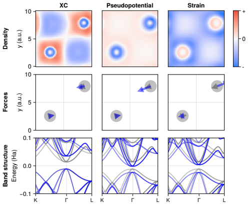
Basis set error estimation in plane-wave density-functional theory

Michael F. Herbst

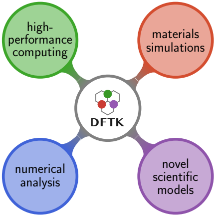
Mathematics for Materials Modelling (`matmat.org`), EPFL

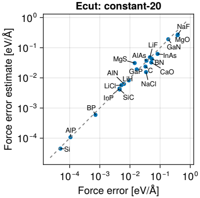
20 April 2026

`https://michael-herbst.com/talks/2026.04.20_DFT_Error_Oslo.pdf`



Gradients of DFT quantities wrt. DFT parameters





Error estimate in atomic forces

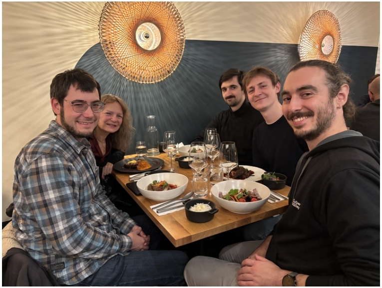
Acknowledgements

MatMat group

Niklas Schmitz

Bruno Ploumhans

- Noe Blassel
- Nathanael Bosch
- Benedikt Menges



Fonds national suisse

MARVEL

NATIONAL CENTRE OF COMPETENCE IN RESEARCH

EPFL MatMat

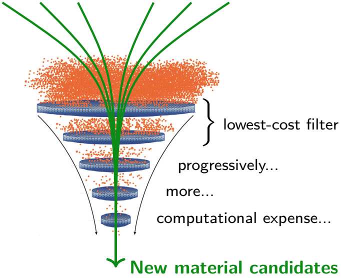
Impact of computational materials discovery

- Current solutions limited by properties of available materials
  - ⇒ Innovation driven by discovering new materials
- Experimental research extremely energy intensive
  - 1 fume hood $\simeq$ 2-3 average households[1]

⇒ Complement experiment by computational materials discovery

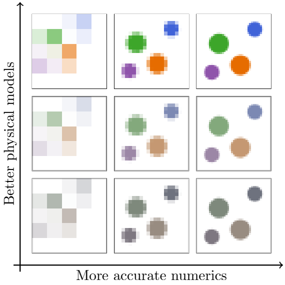
[1] D. Wesolowski *et. al.* Int. J. Sustain. High. Edu. **11**, 217 (2010).

Does everything need to be equally accurate ?



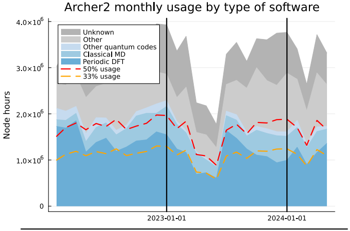
Computational materials design funnel

Involved data-driven multi-physics workflows



Multitude of modelling choices

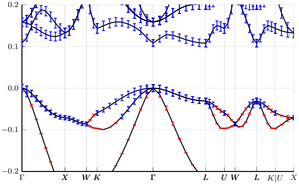
How should we spent our efforts best ?



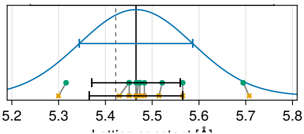
- Notable usage of computational time: >30% usage on key HPC clusters
- DFT computations: Training data for MLIPs
- Heteroscedastic regression models[1]: can tolerate data of varying quality provided error estimate available

[1] K. Fisher, MFH, Y. Marzouk. J. Chem. Phys. **161**, 014114 (2024).

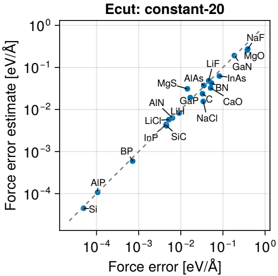
Larger thrust: Quantifying DFT simulation error



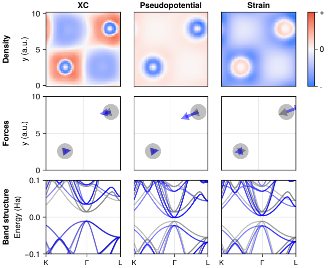
Band structure with guaranteed error bars[1]



DFT lattice constant error distribution[2]



Plane-wave basis error estimates at 20Ha[2,3]



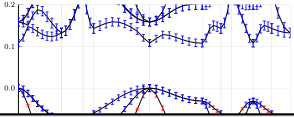
DFT quantity vs. parameter sensitivities[2]

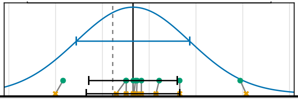
[1] MFH, A. Levitt, E. Cancès. Faraday Discus. **223**, 227 (2020).

[2] N. Schmitz, B. Ploumhans, MFH. npj Computat. Mater. **12**, 6 (2025).

[3] E. Cancès, G. Dusson, G. Kemlin *et. al.* SIAM J. Sci. Comp., **44**, B1312 (2022).

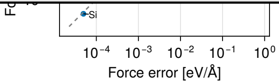
Larger thrust: Quantifying DFT simulation error



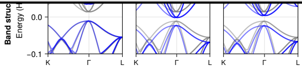


Towards **inexact computation without harm**:

- Quantitative error estimates hardly researched
- Challenge: Turning mathematical analysis into practical tool
- Needs combination of approaches:
  - Efficient algorithms (e.g. randomisation, inexact Krylov)
  - Modern implementation techniques (e.g. AD)
  - Analytic & statistical approaches (e.g. probabilistic) to error estim.
  - Physically-inspired heuristics



Plane-wave basis error estimates at 20Ha[2,3]



DFT quantity vs. parameter sensitivities[2]

[1] MFH, A. Levitt, E. Cancès. Faraday Discus. **223**, 227 (2020).

[2] N. Schmitz, B. Ploumhans, MFH. npj Computat. Mater. **12**, 6 (2025).

[3] E. Cancès, G. Dusson, G. Kemlin *et. al.* SIAM J. Sci. Comp., **44**, B1312 (2022).

Key ingredient: Flexible DFT uncertainty propagation

- θ: Model, structure or discretisation parameters
- DFT energy minimisation yields minimal density D_*

$$D_*(\theta) = \arg\min_{D \in \mathcal{P}} \mathcal{E}_{\mathsf{DFT}}(D, \theta)$$

- From these compute quantities of interest: $Q(D_*)$

- **Question:** Changing θ, how to find change in Q ?
- Approx. **answer:** Linear propagation, e.g. $\delta Q \approx \frac{dQ}{d\theta} \delta\theta$ with

$$\frac{dQ}{d\theta} = \frac{\partial Q}{\partial \theta} + \frac{\partial Q}{\partial D} \frac{\partial D_*}{\partial \theta}$$

- Note: This is DFPT (DFT response theory)

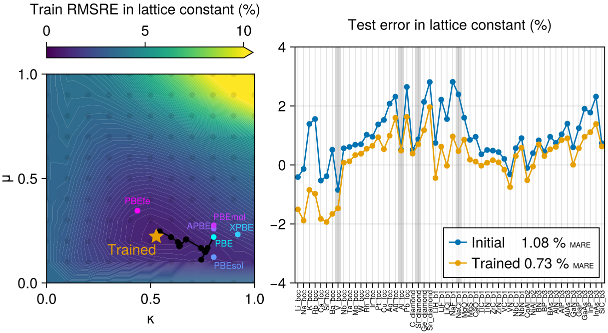
⇒ Need response framework for end-to-end gradient computation
  - **Talk by Niklas Schmitz on Friday** on AD-DFPT[1]
  - Builds on recent improvements on response algorithms[2,3]

[1] N. Schmitz, B. Ploumhans, MFH. npj Computat. Mater. **12**, 6 (2025).

[2] E. Cancès, MFH, G. Kemlin, *et. al.* Lett. Math. Phys. **113**, 21 (2023).

[3] MFH, B. Sun. SIAM J. Sci. Comp. *Efficient Krylov methods for lin. resp. in PW electronic structure*

AD-DFPT teaser slide[1] (more by Niklas on Friday)



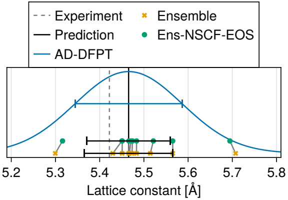
Inverse design of PBE-type model wrt. experimental lattice constant[1]

- Relaxed geometries
- SCF & geometry optimisation loop involved

- Uncertainty in lattice constant

$$a^*(\theta) = \arg\min_a \left(\min_\rho \ \mathcal{E}(\theta, a, \rho) \right)$$

- AD-DFPT compares well to more expensive alternatives



Model uncertainty in optimal lattice constant[1]

[1] N. Schmitz, B. Ploumhans, MFH. npj Computat. Mater. **12**, 6 (2025).

Today: Controlling plane-wave basis error

- Plane-wave DFT: Bloch waves $\psi_{nk}(r) = u_{nk}(r)e^{ik\cdot r}$
- Periodic part $u_{nk}(r)$ expanded in normalised plane-wave basis

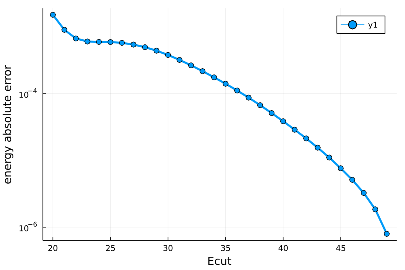
$$\left\{ \frac{1}{\sqrt{\Omega}} e^{iG\cdot r} \,\middle|\, \frac{1}{2} \left\|G + k\right\|^2 \leq E_{\text{cut}} \right\}$$

$\Rightarrow$ Kinetic energy cutoff E_{cut} determines basis set size

- Convergence of DFT quantity Q wrt. E_{cut} depends on
  - **Primarily:** Employed pseudopotential
  - **But also:** Considered system & considered quantity Q

Common strategies to choose E_{cut}

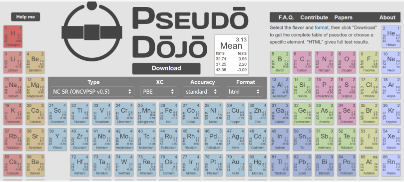
Explicit convergence study



DFTK example (see `docs.dftk.org`)

- Computationally intensive
- Impractical for high-throughput

Tabulated recommendations



`http://www.pseudo-dojo.org/`

- E_{cut} Values per element & pseudo
- Values independent of quantity Q
- Not fully reliable

$\Rightarrow$ Convergence study still recommended

- **Our goal:** Cheap *a posteriori* estimate of discretisation error
  - Adapted to system & property of interest
  - Generically applicable across pseudos & properties

Two-grid approach to error estimation[1]

- Assume plane-wave basis with cutoff E_{ref} yields exact result

⇒ *Fictitious* reference density matrix obtained by minimisation problem in Grassmann manifold; $\mathcal{N}$: size of E_{ref}-basis

$$P_* = \underset{D \in \mathcal{P}^{E_{\text{ref}}}}{\arg\min}\, \mathcal{E}_{\mathsf{DFT}}(D)$$

$$\mathcal{P}^{E_{\text{ref}}} = \left\{ D \in \mathcal{C}^{\mathcal{N}\times\mathcal{N}} \,\middle|\, D^2 = D, \operatorname{tr}(D) = N, D^H = D \right\}$$

- We actually solve $P = \underset{D \in \mathcal{P}^{E_{\text{cut}}}}{\arg\min}\, \mathcal{E}_{\mathsf{DFT}}(D)$ (i.e. in standard basis)
- Corresponding residual due to insufficient basis (manifold geometry):

$$R(P) = [P, [P, H(P)]] = \Pi_{T_P \mathcal{P}} H(P) \quad \text{with } H(P) = \nabla_P \mathcal{E}_{\mathsf{DFT}}(P)$$

- **Goal:** Estimate $\delta P \approx P_* - P$
  - ⇒ Correct density to $\mathfrak{r}(P + \delta P)$ (with $\mathfrak{r}$ retraction to $\mathcal{P}^{E_{\text{ref}}}$)
  - ⇒ Estimate error as $\|\delta P\|$ & propagate as $\big\|Q\big(\mathfrak{r}(P + \delta P)\big)\big\|$

[1] E. Cancès, G. Dusson, G. Kemlin *et. al.* SIAM J. Sci. Comp., **44**, B1312 (2022).

Key idea: Perturbation-based correction

- Consider a $f : \mathbb{R} \to \mathbb{R}$ with $f(x_*) = 0$ and x close to x_*, then:

$$0 = f(x_*) = f(x) + f'(x) \cdot (x_* - x) + O(\|x_* - x\|^2)$$

$$\Rightarrow \quad x_* - x \approx -(f'(x))^{-1} f(x) \qquad \text{(Newton's method)}$$

- In our case:

$$0 = R(P_*) \approx R(P) + dR(P) \cdot (P_* - P)$$

$$\Rightarrow \quad \delta P = P_* - P \approx -(dR(P))^{-1} R(P)$$

- Differentiating the residual $R(P) = [P, [P, H(P)]]$ (with $X \in T_P\mathcal{P}$)

$$dR(P) \cdot X = \underbrace{[X, [P, H(P)]]}_{=0 \text{ at } P_*} + \underbrace{[P, [X, H(P)]]}_{=\Omega(P)\cdot X} + \underbrace{[P, [P, dH(P) \cdot X]]}_{=K(P)\cdot X}$$

$$\approx (\Omega(P) + K(P)) \cdot X \quad \text{(Assume } P \simeq P_* \Rightarrow \text{drop 1st term)}$$

$\Rightarrow$ We obtain $\delta P \approx (\Omega + K)^{-1} R(P)$

Key idea: Perturbation-based correction

- Consider a $f : \mathbb{R} \to \mathbb{R}$ with $f(x_*) = 0$ and x close to x_*, then:

$$0 = f(x_*) = f(x) + f'(x) \cdot (x_* - x) + O(\|x_* - x\|^2)$$

$$\Rightarrow \quad x_* - x \approx -(f'(x))^{-1} f(x) \qquad \text{(Newton's method)}$$

- In our case:

$$0 = R(P_*) \approx R(P) + dR(P) \cdot (P_* - P)$$

$$\Rightarrow \quad \delta P = P_* - P \approx -(dR(P))^{-1} R(P)$$

- Differentiating the residual $R(P) = \big[P, [P, H(P)]\big]$ (with $X \in T_P\mathcal{P}$)

$$dR(P) \cdot X = \underbrace{\big[X, [P, H(P)]\big]}_{=0 \text{ at } P_*} + \underbrace{\big[P, [X, H(P)]\big]}_{=\Omega(P) \cdot X} + \underbrace{\big[P, [P, dH(P) \cdot X]\big]}_{=K(P) \cdot X}$$

$$\approx (\Omega(P) + K(P)) \cdot X \qquad \text{(Assume } P \simeq P_* \Rightarrow \text{drop 1st term)}$$

$\Rightarrow$ We obtain $\delta P \approx (\Omega + K)^{-1} R(P)$

Algorithm: Perturbation-based plane-wave error estimation[1]

- "$_1$": quantities on small basis E_{cut}
- "$_2$": quantities on large basis $E_{\text{ref}} \gg E_{\text{cut}}$

1. Obtain P via an SCF in E_{cut}, transfer to E_{ref}-basis
2. Compute δP with approx. perturbation theory (approximate Newton step)

$$\delta P = -\begin{pmatrix} (\Omega+K)_{11} & (\Omega+K)_{12} \\ 0 & M_{22} \end{pmatrix}^{-1} R(P)$$

   Equivalently:

$$\delta P_2 = -M_{22}^{-1} R_2$$
$$\delta P_1 = -(\Omega+K)_{11}^{-1}\left(R_1 - (\Omega+K)_{12}\delta P_2\right)$$

   - $\Omega + K$: Approx. Riemannian Hessian of $\mathcal{E}(P)$ at P
   - $R(P) = [P, [P, H(P)]]$: Residual, Riemannian gradient
   - $(\Omega+K)_{22} \approx M_{22}$ & $(\Omega+K)_{21} \approx 0$ (approx. high-frequency response)
   - Dominant cost: $(\Omega+K)_{11}^{-1} \Rightarrow$ Only in the **small** basis
3. Propagate error to quantity of interest, e.g. $\delta Q \approx Q\big(\mathfrak{r}(P+\delta P)\big)$

[1]E. Cancès, G. Dusson, G. Kemlin *et. al.* SIAM J. Sci. Comp., **44**, B1312 (2022).

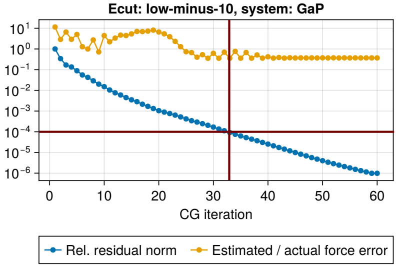
Making perturbative error estimates practical[1]

- Selection of E_{ref}?
  - **Answer:** See ideas on next slides
- This is a linearised correction: Is it valid across pseudopotentials & E_{cut}s?
  - **Answer:** Yes, see numerical benchmarks
- How to propagate: $\delta Q \approx Q\big(\mathfrak{r}(P + \delta P)\big)$ vs. $\delta Q \approx Q(P) + \frac{dQ}{dP}\delta P$
  - **Answer:** $Q\big(\mathfrak{r}(P + \delta P)\big)$ more accurate[1]
- Stopping criterion for inverting $(\Omega + K)_{11}$?
  - **Answer:** Use relative error around 10^{-4} in CG

[1] B. Ploumhans, MFH. *Practical plane-wave error estimation for Kohn-Sham DFT* (in preparation).

Making perturbative error estimates practical[1]

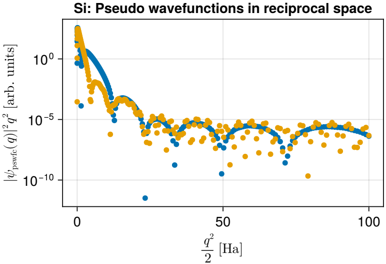
- Selection of E_{ref}?
  - **Answer:** See ideas on next slides
- This is a linearised correction: Is it valid across pseudopotentials & E_{cut}s?
  - **Answer:** Yes, see numerical benchmarks
- How to propagate: $\delta Q \approx Q\big(\mathfrak{r}(P + \delta P)\big)$ vs. $\delta Q \approx Q(P) + \frac{dQ}{dP}\delta P$
  - **Answer:** $Q\big(\mathfrak{r}(P + \delta P)\big)$ more accurate[1]
- Stopping criterion for inverting $(\Omega + K)_{11}$?
  - **Answer:** Use relative error around 10^{-4} in CG



[1] B. Ploumhans, MFH. *Practical plane-wave error estimation for Kohn-Sham DFT* (in preparation).

Choosing E_{ref}

- Requirement to E_{ref}:
  - Sufficiently large, so that E_{ref}-basis $\simeq$ exact result
  - Sufficiently small, so that $(\Omega + K)_{12}$ & $R(P)$ remain tractable
- **Challenge:** Fourier decay complex with numerical pseudopotentials



- **Solution:**
  - Use Fourier decay of $\delta\psi_{nk,2}$ to estimate $\|\delta P_2\|$ (against $E_{\text{ref}} = \infty$) by

$$\Delta(E_{\text{cut}}) \approx C \sqrt{\sum \quad \sum \quad |\widehat{\delta\psi}_{nk,G}|^2}$$

Choosing E_{ref}

- Requirement to E_{ref}:
  - Sufficiently large, so that E_{ref}-basis $\simeq$ exact result
  - Sufficiently small, so that $(\Omega + K)_{12}$ & $R(P)$ remain tractable
- **Challenge:** Fourier decay complex with numerical pseudopotentials
- **Solution:**
  - Use Fourier decay of $\delta\psi_{nk,2}$ to estimate $\|\delta P_2\|$ (against $E_{\text{ref}} = \infty$) by

$$\Delta(E_{\text{cut}}) \approx C \sqrt{\sum_{nk} \sum_{\|G\|^2 > 2E_{\text{cut}}} |\widehat{\delta\psi}_{nk,G}|^2}$$

with $\widehat{\delta\psi}_{nk,G}$ approximated by M_{22}^{-1} times the non-local pseudo residual, i.e.

$$\widehat{\delta\psi}_{nk,G} \approx -M_{nk,GG}^{-1} \sum_{j \in \text{proj}} \widehat{\beta}_{jG} \sum_{k \in \text{proj}} d_{jk} \left\langle \beta_k \middle| \psi_{nk,1} \right\rangle$$

  - Select error reduction factor η (between 5 and 20) and use a E_{ref} s.t.

$$\Delta(E_{\text{ref}}) \leq \frac{1}{\eta} \Delta(E_{\text{cut}})$$

  - Rationale: Non-local projectors usually decay slowest in Fourier

Benchmarking & validation

- Dataset: 21 common insulators
- Reference computed using cutoff of $150\,\text{Ha}$

- Discretisation error estimation parameters:
  - Choose E_{ref} using reduction factor $\eta = 10$
  - Propagate errors as $\delta Q \approx Q\big(\mathfrak{r}(P + \delta P)\big)$
  - Use relative error 10^{-4} when solving $(\Omega + K)_{11}$

- Four E_{cut} settings considered:

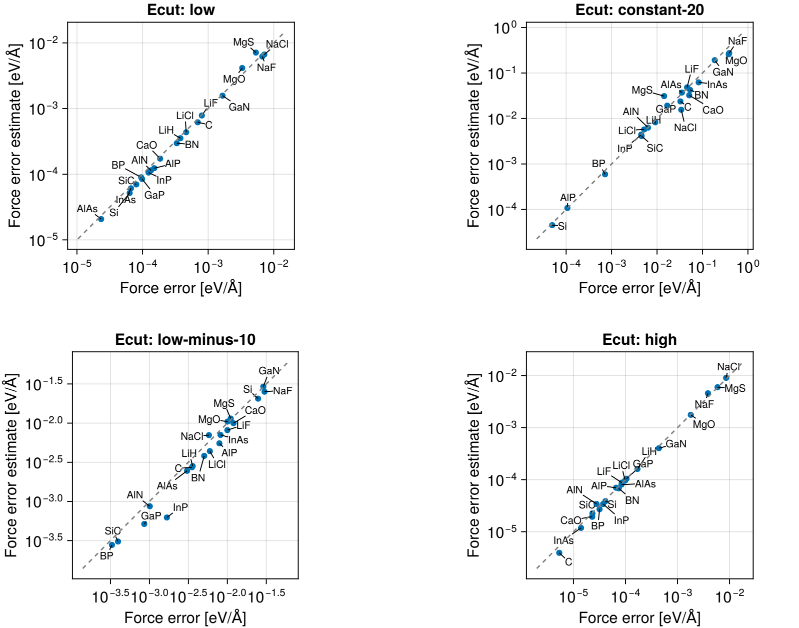
low: Low PseudoDojo recommendation

high: High PseudoDojo recommendation

low-minus-10: Low PseudoDojo recommendation minus $10\,\text{Ha}$

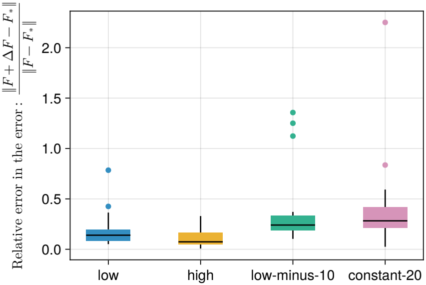
constant-20: Constant $E_{\text{cut}} = 20\,\text{Ha}$

Force error estimate versus “true” error



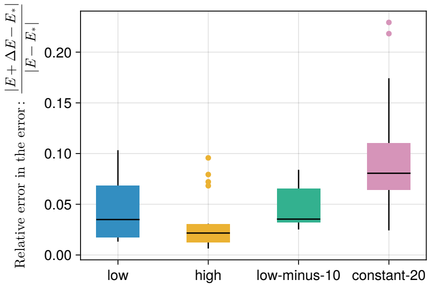
Error estimate accuracy quantiles

Relative accuracy of the force error estimate



⇒ Save outliers, correct order of magnitude

Relative accuracy of the energy error estimate

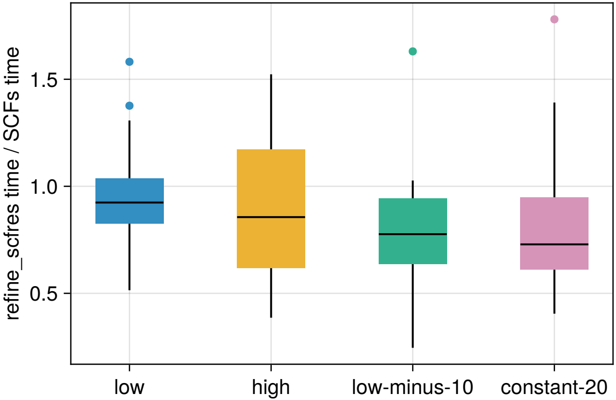


⇒ Save outliers, one quantitative digit

- (Absolute) accuracy of error estimate scales with accuracy of computation

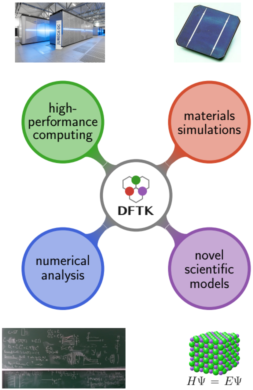
⇒ Error estimate gets more accurate as computed value does

Error estimate timing quantiles



- Cost of error estimation $\lesssim$ cost of SCF

Density-functional toolkit[1] — https://dftk.org

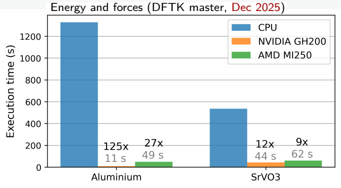


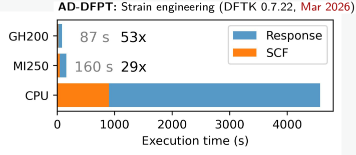
- julia code for cross-disciplinary research:
  - Allows restriction to relevant model problems,
  - *and* scale-up to application regime (1000 electrons)
  - Sizeable feature set in ca. 10k lines of code:[1]
  - Close the gap: Maths ↔ high-throughput:
    MARVEL AiiDA plugin
- Fully composable due to julia abstractions:
  - Algorithmic differentiation (AD)
  - HPC tools: MPI, Nvidia & AMD GPUs
- High-productivity framework & established community:
  - > 50 contributors in 6 years (Maths, physics, CS, ...)
  - Instrumental in a dozen of research works
- Unique features[1]:
  - Self-adapting algorithms
  - Algorithmic differentiation
  - Numerical error estimates (e.g. basis set error in forces)

[1] https://docs.dftk.org/features

DFTK performance

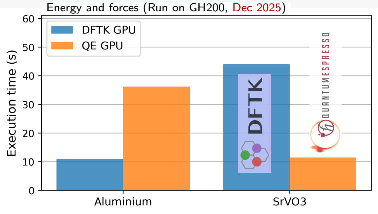
GPU speedups

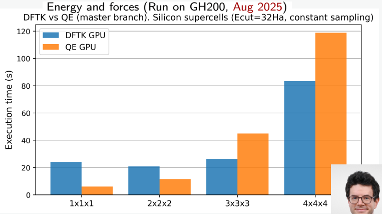




Comparison to QE

All graphs: Smaller is better



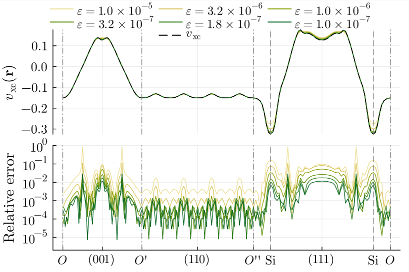


- Ported to Nvidia & AMD GPUs
  - **Performance without bloat**: **DFTK** has ~ 10k lines of code
  - Recently: AD-DFPT on GPU
- Large systems: **DFTK** beats QuantumESPRESSO (std. code)

Augustin Bussy (CSCS)

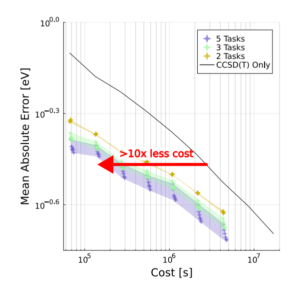
Other ongoing research . . .

Kohn-Sham inversion[1]



- Find exact v_{xc} from exact ρ
- Based on Moreau-Yosida regularised formulation of DFT[1]
- First error bounds & mathematical guarantees[2]

Multi-task learning[3]



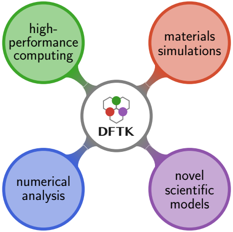
- **Training**: little CCSD(T) & *mixture* of 4 DFT models
- **Predict** at CCSD(T) level
- No accuracy order assumed !
- Towards dataset of opportunity

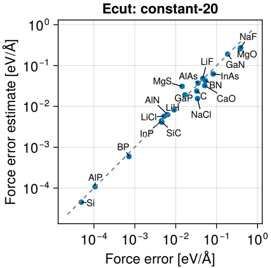
[1]MFH, V. Bakkestuen, A. Laestadius. Phys. Rev. B **111**, 205143 (2025).

[2]M. Penz, MFH, T. Helgaker, A. Laestadius. Electron. Struct. **8**, 022001 (2026).

[3]K. Fisher, MFH, Y. Marzouk. J. Chem. Phys. **161**, 014114 (2024).

Summary: Perturbation-based basis set error estimates





- **Quantitative error control:** Leverage to reduce cost
  - Data-driven materials simulation pipelines can tolerate inaccurate DFT simulations provided their error can be estimated
- **Routine discretisation error control** is feasible
  - Accurate order of magnitude error estimate
  - Efficient with cost $\lesssim$ SCF cost
  - Automatic: default parameters theoretically sound
  - Flexible framework: Extends to other quantities
  - Practical: Readily available in **DFTK**
- **Limitations and outlook**
  - Other quantities of interest: Stresses, geometries, . . .
  - Extend framework to metals
  - Integrate with training of machine-learned interatomic potentials

Questions?

MatMat https://matmat.org

mfherbst @herbst @social.epfl.ch

michael.herbst@epfl.ch

https://michael-herbst.com/talks/2026.04.20_DFT_Error_Oslo.pdf

N. Schmitz, B. Ploumhans, MFH. npj Computat. Mater. **12**, 6 (2025).

B. Ploumhans, MFH. *Practical plane-wave error estimation for Kohn-Sham DFT* (in preparation).

DFTK https://dftk.org