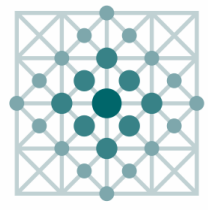
Thoughts on initial guess methods for density-functional theory

Michael F. Herbst

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

29th April 2021



RWTH AACHEN UNIVERSITY

→ https://michael-herbst.com/talks/2021.04.29_dft_guess_overview.pdf

Self-consistent field methods

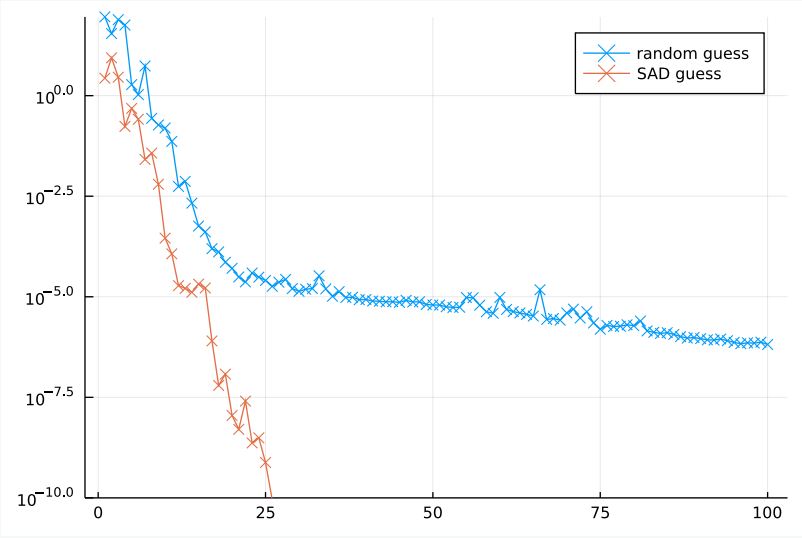
- DFT: Non-linear optimisation problem $\min_\rho E(\rho)$
- Euler-Lagrange equations (DFT):

$$\begin{cases} \hat{\mathcal{F}}_\rho = -\dfrac{1}{2}\Delta + V_{\boldsymbol{\rho}} \\ \boldsymbol{\rho}(\underline{\boldsymbol{r}}) = \displaystyle\sum_i f_{\varepsilon_F}(\varepsilon_i)\,|\psi_i(\underline{\boldsymbol{r}})|^2 \quad \text{with } \hat{\mathcal{F}}_\rho\psi_i = \varepsilon_i\psi_i, \\ \qquad \varepsilon_F \text{ chosen such that } \displaystyle\int \rho\,\mathrm{d}\underline{\boldsymbol{r}} = N, \\ \qquad \text{and } f_{\varepsilon_F}(x) = \left[1+\exp\left(\dfrac{x-\varepsilon_F}{T}\right)\right]^{-1} \end{cases}$$

- Self-consistent field procedure (SCF):
  (1) **Guess initial density $\boldsymbol{\rho}$**
  (2) Build Kohn-Sham operator $\hat{\mathcal{F}}_\rho$
  (3) Diagonalise it to get new $\{\psi_i\}_i$
  (4) Build new $\boldsymbol{\rho}$ go to (2).

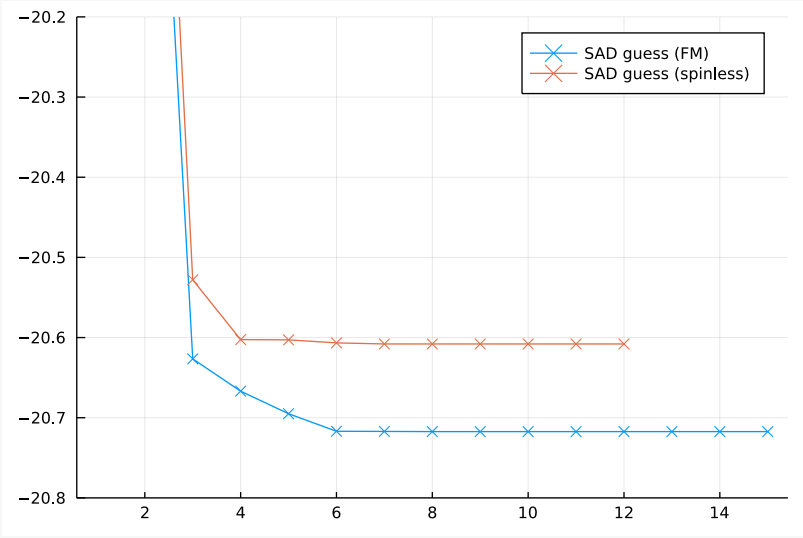
Why the guess matters . . .

Impact on convergence rate



- Lithium-copper alloy system
- Suitable guess suppresses slowly converging modes

Impact on obtained stationary point



- Iron: Non-magnetic, ferromagnetic, antiferromagnetic, . . .
- Guess targets specific state

Review of standard initial guess methods

- Reduced models
  - Extended Hückel, Generalized Wolfsberg-Helmholz, ...
- Precomputed atomic quantities
  - Superposition of atomic densities (SAD)
  - Superposition of atomic potentials (SAP)
- Extrapolation and fitting
  - Reduced-order modelling
  - Machine-learning techniques

Guesses from reduced models

- Core guess ($\rho = 0$)
- Generalized Wolfsberg-Helmholz (GWH)[1]
  - Off-diagonal elements from empirical formula
- Extended Hückel theory (EHT)[2]
  - Empirical Hamiltonian for valence space
  - Recently: Slater rules for core[3]
- Solved in minimal / reduced basis
- Not trivial to employ for real-space / periodic systems
- Still: Default for collinear problems in many molecular codes.

[1] M. Wolfsberg and L. Helmholz J. Chem. Phys. **20**, 837 (1952).

[2] R. Hoffmann J. Chem. Phys. **39**, 1397 (1963).

[3] M. Lee *et. al.* Comput. Theo. Chem. **1062**, 24 (2015).

Guesses from atomic quantities

- Superposition of atomic densities (SAD)
  - Densities from atomic calculations in vacuum
  - Tabulated on grid / reduced basis
  - Different levels of fanciness (fitting / basis size etc.)
  - Assumes: Weak long-range / interatomic interactions
  - Default guess in all major molecular and solid-state codes
- Superposition of atomic potentials (SAP)[1]
  - Tabulated radial potential for each atom
  - Computed on numerical grid
  - Seems to be better than SAD

[1] S. Lehtola J. Chem. Theo. Comp. **15**, 1593 (2019).

Guess from extrapolation

- Guess methods discussed so far are general
- But often related results are already known:
  - Geometry optimisation
  - Molecular dynamics
  - Design space search
- State of the art: Just use previous step

$\Rightarrow$ Clearly one can do better ...

Guess from extrapolation: Methods

- Reduced-order modelling[1]
  - Follow modes of geometry change (i.e. MD / GeoOpt)
  - Reduced-order extrapolation of density matrix
  - Extremely effective: Need only 1–3 steps for convergence!
  - Limited applicability:
    - Mode selection based on prior knowledge
- Machine-learning techniques[2]
  - E.g. ML surrogate to learn difference to standard guess
  - Judge system-specific uncertainty of surrogate
  - Modulate mix-in of surrogate guess

  $\Rightarrow$ Avoid guess deterioration due to surrogate

---

[1] E. Polack *et. al.* Mol. Phys. **118**, 19 (2020).

[2] A. T. Fowler *et. al.* J. Phys. Mater. **2**, 034001 (2019).

Opportunities for improvement

- Most guess methods:
  - Focus on general applicability
  - Few attempts to maximise usage of known results
- Extrapolation methods:
  - Barely explored
  - Promising when treating specific subclasses
  - All approaches I know neglect spin
- Some SCF classes (e.g. collinear): Primitive guess methods
  - Targeting particular spin state by "luck"

  $\Rightarrow$ Suspected reason why these are "hard"
- On large, difficult systems:

  $\Rightarrow$ Importance of targeted high-quality guess increases

Challenges

- Simple systems (i.e. small and easy):
  - Little to improve, even a random guess often good
- Difficult systems: Convergence behaviour
  - . . . dependent on discretisation
  - . . . may change during SCF (i.e. matters where you start)
  - . . . differs between targeted states

  $\Rightarrow$ Risk of making things worse with bad guesses

  $\Rightarrow$ Need ability to target the state of interest
- Considering data-driven approaches . . .
  - Ability to assess guess quality crucial
  - Target and distinguish specific final SCF states
  - Inclusion of sufficient variation wrt. challenging systems