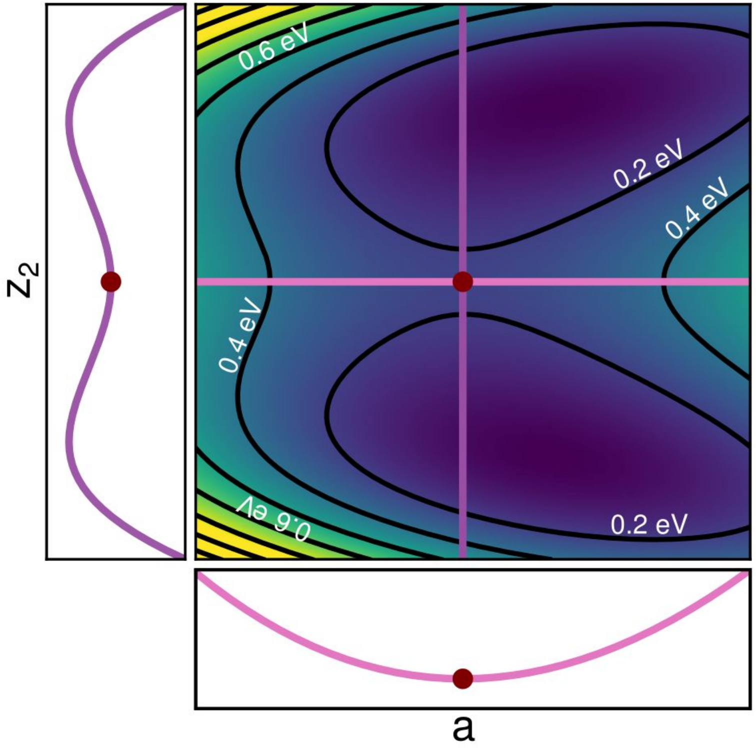
- Volume: Generalizable to any material
- Variety
  - Framework can be repurposed to handle other types of calculation
  - Material choices depend on the user

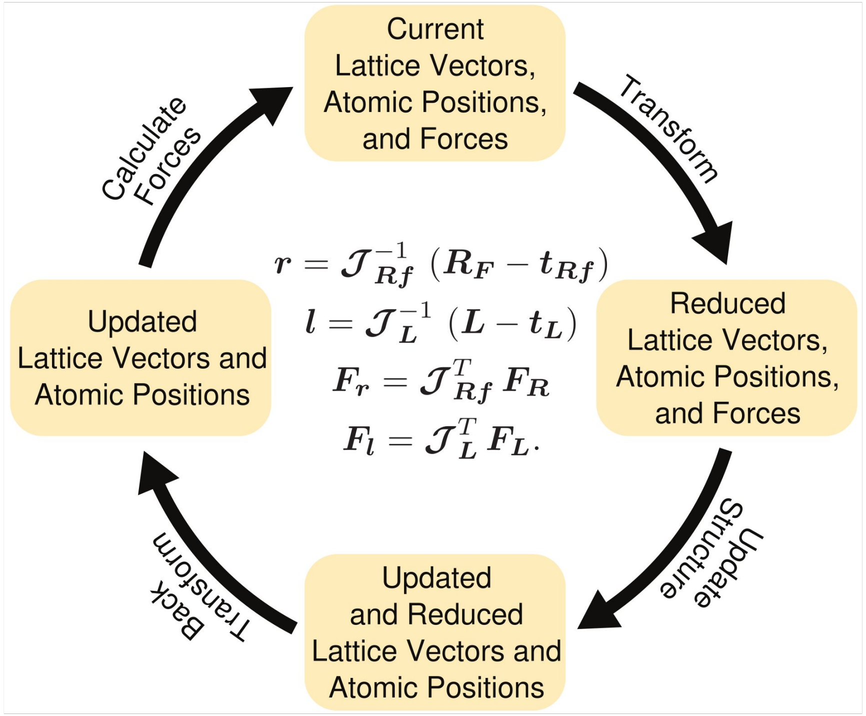
Problem: How to ensure material remains in desired structure?



- Volume: Generalizable to any material
- Variety
  - Framework can be repurposed to handle other types of calculation
  - Material choices depend on the user
- Veracity
  - Will the materials stay in the prototype after relaxation?

New Parametric Constraints in FHI-aims

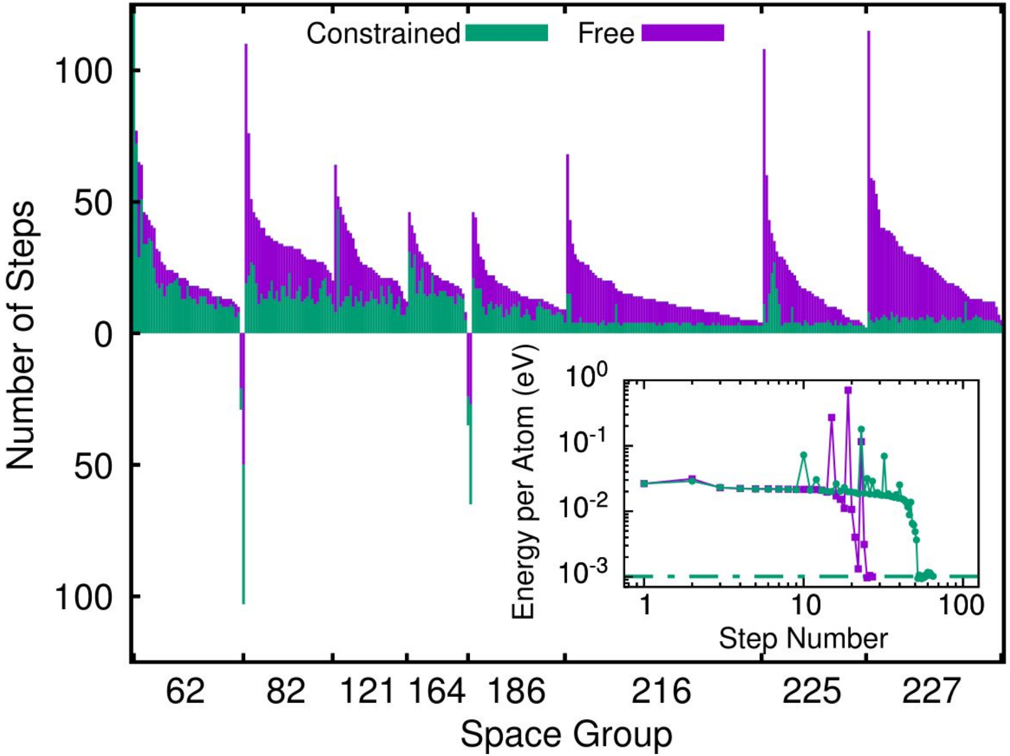




Lenz, M.-O., et al. *submitted to npj Computational Materials*

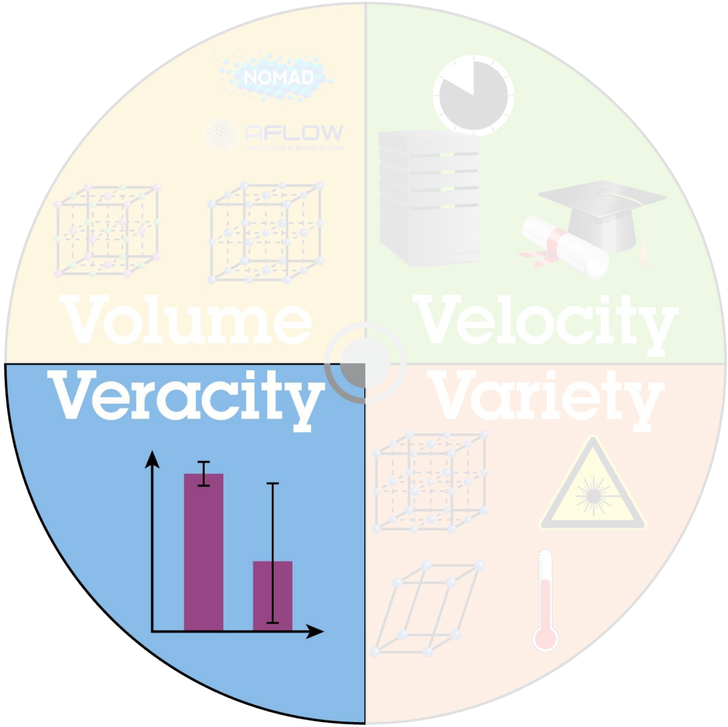
Preserving Structure While Accelerating Relaxations

Strukturbericht	Space Group	# of Materials	Average Savings
B16	62	8	24.61
C25	62	35	18.32
E3	82	35	51.75
$H2_6$	121	29	47.69
C6	164	25	19.68
B4	186	37	41.68
B3	216	35	54.47
$C1_b$	216	42	72.91
$L2_1$	225	11	79.47
C1	225	13	54.20
B1	225	17	51.19
A4	227	3	47.62
$H1_1$	227	50	73.58
Full Dataset		340	49.01



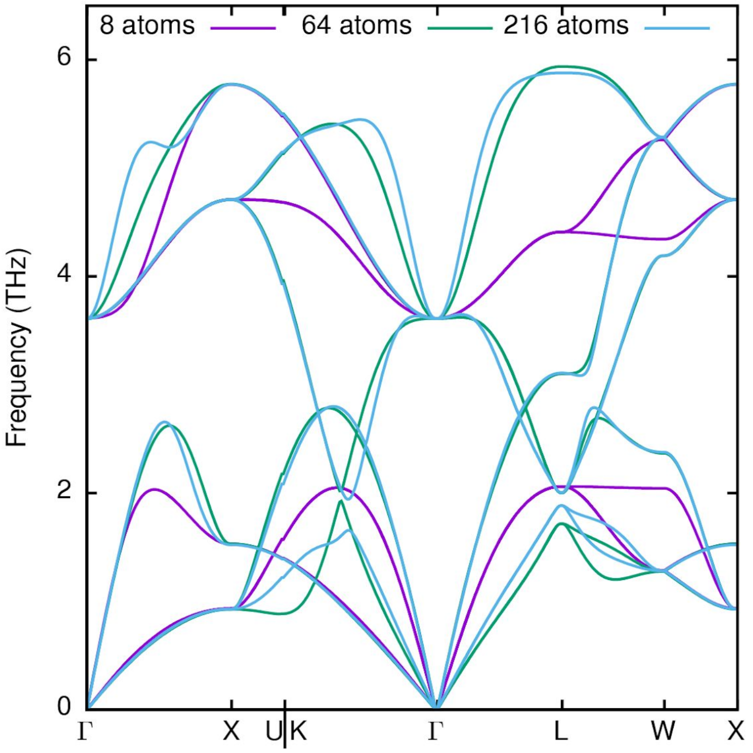
Lenz, M.-O., et al. *submitted to npj Computational Materials*

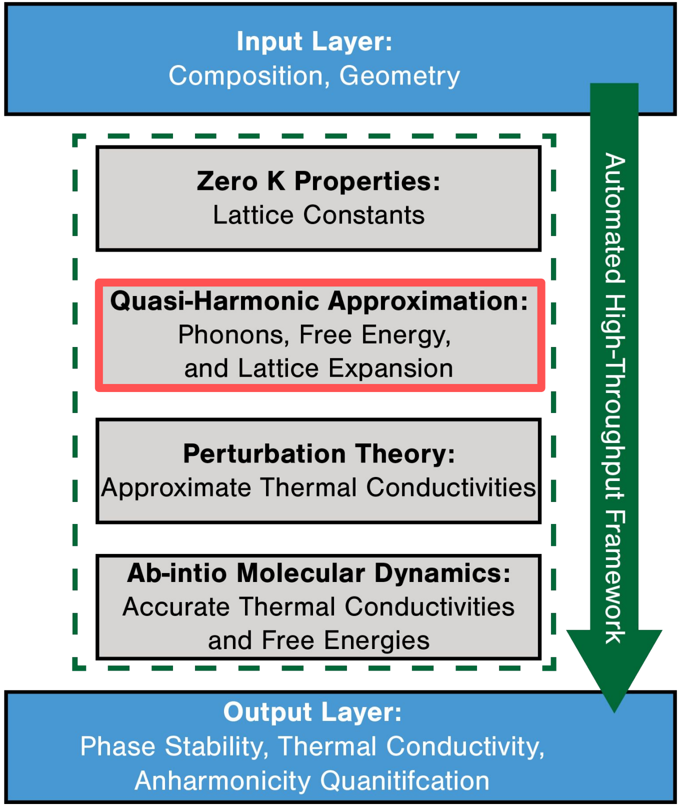
What about converging harmonic forces?



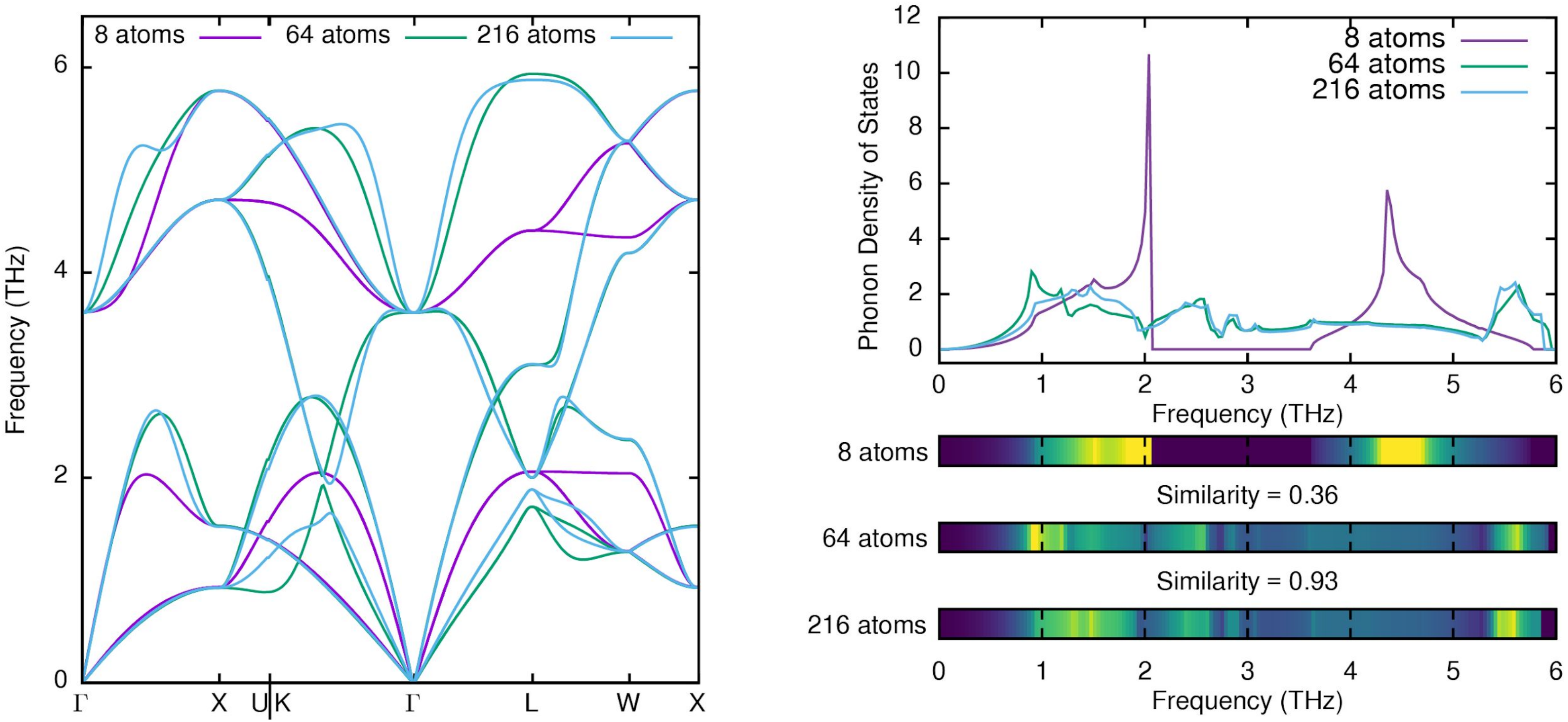
- Volume:
  - Built on top of ASE and FHI-aims
  - Generalizable to all materials
- Variety
  - Framework can be repurposed to handle other types of calculation
  - Material choices depend on the user
- Veracity
  - Will the materials stay in the prototype after relaxation?
  - How to tell if vibrational properties are converged?

The Problem of Unconverged Phonons



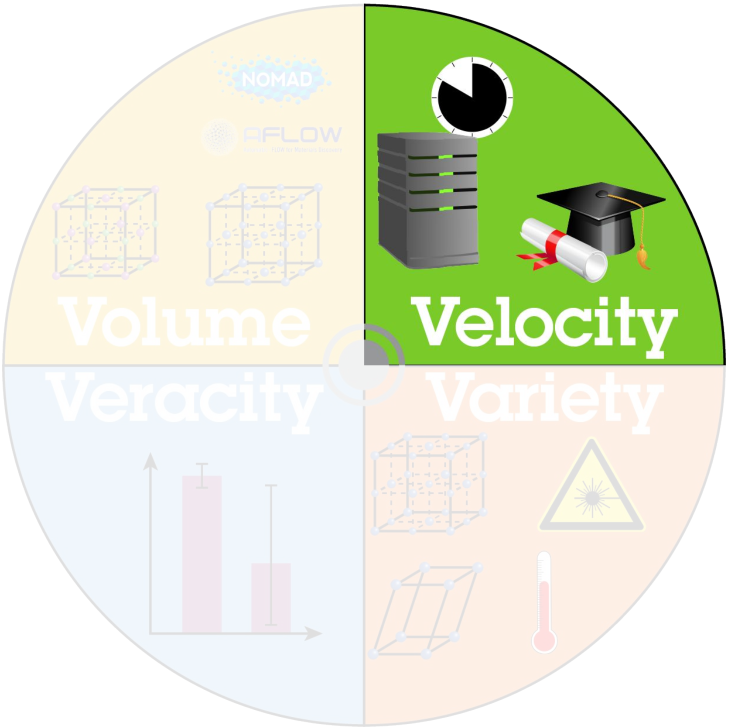


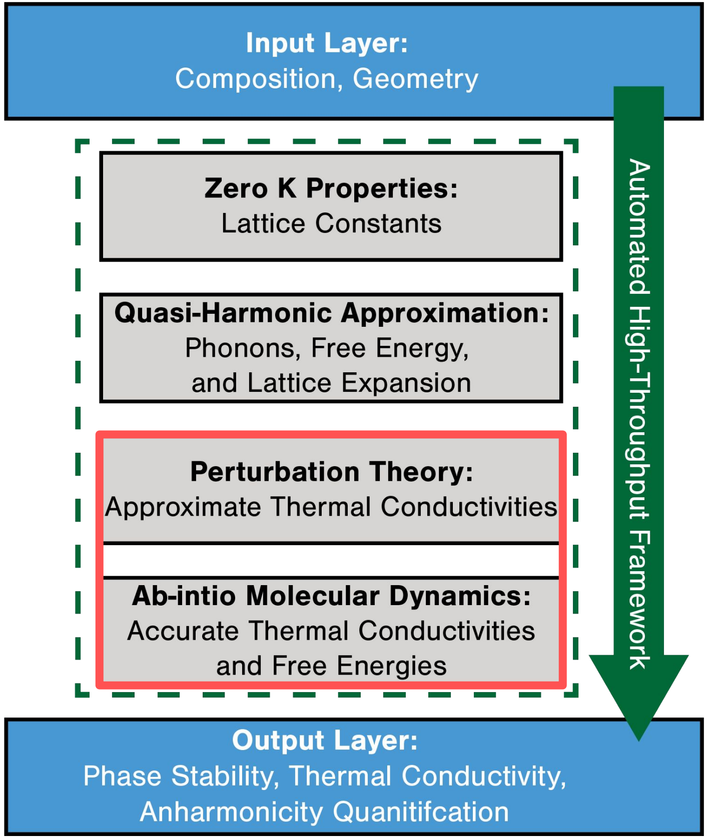
Fingerprinting as a Solution



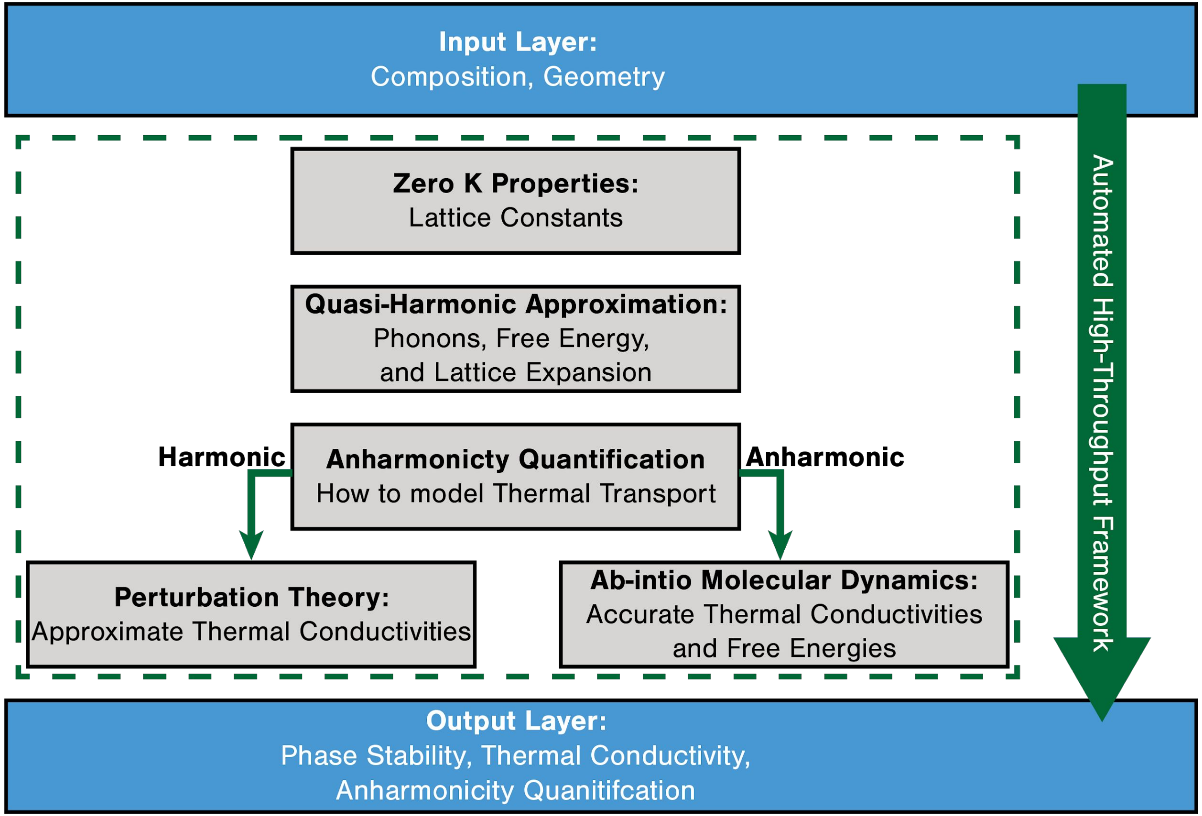
Chem. Mater. **27.** (2015) 735−743

How to calculate thermal conductivity quickly?





Determine which level of theory to use when



The Proposed Basis for the Switch

$$\boldsymbol{F} = \boldsymbol{F}^{\boldsymbol{HA}} + \boldsymbol{F}'$$

$$\langle \boldsymbol{F}' \rangle = \left\langle \boldsymbol{F} - \boldsymbol{F}^{\boldsymbol{HA}} \right\rangle$$

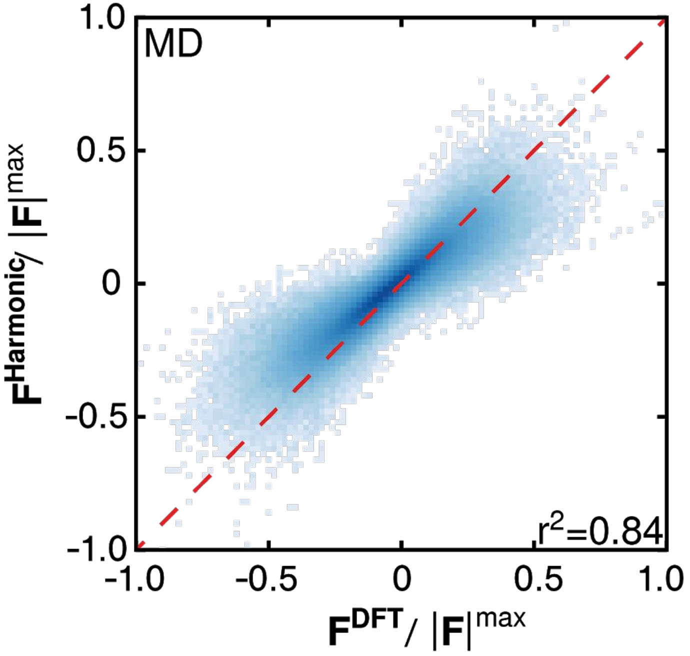
$$\propto 1 - r^2 \left(\boldsymbol{F}, \boldsymbol{F}^{\boldsymbol{HA}} \right)$$

$$r^2 \left(\boldsymbol{F}, \boldsymbol{F}^{\boldsymbol{HA}} \right) = 1 - \frac{SS_{res}}{SS_{tot}}$$

$$SS_{res} = \sum_i \left(F_i - F_i^{HA} \right)^2$$

$$SS_{tot} = \sum_i \left(F_i - \overline{\boldsymbol{F}} \right)^2$$

Molecular Dynamics is the Best Way to Calculate r^2



Can sampling get the same results?

$$\Phi_{ij} = \left.\frac{\partial^2 E}{\partial \mathbf{R}_i \partial \mathbf{R}_j}\right|_{\mathbf{R}^0} = -\left.\frac{\partial}{\partial \mathbf{R}_i}\mathbf{F}_j\right|_{\mathbf{R}^0} \approx -\frac{\mathbf{F}_j\left(\mathbf{R}_i^0 + \varepsilon \mathbf{d}_i\right)}{\varepsilon}$$

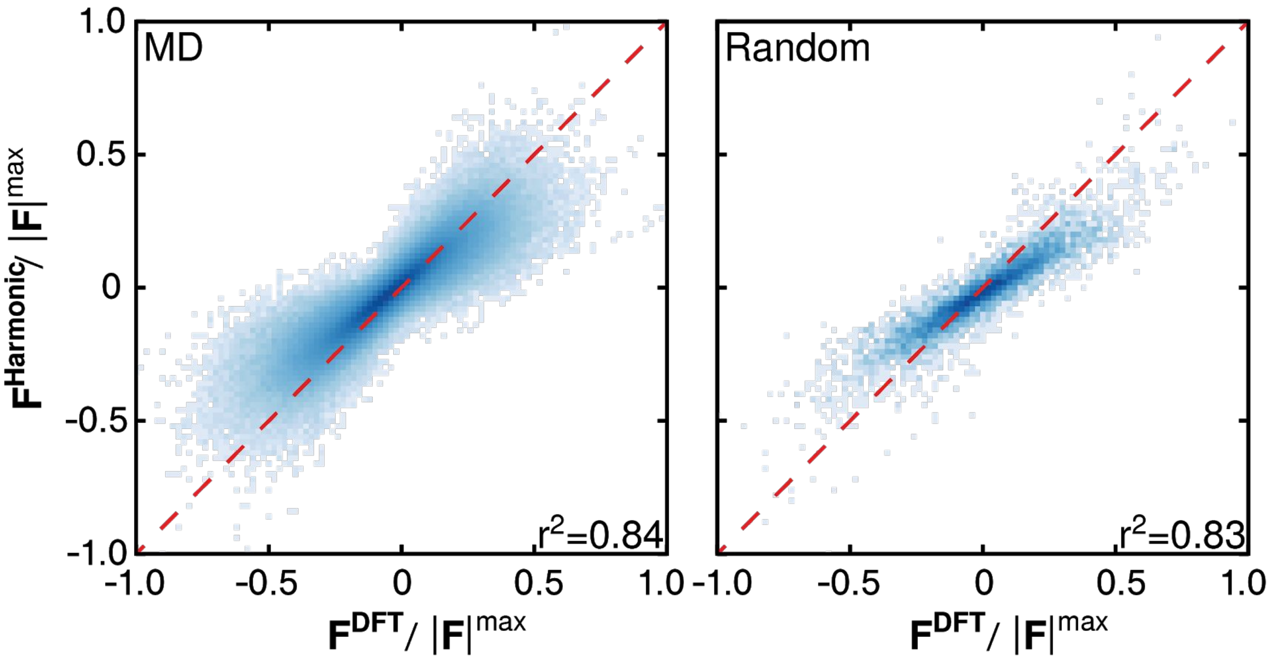
$$D_{i'j'}(\mathbf{q}) = \sum_j \frac{e^{i\mathbf{q}\cdot\left(\mathbf{R}_j^0 - \mathbf{R}_{j'}^0\right)}}{\sqrt{M_{i'} M_{j'}}} \Phi_{i'j}$$

$$\mathbf{D}(\Gamma)\left[\nu_s\left(\Gamma\right)\right] = \omega_s^2\left(\Gamma\right)\left[\nu\left(\Gamma\right)\right]$$

$$A_s = \frac{\sqrt{k_B T}}{\omega_s}\sqrt{-2\,\ln\left(1 - R|_0^1\right)}\,\sin\left(2\pi R|_0^1\right)$$

$$d_s = \sum_{s=3}^{n_\omega} A_s \nu_s\left(\Gamma\right)$$

Random Sampling Gives Similar Results



Can the sampling be done with one supercell?

$$\Phi_{ij} = \left.\frac{\partial^2 E}{\partial \mathbf{R}_i \partial \mathbf{R}_j}\right|_{\mathbf{R}^0} = -\left.\frac{\partial}{\partial \mathbf{R}_i}\mathbf{F}_j\right|_{\mathbf{R}^0} \approx -\frac{\mathbf{F}_j\left(\mathbf{R}_i^0 + \varepsilon \mathbf{d}_i\right)}{\varepsilon}$$

$$D_{i'j'}(\mathbf{q}) = \sum_j \frac{e^{i\mathbf{q}\cdot\left(\mathbf{R}_j^0 - \mathbf{R}_{j'}^0\right)}}{\sqrt{M_{i'} M_{j'}}} \Phi_{i'j}$$

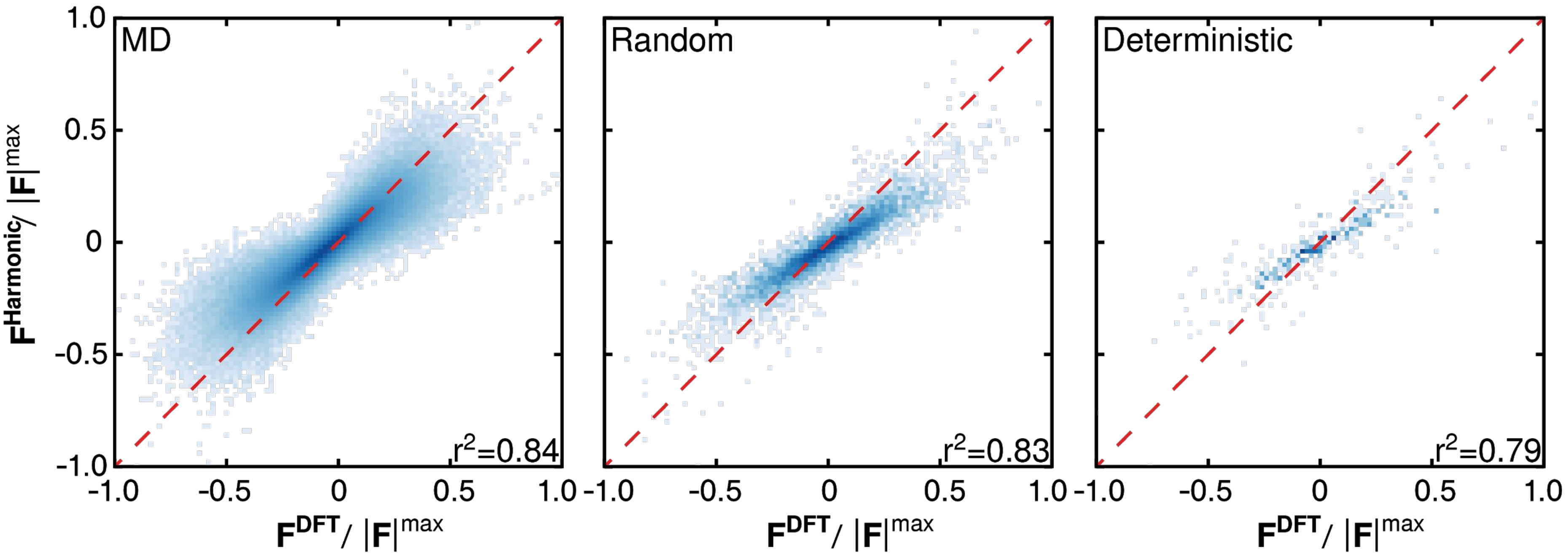
$$\mathbf{D}(\Gamma)\left[\nu_s(\Gamma)\right] = \omega_s^2(\Gamma)\left[\nu(\Gamma)\right]$$

$$A_s = \frac{\sqrt{k_B T}}{\omega_s}$$

$$d_s = \sum_{s=3}^{n_\omega} -1^s A_s \nu_s(\Gamma)$$

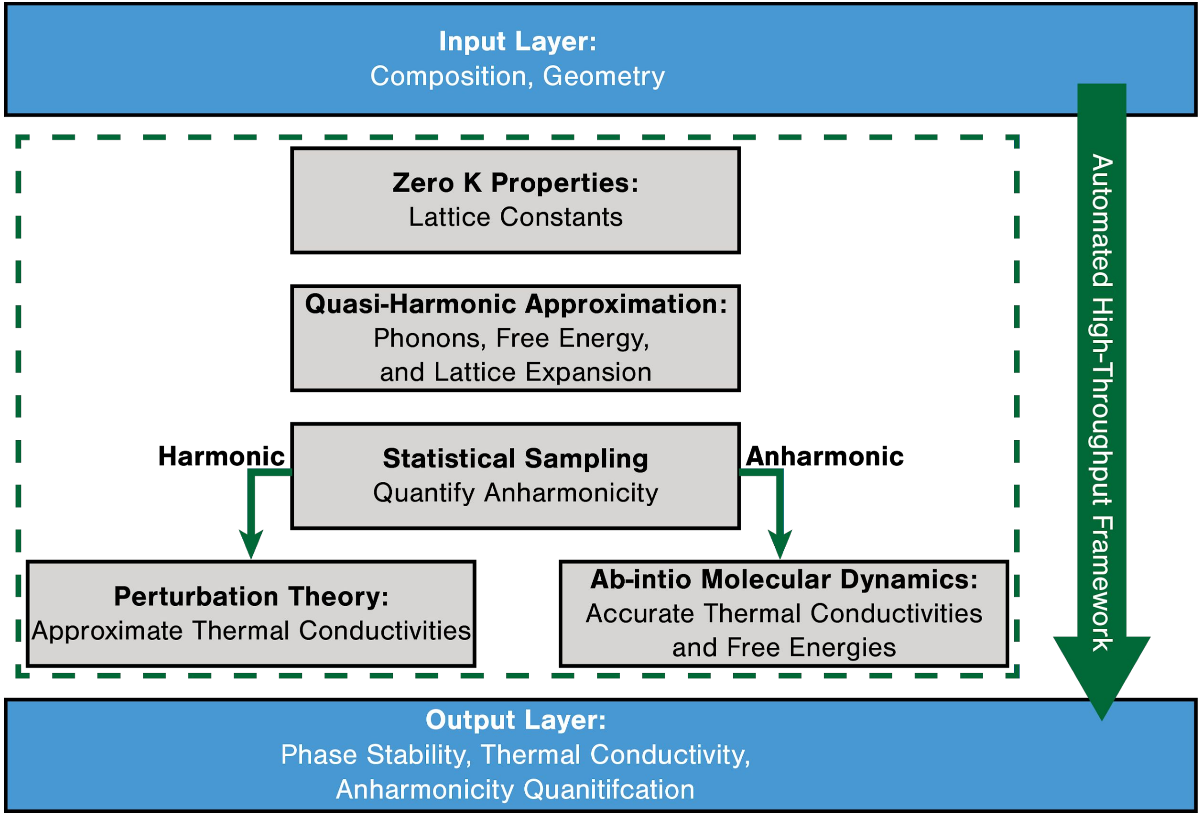
Zacharias, M. and Giustino, F. *Phys. Rev. B.* **94.** (2016) 075125

Deterministic Sampling Quick and Reasonable Results

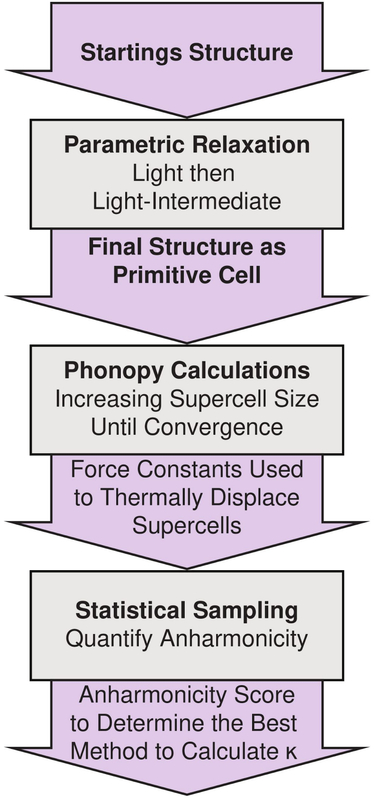


Zacharias, M. and Giustino, F. *Phys. Rev. B.* **94.** (2016) 075125

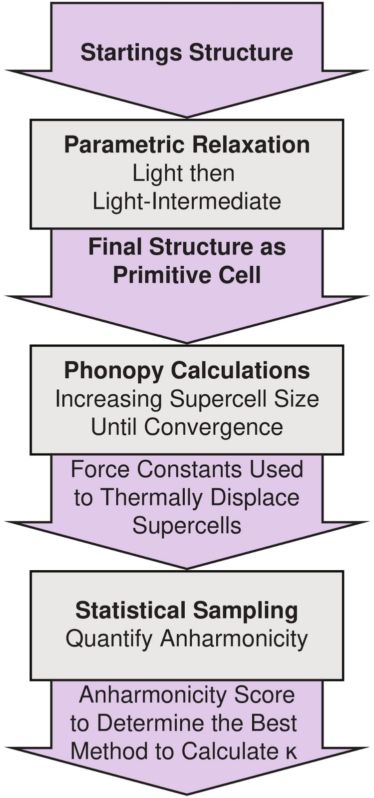
The New General Workflow

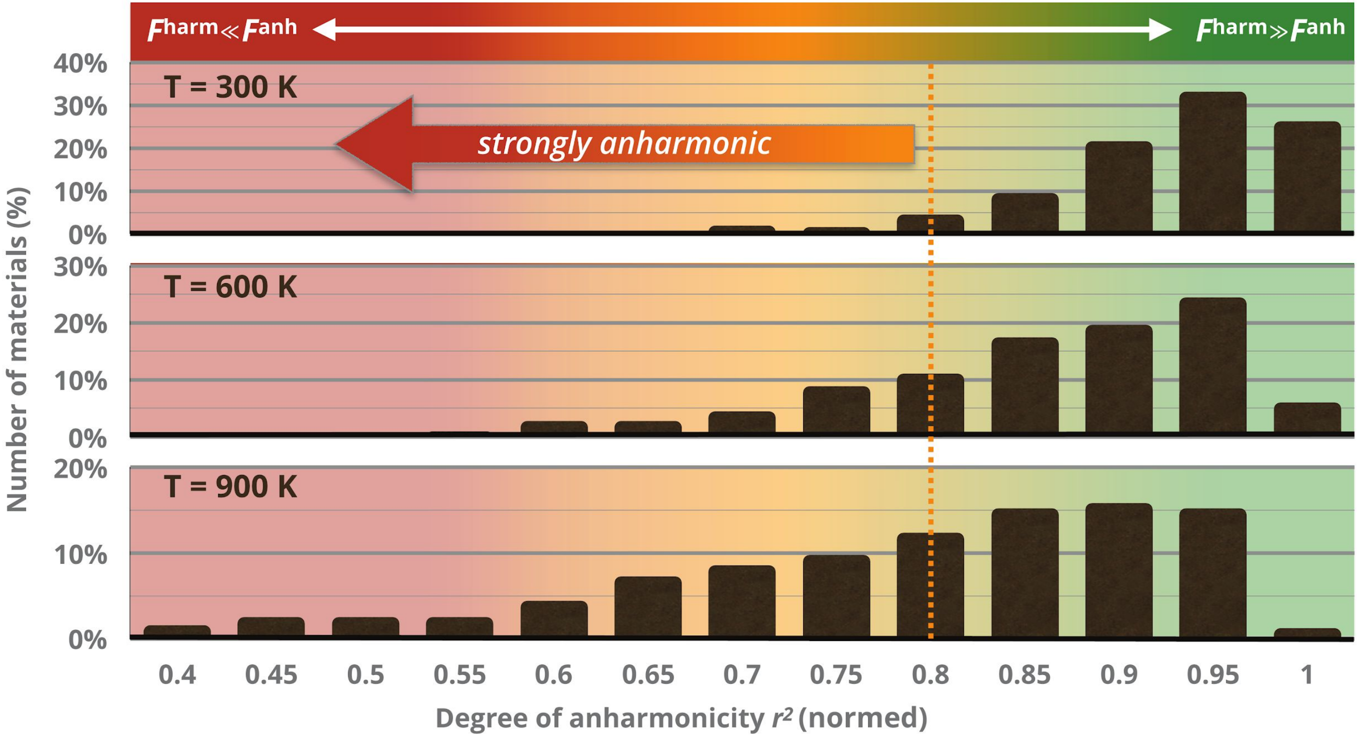


How to Calculate r^2 Efficiently

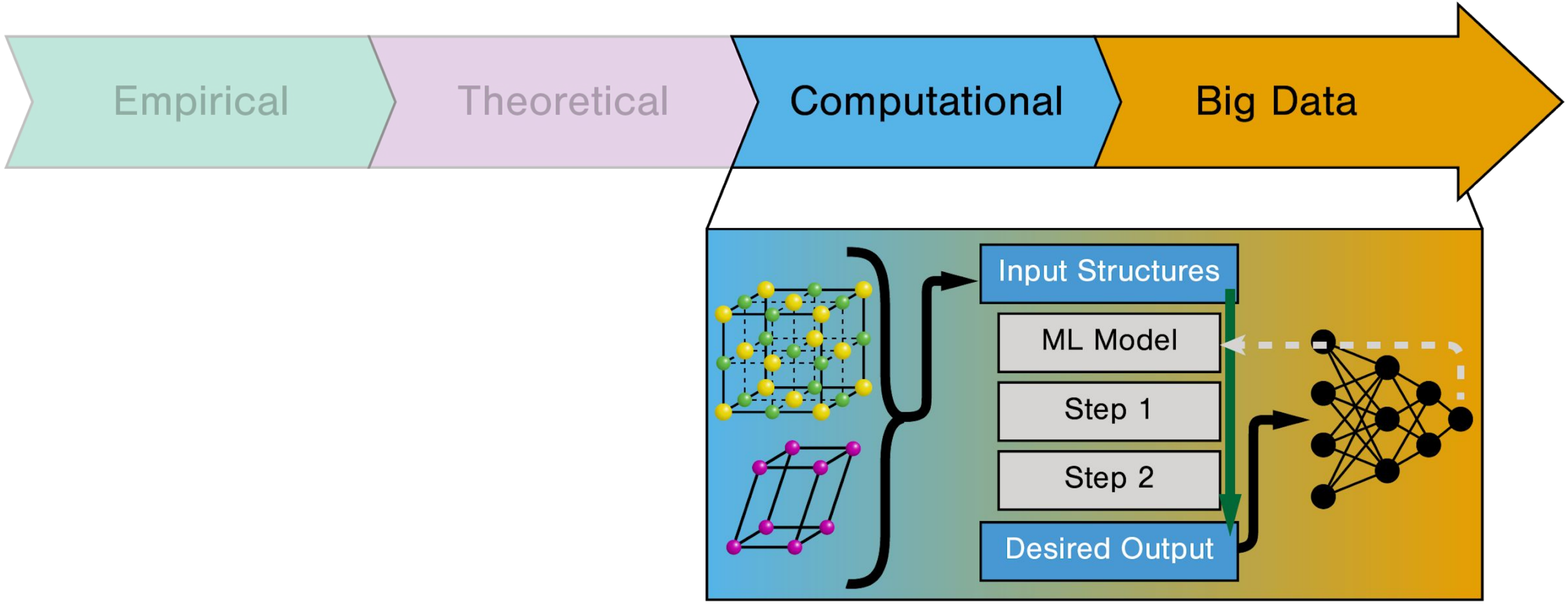


Anharmonic Effects are More Important at Higher Temperatures

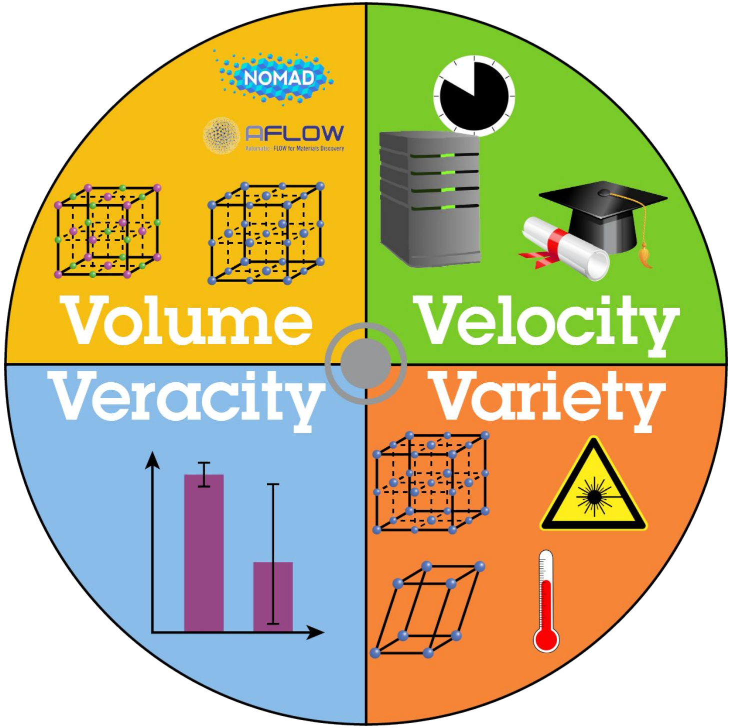


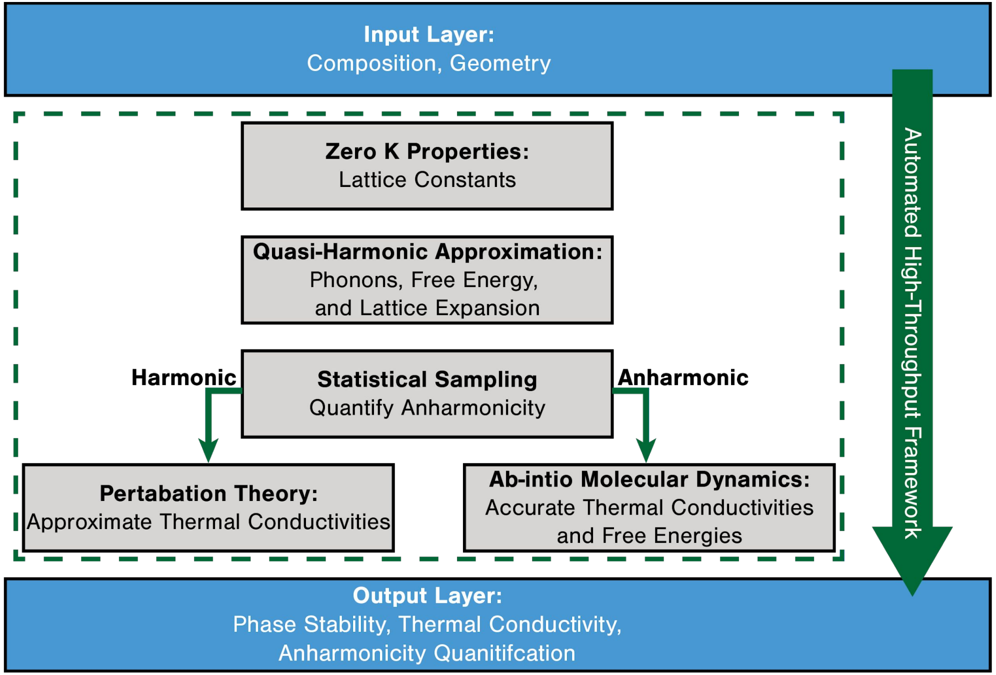


But can we do better?



Summary


NOMAD
AFLOW
Volume
Velocity
Veracity
Variety


Input Layer:
Composition, Geometry
Zero K Properties:
Lattice Constants
Quasi-Harmonic Approximation:
Phonons, Free Energy,
and Lattice Expansion
Harmonic
Statistical Sampling
Quantify Anharmonicity
Anharmonic
Pertabation Theory:
Approximate Thermal Conductivities
Ab-intio Molecular Dynamics:
Accurate Thermal Conductivities
and Free Energies
Automated High-Throughput Framework
Output Layer:
Phase Stability, Thermal Conductivity,
Anharmonicity Quanitifcation

It’s Also Implemented in FHI-vibes