BT5420

Computer Simulations of Biomolecular Systems Project II

Shreeharsha G Bhat | BE21B037
Department of Biotechnology
Indian Institute of Technology Madras

Table of Contents

The paper :

---

1   Objective :

Neural Network Potentials(NNPs) are powerful computational methods for modeling large chemical systems with accuracy of Quantum Mechanical methods at a much smaller computational cost.

The objective of this paper is to extend these NNPs to treat dispersion interactions rigorously by calculating atomic dispersion coefficients through a second set of NNs. The second Neural Network is trained to reproduce the coefficients from the quantum-mechanically derived exchange-hole dipole moment (XDM) model.

1.1   The reason I chose this article

Neural Networks is a topic that interests me very much and given the parameters for choosing a paper, this one seems to check all the boxes. The paper takes an existing NNP and extends the idea of neural networks for calculating potentials on large scale dispersion-driven systems it to get more accurate results with a much lesser computational cost

The network used to calculate the intramolecular forces and potential is imitated to do the same to long range intermolecular forces. The goal of the paper was to create a neural network potential that considered London dispersion forces, which are forces driven by spontaneous disturbances in one atom affecting an atom or molecule close by. The topics this particular article touched upon were aligned with the course content with an addition of neural networks to extract certain constants.

2   Background :

2.1   Neural Network Potentials (NNPs)

NNPs are a class of machine learning models that predict the potential energy of a chemical structure. The accuracy of such a neural network is comparable to the quantum mechanical method. Existing NNPs are very effective for modeling intramolecular interactions, but very few NNP models account for long-range intermolecular interactions. The effects of interactions within a short cutoff radius are considered when modeling these NNPs. Interactions such as electrostatic and dispersion are neglected and are considered as noise for short-range NNPs.

There are several NNPs that describe the long-range electrostatic interactions, but london dispersions are generally ignored

2.2   Dispersion Interactions

London dispersion forces are weak intermolecular van der Waals forces that occur between all atoms and molecules. Occurring due to the fluctuations in the electron distribution within the atom, they show temporary dipoles that induce similar dipoles to its surrounding atoms, creating a weak attractive force.

Dispersion interactions between any two molecules say i and j, is approximated as a multipole expansion of the oscillating electric moments between the two atoms or molecules

$$v_{\text{disp},ij}\left(R_{ij}\right) = -\frac{C_{6,ij}}{R_{ij}^{6}} - \frac{C_{8,ij}}{R_{ij}^{8}} - \frac{C_{10,ij}}{R_{ij}^{10}}\cdots$$

R_{ij} – distance between the molecule i and molecule j
C_6, C_8, C_{10} – dispersion coefficients for each pair of molecules

The dispersion coefficients can be calculated using the exchange-hole dipole moment(XDM).

In the XDM model C_6, C_8, C_{10} are related to the expectation values of the squares of the atomic exchange-hole multipole moments and the atomic polarizability.

$$C_{6,ij} = \alpha_i \alpha_j \frac{\left\langle M_1^2\right\rangle_i \left\langle M_1^2\right\rangle_j}{\alpha_i \left\langle M_1^2\right\rangle_j + \alpha_j \left\langle M_1^2\right\rangle_i},$$

$$C_{8,ij} = \frac{3}{2}\alpha_i\alpha_j \frac{\langle M_1^2\rangle_i\langle M_2^2\rangle_j + \langle M_2^2\rangle_i\langle M_1^2\rangle_j}{\alpha_j\langle M_1^2\rangle_i + \alpha_i\langle M_1^2\rangle_j}$$

$$C_{10,ij} = \alpha_i\alpha_j \frac{2\langle M_1^2\rangle_i\langle M_3^2\rangle_j + 2\langle M_3^2\rangle_i\langle M_1^2\rangle_j + \frac{21}{5}\langle M_2^2\rangle_i\langle M_2^2\rangle_j}{\alpha_j\langle M_1^2\rangle_i + \alpha_i\langle M_1^2\rangle_j}$$

α – polarizability
$\langle M_1^2\rangle$ – expectation value of the square of the exchange-hole dipole moment
$\langle M_2^2\rangle$ – expectation value of the square of the exchange-hole quadrupole moment
$\langle M_3^2\rangle$ – expectation value of the square of the exchange-hole octupole moment

The polarizability is approximated by scaling the free-atom polarizability by the ratio of atom-in-molecule and free-atom Hirshfeld volumes:

$$\alpha_i = \frac{V_i}{V_{i,\text{free}}}\alpha_{i,\text{free}}.$$

When the above C values are substituted to the V_{disp} equation, we get

$$v_{disp,ij}(R_{ij}) = -\sum_i^N\sum_{j>i}\left(\frac{C_{6,ij}}{R_{ij}^6 + R_{vdW,ij}^6} + \frac{C_{8,ij}}{R_{ij}^8 + R_{vdW,ij}^8} + \frac{C_{10,ij}}{R_{ij}^{10} + R_{vdW,ij}^{10}}\right)$$

The Van der Waal(vDW) radii, $R_{\text{vdW, }ij}$, is introduced to damp the dispersion energy at closer distances. This is done by assigning a critical distance $R_{\text{critical, }ij}$, at which various terms in the series expansion become equal.

$$R_{\text{vdW},ij} = a_1 R_{\text{critical},ij} + a_2.$$

a1 and a2 are calculated empirically for each density functional.
The $R_{\text{critical, }ij}$ is calculated by averaging the ratios of C_6, C_8 & C_{10}

$$R_{\text{critical},ij} = \frac{1}{3}\left[\left(\frac{C_{8,ij}}{C_{6,ij}}\right)^{\frac{1}{2}} + \left(\frac{C_{10,ij}}{C_{6,ij}}\right)^{\frac{1}{4}} + \left(\frac{C_{10,ij}}{C_{8,ij}}\right)^{\frac{1}{2}}\right].$$

Hence DFT calculations must be performed first inorder to get the electron distribution function for the XDM

3   Computational Technique :

Accurate Neural Network engine(ANI) NNPs are efficient neural networks providing molecular reactions & conformational energies with the accuracy similar to the level of first principle QM models, they are mainly used due to the advantage of having low computational cost.

To calculate the dispersion coefficients, the article describes the usage of a neural network that is based on those used in ANI-type NNPs. Once the dispersion coefficients are found, they are then used to add dispersion correction to the NNP that predicts the non-dispersion aspects of potential energy. The NNP was trained using a combination of RMS error of the molecular energy and the norm of the gradients as the loss function.

3.1   Construction of the NNP

PBE0 exchange-correlation function is used to train & reproduce the energy and forces. These forces do not take into account the dispersion forces. The total energy of the chemical system is calculated as the sum of the individual atomic energies.

$$\nu_{\text{NNP}} = \sum_{i=1}^{N_{\text{atoms}}} \nu_{\text{NNP},i}.$$

The input chemical environment for all atoms is given in the form of an Atomic Environment Vector (AEV).

3.2   Dispersion coefficient Neural Network

The modified Behler-Parrinello symmetry function AEVs are reused as the features for predictions to maintain consistency. Even the input features and network architecture were conserved between the two networks. The neural network was then used to predict the atomic moments $\langle M_1^2\rangle, \langle M_2^2\rangle, \langle M_3^2\rangle$ and the volume V for the hydrogen, carbon,

nitrogen & oxygen atoms. Inorder to maximize the performance of the Neural network, the moments and volumes are standardized such that the NN predicts the deviation of the quantity from its mean value, scaled by the standard deviation.

3.3 Training Set

The Gaussian 16 program was used to create a large training set of potential energies and forces.
The loss function used while training was a combination of the RMS error of the molecular energy and the norm of the gradient. The Gaussian 16 program is used for modeling electronic structure and ab initio calculations of molecular systems. The final dataset contained QM energies, forces and dispersion coefficients for 2.1 million chemical structures.

4 Advantages & Disadvantages :

4.1 Advantages

This paper provided a model that improves the description of electron correlation. The model takes into account the london dispersion interactions, while most current models stop after considering the electrostatic non-bonded interaction

Exchange-hole Dipole Model(XDM) requires the electron density distributions and its derivatives to be found to calculate the dispersion coefficients. But, this model overrides this and reduces the *computational cost* by billions by running a trained neural network to find the coefficients.

Moreover, the dispersion coefficients are dynamic in nature. The adaptive ability of the coefficients allows the model to be more accurate and true.

The neural network is a significant advantage since the Density functional theory calculations do not need to be performed for the XDM dispersion corrections to be calculated.

4.2 Disadvantages/Constraints

The model performs poorly when compared to the interaction energies calculated by the corrected-wave function method. The performance is better when the coefficients are made adaptive when compared to when it is not, but the error is significant enough to consider it a minor drawback

The neural network architecture and input vectors of the NNP for intramolecular forces and intermolecular forces are the same. This is a drawback since there is no hyperparameter tuning of the network.

If the NNP for dispersion forces were allowed to be tuned when it comes to the network architecture, then the model accuracy and complexity could be improved.

This particular NNP has niche use cases. This model can be used to analyze the interaction in Covalent Organic Frameworks and molecular liquids very accurately, but better models can be used in other cases.

5   Similar Concepts :

In this particular article, the concepts used that were connected to the ones mentioned in the course were.

5.1   Exchange-hole Dipole Moment

The Exchange-hole Dipole Moment or the "XDM" model deals with the question
What is the source of the instantaneous multipole moments that generates the dispersion interaction?
The Van der Waals theory has two variants, one **orbital based** and the other **Density-functional based**. The orbital based theory works slightly better in rare-gas systems and the density functional theory works better in intermolecular complexes
This concept is used to calculate the London dispersion forces and energies. London dispersion forces are non-bonded interaction, which implies that they must have some sort of a cut-off scheme and parameters that include or exclude certain interactions

In the *Hartree-Fock theory* the total energy of a many-electron system is given by

$$E_{HF} = -\frac{1}{2}\sum_{i,\sigma}\int \Psi_{i\sigma}\nabla^2\Psi_{i\sigma}d^3r + \int \rho V_{nuc}\,d^3r + \frac{1}{2}\int\int \frac{\rho(r_1)\rho(r_2)}{r_{12}}d^3r_1 d^3r_2 + E_X$$

Which seems like a lot, but the main focus is the E_x term which signifies the exchange energy of the system.

Exchange Energy is the energy released when electrons in the system with same spin change places in a subshell's degenerate orbitals

A component of E_x will be the spin exchange energy which can be rewritten to contain a term called the "Exchange Hole".

Each electron interacts with a hole whose shape depends on the electron's position. When an electron is at a position r1, the hole measures the depletion of probability w.r.t the total energy density of finding another electron at r2 with the same spin.

Another point to note is that an electron plus its exchange hole always have zero net charge but in general a non-zero dipole moment. This r1 dependent dipole moment is obtained by integrating over r2. Only the occupied orbitals are involved.

5.2 Dispersion interaction : Spherical Atoms
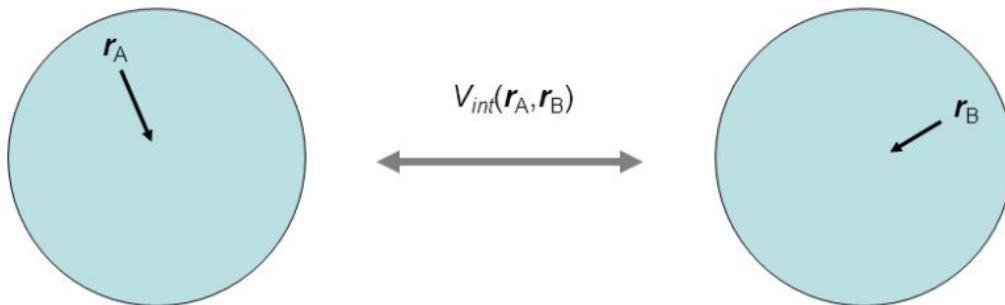
Here, we have two electron systems, at positions r_A and r_B

$V_{int}(r_A, r_B)$ = multipole moments of electron + hole at r_A **interacting with** multipole moments of electron + hole at r_B

To evaluate the ground state energy correction arising from this we would need to evaluate

$\langle V_{int}^2 \rangle$ which would take into account every dipole-dipole, dipole-quadrupole, dipole-octopole, quadrupole-quadrupole and other such interactions.

This will give us the equations that are mentioned above in **section 2.2**

6   Acknowledgements :

I would like to express my sincere gratitude to the course faculty Dr. Hamsa Priya Mohanasundaram for her invaluable guidance and support throughout this course. Her expertise and feedback have been instrumental in shaping this report. I also want to thank my fellow classmates for engaging discussions and helping me understand the course content better.

****