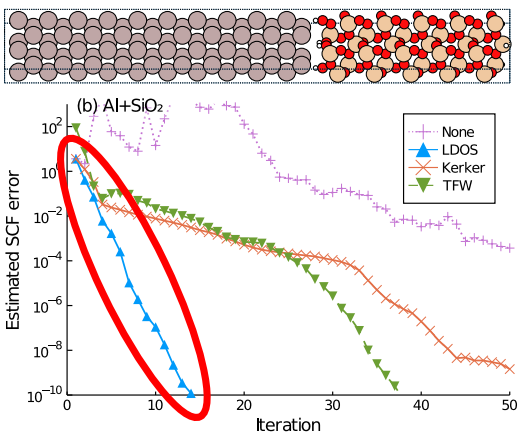
DFTK.jl: An introduction to a multidisciplinary electronic-structure code

Michael F. Herbst

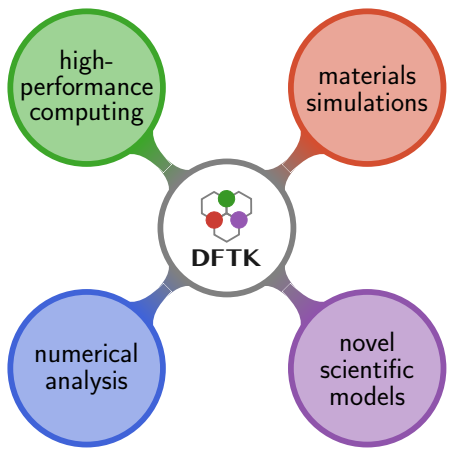
Mathematics for Materials Modelling (matmat.org), EPFL

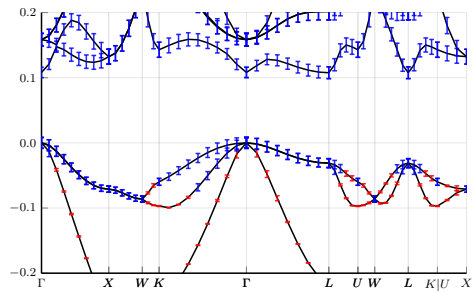
21 October 2024

https://michael-herbst.com/talks/2024.10.21_JuliaMolSim_DFTK.pdf



Novel materials simulation algorithms





Silicon band structure errors

Energy consumption of materials discovery

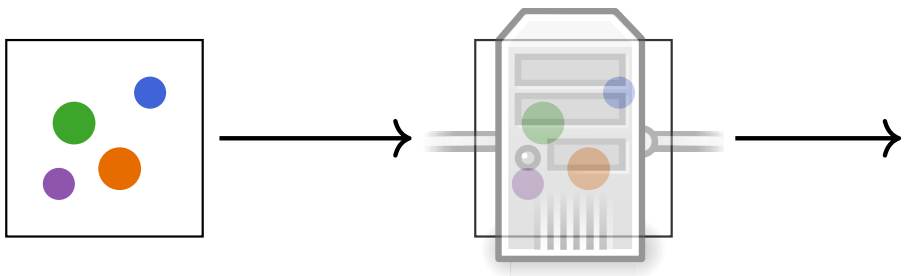
- Current solutions limited by properties of available materials
  - $\Rightarrow$ Innovation driven by discovering new materials
- Experimental research extremely energy intensive
  - 1 fume hood $\simeq$ 2-3 average households[1]

$\Rightarrow$ Complement experiment by computational materials discovery

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

High-throughput materials screening

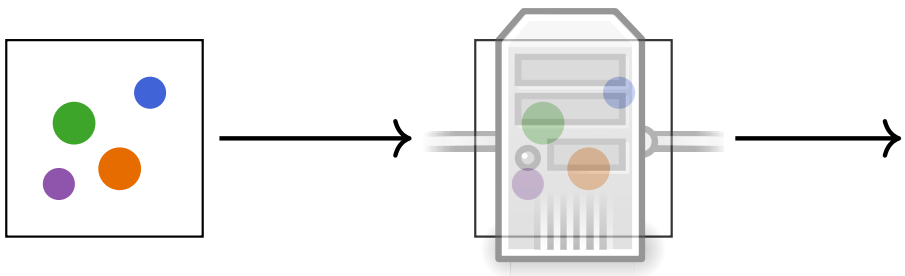
$$\min_{\Psi} \langle \Psi, H\Psi \rangle$$



- Energy consumption ?

High-throughput materials screening

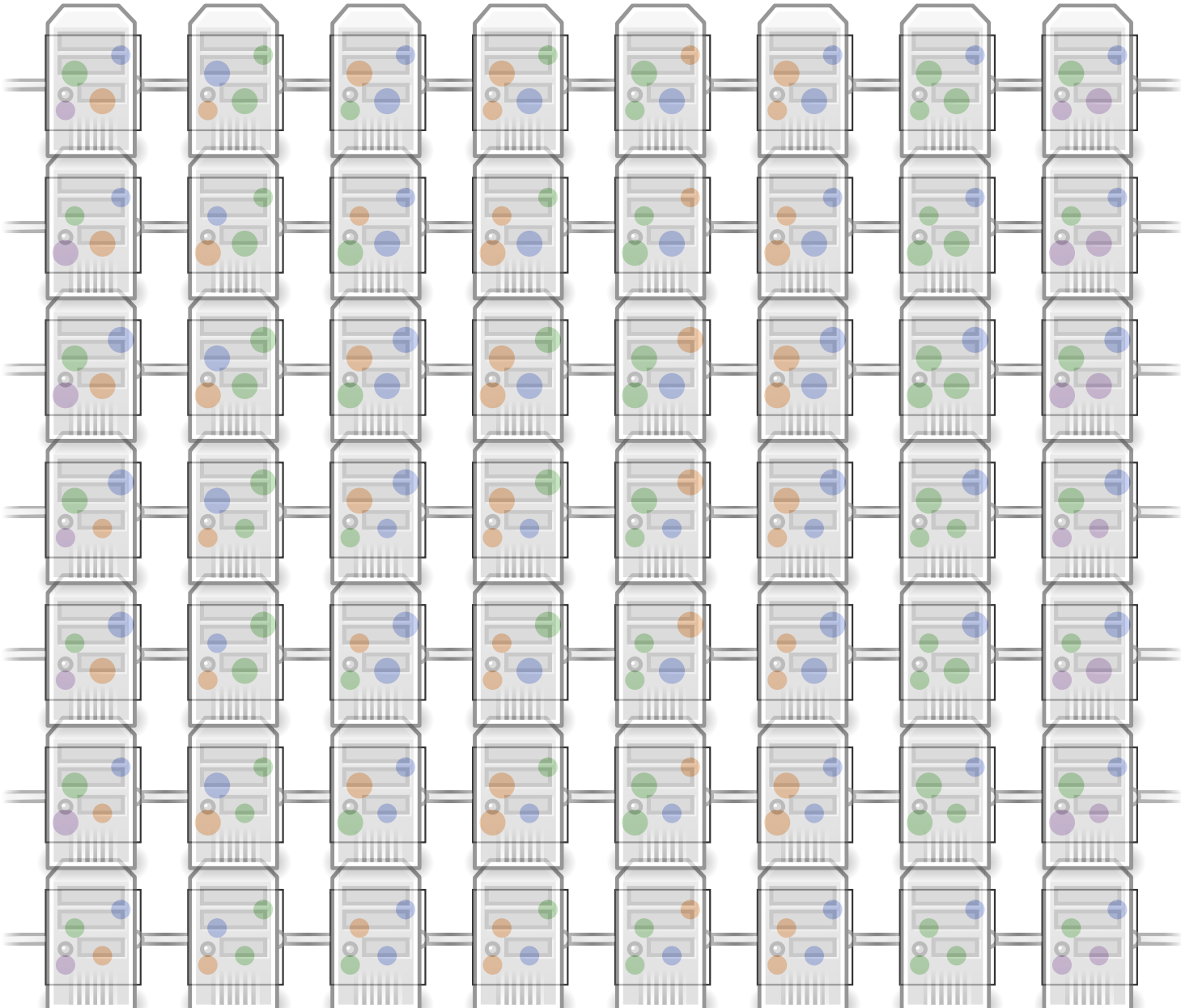
$$\min_{\Psi} \langle \Psi, H\Psi \rangle$$



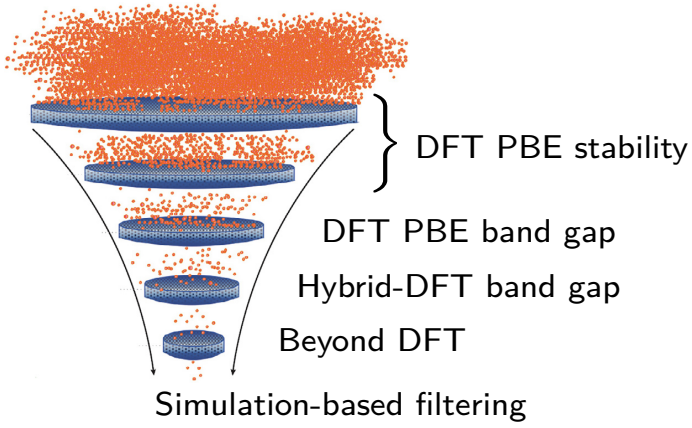
- Energy consumption ?
  - 8h of 36-core processor
    $\simeq$ 4h of average household
    $\simeq$ 1 CHF

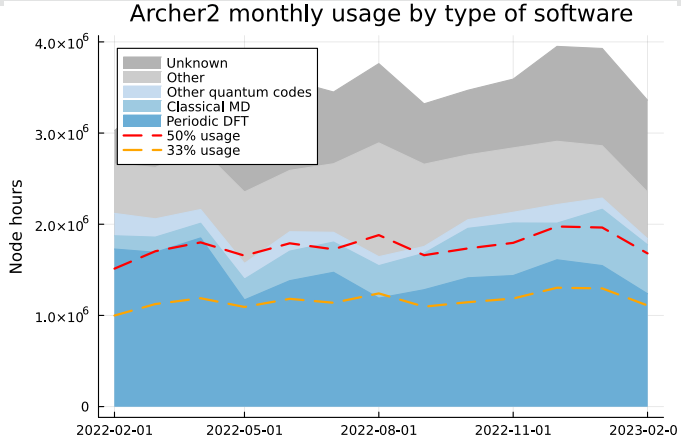
High-throughput materials screening

- We can **fully automate** this !



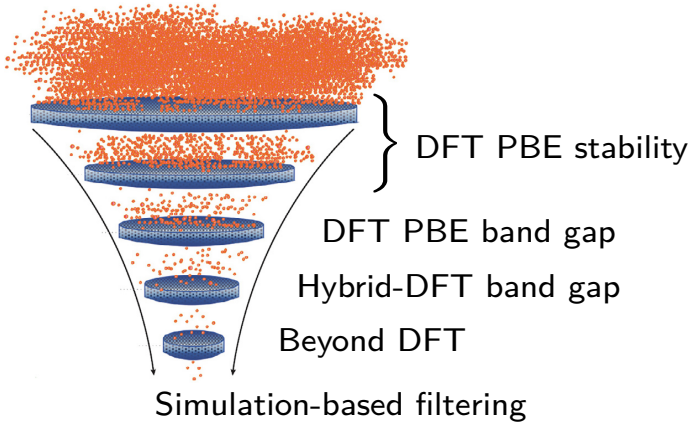
Computational materials discovery

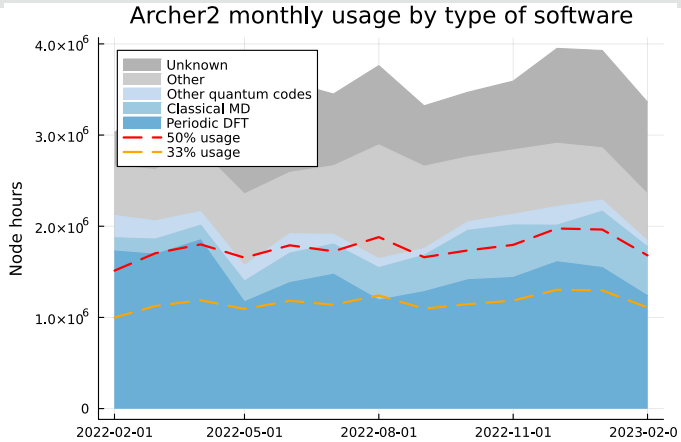




- Goal: Only promising candidates made in the lab
- Systematic simulations on $\simeq 10^4 - 10^6$ compounds
  - Noteworthy share of world's supercomputing resources

Computational materials discovery





- Goal: Only promising candidates made in the lab
- Systematic simulations on $\simeq 10^4 - 10^6$ compounds
  - Noteworthy share of world's supercomputing resources

- Energy consumption of LUMI (one of the most efficient):
  - 60 million kWh / year $\simeq$ 1.5 EPFLs $\simeq$ 14000 households

Challenges of high-throughput regime

- Complexity of multiscale materials modelling
  - Many parameters to choose (algorithms, tolerances, models)
  - Automated workflows & data management software (see above)
- Despite elaborate heuristics: Thousands of failed calculations
  - $\Rightarrow$ Wasted resources
  - $\Rightarrow$ Increased human attention (limits througput)
- Traversing the design space
  - How to best optimise material properties
  - How much accuracy is needed ?
  - How could we explore unusual gradients ?

A focus on robust materials simulations

- **Goal** in ~~MtMat~~ group:
  - Obtain reliable & efficient simulations
  - Develop and employ mathematically sound error indicators
  - Transform empirical wisdom to built-in convergence guarantees

$\Rightarrow$ Understand where and how to spend efforts best

- Practical error indicators:
  - Automatic & robust verification
  - Multi-fidelity statistical surrogates
  - Active learning of missing physics
- Leverage inexactness:
  - Error balancing: Optimal adaptive parameter selection
  - Adaptive tolerances & selective precision

$\Rightarrow$ Multidisciplinary expertise required

A focus on robust materials simulations

- **Goal** in ~~MtMat~~ group:
  - Obtain reliable & efficient simulations
  - Develop and employ mathematically sound error indicators
  - Transform empirical wisdom to built-in convergence guarantees

$\Rightarrow$ Understand where and how to spend efforts best

- Practical error indicators:
  - Automatic & robust verification
  - Multi-fidelity statistical surrogates
  - Active learning of missing physics

- Leverage inexactness:
  - Error balancing: Optimal adaptive parameter selection
  - Adaptive tolerances & selective precision

$\Rightarrow$ Multidisciplinary expertise required

Difficulties of cross-disciplinary research

(A computational science point of view . . .)

- Community conventions . . .
  - Language barriers, publication culture, speed of research, . . .
- . . . that are cemented in software:
  - Priorities differ $\Rightarrow$ What is considered "a good code" differs

Mathematical software

- **Goal:** Numerical experiments
- **Scope:** Reduced models
- High-level **language**: Matlab, python, . . .
- **Lifetime:** 1 paper
- **Size:** $<$ 1k lines
- Does not care about performance

Application software

- **Goal:** Modelling physics
- **Scope:** All relevant systems
- Mix of **languages:** C, FORTRAN, python, . . .
- **Lifetime:** 100 manyears
- **Size:** 100k – 1M lines
- Obliged to write performant code

Difficulties of cross-disciplinary research

(A computational science point of view . . .)

Mathematical software

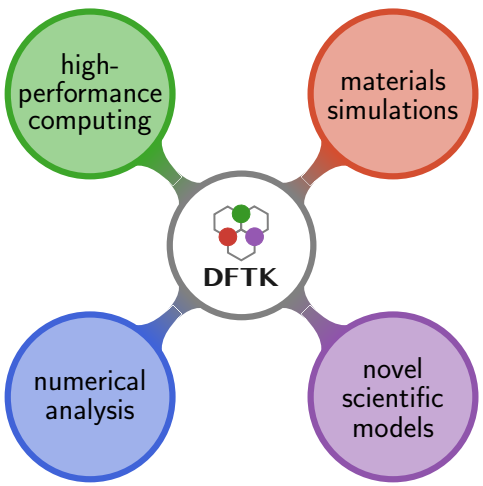
- **Goal:** Numerical experiments
- **Scope:** Reduced models
- High-level **language**:
  Matlab, python, . . .
- **Lifetime:** 1 paper
- **Size:** < 1k lines
- Does not care about performance

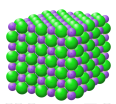
Application software

- **Goal:** Modelling physics
- **Scope:** All relevant systems
- Mix of **languages:**
  C, FORTRAN, python, . . .
- **Lifetime:** 100 manyears
- **Size:** 100k – 1M lines
- Obliged to write performant code

- Working with these codes requires different skillsets
  - $\Rightarrow$ Orthogonal developer & user communities
- Obstacle for knowledge transfer:
  - Mathematical methods never tried in practical setting
    (and may well not work well in the real world)
  - Some issues cannot be studied with mathematical codes
    (and mathematicians may never get to know of them)
- **Hypothesis:** People compose if software composes

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





$H\Psi = E\Psi$

- julia code for cross-disciplinary research:
  - Allows restriction to relevant model problems,
  - *and* scale-up to application regime (1000 electrons)
  - Sizeable feature set in 7500 lines of code
  - Norm-conserving pseudos, mGGA functionals, response
  - Integrated with high-throughput: MARVEL AiiDA
- Fully composable due to julia abstractions:
  - Arbitrary precision (32bit, >64bit, ...)
  - Algorithmic differentiation (AD)
  - HPC tools: GPU acceleration, MPI parallelisation
- Accessible high-productivity research framework:
  - Key contributions by undergrads (AD, GPU, Pseudos, ...)
  - Over 30 contributors in 5 years (Maths, physics, CS, ...)

[1] MFH, A. Levitt, E. Cancès. JuliaCon Proc. **3**, 69 (2021).

Contents

EPFL MATMAT

Density-functional theory (insulators)

- Goal: Understand electronic structures (Many-body quantum system)
- DFT approximation: Effective single-particle model

$$\begin{cases} \forall i \in 1 \dots N : \left(-\frac{1}{2}\Delta + V\left(\rho_\Phi\right) \right) \psi_i = \varepsilon_i \psi_i, \\ V(\rho) = V_{\text{ext}} + V_{\text{Hxc}}(\rho), \\ \rho_\Phi = \sum_{i=1}^{N} |\psi_i|^2 , \end{cases}$$

- Self-consistent field (SCF) fixed-point problem

$$\rho\big(V(\rho)\big) = \rho$$

- Density mixing (preconditioner P, damping α)

$$\rho_{n+1} = \rho_n + \alpha\, P^{-1} \left[\rho\big(V(\rho_n)\big) - \rho_n \right]$$

- Best P & α highly system dependent (metal, insulator, . . .)
  - Usually chosen by trial and error (Impact on energy consumption . . .)

Self-consistent field problem

- Density-mixing SCF procedure (preconditioner P, damping α)

$$\rho_{n+1} = \rho_n + \alpha\, P^{-1} \left[\rho\big(V(\rho_n)\big) - \rho_n\right]$$

- Near a fixed-point the error goes as

$$e_{n+1} \simeq \left[1 - \alpha P^{-1}\varepsilon^\dagger\right] e_n$$

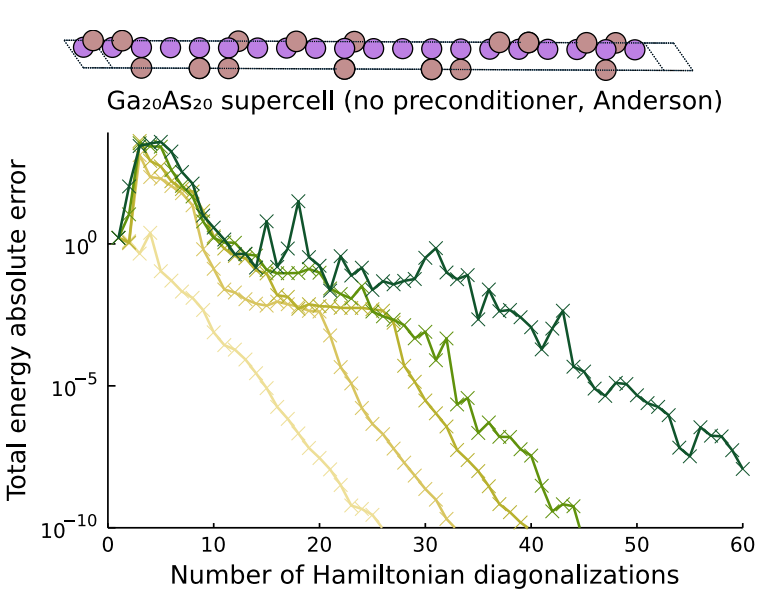
with dielectric matrix $\varepsilon^\dagger = (1 - \chi_0 K)$, $K(\rho) = V'(\rho)$, $\chi_0(V) = \rho'(V)$

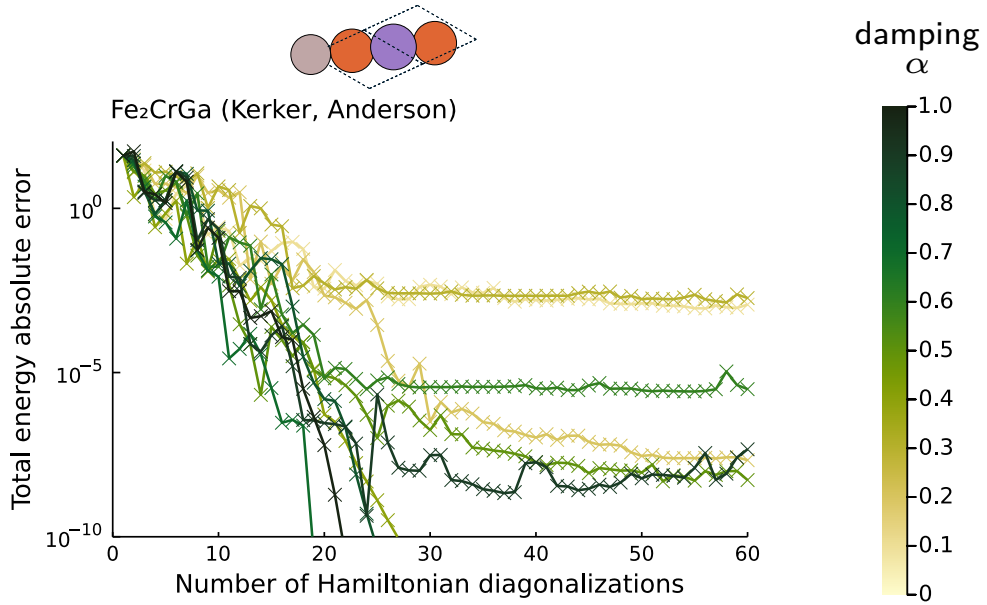
- Convergence iff $-1 < \left[1 - \alpha P^{-1}\varepsilon^\dagger\right] < 1$
  - Dielectric matrix $\varepsilon^\dagger$: Depends on physics (conduction, screening)
  - By second-order conditions: $\varepsilon^\dagger \geq 0$ (near fixed point)

$\Rightarrow$ Ideal preconditioner has $P^{-1}\varepsilon^\dagger \approx I$

  - Note: P needs to adapt to physics of unknown system!
  - No such P available: Choose α appropriately (Trial and error)

Illustration: Guessing a suitable damping α can be hard





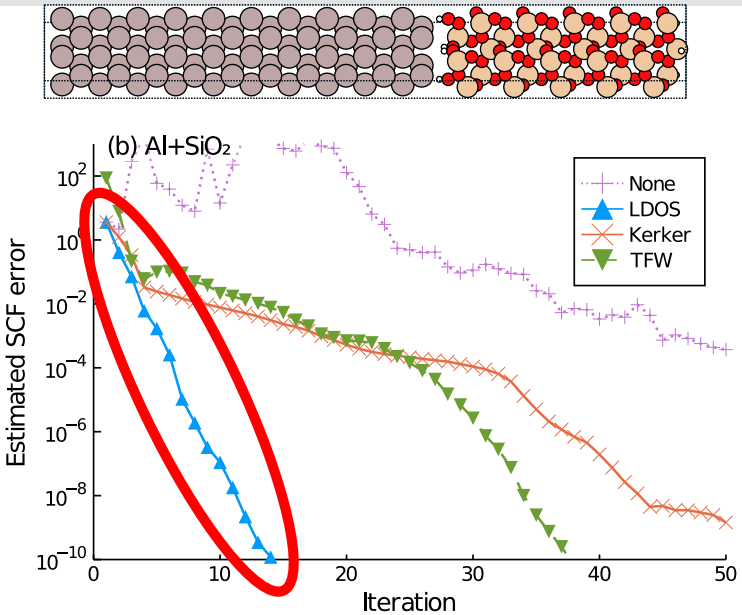
- Inefficient standard damping $(0.6 - 0.8)$
- Surprisingly small damping for smooth convergence

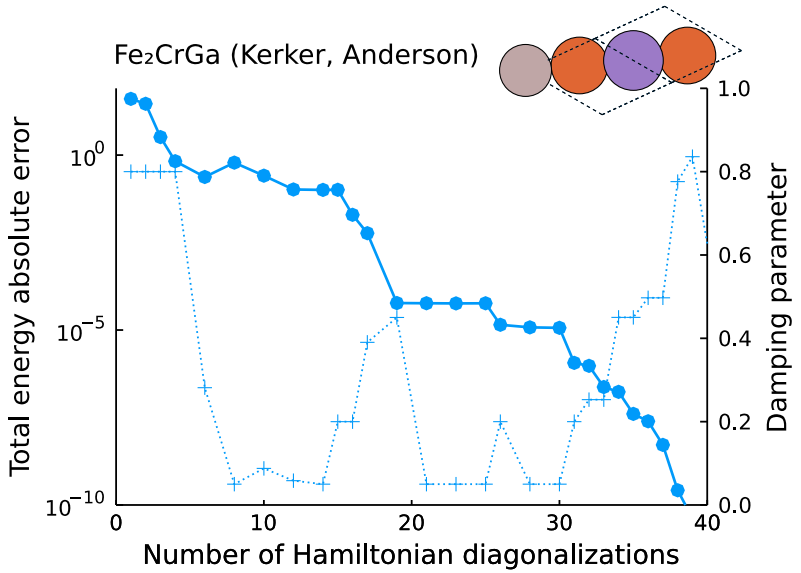
- Heusler alloy: Materials class with unusual magnetic properties

$\Rightarrow$ Numerically challenging behaviour

- SCF irregular: α versus convergence
- Usual heuristics breaks: Larger damping is better

Self-adapting black-box algorithms





- Preconditioning inhomogeneous systems (surfaces, clusters, . . .)
- LDOS preconditioner[1]: Parameter-free and self-adapting
- ca. 50% less iterations

- Damping α adapted *in each step* (using tailored quadratic model)
- Avoids trial and error (but may have a small overhead)
- Safeguard with theoretical guarantees[2]

⇒ Maths / physics collaboration:
Exchange of ideas between simplified & practical settings crucial

[1] MFH, A. Levitt. J. Phys. Condens. Matter **33**, 085503 (2021).

[2] MFH, A. Levitt. J. Comput. Phys. **459**, 111127 (2022).

Response, properties and algorithmic differentiation

- DFT properties: Response of system to external changes:
  - Connection Theory $\Leftrightarrow$ Experiment
  - Modelling: Potential $V(\theta, \rho)$ depends on parameters θ
    (e.g. atomic positions, el. field)

- SCF procedure yields fixed-point density ρ_*

$$0 = \rho\Big(V(\theta, \rho_*)\Big) - \rho_*$$

$\Rightarrow$ Defines implicit function $\rho_*(\theta)$

- Properties are derivatives:
  - **Forces** (energy wrt. position), **dipole moment** (energy wrt. el. field), **elasticity** (energy cross-response to lattice deformation), phonons, electronic **spectra**, ...

$\Rightarrow$ Great application for algorithmic differentiation !

- Byproduct: Arbitrary derivatives
  - Sensitivities, improved training of surrogates ...

AD for stresses keeps code accessible

$$\text{Stress} = \frac{1}{\det(\mathbf{L})} \frac{\partial E\left[P_*, (I + \mathbf{M})\,\mathbf{L}\right]}{\partial \mathbf{M}}\bigg|_{\mathbf{M}=0}$$

```
# Run SCF, get P*
scfres = self_consistent_field(basis)
     L = basis.model.lattice
stress = 1/det(L) * gradient(
    M -> recompute_energy(
                 scfres, (I + M) * L),
    zero(L)
)
```

- Stress computation (Definition *vs.* julia code)[1]
- Post-processing step ⇒ Not performance critical
- Comparison of implementation complexity:
  - **DFTK**: 20 lines[1] (forward-mode algorithmic differentiation)
  - Quantum-Espresso: 1700 lines[2]
  - Initial version: ≃ 10-week GSoC project

[1] https://github.com/JuliaMolSim/DFTK.jl/blob/master/src/postprocess/stresses.jl

[2] https://github.com/QEF/q-e/blob/develop/PW/src

Arbitrary derivatives: Need efficient response

- Full DFT equivalent is density-functional perturbation theory

$$\frac{\partial \rho_*}{\partial \theta} = \left[1 - \chi_0 K\right]^{-1} \chi_0 \frac{\partial V}{\partial \theta} \tag{2}$$

- Challenge: Need *many* applications of χ_0:
  - Each requires solving N Sternheimer equations

$$\left(\tilde{H} - \varepsilon_i\right) \delta\psi_i = -P\, \delta V \psi_i \qquad \forall i = 1, \ldots, N$$

$H = -\frac{1}{2}\Delta + V$, $\tilde{H} = PHP$ and P some projector (ε_i, ψ_i) eigenpairs of H

$\Rightarrow$ Nested iterative problem . . . which can be ill-conditioned

Sternheimer equations

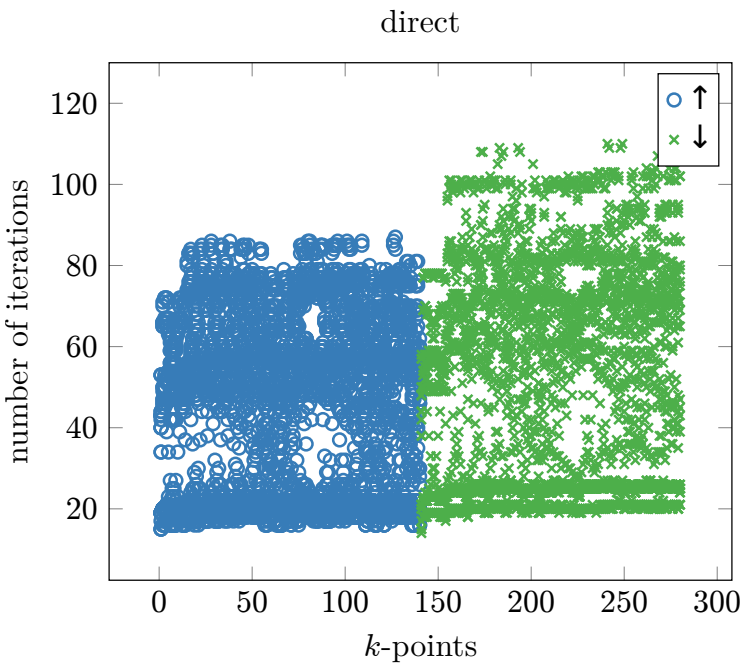
- Product $\chi_0 \delta V$ requires solving Sternheimer equations

$$\left(\tilde{H} - \varepsilon_i\right) \delta\psi_i = -P\,\delta V \psi_i \qquad \forall i = 1, \ldots, N$$

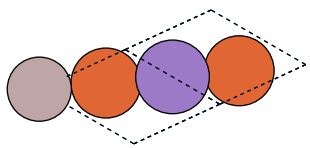
$H = -\frac{1}{2}\Delta + V$, $\tilde{H} = PHP$ and P some projector

(ε_i, ψ_i) eigenpairs of H

$\Rightarrow$ Badly conditioned for metallic systems (ε_i near eigenvalue of $\tilde{H}$)



(Number of iterations for various i)



Fe_2MnAl Heusler alloy

Schur complement approach to response[1]

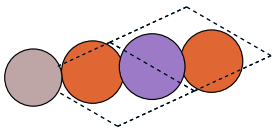
- Numerics of eigensolver:
  We have N_{ex} "extra" bands
- Use these to partition $\tilde{H}$:

$$\tilde{H} = \begin{pmatrix} E_{\text{ex}} & \mathbf{C} \\ \mathbf{C}^\dagger & \mathbf{R} \end{pmatrix}$$

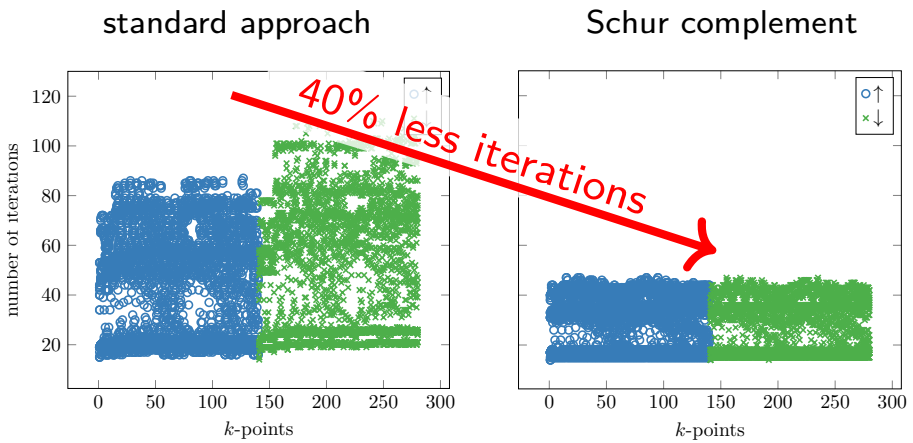
$E_{\text{ex}} = \text{diag}(\varepsilon_{N+1}, \ldots, \varepsilon_{N+N_{\text{ex}}})$
& $\mathbf{C}$, $\mathbf{R}$ projections of $\tilde{H}$

$\Rightarrow$ Use Schur complement:
Better-conditioned systems

$$(\mathbf{R} - \varepsilon_i)x = b$$



Fe_2MnAl Heusler alloy

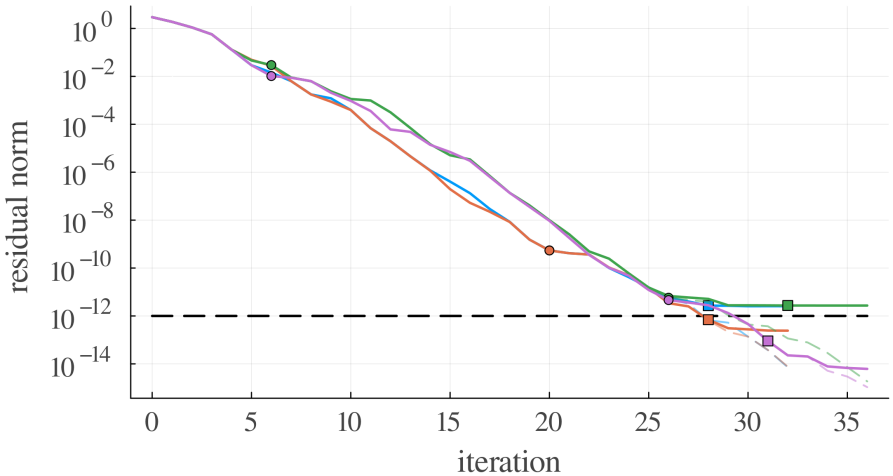


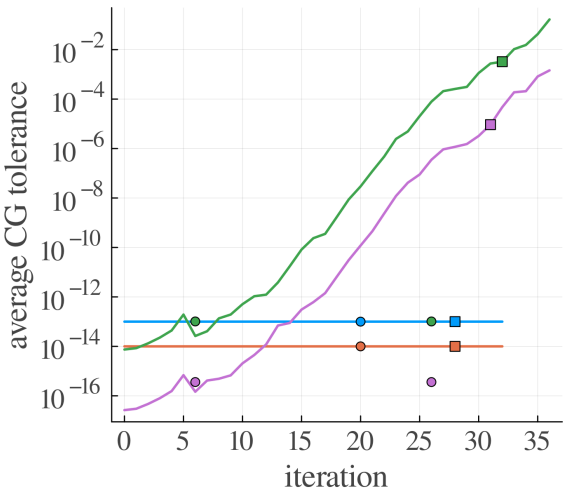
- Schur-based approach tames CG
- ca. 40% less iterations
- Development guided using a "real material"

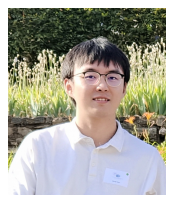
[1] E. Cancès, MFH, G. Kemlin, *et. al.* Lett. Math. Phys. **113**, 21 (2023).

WIP: Inexact Krylov methods

- DFPT + Sternheimer: Nested linear problems
- Inexact Krylov methods:[1] Framework to tolerate *less tight* solutions of Sternheimer
- First results indicate 25%–50% less Hamiltonian applications (the expensive step)



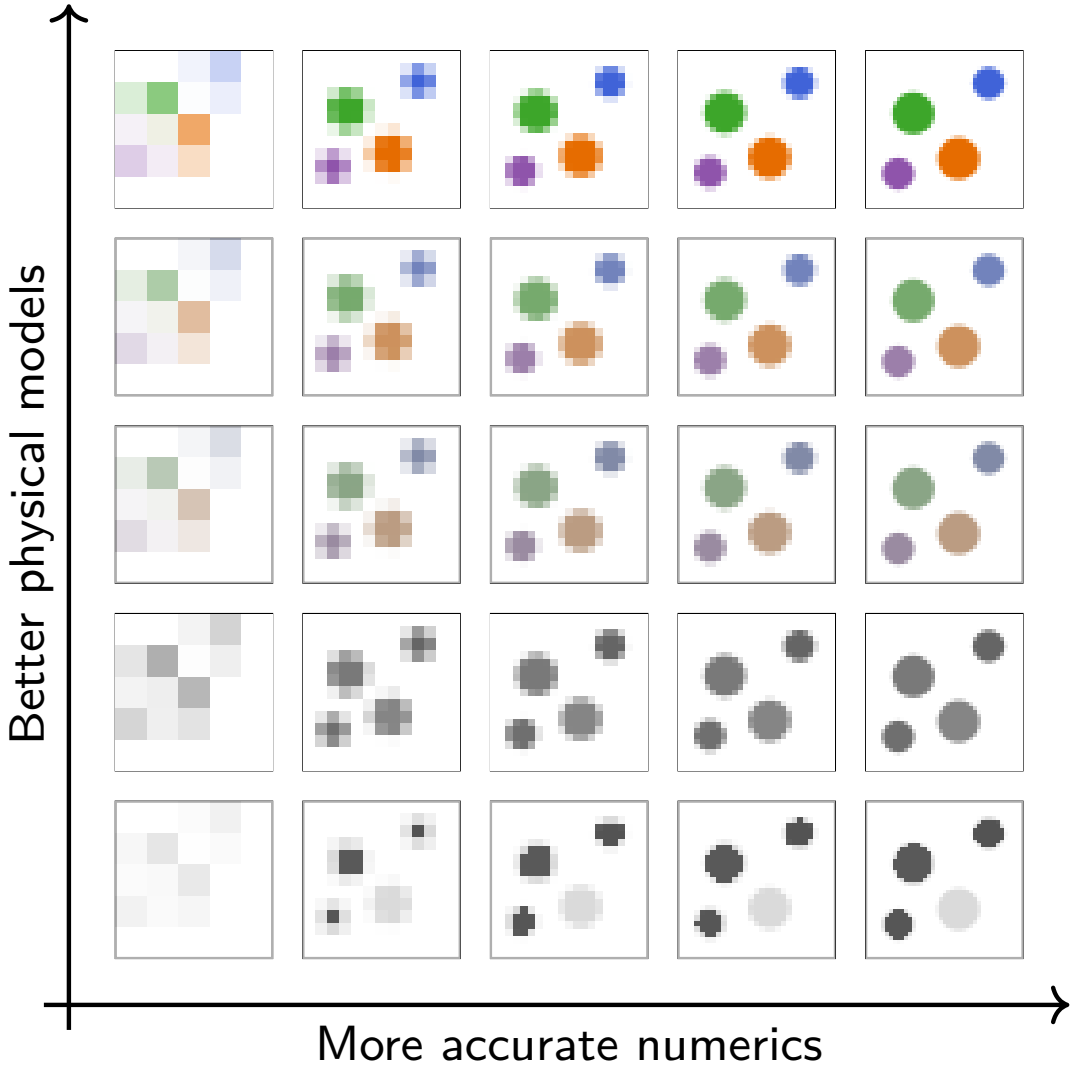




Bonan Sun

[1] V. Simonicini, D. Szyld. SIAM J. Sci. Comput., **25**, 454 (2003).

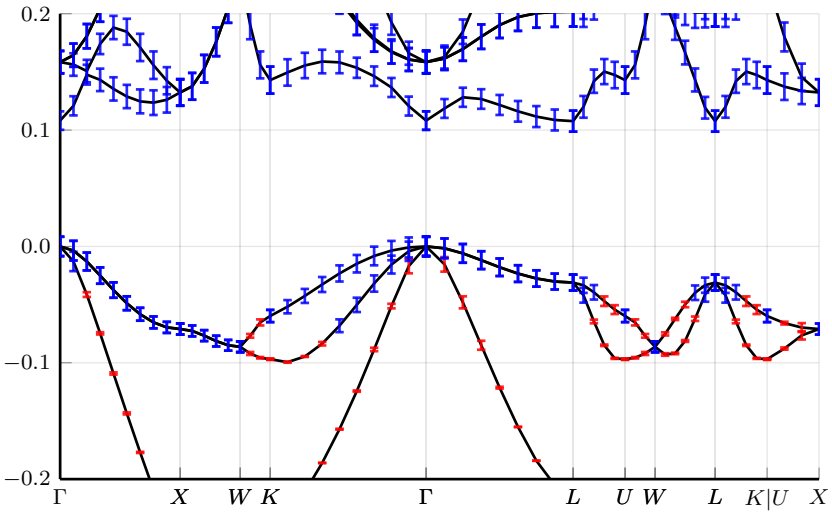
Case for error control: Error comes in different flavours



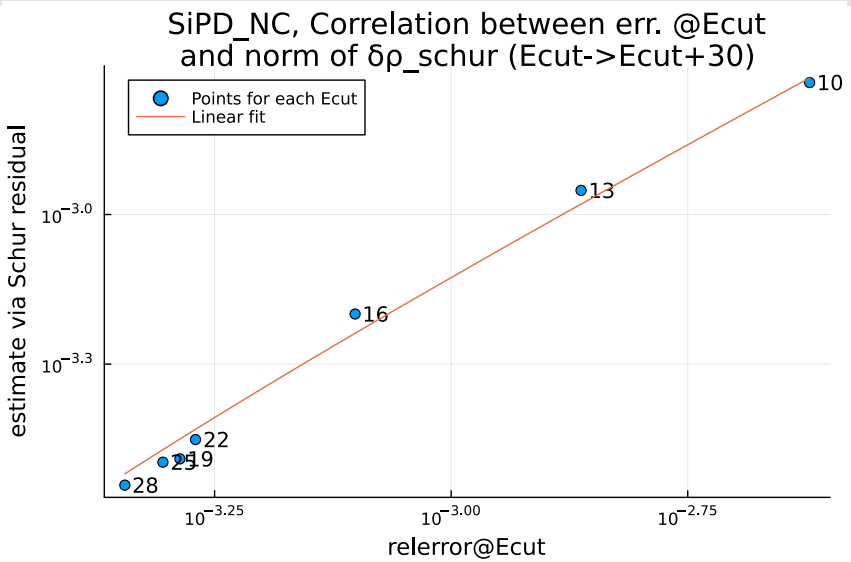
- Ideally want to balance errors

$\Rightarrow$ Need reliable error indicators !

Numerical error: Analytical techniques



Band structure with guaranteed errors[1]



Estimation of basis set error in ρ^2

- Momentum towards numerical error estimators for DFT
  - Focus on basis set error (some also tackle floating-point, SCF convergence)
- Results promising, but many challenges & caveats remain
  - Numerical experiments & problem simplifications crucial
  - ⇒ **DFTK** is major research tool for this development[1-4]
- Techniques for DFT error less developed (and hard to tackle analytically)

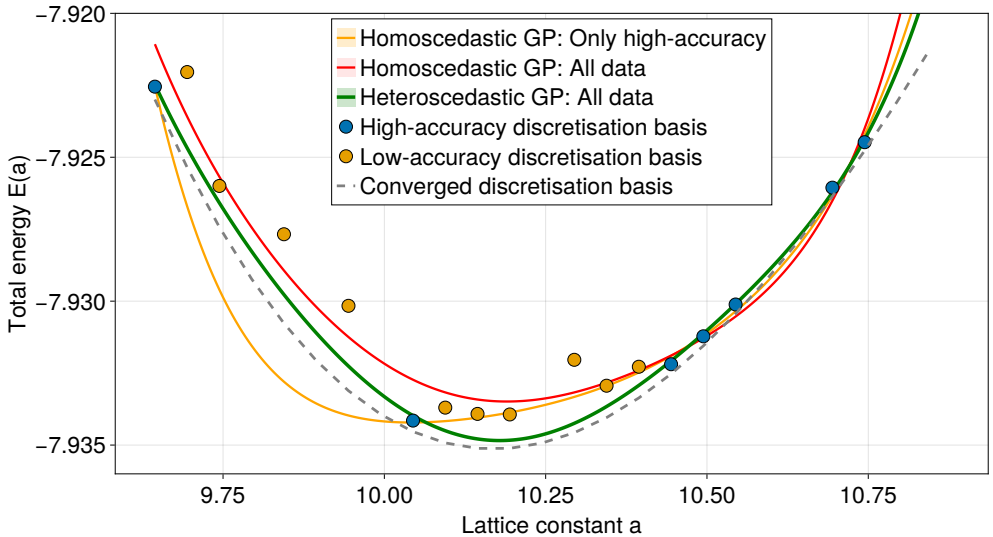
[1] MFH, A. Levitt, E. Cancès. Faraday Discus. **223**, 227 (2020).

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

[3] E. Cancès, G. Kemlin, A. Levitt. J. Matrix Anal. Appl., **42**, 243 (2021).

[4] E. Cancès, G. Kemlin, A. Levitt. J. Sci. Comput., **98**, 25 (2024)

WIP: Heteroscedastic regression models



- 1D proof of principle: energy-volume curve (Equation of state)
- High-dimensional regression problems: Data is scarce
- Error δ_i can be estimated $\Rightarrow$ supply to GP
- For example: Heteroscedastic model:

$$E_i = \mathrm{DFT}(a_i) + \varepsilon_i \qquad \varepsilon_i \sim \mathcal{N}(0, \delta_i)$$

Anna Paulish

DFT error: Computing model sensitivities

- Consider model sensitivity of force $\mathcal{F}(\rho_*(\theta))$:

$$\frac{d\mathcal{F}}{d\theta} = \frac{\partial \mathcal{F}}{\partial \rho_{\text{SCF}}} \frac{\partial \rho_*}{\partial \theta} \tag{1}$$

- Computed by response theory (we've seen this before !):

$$\frac{\partial \rho_*}{\partial \theta} = [1 - \chi_0 K]^{-1} \chi_0 \frac{\partial V}{\partial \theta}$$

- Parameters appear in innermost layer (model definition)
  - Each DFT model: Different derivatives $\frac{\partial V}{\partial \theta}$ (can be horrible)
  - Each quantity of interest: Different sensitivity expression (1)

  $\Rightarrow$ Combinatorial explosion

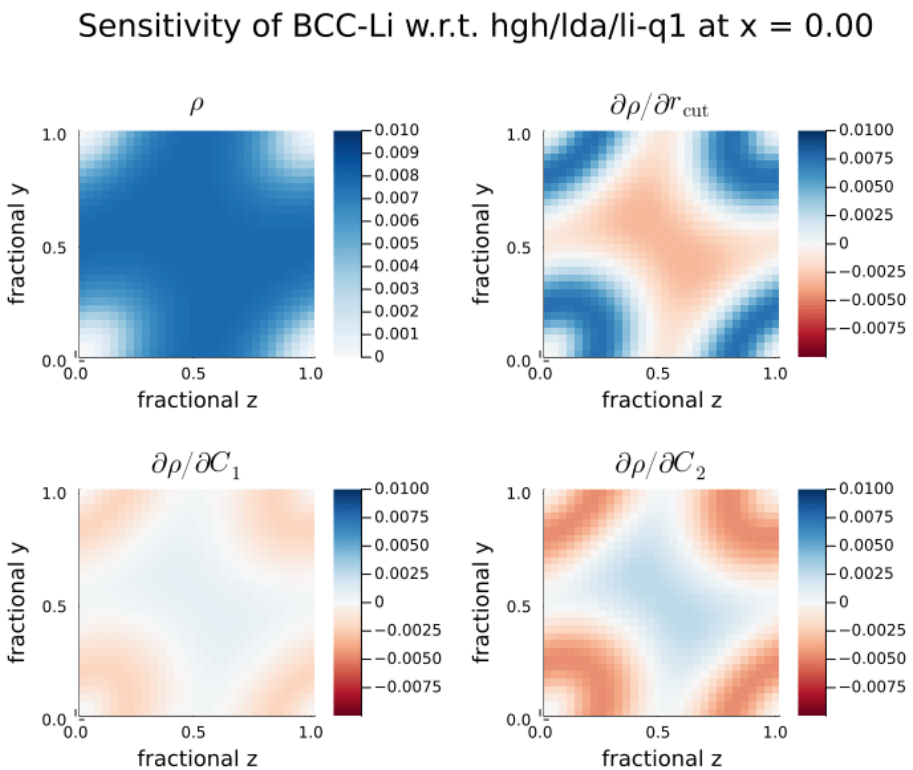
WIP: Sensitivity analysis in one line of code

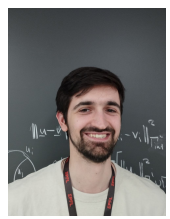
- **DFTK**: Algorithmic differentiation (AD)
  - Generic framework for derivatives: Request gradient, AD delivers

  ⇒ New properties/derivatives by non-DFT experts!

⇒ Setting for uncertainty quantification:

- Pseudopotential sensitivity of electronic density





Niklas Schmitz

High-level structure of density-functional theory

- Energy minimisation problem (discretised setting):
$$\min_{D\in\mathcal{P}} E(D) = \min_{D\in\mathcal{P}} \left[\mathrm{tr}(H_0 D) + E_{\mathsf{Hxc}}(D)\right]$$
  - Non-linear, non-convex Riemannian optimisation ($\mathcal{P}$: Grassmanian)

- What we care about: Illustration on model problem $x_* = \min_{x\in\mathbb{R}^N} E(x)$

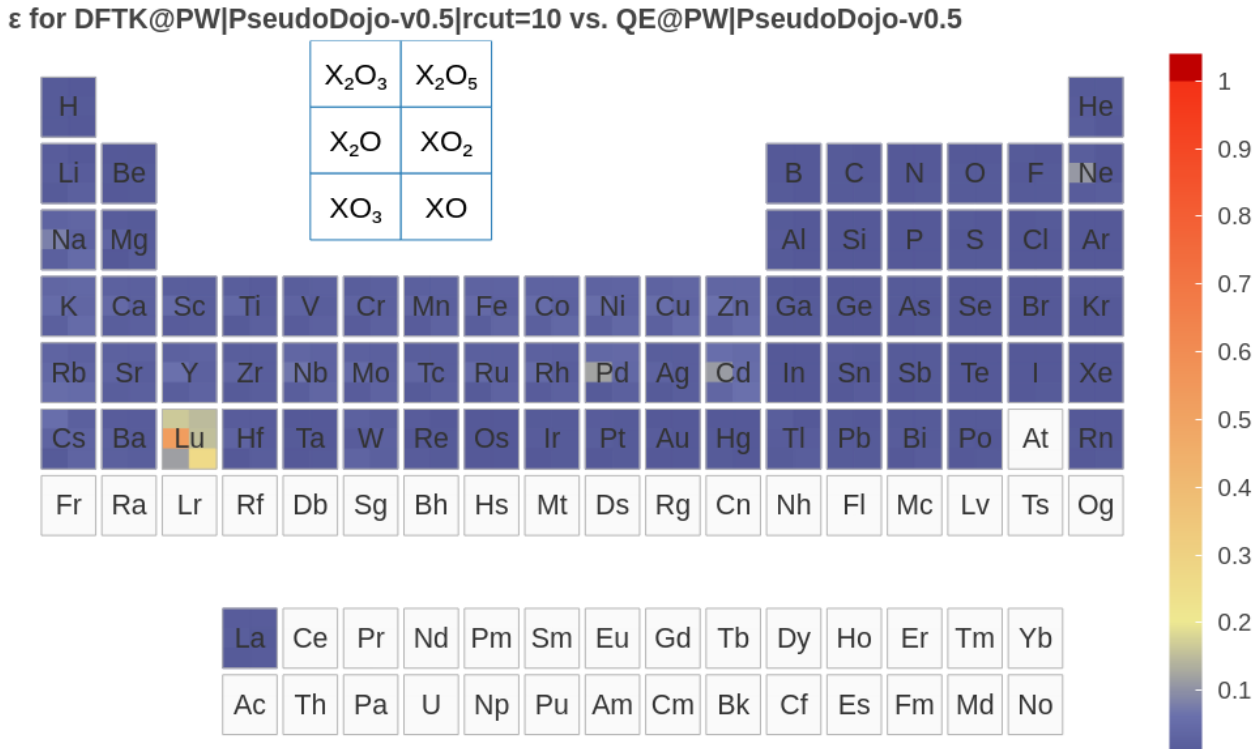
- Numerical methods: $x_{k+1} = x_k - \alpha\nabla E(x_k)$ (SCF, direct min.)
  - Convergence depends on $1 - \alpha\nabla^2 E(x_*)$
  - Need to understand $\nabla^2 E(x_*)$ for preconditioning

- Response, properties and algorithmic differentiation
  - Solution to $\min_x E(x,\theta)$ satisfies
    $x(\theta) \approx x_* - \theta\,\nabla^2 E(x_*,0)^{-1}\,\frac{\partial}{\partial\theta}\nabla E(x_*,0)$
  - Changing θ: This is how experiments explore physics
  - Sensitivities & model uncertainties

- A posteriori error: $x - x_* \approx -\nabla^2 E(x_*)^{-1}\,\nabla E(x)$
  - Estimate accuracy of simulations

WIP: Integration into high-throughput frameworks

- Algorithms are only useful if they work in practice !
- DFTK plugin for AiiDA workflow manager
- **Goal:** Automated testing of algorithms and error estimates
- Verification study Quantum-Espresso vs. DFTK
  - ⇒ Results agree, algorithms can outperform QE



Bruno Ploumhans

Contents

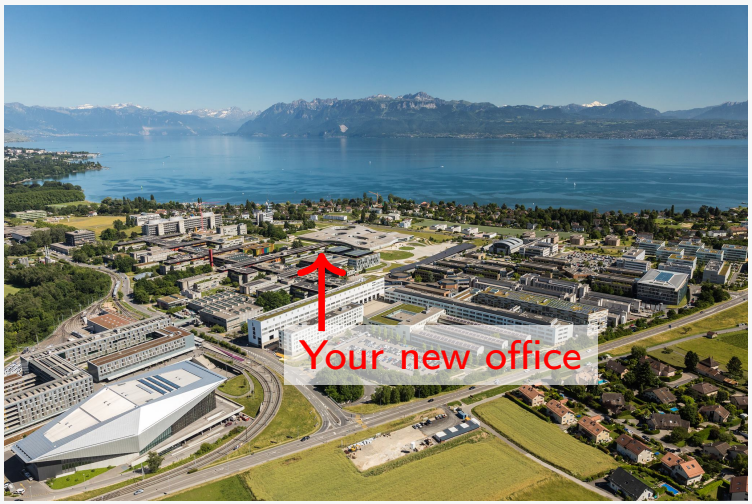
EPFL MX t Mat

DEMO

DFTK interface and ecosystem integration

Advertisement break

Open PostDoc in the EPFL MatMat group



Topic: Efficient inverse materials design

- Bayesian optimisation
- AD & gradient approaches
- Interdisciplinary environment of MARVEL:
  Reproducible workflows, sustainable software,
  computational materials discovery, statistical learning
- See `https://matmat.org/jobs/`

Psi-k workshop *(M. F. Herbst, A. Levitt, J. Haegeman)*:

"Julia for numerical problems in quantum and solid-state physics"

- **26–28 November 2024** at EPFL, CECAM-HQ, Lausanne
- Targets: Linear algebra, physics and julia communities

⇒ `https://www.cecam.org/workshop-details/1355` (Deadline: 20th Sep)

Summary and outlook

- Current state of DFTK:
  - Unique robust material-adapting DFT algorithms
  - `ForwardDiff` to setup & solve response problems
  - Reduced settings (error analysis) and high-throughput testing
- Future work:
  - Explore error control & sensitivity (inverse design, surrogates)
  - Employ as frontend for domain-specific libraries (SIRIUS)
  - Composability with JuliaMolSim (structure opt., surrogates, . . .)
  - Bring methods to AiiDA (for adoption and testing !)
- Where you can help:
  - Improve GPU performance (Hackaton anyone ?)
  - Parallelisation & performance bottle necks in AD / response
  - Explore alternative AD backends (Enzyme)
  - Use DFTK & JuliaMolSim, report bugs, enhance docs

Acknowledgements

MatMat group

- Bruno Ploumhans (MatMat)
- Anna Paulish (MatMat)
- Niklas Schmitz (MatMat)
- Cédric Travelletti (MatMat)

Robust algorithms

- Eric Cancès (École des Ponts)
- Gaspard Kemlin (Université de Picardie)
- Antoine Levitt (Université Paris-Saclay)
- Benjamin Stamm (Stuttgart)
- Bonan Sun (EPFL)

Aiida interface & verification

- Giovanni Pizzi (PSI)
- Junfeng Qiao (EPFL)
- Yihan Wu (EPFL)
- Austin Zadoks (EPFL)

- All > 40 DFTK contributors

Fonds national suisse

MARVEL
NATIONAL CENTRE OF COMPETENCE IN RESEARCH

EPFL MatMat

Questions?

https://matmat.org

mfherbst

michael.herbst@epfl.ch

https://michael-herbst.com/talks/2024.10.21_JuliaMolSim_DFTK.pdf

DFTK https://dftk.org