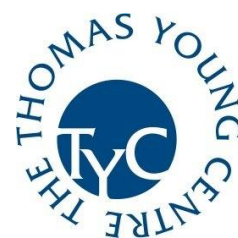# Thomas Young Centre Student Day 2024

Wednesday 22nd May 2024
09.30 to 18.00
Lecture Theatre 2.28, Royal School of Mines, Imperial College London

Time	
09.30 – 10.00	Welcome with tea & coffee
10.00 – 10.05	Opening remarks
	Student presentations
10.05 – 10.20	Understanding the electronic properties and mechanisms of formation of 1D defects observed in 2D MoS_2 *Daria Kieczka, University College London*
10.20 – 10.35	New insights into methane conversion to graphene mesosponge *Qi Zhao, Queen Mary University of London*
10.35 – 10.50	Towards modelling realistic $WS_2/H_2O/SiO_2$ interfaces *Katherine Milton, University College London*
10.50 – 11.05	Ab Initio study of the onset of Al corrosion *Rashid Al-Heidous, Imperial College London*
11.05 – 11.30	**Tea & coffee**
11.30 – 11.45	MDAutoMut: an automated Python library for assessing the effects of mutations on protein dynamics and developing predictive machine learning models *Namir Oues, Brunel University*
11.45 – 12.00	Free energy surfaces and their convergence from sets of asynchronous molecular dynamics simulations subject to multiple biases *Antoniu Bjola, University College London*
12.00 – 12.15	Insights from molecular dynamics and meta dynamics simulations into ligand unbinding kinetics in glycine receptors *Guangpeng Xue, King's College London*
12.15 – 12.30	Anharmonic phonons with Gaussian processes *Keerati Keeratikarn, Imperial College London*
12.30 – 13.30	**Lunch**
13.30 – 13.45	Addressing fermionic complexity: advances in variational Monte Carlo techniques *Massimo Bortone, King's College London*
13.45 – 14.00	Machine learning optimisation and structural dynamics of hybrid halide perovskites *Xia Liang, Imperial College London*
14.00 – 14.15	Many body physics with quantum computers *Araf Haque, King's College London*
14.30 – 15.30	**Posters & refreshments (upstairs in room 301D/301E)**
	Plenary talks
15.30 – 16.05	Multiscale materials simulation for engineering a topological qubit stack: Band offsets at the semiconductor/superconductor interface **Fabiano Corsetti, Microsoft**
16.05 – 16.40	Correlated disorder in functional materials **Prof Andrew Goodwin, Oxford University**
16.40 – 17.00	Prize announcement & closing remarks
17.00 – 18.00	**Reception**

Chairs: Ruiqi Wu, Imperial College London & Aadi Konidena, Kings College London

Talk & poster prizes kindly contributed by

Microsoft Corporation is a global technology company renowned for its pioneering contributions to the computing industry. Since its inception, Microsoft has contributed to a multitude of technological areas. Microsoft creates platforms and tools powered by AI to deliver innovative solutions and is committed to making AI available broadly and doing so responsibly.

Microsoft aims to solve global grand challenges, such as climate change and food insecurity, which requires scientific breakthroughs in chemistry and materials science. Advancements in this space will affect 96 percent of products and 100 percent of humanity. Azure Quantum applications have been designed to help quantum chemists and other scientists rapidly scale and speed up research.

Microsoft continues to shape the future of technology through its groundbreaking innovations and enduring commitment to excellence. Microsoft has generously agreed to sponsor the Thomas Young Centre's Postgraduate Student Day 2024.

Credit: Microsoft.com/en-gb/solutions/quantum-computing/

Final year student presentations

Abstracts

Understanding the electronic properties and mechanisms of formation of 1D defects observed in 2D MoS$_2$

Daria Kieczka, University College London

Defects in 2D semiconductors play an important role in their electronic properties. In 2D materials, defects are often present at high concentrations of 10^{13} cm^{-2} and therefore are an important area of research for understanding the properties of 2D materials. An effective way to observe defects in 2D TMDs such as MoS$_2$ is by using a four-dimensional scanning transmission electron microscope (4D STEM), which allows access to atomic resolution. In this work, we combine ptychography-reconstructed four-dimensional-scanning transmission electron microscopy (4D-STEM) datasets of defects in 2D MoS$_2$ and use density functional theory (DFT) to understand the origin and dynamics of these defects. The datasets show that point defects aggregate into one-dimensional defect lines. We use density functional theory (DFT) with a non-local PBE0-TC-LRC functional in CP2K to calculate the electronic properties of one-dimensional defects, including sulfur vacancy lines, kinks, and rhenium (which substitutes molybdenum) lines. We find that vacancy aggregation results in areas of increased electron density localized at the line defect sites and observe an increase in the number of defect states present when the 'kink' limit is reached. The rhenium dopant lines have potential for high spin states (quartet) which could induce ferromagnetism in the material. We observe a relatively small energy gain from the clustering of sulfur vacancies into lines (of around 0.1 eV). Using nudged elastic band (NEB) calculations with the PBE functional, we calculate the most likely pathways for vacancy diffusion including effects of different charge states on the dynamics of the vacancies. We find that charged di-vacancies have a very low barrier for diffusion which can allow for rapid diffusion as observed with 4D STEM. This work gives insights into the dynamics of defects in 2D MoS$_2$ and their potential for use in doping.

New insights into methane conversion to graphene mesosponge

Qi Zhao, Queen Mary University of London

Graphene mesosponge (GMS) is a new class of mesoporous carbon consisting mainly of single-layer graphene walls. Thanks to its excellent and unique features, GMS outperforms conventional carbon materials: they have high electric conductivity, elastic and flexible nature and unprecedentedly high electrochemical stability, which originates from their three-dimensionally continuous and seamless nanostructure with a large surface area. GMS can be synthesised via chemical vapour deposition (CVD) of methane onto a template of metal oxides such as magnesium oxide and alumina, which catalyses methane conversion at high temperatures. Here, we report a computational investigation of the early stage of CH$_4$ activation toward the porous nanocarbon formation on different metal oxides. We show the role of oxygen vacancies the catalytic process without introducing any transition metal reaction centres used in conventional methane activation catalysts for carbon nanotube growth, oxidative coupling ethylene and partial oxidation to methanol. The

computational results will be presented alongside experimental insights obtained via reaction kinetics using thermogravimetry-mass spectrometry and surface analysis using annular dark-field scanning transmission electron microscopy and in-situ infrared spectroscopy. Understanding the reaction steps controlling the CH_4-CVD process reported in this study will support the rational synthesis of sophisticated mesoporous carbon materials by using tastefully designed templates at lower temperatures via CH_4-CVD chemistry on oxide nanoparticles.

Towards Modelling Realistic WS₂/H₂O/SiO₂ Interfaces

Katherine Milton, University College London

Transition metal dichalcogenides (TMDs) are studied extensively in electronic devices due to their 2D nature, high electron mobility, and semiconducting properties. Despite extensive study, the influence of substrates on TMD film properties remains inadequately understood. Particularly, the interaction between TMDs and SiO_2 substrates, where water layers are prevalent under ambient conditions, presents a complex interface deserving further investigation. Our research aims to elucidate the effects of water thickness on the electronic structure and interfacial properties.

In this work, we use density functional theory (DFT), classical molecular dynamics (MD), and ab initio molecular dynamics (AIMD) to investigate the interface of SiO_2/WS_2 with different thicknesses of interfacial water layers.

Our findings reveal a type I band alignment at the interface, with a calculated band offset between SiO_2 conduction band minimum (CBM) and WS_2 valence band maximum (VBM) of 5.13 eV. This is consistent with previous reports for XPS but shows a larger offset compared to IPE values. Additionally, we observe in-gap water states above the water VBM; the prevalence of these gap states increases with the addition of more water layers. These results underscore the significance of interfacial water layers in modulating the electronic properties of TMD-based devices.

Ab Initio study of the onset of Al corrosion

Rashid Al-Heidous, Imperial College London

Aluminium (Al) is an essential element in several industries such as aeronautics, wastewater treatment, and batteries. All the mentioned industries have limited their use of Al and its alloys, due to pitting corrosion. In this work we investigate the onset of Al corrosion and surface re-oxidation. Using ab initio, realistic, dynamical representation of the Al (111)-water interface, we have compared the morphology of clean and oxidised Al (111) substrates, and revealed a wealth of interfacial features, including interfacial solvation structures, potential drops, mass, angle, and charge distributions. In addition, we have quantified equilibrium coverages and adsorption strength for O and OH species on the metal surface, as a function of the electrode potential; Our findings represent the base for a deeper fundamental understanding of Al corrosion, which will ultimately enable a molecular based engineering of targeted solutions to counteract it.

MDAutoMut: an automated Python library for assessing the effects of mutations on protein dynamics and developing predictive machine learning models

Namir Oues, Brunel University

Molecular dynamics (MD) simulations have become routine tools for studying protein dynamics and function. Advances in methods for determining how mutations affect protein structure have made it possible to gain insights into protein function and inform drug development. However, there is a lack of automated workflows for large-scale assessment of the effects of mutations on conformational changes in proteins.

We present MDAutoMut, an object-oriented Python library designed to automatically run MD simulations, engineer mutations, and assess their impact on protein dynamics. The toolkit integrates with MDSubSampler [1] and MDAnalysis [2] and extends their usability by providing a protocol to use machine learning techniques to predict mutations that lead to desired changes in protein dynamics. MDAutoMut includes two core classes: the Protein class, representing all the information of the protein system, and the Mutation class, representing the mutation engine directly linked to PyRosetta [3]. System preparation and MD simulations are performed in GROMACS via gmxapi. MDAutoMut can run automated workflows to explore a list of single or double-site mutations and perform simulations of the mutated structures.

The workflow estimates changes in relevant properties between the wild-type and mutated structures. Mutations that cause a desired change in protein dynamics can be automatically detected. The software is designed and implemented for three different user demographics. Novice users with limited experience in software development can execute preprepared Python scripts. Advanced users can interact with a Unix-like command line interface. Scientific software developers can benefit from a set of reusable Python classes for ad-hoc customization and integration into existing workflows.

In summary, MDAutoMut provides an automated approach for investigating mutations and their effects on conformational changes in proteins.

References.

[1] N. Oues et al., Bioinformatics (2023), 39, btad427

[2] N. Michaud-Agrawal et al., J. Comput. Chem., (2011), 32, 2319-2327

[3] S. Chaudhury et al., Bioinformatics (2010), 26(5), 689-691

Free energy surfaces and their convergence from sets of asynchronous molecular dynamics simulations subject to multiple biases

Antoniu Bjola, University College London

Molecular dynamics (MD) simulations are a powerful tool to study and predict the thermodynamics of molecular systems. They allow scientists to develop atomistic insights into the behaviour of complex systems and provide a route to obtain average properties of interest. However, molecular systems often contain multiple metastable states, separated by high free energy barriers, rendering the simulation of some

systems unfeasible. To address this issue, numerous methods have been proposed. Two widely used methods are umbrella sampling (US) and metadynamics (MetaD), which introduce external bias potentials that perturb the system's Hamiltonian to overcome the energy barrier and improve the sampling of the molecular system.

We present mean force integration (MFI), a new method to compute the local mean force, which is integrated to obtain the free energy surface (FES). Using MFI, one can obtain the FES from sets of multiple asynchronous simulations subject to various bias potentials, including MetaD, US and/or irregular potentials tailored for specific simulations. Furthermore, the statistical error of the average mean force grants a real-time estimation of the convergence, and the accuracy of the FES is validated via bootstrapping analysis.

These features are exploited through a parallel-serial simulation approach, where several simulations are initiated simultaneously. As the simulations progress, the generated data is analysed in real-time, providing information about the exploration and convergence of the FES. A set of automated rules determines when an active simulation should be terminated and how to initiate the next simulation.

We demonstrate MFI's effectiveness for a range of systems, including analytical Langevin models, the nucleation of a liquid phase from a supersaturated Lennard-Jones vapour, and the crystallisation of a colloidal system via a two-step mechanism. Finally, we show that with MFI, it is possible to systematically improve the sampling and convergence of large-scale biased MD simulations.

Insights from Molecular Dynamics and Meta dynamics Simulations into Ligand Unbinding Kinetics in Glycine Receptors

Guangpeng Xue, King's College London

The glycine receptor, a member of Pentameric Ligand-gated Ion Channels, plays a vital role in inhibitory neurotransmission, reflexes, and neuronal excitability. We built an extra cellular domain (ECD) model for the glycine receptor from cryo-EM data, concentrating on ligand unbinding processes within the ECD domain. The binding or unbinding of ligands initiates a series of conformational re-arrangements that culminate in the opening or closing of the ion channel in the cell membrane. We investigated binding site dynamics using Molecular Dynamics simulations, using models with different ligand occupancy. The goal of this study is to evaluate the potential cooperativity between binding pockets, as experiments suggest that two or three bound ligands are sufficient to activate the receptor. We found that the motions of neighbouring pockets have an impact on binding pocket due to the flexibility of loop C which forms the pocket with other loops.

Our use of the enhanced sampling method Metadynamics to analyse ligand unbinding kinetics showed that ligand ejection requires the pocket to open, pointing to a reverse binding mechanism. Our simulations shed light on the mechanisms of ligand unbinding providing details at the atomic scale. Accurate force parameters are crucial for assessing the dynamical features of an event. The parameterization of ligand in zwitterionic state was refined by comparing the dihedral angle energy with quantum calculations to enhance the accuracy of the kinetics results. An enhanced

protocol was established, which can be applied to assign accurate charges to atoms and forces for dihedral angles to any standard or non-standard residues.

Anharmonic phonons with Gaussian processes

Keerati Keeratikarn, Imperial College London

The finite temperature properties of matter require understanding thermal motion. In crystals this can be described by representing the collective excitation of the centre of mass motion as a phonon. Phonon Energies in the harmonic approximation come from the second order force constants of the potential energy surface (PES). The standard approach is to use a finite displacement method (FDM). Anharmonic contributions (required for finite thermal conductivity) require higher order force constants. FDM-based calculations scale poorly both with the size of the system, and the order of the force constants. As an alternative approach, we can use a more sophisticated surrogate method of describing the energy surface than using a Taylor expansion fitted to finite displacements. Gaussian processes (GPs) are a machine learning method which can describe an arbitrary function. GPs are naturally Bayesian (probabilistic). Our reference data is the electronic structure from which conditions the model (training). Due to the underlying Gaussian form of a GP, the model is infinitely differentiable. This allows the model to be trained directly on forces (the derivative of PESs), reducing the calculations required for a given accuracy of PES evaluation. This differentiation can be extended to compute the second and the third derivative of PESs, FCs, by using automatic differentiation (AD). By performing linear operations between arbitrary derivative orders of the GP, the covariance functions among PESs, forces and those FCs can be achieved and are in close agreement with more standard approaches of FDM (in Phonopy and Phono3py).

Addressing Fermionic Complexity: Advances in Variational Monte Carlo Techniques

Massimo Bortone, King's College London

Modelling the ab initio electronic structure wavefunction of molecular systems presents significant challenges, primarily due to the exponential scaling of computational complexity and the need for high accuracy in capturing electron correlations.

Variational Monte Carlo (VMC) has emerged as a powerful method to address these challenges by using a stochastic approach to approximate the ground state wavefunction with a parametrized model optimized via gradient descent methods.

Crucially, the effectiveness of VMC relies on the expressiveness of the model, the efficiency of the sampling process and the robustness of the optimization scheme.

Over the years, physics-inspired models with limited number of parameters have been increasingly replaced by complex, machine learning-inspired models, featuring potentially hundreds of thousands of parameters.

These models have proven transformative in studying quantum spin systems, unlocking previously inaccessible ground states, and revealing new phases in frustrated quantum magnets.

However, similar breakthroughs in the study of the electronic structure of molecular systems with VMC remain elusive.

In this talk, I will explore the unique problems encountered in modelling fermionic systems, such as the antisymmetry requirement, the sampling efficiency, and the scaling of the local energy calculation, while also highlighting recent advancements such as autoregressive and backflow models.

Lastly, I will discuss promising future opportunities provided by integrating VMC with quantum embedding methods, such as Density Matrix Embedding Theory, which could offer a potential bridge to tackle larger molecular systems more effectively.

Machine Learning Optimisation and Structural Dynamics of Hybrid Halide Perovskites

Xia Liang, Imperial College London

This research aims to advance the efficiency and stability of perovskite photovoltaic (PV) devices through an in-depth exploration of the structural effects of their constituent components and the underlying phase transition processes. Molecular dynamics (MD) simulations stand out as a suitable method for probing these behaviours at the atomic level, while density functional theory (DFT) provides a high-accuracy approach for modelling the atomic structure of materials. However, the computational demands of DFT calculations impose limitations, particularly for large-scale and long-term simulations, especially in systems with slow rotational dynamics of organic molecular cations. To overcome these challenges, one promising avenue involves leveraging machine learning potentials (MLP). These MLPs offer a means to approximate the energy landscape, enabling more efficient simulations without sacrificing accuracy. A significant aspect of this research lies in the development and implementation of a real-space structural analysis method. This method offers a direct and intuitive means to comprehend the dynamic structure of perovskites, aiding in the interpretation of simulation results and real-space experimental measurements. Central to this analysis framework is an octahedral structure matching kernel, which effectively dissects the effects of octahedral tilting and distortion, providing a compact and exact descriptor of their status within the dynamic structure. Moreover, the method includes the quantification of organic molecule orientations, lattice spacing, and a series of correlation functions to assess their temporal and spatial correlations. Initial simulations focus on single-composition and mixed halide perovskite structures, successfully replicating experimentally identified phases. These simulations reveal additional symmetry-breaking effects near phase transitions, particularly in the octahedral tilting of halide perovskites with organic A sites. Looking ahead, the aim is to extend this workflow to more complex perovskite material systems featuring mixed chemical species in both A and X sites.

Many Body Physics with Quantum Computers

Araf Haque, King's College London

Contemporary research into strongly correlated materials relies on quantum embedding techniques to resolve emergent phenomena beyond the capabilities of mean field methods alone. Quantum embedding involves optimizing a local impurity within a strongly correlated system while considering its interaction with the surrounding environment. Energy-weighted Density Matrix Embedding Theory (EwDMET) has emerged as a method to systematically control the resolution of fragment-environment coupling, enabling the self-consistent optimization of quantum fluctuations within a static framework. However, for large systems, the exponential growth of the many-particle Hilbert space presents a challenge when considering any embedding method. Quantum computers offer a promising avenue for to deal with such large Hilbert spaces, given that the number electronic orbitals have a one-to-one mapping with the number of qubits. In this study, we leverage two variational quantum algorithms to calculate the energy-weighted density matrix, referred to as the 'moment'. These calculations will facilitate the calculation of potential energy surfaces, self-energies, and spectral functions of Hubbard models and abinitio systems.

Plenary Speakers

Multiscale materials simulation for engineering a topological qubit stack: Band offsets at the semiconductor/superconductor interface

Dr Fabiano Corsetti, Microsoft

The realisation of a topological qubit device for quantum computation requires an exceptional level of understanding and control of the underlying material platform. Materials modelling plays a key role in the design of the devices, with different levels of theory being able to access different properties. In this context, we discuss the challenge of determining the band offset at the interface between the semiconductor and metal in semiconductor/superconductor heterostructure devices, an important parameter for controlling the topological phase. We show how first principles calculations can be used to accurately predict the band offset, and how this parameter then feeds into larger-scale models.

Correlated Disorder in Functional Materials

Professor Andrew Goodwin, Oxford University

All materials are disordered at finite temperatures. Sometimes this disorder is random; more frequently it's not. This talk will explore some key examples where correlated (non-random) disorder is crucial for material function. Such systems pose a number of important and interesting challenges for experiment, computation, and theory alike, and the talk will also cover some of the open questions in the field.

Poster presentations

Poster Number	Name	Poster Title
1	Daria Kieczka	Understanding the electronic properties and mechanisms of formation of 1D defects observed in 2D MoS2
2	Namir Oues	MDAutoMut: an automated Python library for assessing the effects of mutations on protein dynamics and developing predictive machine learning models
3	Katherine Milton	Towards Modelling Realistic WS2/H2O/SiO2 Interfaces
4	Antoniu Bjola	Free energy surfaces and their convergence from sets of asynchronous molecular dynamics simulations subject to multiple biases
5	Massimo Bortone	Addressing Fermionic Complexity: Advances in Variational Monte Carlo Techniques
6	Rashid Al-Heidous	Ab Initio study of the onset of Al corrosion
7	Guangpeng Xue	Insights from Molecular Dynamics and Metadynamics Simulations into Ligand Unbinding Kinetics in Glycine Receptors
8	Qi Zhao	New insights into methane conversion to graphene mesosponge
9	Huanyu Zhou	Ab initio Simulation of Molecular Crystal Regrowth from Solution
10	Niamh Smith	Al-O-Al defect complexes as possible candidates for channel electron mobility reducing trapping centres in 4H-SiC MOSFETs
11	Wan Tong Lou	Neural Wave Functions for Superfluid's.
12	Margaret-Ann Withington	Molecular Dynamics Simulations of Liquids: Thermophysical Properties
13	Christopher Tat Shun Cheung	Charge density waves and moiré reconstruction in twisted NbSe2/NbSe2 bilayers
14	Wenyi Feng	Atomistic molecular dynamics simulations of modified chitosan polymer GCPQ suggest self- assembly into a nano-sized gel matrix
15	Lauren Cook	Simulating Nonadiabatic Dynamics Using the Meyer-Miller-Stock-Thoss Hamiltonian: A Comparison of Algorithms
16	Kit Joll	Machine learning the electric field response of condensed phase systems using perturbed neural network potentials
17	Wei Dai	Cheminformatics and Machine Learning approaches for GPCR Computer-Aided Drug Design
18	Aaron Dines	Stabilized coupled trajectory mixed quantum–classical algorithm with improved energy conservation: CTMQC-EDI
19	Shambhu Bhandari Sharma	Ab-intio melting curve of aluminium and Al-Mg alloy
20	Md Maruf Mridha	Structural and electronic properties of g-C3N4
21	Michael Ingham	Point charge embedding approaches for simulating excited states in MOFs
22	Hannah Turney	Software toolkits for in silico screening of polymer excipients used in small molecule formulation and drug delivery.
23	Yu-Yuan (Stuart) Yang	Deep learning for binding site segmentation in protein ensembles
24	András Pethő	Investigating conductive properties of single-heme cytochrome

25	Gowan Whalley	Data-Driven Design of Materials for CO_2 Sequestration
26	Marcus Allen	A novel 'moment'-based GW algorithm for molecules and solid
27	Kanyarat Rueangboon	Understanding Organic Cocrystals for Room Temperature Phosphorescence
28	Shirui Wang	Computational insights into emerging chalcogenide perovskite photovoltaics
29	Isaac Mackley	Effect of Nitrogen Incorporation on Charge Trapping in Amorphous HfO_2
30	Yao Wei	Rational Design of Molybdenum Transition-Metal sub nanoclusters catalysts with Particle Swarm Optimization
31	Stefani Alexandria Setiono	MOFs for water diffusion
32	Ka Fu Man	The Effect of Missense Mutations on the Binding Pocket Dynamics of Skeletal Myosin
33	Wei Dai	Cheminformatics and Machine Learning approaches for GPCR Computer-Aided Drug Design
34	Aadi Konidena	Non-Equilibrium Greens Functions and Quantum Transport
35	Henry Tomlins	Transport beyond the Semi-Classical Regime and Scaling Breakdown in Band-Tuned Metal-Insulator Transitions.
36	Delchere DON-TSA	Predictive Models for Inorganic Materials Thermoelectric Properties with Machine Learning
37	Felix Mildner	K-Resolved Photoemission Modelling - Aiding and Guiding Photocathode Research for Next-Generation Accelerators
38	Ruiqi Wu	Tilt-induced charge localisation in phosphide antiperovskite photovoltaics
39	Xia Liang	Machine Learning Optimisation and Structural Dynamics of Hybrid Halide Perovskites
40	Araf Haque	Many Body Physics with Quantum Computers
41	Keerati Keeratikam	Anharmonic phonons with Gaussian processes
42	Chaiyawat Kaewmeechai	Investigation of degradation mechanism in crystalline and amorphous Ga_2O_3
43	Teo Cobos	Molecular modelling degradation in modern microelectronic devices.
44	Akshayini Muthu	Electrocatalytic CO_2 Reduction on Amorphous Copper: A Combined Machine Learning Forcefield and DFT Modelling Approach
45	Cambrin Kemble-Diaz	Investigating Epicatechin's Cardiovascular Benefits: Computational and Experimental Insights into FAK Interactions
46	Fei Gao	Molecular Level Thermodynamics and Properties of the Aqueous Hematite Interface from First Principles