"Quantum Machine Learning for determining Exchange-Correlation Potentials in Many-Body Systems"

Asmita Datta

BS-MS at Department of Physics, IISER Bhopal, Madhya Pradesh-462066, India.

Email: asmita18@iiserb.ac.in

Advisor: Dr Kuntal Roy, EECS Department, IISER Bhopal.

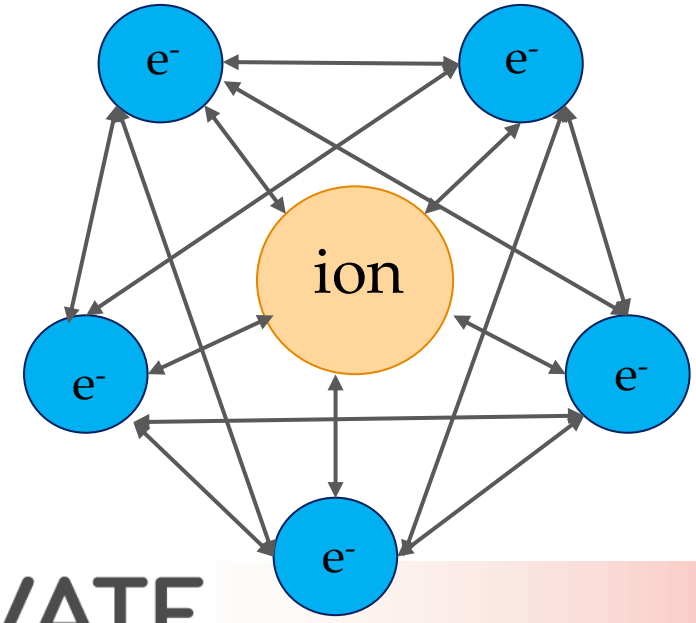
Funded by the European Union

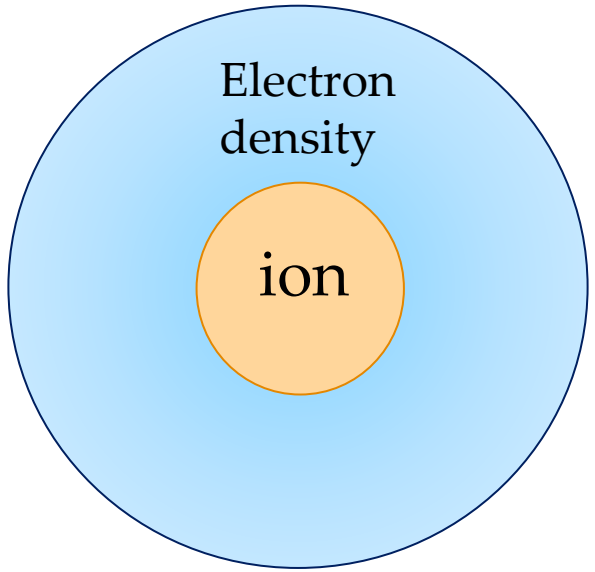
AQTIVATE

The *AQTIVATE* project receives funding from the European Union's HORIZON MSCA Doctoral Networks programme, under Grant Agreement No. 101072344.

Introduction

- **Aim:** Applying Quantum Machine Learning algorithms to obtain the hybrid Heyd-Scuzeria-Ernzerhof (HSE) exchange-correlation functional parameters in Density Functional Theory (DFT).
- **Many-body Theory:** Solving the N-body Schrodinger eqn. to obtain the wavefunction (Ψ).
- Electron-electron interaction term in the Schrodinger Hamiltonian makes it extremely difficult to obtain Ψ.





$$\hat{H} = -\frac{\hbar^2}{2m}\sum_i^N \nabla_i^2 - \sum_{i,I}^{N,M}\frac{Z_I e^2}{|\mathbf{r}_i - \mathbf{R}_I|} + \frac{1}{2}\sum_{i\neq j}\frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_I^M \frac{\hbar^2}{2M_I}\nabla_I^2 + \frac{1}{2}\sum_{I\neq J}\frac{Z_I Z_J e^2}{|\mathbf{R}_I - \mathbf{R}_J|}$$

Density Functional Theory (DFT)

- Treats electron density $n(\boldsymbol{r})$ as the basic variable.

Hohenberg-Kohn theorem: $\Psi(\boldsymbol{r_1}, \boldsymbol{r_2}, \cdots, \mathbf{r_N}) \Leftrightarrow n(\boldsymbol{r})$

Kohn-Sham (KS): $\left(-\frac{\hbar^2\nabla^2}{2m} + v_{KS}(\mathbf{r})\right)\boldsymbol{\varphi}_{\mathrm{k}}(\mathbf{r}) = \epsilon_{\mathrm{k}}(\boldsymbol{\varphi}_{\mathrm{k}}(\mathbf{r}))$

Eff. Pot.: $v_{KS}(\mathbf{r}) = \left(v_{ext}(\mathbf{r}) + v_H(\mathbf{r}) + v_{xc}(\mathbf{r})\right)$

- **Exchange-correlation (xc) functionals:** explicitly unknown contributions to the energy functional.

$$E_{xc}[n] = (T[n] - T_s[n]) + (W[n] - E_H[n])$$

Local Density Approximation (LDA), Generalized Gradient Approximation (GGA)

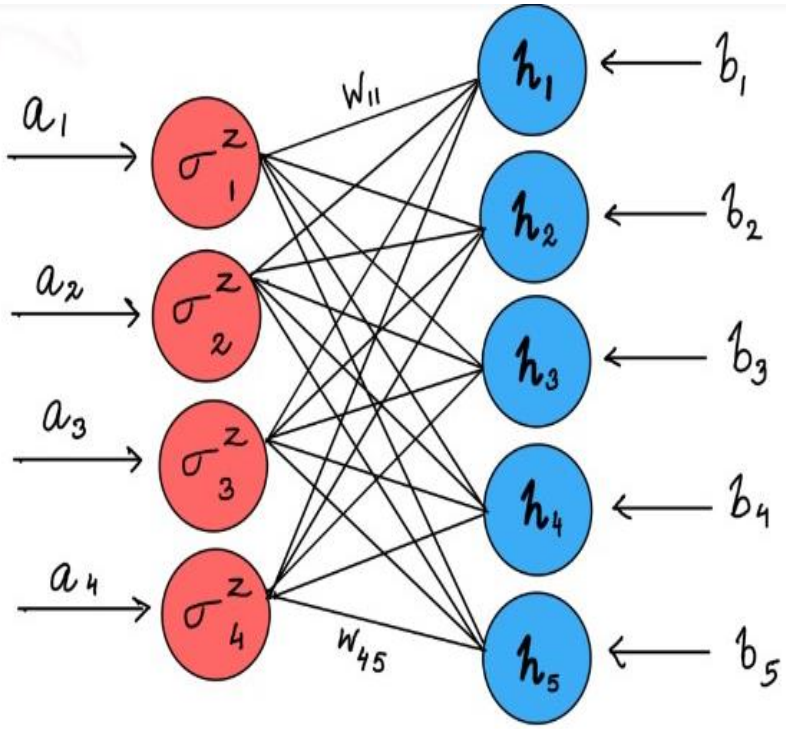
- **Heyd-Scuzeria-Ernzerhof (HSE) hybrid:** Fock Exchange with GGA with a screened Coulomb potential applied to the exchange part

$$\frac{\mathbf{1}}{\boldsymbol{r}} = \frac{\mathbf{erfc}(\boldsymbol{\omega r})}{\boldsymbol{r}} + \frac{\boldsymbol{erf}(\boldsymbol{\omega r})}{\boldsymbol{r}}$$

$$\boldsymbol{E_{xc}^{HSE}[n(r)] = \alpha E_x^{HF,SR}(\omega) + (1-\alpha)E_x^{PBE,SR}(\omega) + \alpha E_x^{PBE,LR}(\omega) + E_c^{PBE}}$$

Quantum Machine Learning (QML)

- **Quantum Machine Learning:** algorithms based on the quantum principles to speedup ML algorithms, give better accuracy, solve classically hard computations



Restricted Boltzmann Machine (RBM)

Single input and hidden layer, no connections between nodes in the same layer, all-to-all connections b/w input and hidden neurons

N-body spin configuration as input

$$S = (\sigma_1^z, \sigma_2^z, \dots \dots \sigma_N^z)$$

M auxiliary spin variables

$$H_L = (h_1, h_2, \dots \dots, h_M)$$

Variational states

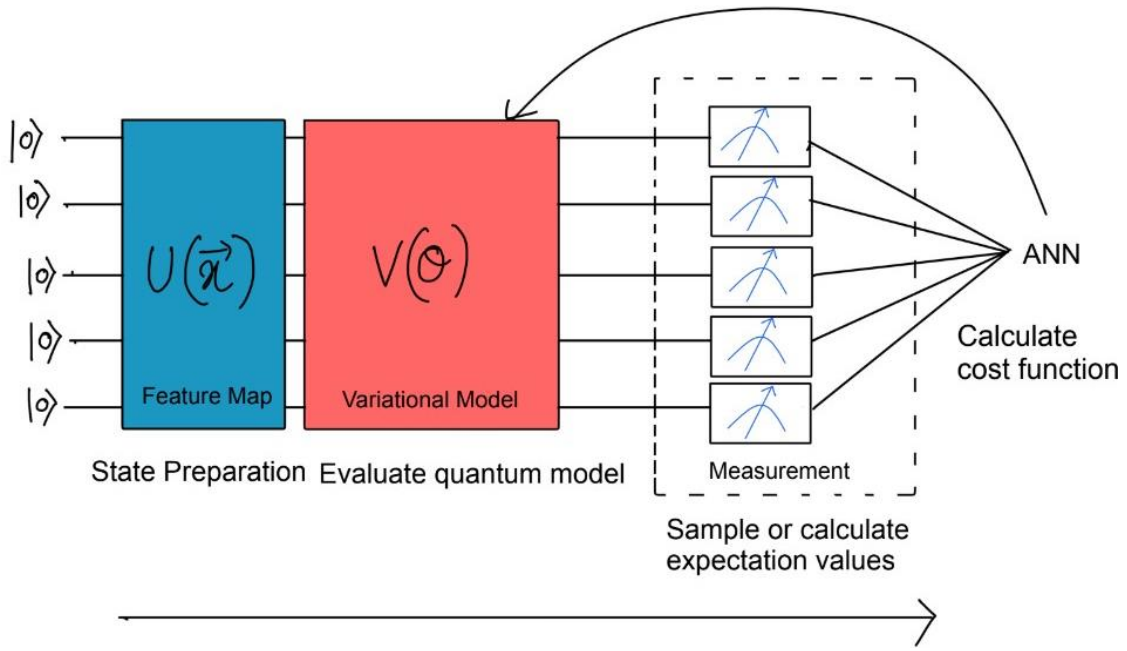
$$\Psi(S; W) = \sum_{h_i} e^{\sum_j a_j \sigma_j^z + \sum_i b_i h_i + \sum_{ij} w_{ij} h_i \sigma_j^z}$$

$$E(W) = \langle \Psi | H | \Psi \rangle / \langle \Psi | \Psi \rangle$$

$\{a_i, b_i\} \rightarrow$ biases of the neural network

$w = \{w_{ij}\} \rightarrow$ weights of the neural network

Quantum Machine Learning (QML)



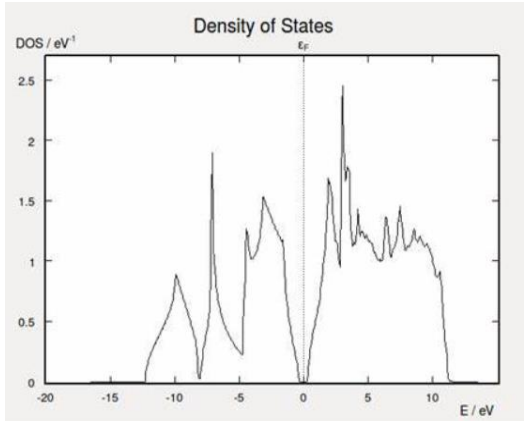
Ansatz – quantum circuit that encodes the parameters $\Theta = [\theta_1, \theta_2, \dots\dots, \theta_m]$ that are optimized by the classical optimizer

- Task encoded into a parameterized quantum circuit -> evaluated using a quantum computer -> classical optimizer trains the parameters in the VQA
- **Feature Map –** encode classical data into quantum states by unitary transformations

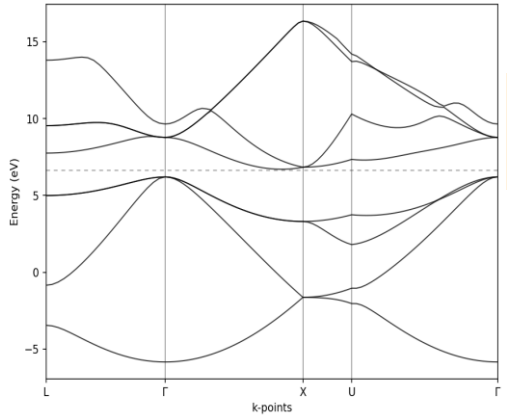
  $\Psi : \chi \to \mathcal{H}$; where $x \in \chi$ is the input vector in the feature space χ; $|\Psi(x)\rangle$ is the mapped state
- **Measurement –** final state of the circuit is measured -> collapses to one of the computational n-qubit basis states -> classical value given to the optimizer

Results - DFT

Quantum Espresso (QE)



Density of States (DOS) of Si using GGA (PBE)

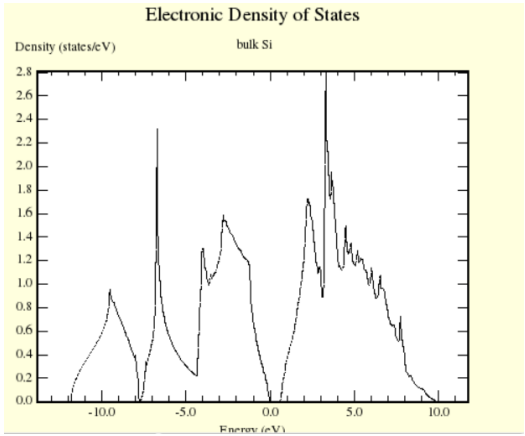


Band structure of Si using GGA (PBE)

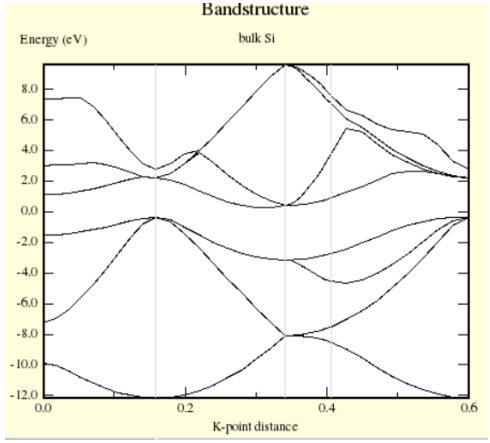
Bandgap (QE) : 0.6 eV

Bandgap (VASP) : 0.65 eV

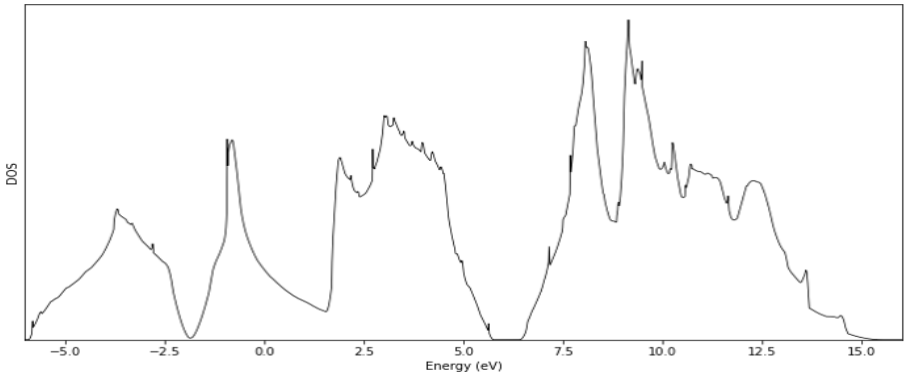
VASP



Density of States (DOS) of Si using GGA (PBE)



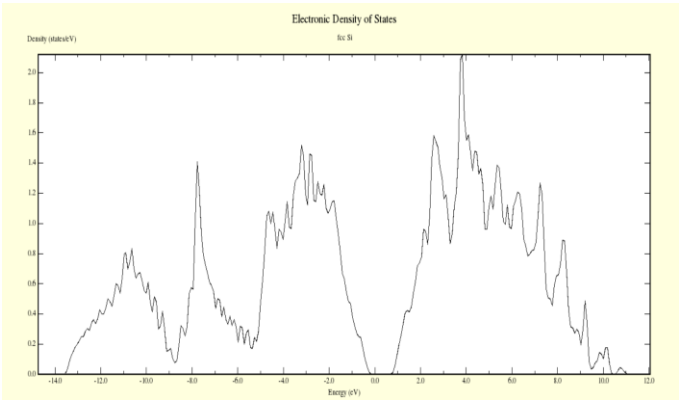
Band structure of Si using GGA (PBE)



DOS of Si using hybrid (HSE)

Bandgap (QE) : 1.21 eV

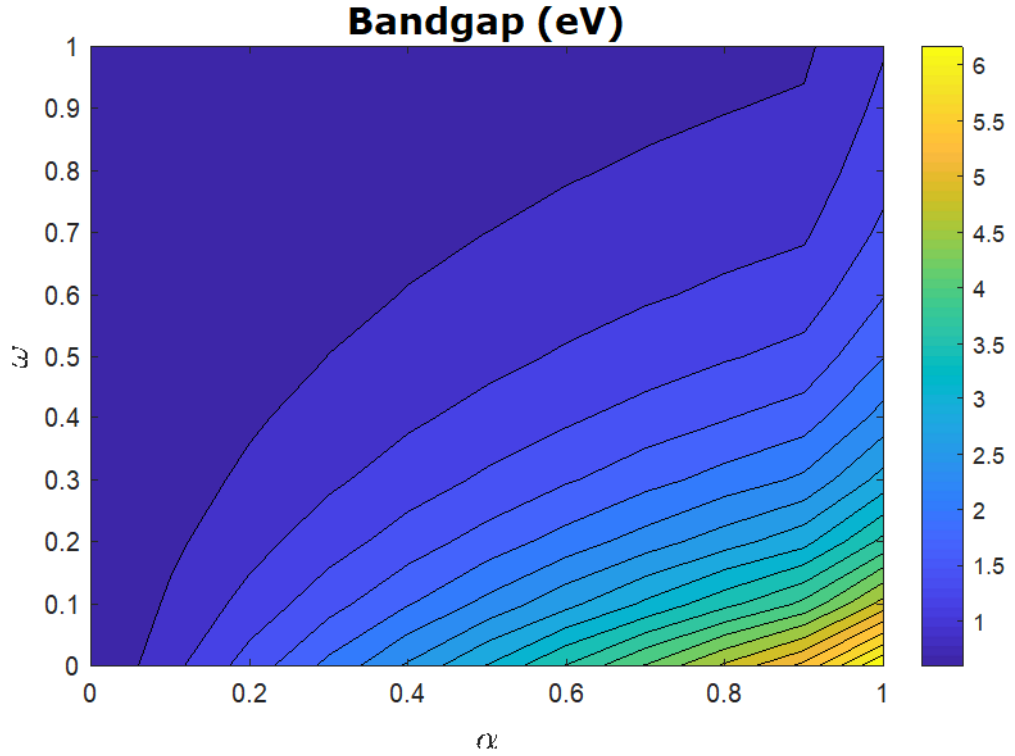
Bandgap (VASP) : 1.1925 eV



DOS of Si using hybrid (HSE)

Thus, we see that HSE calculates the bandgap almost accurately.

Results - HSE



- Motivation to tune these parameters to obtain an accurate HSE functional arises

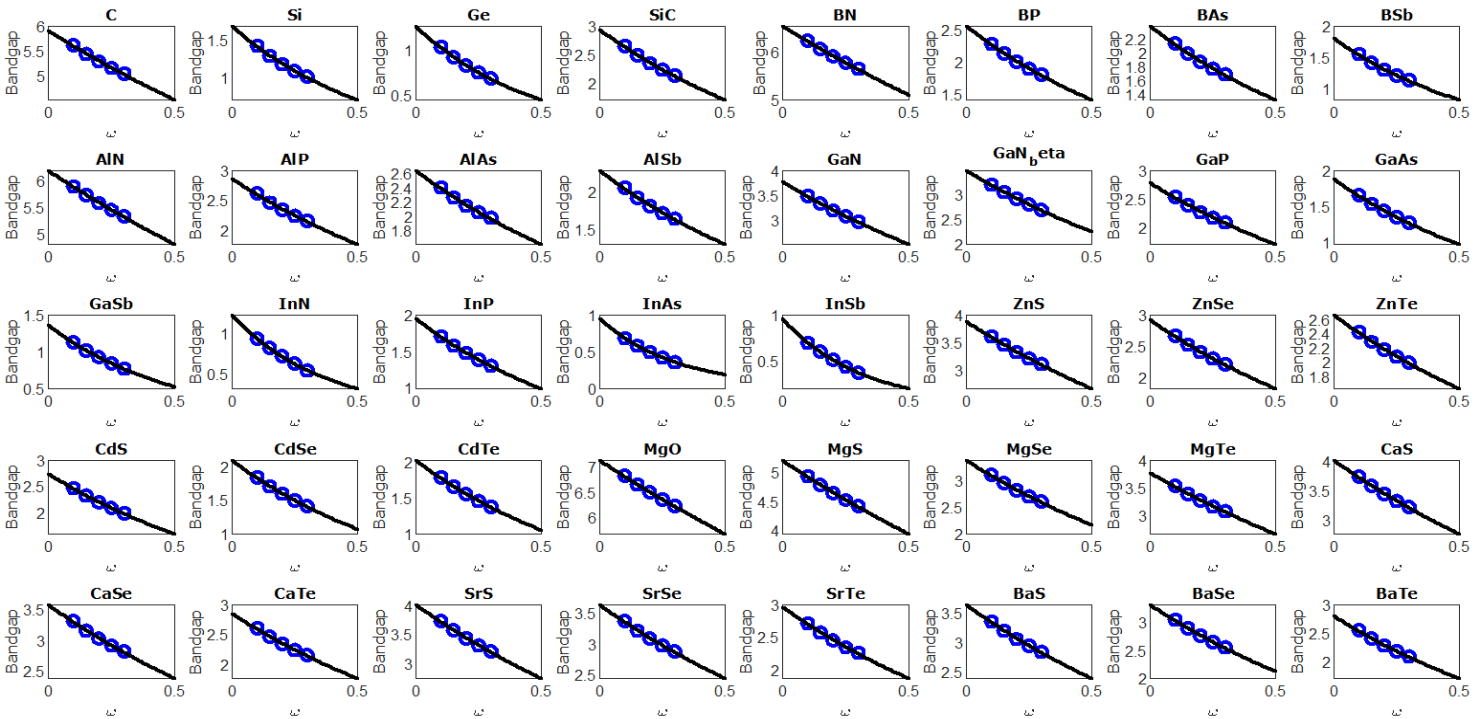
- HSE functional has two parameters:
  $\alpha - fraction\ of\ exact\ (Fock)\ exchange$
  $\omega - range\ separation\ parameter$

- Usually they are fixed;
  α=0.25; from many-body perturbation theory,
  ω=0.20; (used as default in HSE06)

  ω=0 => PBE0 hybrid functional
  $\omega \rightarrow \infty$ => asymptotically tending to PBE

- Unable to predict bandgaps of most materials to a high accuracy

Results - HSE



Materials	row	valence_1	valence_2	Atomic number	Atomic mass	Ionization energy	Packing fraction	Thomas-fermi length	omega_fit
C	2	4	4	6	12.0107	11.26029	0.34	0.38851	0.1415
Si	3	4	4	14	28.0855	8.15168	0.34	0.479347	0.2096
Ge	4	4	4	32	72.64	7.899435	0.34	0.397104	0.2587

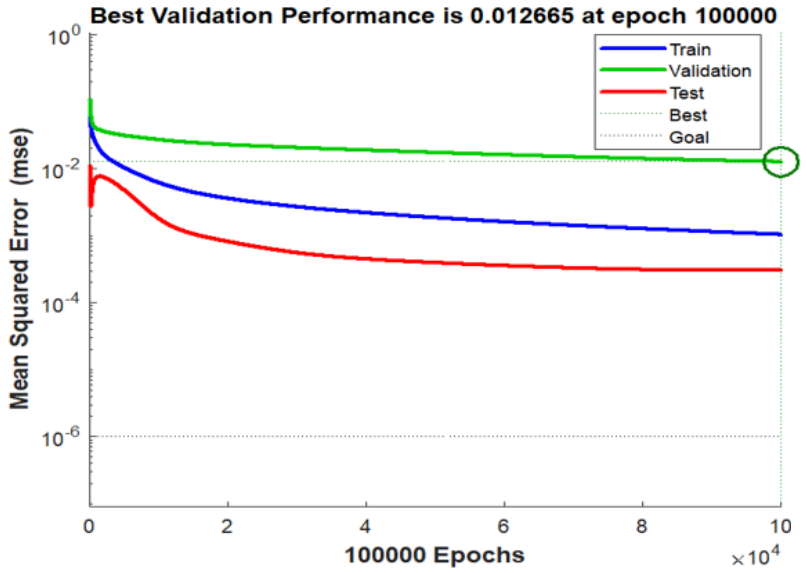
Considered a set of 40 diverse materials; identified a set of properties that dictate the band structure

Performed HSE simulations in VASP in HPC clusters to acquire data and understand patterns in them

ω_{fit} at constant α obtained from HSE calculations by computing bandgaps of each material for 5 'ω' values, then fitting the data using the function $Ae^{-x/\lambda}$; A, λ are the fitting parameters

AQTIVATE

Results – Artificial Neural Network (ANN)



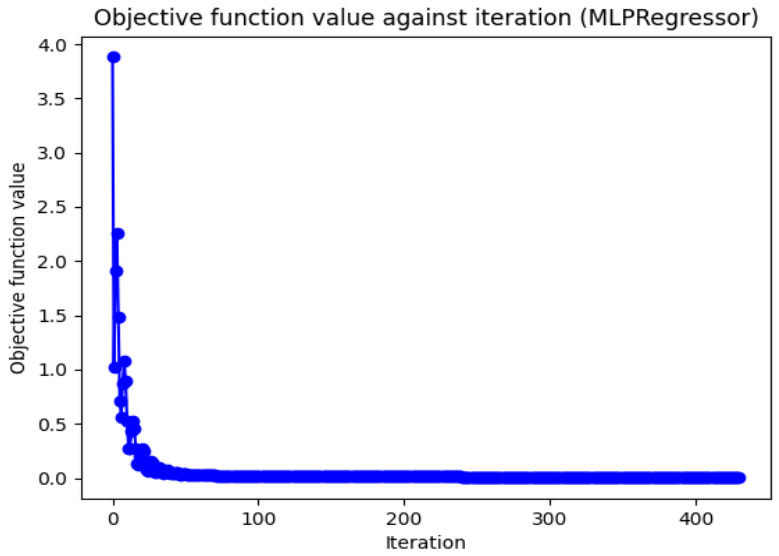
- Multi-Layer Perceptron (MLP) trained with 4 features (left), 8 features (bottom left) and 18 features (bottom right)
- Mean Squared Error (MSE):

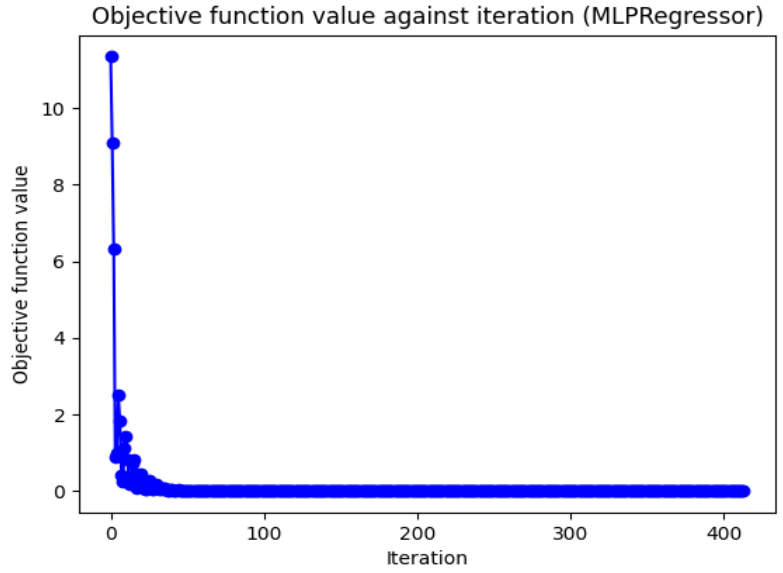
  2×10^{-2} (4 features)

  3.8×10^{-3} (8 features)

  4.3×10^{-4} (18 features)

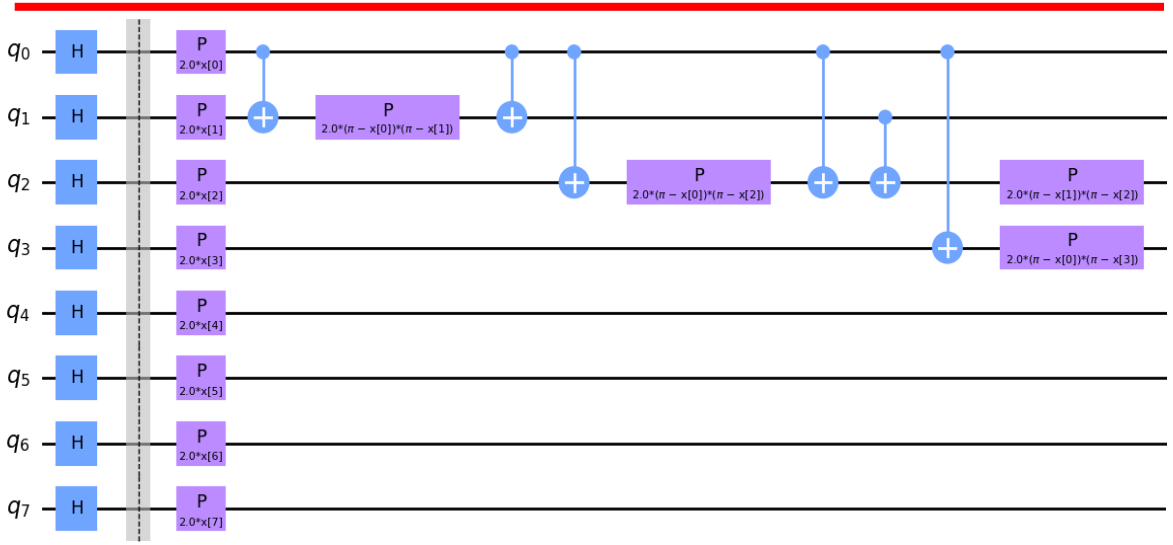
Feed-forward NN trained with multiple hidden layers and with the omega fit values as the output of the NN. The data was divided into three sets for training, validation and testing; gradient converges to a very small value





AQTIVATE

Results – Quantum Neural Network (QNN)



RealAmplitudes Ansatz –

- $\theta = [, \theta_2, \ldots\ldots, \theta_m]$

- alternating sets of RY gates and CNOT gates

$$RY(\theta) = e^{(-\frac{i\theta}{2}Y)} = \begin{pmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix} \qquad CNOT = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

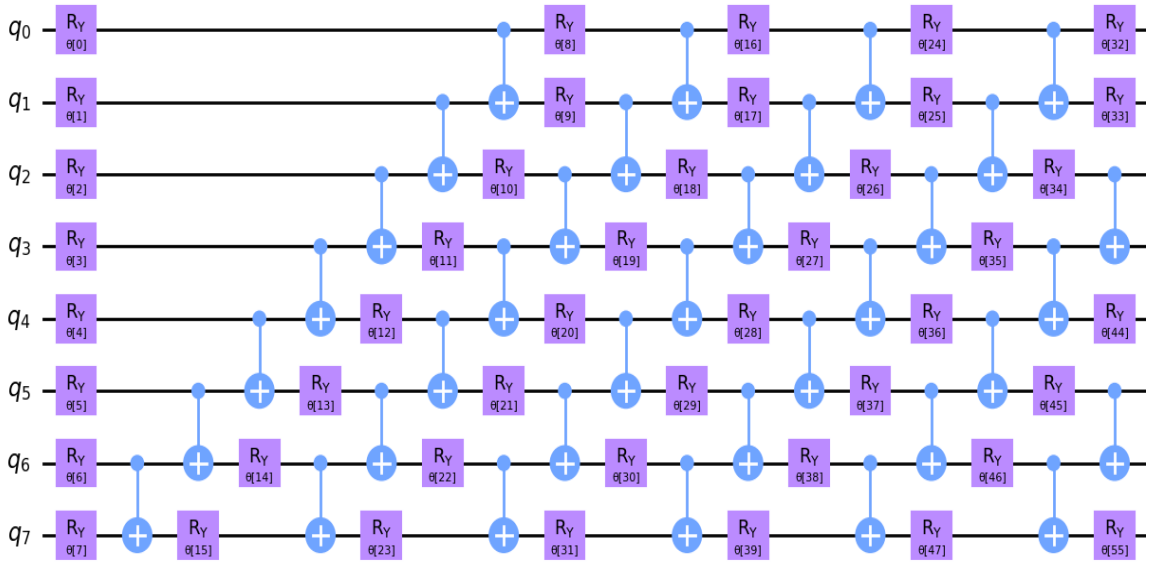
ZZFeature Map –

- second order Pauli-Z evolution circuit

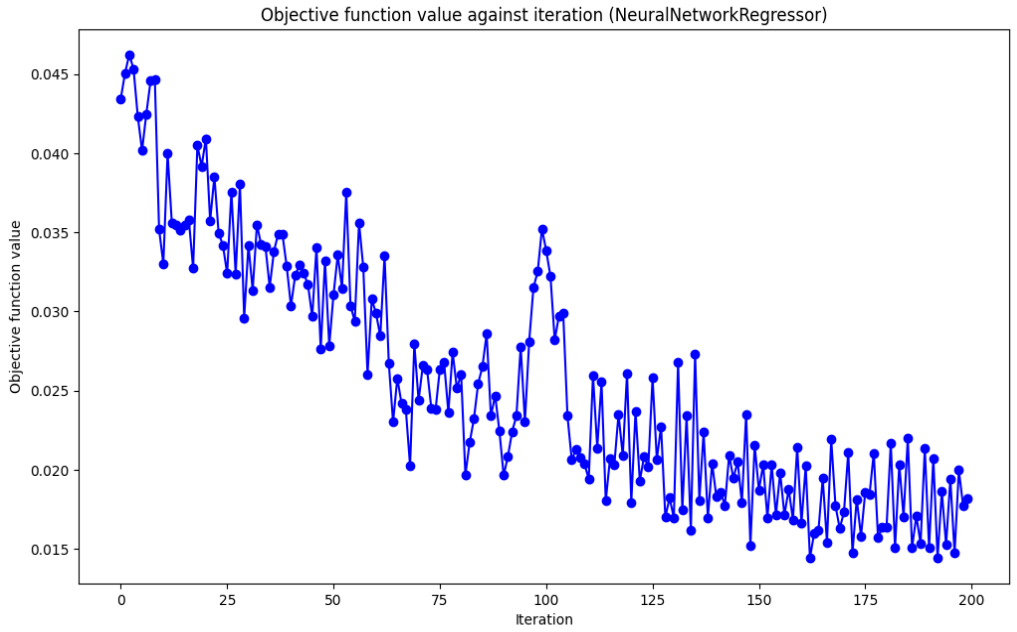
$$U_{\Phi(\vec{x})} = \exp\left(i \sum_{S \subseteq [n]} \phi_S(\vec{x}) \prod_{i \in S} P_i\right)$$

$$\phi_S(\vec{x}) = \begin{cases} x_0 \text{ if } S = \{0\} \\ \prod_{j \in S} (\pi - x_j) \text{ otherwise} \end{cases}$$

- taking $S = \{0\}; P_{0,1} = ZZ$

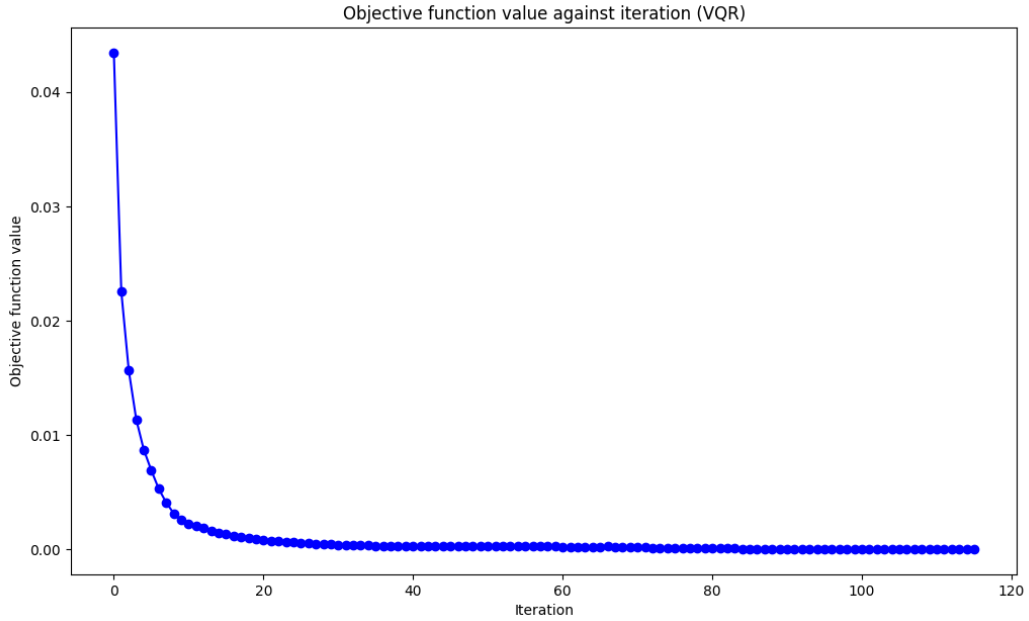


Neural Network Regressor and Variational Quantum Regressor



- arguments taken - user-defined feature map and ansatz, observable to be measured, L2 loss function, L_BFGS_B optimizer and callback graph
- after training for max 200 iterations, convergence is reached before 120 iterations and MSE = 3.3632×10^{-6}

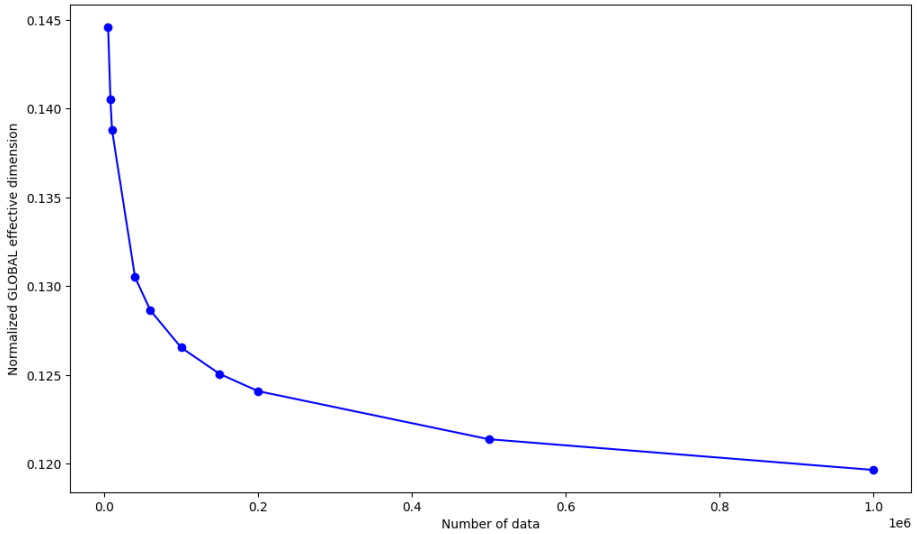
- arguments taken - input values and initial values of trainable parameters, L2 or squared_loss as cost function, COBYLA optimizer and callback graph
- after training for 200 iterations, the Mean Squared Error (MSE) = 1.4409×10^{-2}



AQTIVATE

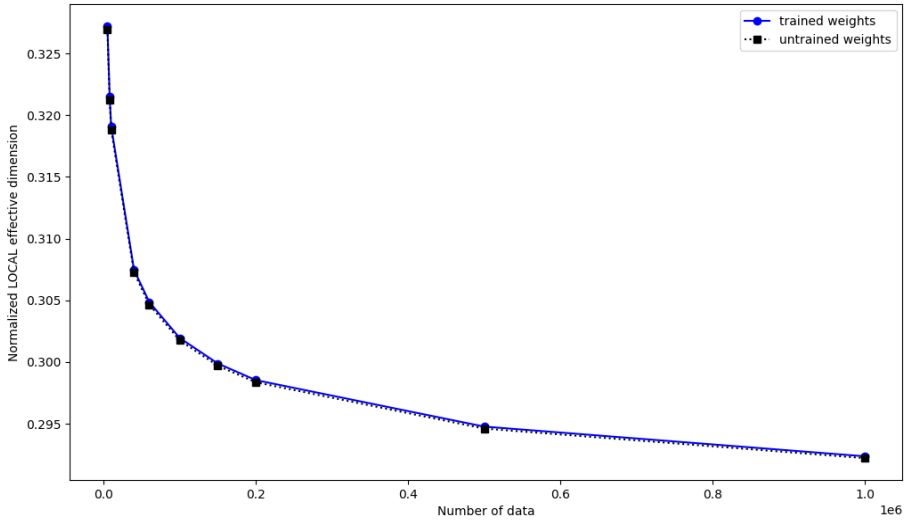
Effective Dimension

Global effective dimension



Global effective dimension is calculated over the whole parameter space while local effective dimension is calculated on a small, finite subset of parameters and depends on the training algorithm.

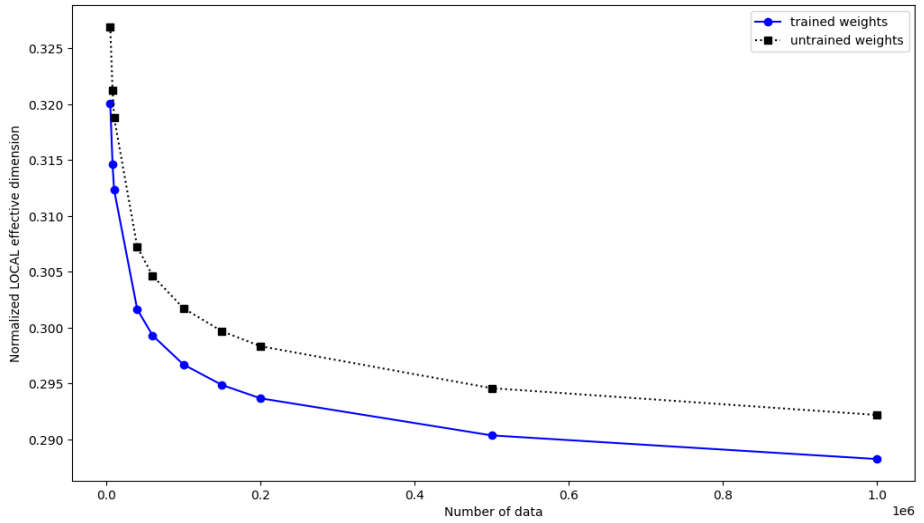
Local effective dimension



- Fisher information matrix – information gained with different parameterizations

$$\tilde{F}_k(\theta) = \frac{1}{k}\sum_{j=1}^{k} \frac{\partial}{\partial\theta}\log p\left(x_j, y_j; \theta\right) \frac{\partial}{\partial\theta}\log p\left(x_j, y_j; \theta\right)^{\mathrm{T}}$$

- Effective dimension – a measure of the space occupied by a parameterized model depending on the size of parameter and sample space.



AQTIVATE

Summary and Future Works

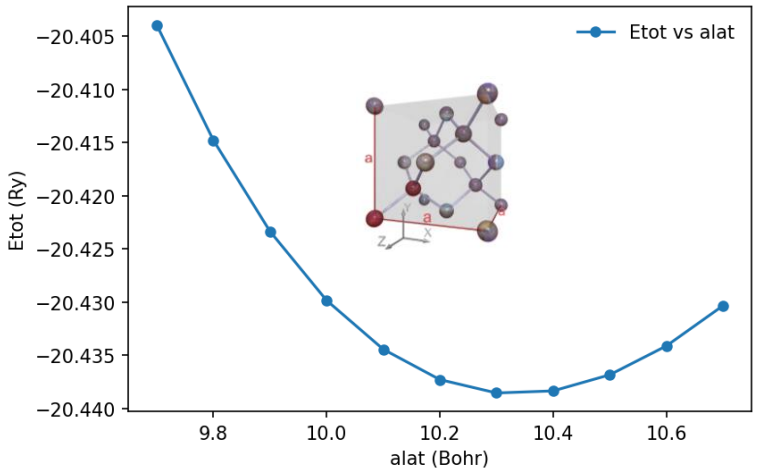
- Executed DFT based calculations in our HPC clusters to acquire the required data to form the neural network for a material database.
- Analyzed the underlying patterns, e.g., the trend in energy band gaps in crystalline materials and compounds.
- Used NN to perform regression on data since they are capable of finding complex patterns and handle multiple inputs
- Used QNN to show that it achieves higher accuracy and better capacity when trained on the same set of data as ANN
- Showed the importance of having a good feature map and ansatz (as in VQR) that can speed up optimization and avoid barren plateaus.
- Might reduce the high number of computations to be performed to obtain the correct HSE parameters particular to each material and compound

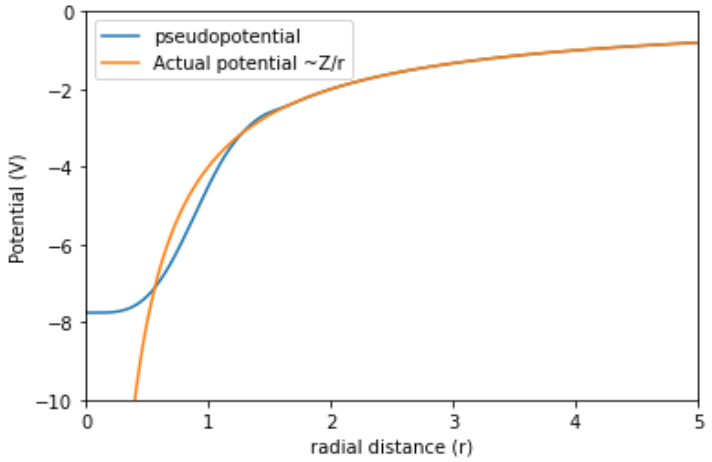
- Different quantum algorithms with relevant understandings to outperform the classical neural networks may be possible, however, that needs to be substantiated
- Inclusion of a wider database of materials with varying crystal structures and properties can be generated with the help of the scripts
- We will also plan to deploy these quantum algorithms on actual quantum hardware

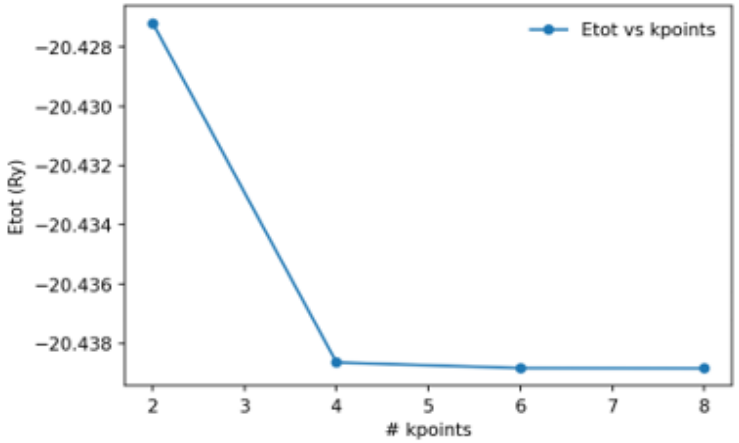
Thank you

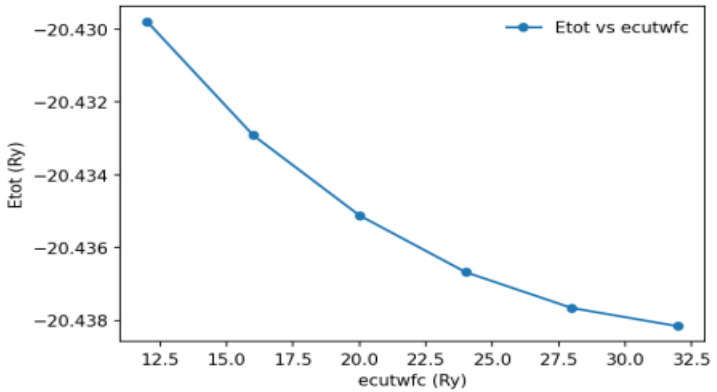
Results - DFT

Calibration of Input Parameters

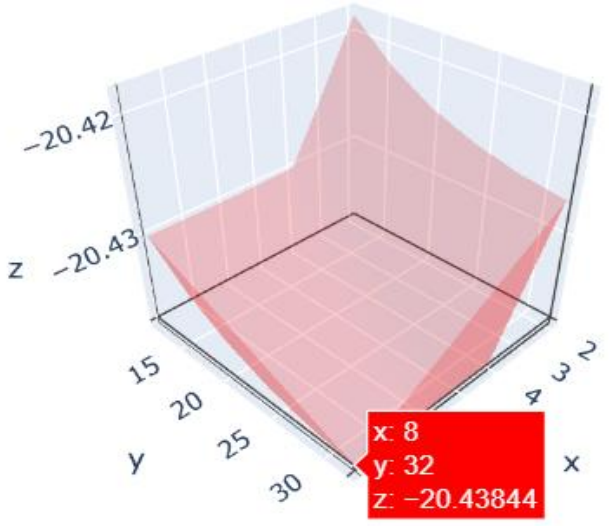








The obtained value of the lattice parameter is indeed optimized and the energy is optimized at a k-point mesh size of 8 × 8 × 8 with a corresponding energy-cutoff value of 32 Bohr which we used as our inputs to perform the electronic structure calculations of bulk Si.



AQTIVATE