A POSTERIORI ESTIMATORS FOR NONLINEAR KOHN-SHAM MODELS

Contents

1. Settings

We are aiming here at developing guaranteed *a posteriori* error estimators for nonlinear models used in DFT, such as Kohn-Sham DFT. To this aim, we will use the framework introduced in [1]: we consider a nonlinear energy of the form $E(P) = \mathrm{Tr}(H_0 P) + E_{\mathrm{nl}}(P)$, with P a rank N orthogonal projector. We will use in particular two objects:

- K_*, the Hessian of the energy at P_* projected onto the tangent space (note that it is 0 for linear problems);
- Ω_*, an operator acting on the tangent space as

$$\Omega = \sum_{i=1}^{N} \sum_{a=N+1}^{N_b} (\varepsilon_a - \varepsilon_i) \left(|\phi_i \otimes \phi_a\rangle \langle \phi_i \otimes \phi_a| + |\phi_a \otimes \phi_i\rangle \langle \phi_a \otimes \phi_i| \right), \tag{1}$$

  where $(\varepsilon_i)_{1 \leqslant i \leqslant N}$ are the energy levels of the occupied orbitals and $(\varepsilon_a)_{N < a \leqslant N_b}$ are the energy levels of the virtual orbitals (they all are eigenvalues of $H(P_*)$).

We assume that E has a nondegenerate local minimizer P_* in the sense that, at this point, there exists $\eta > 0$ such that $\Omega_* + K_* \geqslant \eta$ (seen as an operator on the tangent space). We can find numerically this minimizer either by solving a constrained minimization problem or by using SCF cycles.

In this framework, we want to cancel the projection of the gradient of the energy on the tangent plane, that is to say $[P_*, [P_*, H(P_*)]] = 0$. Moreover, as the Jacobian of the map $P \mapsto [P, [P, H(P)]]$ at P_* is $\Omega_* + K_*$, we have the following relation between the residual $[P, [P, H(P)]]$ and the error $P - P_*$:

$$[P, [P, H(P)]] = [P_*, [P_*, H(P_*)]] + (\Omega_* + K_*)(P - P_*) + O\left((P - P_*)^2\right). \tag{2}$$

Hence, we have the following approximation (up to smaller order terms), on which we will rely to compute bounds,

$$P - P_* \approx (\Omega_* + K_*)^{-1} [P, [P, H(P)]]. \tag{3}$$

Therefore, all the work lies in the approximation of the operator norm of $(\Omega_* + K_*)^{-1}$, and how it behaves for different norms.

2. First bounds

2.1. L^2 norm.

We have the following decomposition

$$P - P_* = (\Omega_* + K_*)^{-1} [P, [P, H(P)]] = (1 + \Omega_*^{-1} K_*)^{-1} \Omega_*^{-1} [P, [P, H(P)]], \tag{4}$$

which enables us to work on two separate operators:

- $(1 + \Omega_*^{-1} K_*)^{-1}$, which we know, empirically, that its norm, depending on the one we choose and the model, is between 0.8 and 1.2 for some systems without strong exchange-correlation we tested;
- Ω_*^{-1}, whose operator norm computation is the real challenge.

We propose a first approach to the computation of the operator norm of Ω_*^{-1}. Given (1), we have

$$\Omega_*^{-1} = \sum_{i=1}^{N} \sum_{a=N+1}^{N_b} \frac{1}{\varepsilon_a - \varepsilon_i} \left(|\phi_i \otimes \phi_a\rangle \langle \phi_i \otimes \phi_a| + |\phi_a \otimes \phi_i\rangle \langle \phi_a \otimes \phi_i| \right). \tag{5}$$

Then, given $k > 1$ an integer, we can decompose the set of basis functions of the tangent plane into

$$\{|\phi_i \otimes \phi_a\rangle \mid 1 \leqslant i \leqslant N < a \leqslant N_b\} = \{|\phi_i \otimes \phi_a\rangle \mid 1 \leqslant i \leqslant N < a \leqslant N+k\} \cup \{|\phi_i \otimes \phi_a\rangle \mid 1 \leqslant i \leqslant N \text{ and } N+k+1 \leqslant a \leqslant N_b\}$$

The first set corresponds to the eigenfunction we actually compute: for instance we need the first N-th eigenfunctions but we compute the first $(N+k)$-th to avoid any difficulties linked to the gap between the N-th and $(N+1)$-th energy levels. Thus, on this set we can compute Ω_*^{-1}, while on the second one we will need to make approximations. Given this decomposition, we have

$$\Omega_*^{-1} = \underbrace{\sum_{i=1}^{N} \sum_{a=N+1}^{N+k} \frac{1}{\varepsilon_a - \varepsilon_i} \left(|\phi_i \otimes \phi_a\rangle \langle \phi_i \otimes \phi_a| + |\phi_a \otimes \phi_i\rangle \langle \phi_a \otimes \phi_i| \right)}_{\Omega_{*,1}^{-1}} + \underbrace{\sum_{i=1}^{N} \sum_{a=N+k+1}^{N_b} \frac{1}{\varepsilon_a - \varepsilon_i} \left(|\phi_i \otimes \phi_a\rangle \langle \phi_i \otimes \phi_a| + |\phi_a \otimes \phi_i\rangle \langle \phi_a \otimes \phi_i| \right)}_{\Omega_{*,2}^{-1}} \tag{6}$$

$\Omega_{*,1}^{-1}$ can be computed numerically and, in the L^2 norm setting, we have the following upper bound for $\Omega_{*,2}^{-1}$:

$$\Omega_{*,2}^{-1} \leqslant \frac{1}{\varepsilon_{N+k+1} - \varepsilon_N}.$$

We see here that computing a bigger set of orbitals than the one we actually need reduces the bound on Ω_* (which would be otherwise controlled by $1/\nu$ with ν the gap of the system).

2.2. **Changing the norm.** One can also think of changing the norm in which we want to compute these operator norms. To this end, we introduce two operators to perform this change of metric: M_1 and M_2 (a first guess choice could be to take $M_1 = M_2 = 1 - \Delta$ to compute H^1 norms). Then, (4) becomes, by introducing M_1, M_2 and M_3,

$$\begin{aligned} M_1^{-1/2}(P - P_*) &= M_1^{-1/2}(\Omega_* + K_*)^{-1}[P,[P,H(P)]] \\ &= M_1^{-1/2}(1 + \Omega_*^{-1}K_*)^{-1} M_2^{-1/2} M_2^{1/2} \Omega_*^{-1} M_3^{1/2} M_3^{-1/2}[P,[P,H(P)]]. \end{aligned}$$

From this, we infer that

$$\left\| M_1^{1/2}(P - P_*) \right\| \leqslant \left\| M_1^{-1/2}(1 + \Omega_*^{-1}K_*)^{-1} M_2^{-1/2} \right\| \left\| M_2^{1/2} \Omega_*^{-1} M_3^{1/2} \right\| \left\| M_3^{-1/2}[P,[P,H(P)]] \right\|,$$

where:

- the error and the residual can be computed in different metrics M_1 and M_3;
- the first term of the right-hand side is empirically known to be close to one;
- the term we need to work on is $\left\| M_1^{1/2} \Omega_*^{-1} M_2^{1/2} \right\|$.

Those two operators are *super*-operators from the tangent plane to itself. Here are some examples (note that we don't need *a priori* to chose the same for both of them):

- I;
- Ω_*;
- $H(P_*) - \mu$ where μ is a given shift;
- $H(P_*) - T$ where T is acting on $|\phi_i \otimes \phi_a\rangle$ or $|\phi_a \otimes \phi_i\rangle$ as a multiplication by $T_i \in \mathbb{R}$, $(T_i)_{1 \leqslant i \leqslant N}$ being a nondecreasing sequence (in particular, if $T_i = \varepsilon_i$ we recover Ω_*);
- $-\Delta - \mu$ where μ is a given shift;
- $-\Delta - T$ with T as above;
- ...

2.2.1. *A first example.* For a first example, we can consider a Hamiltonian of the form $-\Delta + C$. Let $(\varepsilon_k)_{k \geqslant 0}$ be the eigenvalues of $-\Delta$. Let also $M_1 = M_2 = M$ where $M := T - \Delta$ is acting as

$$M(|\phi_i \otimes \phi_a\rangle \langle \phi_i \otimes \phi_a| + |\phi_a \otimes \phi_i\rangle \langle \phi_a \otimes \phi_i|) = (T_i + \varepsilon_a)(|\phi_i \otimes \phi_a\rangle \langle \phi_i \otimes \phi_a| + |\phi_a \otimes \phi_i\rangle \langle \phi_a \otimes \phi_i|),$$

where $(T_i)_{1 \leqslant i \leqslant N}$ is a nondecreasing sequence. Thus, we have

$$M^{1/2} \Omega_{*,2}^{-1} M^{1/2}(|\phi_i \otimes \phi_a\rangle \langle \phi_i \otimes \phi_a| + |\phi_a \otimes \phi_i\rangle \langle \phi_a \otimes \phi_i|) = \frac{T_i + \varepsilon_a}{\varepsilon_a - \varepsilon_i}(|\phi_i \otimes \phi_a\rangle \langle \phi_i \otimes \phi_a| + |\phi_a \otimes \phi_i\rangle \langle \phi_a \otimes \phi_i|).$$

We can then compute the operator norm of $M^{1/2}\Omega_{*,2}^{-1}M^{1/2}$ by computing

$$\sup_{\substack{1\leqslant i\leqslant N\\ N+k+1\leqslant a\leqslant N_b}} \frac{T_i+\varepsilon_a}{\varepsilon_a-\varepsilon_i} = \sup_{N+k+1\leqslant a} \frac{T_N+\varepsilon_a}{\varepsilon_a-\varepsilon_N} = \sup_{N+k+1\leqslant a} 1+\frac{T_N+\varepsilon_N}{\varepsilon_a-\varepsilon_N} = \max\left(1+\frac{T_N+\varepsilon_N}{\varepsilon_{N+k+1}-\varepsilon_N},1\right).$$

In the end, we get

$$M^{1/2}\Omega_{*,2}^{-1}M^{1/2} \leqslant \max\left(\frac{T_n+\varepsilon_{N+k+1}}{\varepsilon_{N+k+1}-\varepsilon_N},1\right).$$

3. Extension to observable error computations

From now on, we denote by Res the residual $[P,[P,H(P)]]$. Thus, relation (3) becomes

$$\delta P \approx (\Omega_*+K_*)^{-1}\mathrm{Res}.$$

Let A be a local observable. A direct Cauchy-Schwarz inequality yields

$$\mathrm{Tr}(A\delta P) \leqslant \|A\|_{P^*}\|\delta P\|_P \leqslant \|A\|_{P^*}\left\|(\Omega+K_*)^{-1}\right\|_{\mathrm{Res}\to P}\|\mathrm{Res}\|_{\mathrm{Res}}, \tag{7}$$

where $\|\cdot\|_{P_*}$ denotes the dual norm of the norm $\|\cdot\|_P$ we use to measure the error and $\|\cdot\|_{\mathrm{Res}\to P}$ is the operator norm from the residual to the error norms.

We also introduce two operators:

- an extension operator E which, given a density $\rho(r)$, builds the density matrix $(E\rho)(r,r') = \delta_{r,r'}\rho(r)$;
- a restriction operator R which, given a density matrix $O(r,r')$, builds the density $(RO)(r) = O(r,r)$.

In this framework, the linear response operator is $\chi_0 = -R\Omega_*^{-1}E$ and the dielectric operator is $\varepsilon^{-T} = R(1+\Omega_*^{-1}K)^{-1}E$. Therefore, the relation between the residual and the error on the density $\delta\rho$ is

$$\delta\rho = R\left(\left(1+\Omega_*^{-1}K\right)^{-1}ER\Omega_*^{-1}\mathrm{Res}\right) = \varepsilon^{-T}R\Omega_*^{-1}\mathrm{Res}.$$

As A is a local observable, there exists a function a such that

$$\mathrm{Tr}(A\delta P) = \int_{\mathbb{R}^3} a\delta\rho.$$

Therefore, Cauchy-Schwarz inequality yields (with the same notations than in (7))

$$\int_{\mathbb{R}^3} a\delta\rho \leqslant \|a\|_{\rho_2^*}\left\|\varepsilon^{-T}\right\|_{\rho_1\to\rho_2}\left\|R\Omega_*^{-1}\right\|_{\mathrm{Res}\to\rho_1}\|\mathrm{Res}\|_{\mathrm{Res}}.$$

The main question now is: how to evaluate the error we make by this Cauchy-Schwarz estimate ? It is likely that a will be composed of low frequencies while $\delta\rho$ is composed of high frequencies so that the this estimate will not be optimal. The next step is to evaluate for some systems how this bound is far from being optimal.

3.1. Estimates on χ_0.

The upper bound

$$\left\|\sum_{1\leqslant i\leqslant N<a}\frac{|\phi_i\phi_a\rangle\langle\phi_i\phi_a|}{\varepsilon_a-\varepsilon_i}\right\| \leqslant \sum_{1\leqslant i\leqslant N<a}\frac{\|\phi_i\phi_a\|^2}{\varepsilon_a-\varepsilon_i}$$

tells us that χ_0 can explode in two different ways: either $\|\phi_i\phi_a\|\to+\infty$ or $\varepsilon_a-\varepsilon_i\to 0$.

On the other hand, let v a potential such that $\|v\|_{\mathrm{L}^2(\mathbb{R}^3)}$. Then

$$\begin{aligned}
\|\chi_0 v\|_{L^2}^2 &= \left\|\sum_{1\leqslant i\leqslant N<a}\frac{|\phi_i\phi_a\rangle\langle\phi_i\phi_a|v\rangle}{\varepsilon_a-\varepsilon_i}\right\|_{L^2}^2 = \sum_{1\leqslant i,j\leqslant N<a,b}\frac{\langle\phi_i\phi_a|v\rangle\langle\phi_j\phi_b|v\rangle\langle\phi_i\phi_a|\phi_j\phi_b\rangle}{(\varepsilon_a-\varepsilon_i)(\varepsilon_b-\varepsilon_j)}\\
&\leqslant \left(\sum_{1\leqslant i,j\leqslant N<a,b}\frac{\langle\phi_i\phi_a|v\rangle}{\varepsilon_a-\varepsilon_i}\|\phi_i\phi_a\|_{L^2}\right)^2\\
&\leqslant M\left(\sum_{1\leqslant i\leqslant N<a}\frac{1}{(\varepsilon_a-\varepsilon_i)^2}\right)\left(\sum_{1\leqslant i\leqslant N<a}|\langle\phi_a|v\phi_i\rangle|^2\right)\\
&\leqslant M^2\left(\sum_{1\leqslant i\leqslant N<a}\frac{1}{(\varepsilon_a-\varepsilon_i)^2}\right),
\end{aligned}$$

where $M := \max_{1 \leqslant i \leqslant N} \|\phi_i\|_{\mathrm{L}^\infty}$. As v is of norm 1, we have

$$\|\chi_0\|_{\mathrm{L}^2 \to \mathrm{L}^2} \leqslant M^2 \left(\sum_{1 \leqslant i \leqslant N < a} \frac{1}{(\varepsilon_a - \varepsilon_i)^2} \right)^{1/2}.$$

References

[1] E. Cancès, G. Kemlin, and A. Levitt. Convergence analysis of direct minimization and self-consistent iterations. Preprint, Apr. 2020.