doped & ShakeNBreak: python packages for solid-state defect calculations and structure-searching

S. R. Kavanagh*[1], I. Mosquera-Lois[1], A. G. Squires[2], A. T. J. Nicolson[2], B. Zhu[3], D. O. Scanlon[2] and A. Walsh[1]

[1]Department, University, City, Country

[2]Department, University, City, Country

[1]Thomas Young Centre and Department of Materials, Imperial College London, Exhibition Road, London SW7 2AZ, U.K.

[2]Department of Chemistry, University of Birmingham, Birmingham, B15 2TT, U.K.

[3]Thomas Young Centre and Department of Chemistry, University College London, 20 Gordon Street, London WC1H 0AJ, U.K.

* Contact: s.kavanagh19@imperial.ac.uk

Computational methods are widely used to predict defect behaviour in solids, before combining and comparing theoretical predictions with experimental measurements. However, there are many critical stages in the computational workflow for defects, which, when performed manually, not only leave room for human error but also consume significant researcher time and effort. Moreover, there are growing efforts to perform high-throughput defect investigations, necessitating robust, user-friendly and efficient software implementing this calculation workflow.[1-4]

Here we report *doped*, our python package for the full generation, calculation setup, post-processing and analysis of defect supercell calculations.[4-7] The generation and thermodynamic analysis (i.e. defect formation energy diagrams, chemical potentials, doping analysis etc.) are agnostic to the underlying first-principles software, while input file generation is supported for several of the most widely-used DFT codes, including VASP, FHI-aims, CP2k, Quantum Espresso and CASTEP. A defect charge state prediction algorithm is implemented, which is shown to significantly outperform previous oxidation-state approaches in terms of both efficiency and completeness. Moreover, *doped* is built to be compatible with other computational toolkits for advanced defect characterisation, including *ShakeNBreak*[8] for defect structure-searching (which will also be briefly introduced), *py-sc-fermi*[9] for in-depth concentration, doping and Fermi level analysis, and *CarrierCapture.jl*[10]*/nonrad*[11] for non-radiative recombination calculations. Its object-oriented python framework make it readily-usable in high-throughput architectures such as *atomate(2)* or *AiiDA*, with examples included in the documentation.

I will discuss the key features of *doped* and *ShakeNBreak* for computational defect workflows. We anticipate that these codes will serve as a highly useful tool for computational defect researchers, being an efficient platform for conducting reproducible calculations of solid-state defect properties.

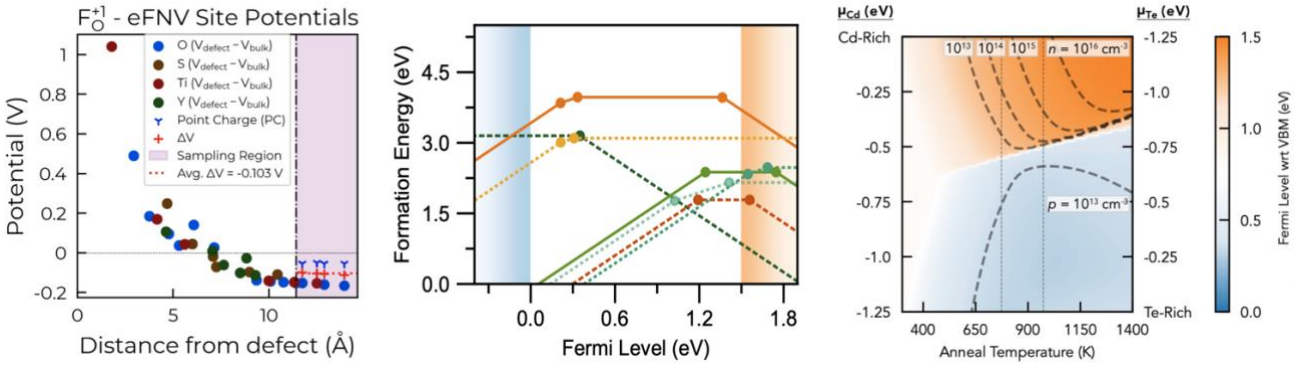

Figure 1. Example outputs from doped: (Left) Automated Kumagai-Oba (eFNV) finite-size charge correction analysis. (Middle) Defect formation energy diagram. (Right) Heatmap of carrier concentration versus temperature and chemical potential (growth conditions).

[1] D. Dahliah, G. Brunin, J. George, V.-A. Ha, G.-M. Rignanese and G. Hautier, High-throughput computational search for high carrier lifetime, defect-tolerant solar absorbers, *Energy Environ. Sci.*, 2021, **14**, 5057–5073.

[2] pymatgen-analysis-defects — pymatgen-analysis-defects, https://materialsproject.github.io/pymatgen-analysis-defects/intro.html, (accessed 18 October 2023).

[3] D. Broberg, K. Bystrom, S. Srivastava, D. Dahliah, B. A. D. Williamson, L. Weston, D. O. Scanlon, G.-M. Rignanese, S. Dwaraknath, J. Varley, K. A. Persson, M. Asta and G. Hautier, High-throughput calculations of charged point defect properties with semi-local density functional theory—performance benchmarks for materials screening applications, *npj Comput Mater*, 2023, **9**, 1–12.

[4] W. D. Neilson and S. T. Murphy, Computational Materials Science, 2022, 210, 111434.

[5] S. R. Kavanagh, A. Walsh and D. O. Scanlon, Rapid Recombination by Cadmium Vacancies in CdTe, *ACS Energy Lett.*, 2021, **6**, 1392–1398.

[6] Y.-T. Huang, S. R. Kavanagh, M. Righetto, M. Rusu, I. Levine, T. Unold, S. J. Zelewski, A. J. Sneyd, K. Zhang, L. Dai, A. J. Britton, J. Ye, J. Julin, M. Napari, Z. Zhang, J. Xiao, M. Laitinen, L. Torrente-Murciano, S. D. Stranks, A. Rao, L. M. Herz, D. O. Scanlon, A. Walsh and R. L. Z. Hoye, Strong absorption and ultrafast localisation in $NaBiS_2$ nanocrystals with slow charge-carrier recombination, *Nat Commun*, 2022, **13**, 4960.

[7] S. R. Kavanagh, D. O. Scanlon, A. Walsh and C. Freysoldt, Impact of metastable defect structures on carrier recombination in solar cells, *Faraday Discuss.*, 2022, **239**, 339–356.

[8] A. Nicolson, S. R. Kavanagh, C. N. Savory, G. W. Watson and D. O. Scanlon, Cu2SiSe3 as a promising solar absorber: harnessing cation dissimilarity to avoid killer antisites, *J. Mater. Chem. A*, 2023, **11**, 14833–14839.

[9] I. Mosquera-Lois, S. R. Kavanagh, A. Walsh and D. O. Scanlon, ShakeNBreak: Navigating the defect configurational landscape, *Journal of Open Source Software*, 2022, **7**, 4817.

[10] A. G. Squires, D. O. Scanlon and B. J. Morgan, py-sc-fermi: self-consistent Fermi energies and defect concentrations from electronic structure calculations, *Journal of Open Source Software*, 2023, **8**, 4962.

[11] S. Kim, S. N. Hood, P. van Gerwen, L. D. Whalley and A. Walsh, CarrierCapture.jl: Anharmonic Carrier Capture, *Journal of Open Source Software*, 2020, **5**, 2102.

[12] M. E. Turiansky, A. Alkauskas, M. Engel, G. Kresse, D. Wickramaratne, J.-X. Shen, C. E. Dreyer and C. G. Van de Walle, Nonrad: Computing nonradiative capture coefficients from first principles, *Computer Physics Communications*, 2021, **267**, 108056.