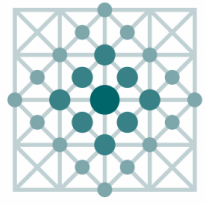
A novel black-box preconditioning strategy for high-throughput density-functional theory

Michael F. Herbst* and Antoine Levitt

*Applied and Computational Mathematics, RWTH Aachen University
https://michael-herbst.com

17th March 2021



Applied and Computational Mathematics

RWTH AACHEN UNIVERSITY

→ https://michael-herbst.com/talks/2021.03.17_gamm_scf_preconditioning.pdf

Societal challenges of 21st century

- Renewable energy
- Green chemistry and catalysts
- Drug design
- Transportation
- Data storage and communication

$\Rightarrow$ Need for novel materials

$\Rightarrow$ High-throughput computational screening

$\Rightarrow$ Common approach: Density-functional theory (DFT)

Typical scale

- One DFT calculation: $\mathcal{O}(\text{hours})$ to $\mathcal{O}(\text{days})$
- E.g. Open Catalyst Project[1]
  - 1.3 million DFT calculations
  - > 250 million DFT energy evaluations
  - Workflow success rate: $\simeq 50\%$[2]

$\Rightarrow$ Need high degree of automation

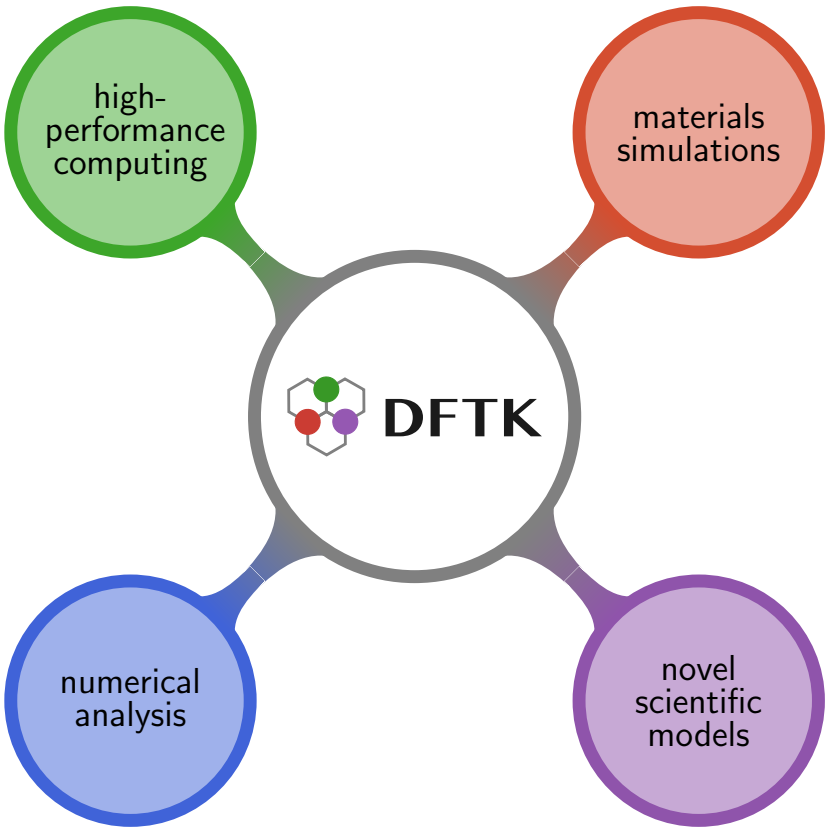
$\Rightarrow$ Reliability needs to be improved!

- Multidisciplinary research problem

[1] L. Chanussot *et. al.* The Open Catalyst 2020 (OC20) Dataset, 2020, arXiv 2010.09990.

[2] Z. Ulissi, private communication in ARPAE differentiate group seminar, Dec 2020.

Density-functional toolkit (DFTK)



- `https://dftk.org`
- 2 years of development
- Pure julia code

- Supports mathematical developments *and* scale-up to regime relevant to applications
- Low entrance barrier: Only 6k lines of code!

Current research with DFTK

⇒ Vision: Improve high-throughput workflows:

- Use physics: Reliable black-box SCF algorithms (this work)[1]
- Use maths: Error estimates and automatic error balancing[2]
- Better algorithms: Numerical analysis of SCF methods[3]

⇒ Reliable automatic selection of parameters in DFT workflows

- International and interdisciplinary user base:
  - RWTH Aachen, ENPC Paris, MIT, CMU, ...
- Involved in multidisciplinary research projects:
  - ACED-DIFFERENTIATE, EMC2 erc synergy, CESMIX

[1] M. F. Herbst, A. Levitt. J. Phys. Condens. Matter **33**, 085503 (2021).
[2] M. F. Herbst, A. Levitt and E. Cancès. Faraday Discuss. **224**, 227 (2020).
[3] E. Cancès, G. Kemlin, A. Levitt. arXiv 2004.09088 (2020).

Density-functional theory

- Coupled set of non-linear elliptic partial differential equations:

$$\Big(-\frac{1}{2}\Delta + \mathcal{V}(\rho)\Big)\psi_i = \varepsilon_i \psi_i, \quad \int \psi_i^* \psi_j = \delta_{ij}$$

$$\rho = \sum_{i=1}^{\infty} f\left(\frac{\varepsilon_i - \varepsilon_F}{T}\right) |\psi_i|^2, \quad \text{with } \varepsilon_F \text{ such that } \int \rho = N$$

- Density-dependent potential

$$\mathcal{V}(\rho) = V_{\text{Nuc}} + \int (v_C \rho) + V_{\text{xc}}(\rho)$$

with Coulomb kernel $v_C(\underline{\boldsymbol{r}}, \underline{\boldsymbol{r}}') = \|\underline{\boldsymbol{r}} - \underline{\boldsymbol{r}}'\|^{-1}$

- Fermi-Dirac function $f(x) = 1/(1+e^x)$
- Temperature T, electron count N
- Exchange-correlation potential V_{xc}, nuclear attraction V_{Nuc}

Self-consistent field (SCF) approach

- Here: Periodic boundary conditions, plane-wave basis

- Density-dependent potential

$$\mathcal{V}(\rho) = V_{\text{Nuc}} + \int (v_C \rho) + V_{\text{xc}}(\rho)$$

- Potential-to-density map F

$$F(V) = \sum_{i=1}^{\infty} f\left(\frac{\varepsilon_i - \varepsilon_F}{T}\right) |\psi_i|^2$$

  with (ε_i, ψ_i) eigenpairs of $-\frac{1}{2}\Delta + V$.

$\Rightarrow$ SCF solves $\rho = F(\mathcal{V}(\rho))$

- Key computational step of density-functional theory (DFT)

The SCF Jacobian

- SCF solves $\rho = F(\mathcal{V}(\rho))$
- Consider damped fixed-point scheme:

$$\rho_{n+1} = \rho_n + \alpha \left[F(\mathcal{V}(\rho_n)) - \rho_n\right]$$

- Near a fixed-point the error goes as

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

  where $\epsilon^\dagger = 1 - \chi_0(v_C + f_{\text{xc}})$

- χ_0: Independent-particle susceptibility (derivative of F)
- f_{xc}: XC kernel (derivative of V_{xc}, usually small)

$\Rightarrow$ Jacobian $J_\alpha = 1 - \alpha\epsilon^\dagger$ contains dielectric matrix ϵ

$\Rightarrow$ SCF convergence linked to dielectric properties

How bad can it get?

- In Fourier space: $\widehat{(v_C \rho)}(\underline{\boldsymbol{q}}) = \frac{4\pi \hat{\rho}(\underline{\boldsymbol{q}})}{|q|^2}$
- Smallest $q \sim 1/L$ where L is crystal length

$\Rightarrow \lambda_{\text{max}}(v_C) \sim L^2$

- Can imply $\lambda_{\text{max}}(\epsilon^\dagger) \sim L^2$ (e.g. in metals, charge sloshing)

$\Rightarrow$ SCF condition number grows as L^2

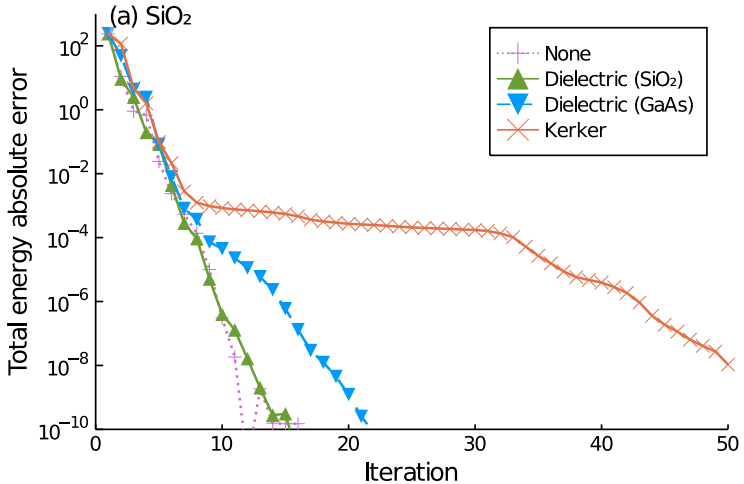
$\Rightarrow$ Infeasibly small damping α needed for large systems

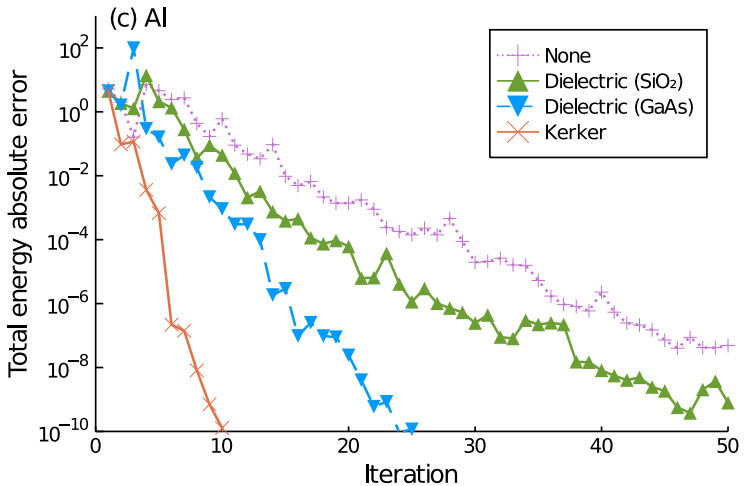
- Mixing schemes: Preconditioned quasi-Newton updates

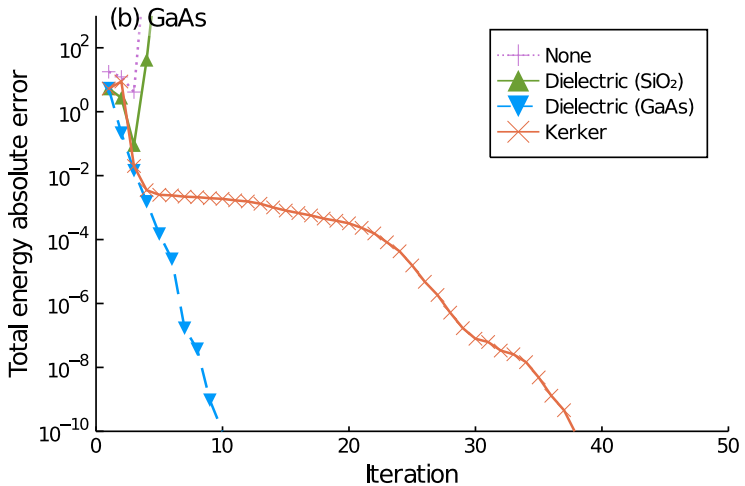
$$\rho_{n+1} = \rho_n + \alpha P^{-1} \left[F(\mathcal{V}(\rho_n)) - \rho_n \right]$$

where $P^{-1} \approx \left(\epsilon^\dagger\right)^{-1}$.

Convergence results for bulk materials[1]







- silica (SiO_2) insulator
- gallium arsenide (GaAs) semiconductor
- aluminium (Al) metal

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

Problems with established mixing methods

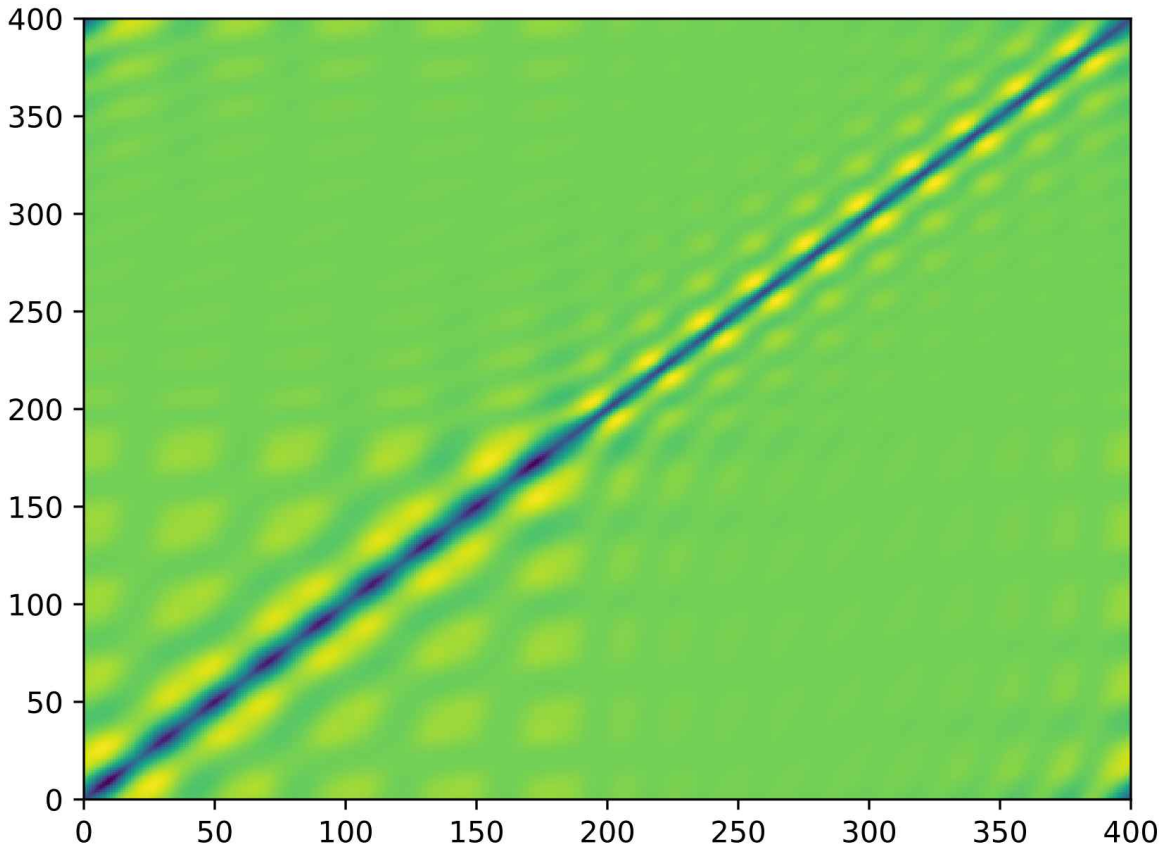
- $\epsilon^\dagger = 1 - \chi_0(v_C + f_{\text{xc}})$
- State-of-the art mixing models
  - Directly approximate $P^{-1} = \left(\varepsilon^\dagger\right)^{-1}$
  - "Neglect of local field effects" (to get closed expression in q)
  - Either highly material specific or feature many parameters

  $\Rightarrow$ Need to know properties *a priori* and manually chose

- Our approach: Approximate χ_0 directly
- Try a *non-local* approximation $\widetilde{\chi_0}(\underline{\boldsymbol{r}}, \underline{\boldsymbol{r}}') \simeq \chi_0(\underline{\boldsymbol{r}}, \underline{\boldsymbol{r}}')$
- Apply preconditioner iteratively

$$P^{-1}\rho_n = (1 - \widetilde{\chi_0} v_C))^{-1}\,\rho_n$$

Plot of (exact) χ_0



- 1D system (Chain of 10 Sodium atoms and 10 helium atoms)

Local density of states (LDOS) approximation for χ_0[1]

- Main interest: Large-scale variations in $\chi_0(\underline{\boldsymbol{r}}, \underline{\boldsymbol{r}}')$

$\Rightarrow$ Argue $\underline{\boldsymbol{r}}' \mapsto \chi_0(\underline{\boldsymbol{r}}, \underline{\boldsymbol{r}}')$ more localised around $\underline{\boldsymbol{r}}$ than $V(\underline{\boldsymbol{r}})$.

- "Row-sum mass lumping":

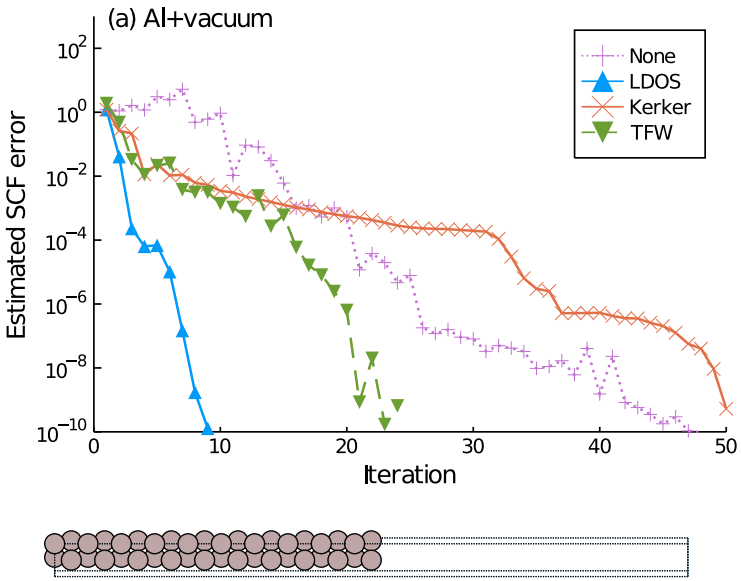
$$\begin{aligned}\int \chi_0(\underline{\boldsymbol{r}}, \underline{\boldsymbol{r}}') V(\underline{\boldsymbol{r}}') \,\mathrm{d}\underline{\boldsymbol{r}}' &\simeq V(\underline{\boldsymbol{r}}) \int \chi_0(\underline{\boldsymbol{r}}, \underline{\boldsymbol{r}}') \,\mathrm{d}\underline{\boldsymbol{r}}' \\ &= -V(\underline{\boldsymbol{r}}) D_{\mathrm{loc}}(\underline{\boldsymbol{r}})\end{aligned}$$

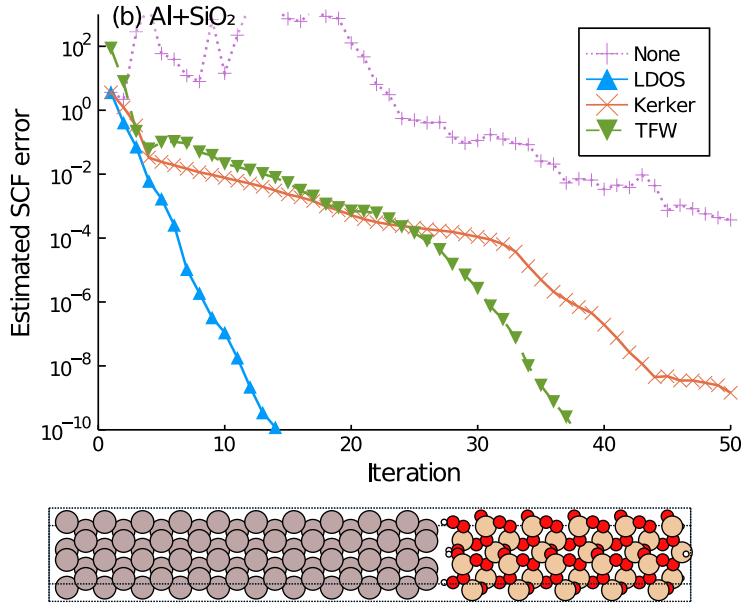
with local density of states

$$D_{\mathrm{loc}}(\underline{\boldsymbol{r}}) = \frac{1}{T} \sum_i f'\left(\frac{\varepsilon_i - \varepsilon_F}{T}\right) |\psi_i(\underline{\boldsymbol{r}})|^2$$

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

LDOS preconditioning (examples)





- 20 repeats of aluminium + 20 repeats vacuum / silica
- TFW: local Thomas-Fermi-von Weizsäcker mixing[1]
- LDOS automatically interpolates between Kerker mixing (in the metallic region) and no mixing (insulating region)

$\Rightarrow$ Parameter-free and black-box

[1] D. Raczkowski, A. Canning, L. W. Wang, Phys. Rev. B. **64**, 121101 (2001).

LDOS preconditioning results

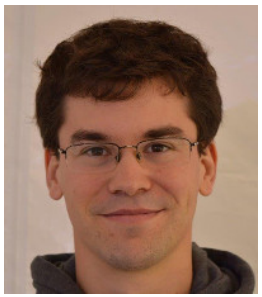
		None		Dielectric		Kerker		LDOS		LDOS+ Dielectric	
	$\mathcal{N}$	it	κ	it	κ	it	κ	it	κ	it	κ
SiO_2+vacuum	10	11	3.3	26	19.7	50	95.7	11	3.3	26	19.7
	20	12	3.4	30	24.4	n.c.	351.5	12	3.4	30	21.7
GaAs+vacuum	10	17	13.4	18	6.2	23	67.0	17	12.4	18	10.4
	20	20	15.5	22	12.9	n.c.	312.2	20	15.5	22	12.9
Al+vacuum	10	19	51.5	24	44.3	22	64.4	9	3.7	16	10.3
	20	47	170.8	49	168.5	n.c.	323.9	9	3.5	20	10.5
GaAs+SiO_2[a]	10	45	13.7	19	8.9	34	52.4	45	13.4	19	8.8
	20	n.c.	18.2	20	10.2	n.c.	170.1	n.c.	18.2	20	10.2
Al+SiO_2	10	43	93.1	29	33.6	30	50.9	17	6.1	20	9.2
	20	n.c.	316.6	n.c.	118.4	n.c.	159.4	14	5.4	20	10.1
Al+GaAs	10	n.c.	144.0	24	22.4	16	9.0	15	7.2	11	3.5
	20	n.c.	485.0	40	59.0	26	28.8	26	21.4	13	5.0
Al+GaAs+SiO_2	10	n.c.	149.5	34	50.4	36	62.9	26	21.5	19	9.0

- Coloured: Condition number κ less than doubled on doubling system size

Summary and outlook

- LDOS preconditioner:
  - Adaptive preconditioning for inhomogeneous systems
  - Parameter-free $\Rightarrow$ Highly suitable for high-throughput
- **DFTK** usage:
  - First develop LDOS scheme on test systems (1D, toy problems)
  - Test scheme on > 800 electrons (in the same code!)
- Next steps for a full black-box SCF preconditioner:
  - Spin / XC term
  - Black-box model for semiconductors
  - Localised states

Acknowledgements



Antoine Levitt

Eric Cancès
Xavier Gonze
Phil Hasnip
Lin Lin
Chao Yang

Other DFTK contributors:
E. Berquist, G. Kemlin, L. Ponet, D. Santra, N. Schmitz, S. Siraj-Dine, Z. Tóth

Questions?

DFTK https://dftk.org

julia: https://michael-herbst.com/learn-julia

mfherbst

https://michael-herbst.com/blog

herbst@acom.rwth-aachen.de