Adaptive Gaussian basis for (post) Hartree-Fock calculation

Jiaxi Zhao

December 3, 2024

Chemists' notation

Slater determinant

The many-body wave function of the Hartree-Fock theory is given by a single Slater determinant:

$$|\Psi_{\mathrm{HF}}\rangle = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(1) & \phi_1(2) & \cdots & \phi_1(N) \\ \phi_2(1) & \phi_2(2) & \cdots & \phi_2(N) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_N(1) & \phi_N(\mathbf{r}_2) & \cdots & \phi_N(\mathbf{r}_N) \end{vmatrix},$$

Hartree-Fock energy

Given a Slater determinant $|\Psi_{\mathrm{HF}}\rangle$, the energy is given by:

$$
\begin{aligned}
E[\Psi^{HF}] &= \left\langle \Psi^{HF} | \hat{H}^e | \Psi^{HF} \right\rangle \\
&= \sum_{i=1}^{N} \int \mathrm{d}\mathbf{x}_i \, \phi_i^*(\mathbf{x}_i) \hat{h}(\mathbf{x}_i) \phi_i(\mathbf{x}_i) \\
&+ \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \int \mathrm{d}\mathbf{x}_i \int \mathrm{d}\mathbf{x}_j \phi_i^*(\mathbf{x}_i) \phi_j^*(\mathbf{x}_j) \frac{1}{|\mathbf{x}_i - \mathbf{x}_j|} \phi_i(\mathbf{x}_i) \phi_j(\mathbf{x}_j) \\
&- \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \int \mathrm{d}\mathbf{x}_i \int \mathrm{d}\mathbf{x}_j \phi_i^*(\mathbf{x}_i) \phi_j^*(\mathbf{x}_j) \frac{1}{|\mathbf{x}_i - \mathbf{x}_j|} \phi_i(\mathbf{x}_j) \phi_j(\mathbf{x}_i)
\end{aligned}
$$

Minimizing the total energy will give the ground state configuration which is consistent with the self-consisten field approach.

Basis set

Classically, basis set of increasing complexity is used to

1. STO-n
2. 6-311G*
3. cc-pVDZ

In this work, we propose to use the adaptive Gaussian basis with tunable mean and covariance:

$$\phi_i = \sum_{j \in I} c_{ij} \mathcal{N}(\mathbf{r}; \mu_j, \Sigma_j).$$

In general, all the optimizable parameters are $\mu \in \mathbb{R}^{|I| \times 3}, \Sigma \in \left(\mathbb{S}^3_+\right)^{|I|}, C \in \mathbb{C}^{|I| \times N}$. $|I|$ is the size of the basis set, N is the number of electrons (ground state + excited state).

orthognormality and overlap matrix

The orthgnormality of the orbitals in SCF is guaranteed by the eigensolver. In total energy minimization, we have to handle this explicitly.

As the orbitals are supposed to be orthognormal, it can be written as

$$\mathsf{C}^\dagger \mathsf{S}\, \mathsf{C} = \mathsf{I}\,.$$

where S is the overlap matrix defined as:

$$\mathsf{S}_{ij} = \langle \mathcal{N}\left(\mathbf{r}; \mu_{\mathbf{1}}, \Sigma_1\right) |\, \mathcal{N}\left(\mathbf{r}; \mu_{\mathbf{2}}, \Sigma_2\right) \rangle\,.$$

The orthognormality is enforced by QR decomposition modified by the Cholesky factorization of S.

Coulumb and external energy

Our primary idea is to approximate the Coulumb kernel by a series of Gaussian modes:

$$\frac{1}{r} = \sum_i c_i e^{-\alpha_i r^2},$$

which is done by minimizing the

$$\min_{c_i,\alpha_i} \int_{B_M} \left(\frac{1}{|\mathbf{r}|} - \sum_i c_i e^{-\alpha_i |\mathbf{r}|^2} \right)^2 d\mathbf{r} = \min_{c_i,\alpha_i} 4\pi \int_0^M \left(1 - \sum_i c_i r e^{-\alpha_i r^2} \right)^2 dr.$$

Parameterization of Gaussian basis

We parameterize the Gaussian basis as

$$\Sigma_1 = \mathrm{U}_1 \, \mathrm{D}_1 \, \mathrm{U}_1^T, \quad \Sigma_2 = \mathrm{U}_2 \, \mathrm{D}_2 \, \mathrm{U}_2^T,$$

where U_i is obtained from the QR decomposition and the positivity of the diagonal element of D_i is enforced by the softplus function.

Numerical experiments

Table: .

	e_kin + e_ext	e_kin	e_ext	e_nuc	e_coul + e_exc	e_coul	e_exc	e_tot
OUR (H_2), 10	-2.5062	1.1254	-3.6316	0.7138	0.6591	1.3182	-0.6591	-1.1334
HF (STO-3G, 2)	-2.5049	*	*	0.7138	0.6745	*	*	-1.1167
HF (6-31G, 4)	-2.4902	*	*	0.7138	0.6497	*	*	-1.1267
HF (6-311G, 6)	-2.4924	*	*	0.7138	0.6507	*	*	-1.1280
OUR (CH_4), 22	-79.6713	40.0987	-119.7700	13.4477	26.1010	32.6936	-6.5926	-40.1225
HF (STO-3G, 9)	-79.3617	*	*	13.4477	26.1872	*	*	-39.7267
HF (6-31G, 17)	-79.6901	*	*	13.4477	26.0619	*	*	-40.1804
HF (6-311G, 25)	-79.6872	*	*	13.4477	26.0515	*	*	-40.1880
OUR (H_2O), 22	-122.9802	75.9462	-199.0371	9.1895	37.8786	46.8579	-8.9601	-76.0036
HF (STO-3G, 7)	-122.3614	*	*	9.1895	38.2089	*	*	-74.9630
HF (6-31G, 13)	-122.9701	*	*	9.1895	37.7966	*	*	-75.9839
HF (6-311G, 19)	-123.0146	*	*	9.1895	37.8157	*	*	-76.0094

Next step

1. Compare the scalability of the double-electron integral calculation over the number of basis functions with the SOTA method, e.g. MD, Rys, HDG.
2. Evidence from numerical experiments to show that the adaptive Gaussian basis can achieve the same accuracy with a smaller number of basis.
3. Implement post HF method, e.g. CCSD, CCSD(T).

Q & A

Motivation

The effect cutoff used by the plane-wave method to solve large system is unaffordable. Moreover, given the nature of the electron density which will be concentrated around the nucleus, while away from the nucleus which is the case of majority of the computational region, the electron density is relatively smooth, the plane-wave method is not efficient. We need to survey the literature for some multiscale basis such as the wavelet basis.

Literature

In literature, there has been lots of work on exploring the adaptive basis for solving DFT.

- Use fixed analytic transformation mapping depending on the position of the atom. (Similar to the r type mesh refinement in numerical PDE.)
- Solve some OT-type problems to predetermine the basis before the SCF loop.

However, our method is different from the above method in several aspects:

- Our basis is continuously optimized during the optimization porcess of DFT.
- We use Monte-Carlo method to estimate the density related energy, differs from the traditional PDE solving method.

Advantages

There are several advantages of using the distorted plane-wave method.

The orthognormality of the basis is preserved.

Calculation of the density related energy can be estimated efficiently by the Monte-Carlo method instead of solving the PDE.

References

M. Benjamin, S. Domino, and G. Iaccarino
Neural Networks for Large Eddy Simulations of Wall-bounded Turbulence: Numerical Experiments and Challenges
The European Physical Journal E

J. Zhao and Q. Li (2024)
Mitigating Distribution Shift in Machine Learning-augmented Hybrid Simulation
Arxiv preprint https://arxiv.org/pdf/2401.09259

S. Arisaka and Q. Li (2024)
Accelerating Legacy Numerical Solvers by Non-intrusive Gradient-based Meta-solving
International Conference on Machine Learning 2024