# Machine learning generalised DFT+*U* projectors to model polarons in catalyst and battery materials

Amit Chaudhari *[a], Kushagra Agrawal [a], Andrew Logsdail [a]
[a] Cardiff Catalysis Institute, School of Chemistry, Cardiff University, Cardiff CF10 1AT, UK
* chaudharia@cardiff.ac.uk

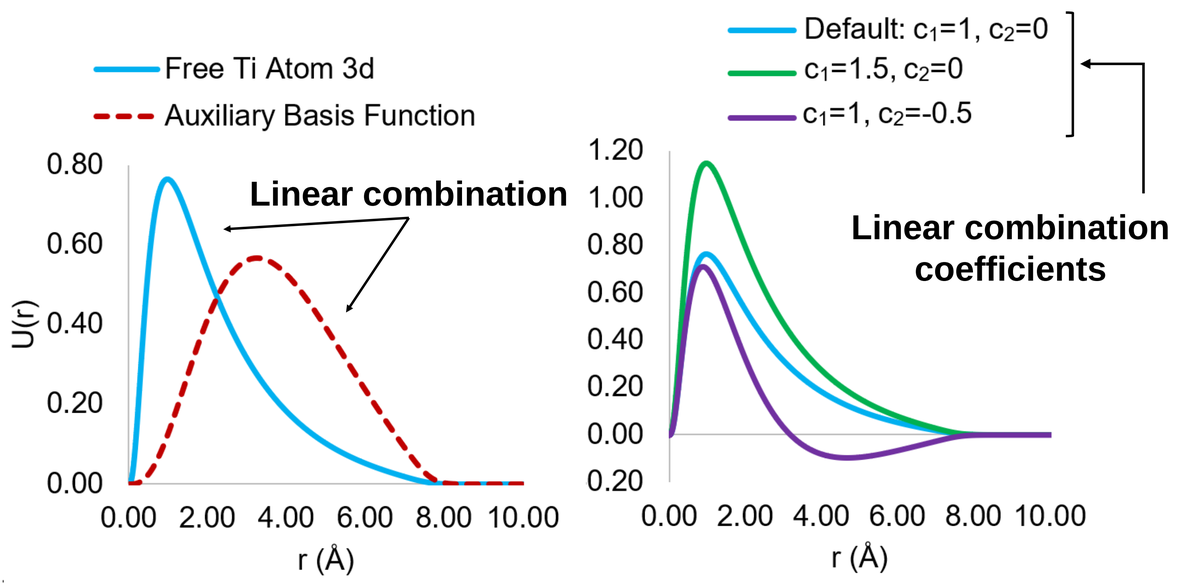
Thomas Young Centre 7th Energy Materials workshop: from data to discovery of new energy materials, July 2024, London

## Simulating Polarons in Strongly Correlated Metal Oxides using Hubbard corrected DFT+*U*

- We need **accurate and computationally affordable beyond-DFT methods** that can correct for the **self-interaction error** in transition metal and rare-earth metal oxides (*i.e.*, correlation between localised *d* or *f* orbital electrons which cannot be accurately modelled using DFT)
- We are using **Hubbard corrected density functional theory (DFT+*U*)** in a **numeric atom-centred orbital framework as implemented in FHI-aims**
- The **default atomic Hubbard projector function prevents the numerically stable simulation** of point defects
- We are using **SISSO (regression)** and **support vector machines (classification)** to optimise both the Hubbard *U* value and projector function enabling the **high-throughput screening of defects in bulk anatase and rutile $TiO_2$**

## Bayesian Optimisation of the Ti 3*d* Hubbard projectors in FHI-aims

**Hubbard projectors are easily tuned in FHI-aims (linear combination of numeric atom-centred orbital basis functions)**



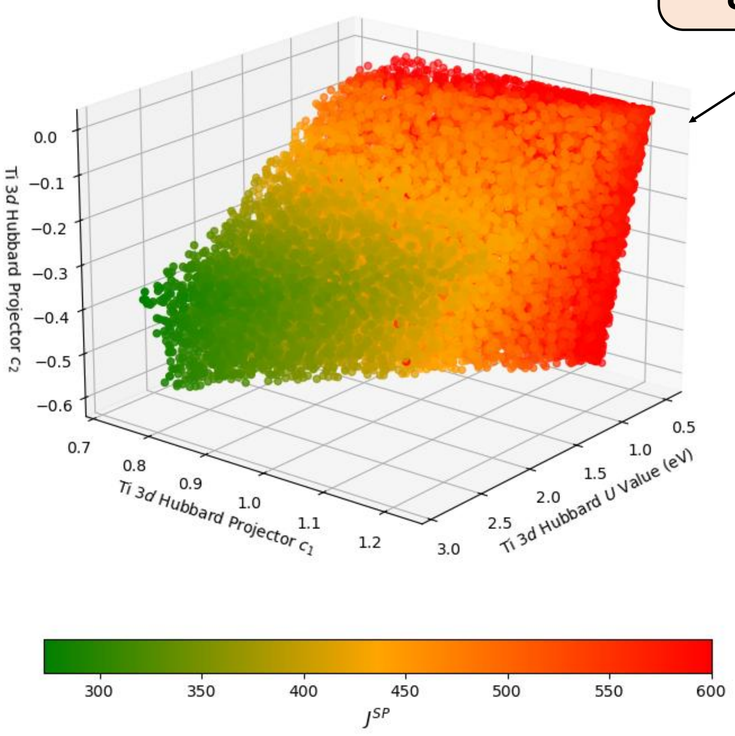

| *U* | $c_1$ | $c_2$ | Optimised Unit Cell | | | | | Optimised Supercell |
|---|---|---|---|---|---|---|---|---|
| | | | $E_{BG}$ (eV) | $v_0$ (Å$^3$) | *Tr[n(3d)]* | *Tr[n(2p)]* | $E_{Total}$ (eV) | Defect Calc. |
| 1 | 1 | 0 | 2.72 | 137.76 | 1.06 | 5.06 | … | Ground State |
| 3 | 1 | 0 | 3.13 | 137.07 | 0.79 | 5.39 | … | Unphysical |
| 3 | 0.9 | -0.5 | 3.06 | 137.79 | 0.99 | 4.68 | … | Charge Sloshing |

**Training Set**

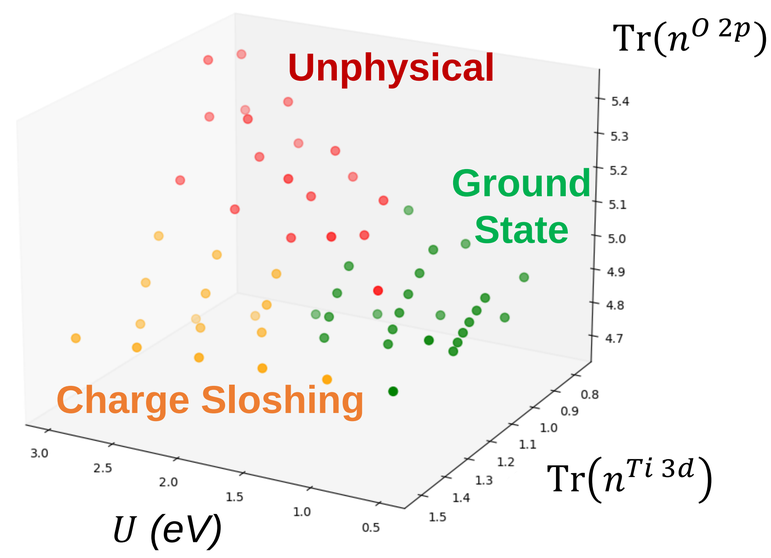
**SISSO Algorithm**
Predict $E_{BG}$, $v_0$, *Tr[n(3d)], Tr[n(2p)]* and $E_{Total}$ in terms of *U*, $c_1$ and $c_2$

**Support Vector Machines**
Classify defect calculation stability/accuracy in terms of *U*, *Tr[n(3d)],Tr[n(2p)]*, $c_1$ or $c_2$

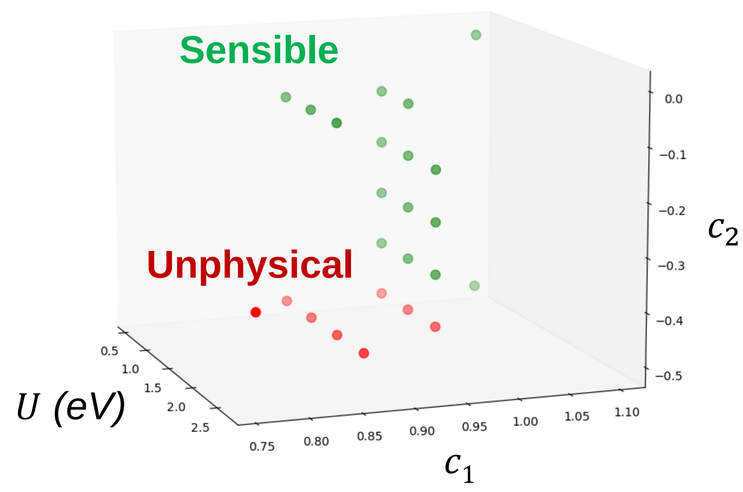
**Bayesian Optimisation for *U*, $c_1$ and $c_2$**




**Numerical instability in SCF optimisation of electron density**




**Unphysical defect energies at predicted ground state**




**Validation**

| $TiO_2$ Polymorph | Projectors | Bulk Defect | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Oxygen Vacancy | Nb | W | Co | Mn | Au | Pd | Pt |
| Anatase | Default | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Rutile | Default | ✗ | ✗ | ✗ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Anatase | ML | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Rutile | ML | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

✗ **DFT+*U* calculation fails in 3-4 SCF cycles**

✓ **Converged DFT+*U* geometry optimisation calculation**

**Acknowledgements**