

## Welcome

Dear Participants,

The Organising Committee are delighted to welcome you to the University of Bath's 7<sup>th</sup> Annual HPC Symposium.

This year, we have four outstanding speakers, Dr Timothy Lanfear (NVIDIA), Prof Giovanni Montana (University of Warwick), our own Prof Tina Düren (Chemical Engineering) and Jess Jones (Cray UK Ltd). Timothy Lanfear will talk about how NVIDIA's technologies can help to deliver Data Centre infrastructure for the AI revolution that is underway, Giovanni Montana will talk about his work applying Machine Learning to interpret chest X-rays, while Tina Düren will talk about how her group uses Balena to study the properties of porous materials for a range of applications. Former Bath student, Jess Jones, will also deliver a presentation on the conception and installation of the current MetOffice Supercomputer by Cray.

We have a host of contributions from researchers across the University, delivered as talks and quick-fire introductions to posters which can be viewed during the breaks (watch out for this year's interactive poster!). The line-up includes contributions from Biology and Biochemistry, Chemistry, Computer Science, Mathematical Sciences, Physics, Chemical, Electrical and Mechanical Engineering, and covers topics including machine learning, software development, applications of quantum and molecular simulations and high-level finite-element methods to solve topical research problems. This year's symposium has a specific Machine Learning theme and a session showcasing research at Bath will follow the first two keynote speakers.

The University's second-generation HPC cluster, Balena, has been in service for over three years now, is seeing a consistently high utilisation across the University from a growing number of users and research groups. It has recently received an upgrade with 17 new compute and development nodes with the latest generation Intel Skylake processors, new GPU cards, each of which have half the theoretical compute power of Balena's predecessor, and increased storage. Additionally, Isambard, the GW4 Tier-2 HPC service is due to be delivered soon to undergo acceptance testing. Once in production Bath researchers will be able to gain access through a lightweight Resource Allocation Process. Alongside this the University has also invested in a central Research Software Engineer to support users to make the best use of Isambard, Balena, as well as provide training in software development, profiling and benchmarking their codes/calculations.

The Annual HPC Symposium showcases the breadth of research being carried out by the University's HPC community, and has evolved over its last six iterations into a forum for discussion and new collaborations.

We look forward to joining you for what promises to be an interesting and enjoyable event.

Yours faithfully,

The Organising Committee

Gaël Donval, Chemical Engineering  
James Grant, Computing Services  
Jonathan Skelton, Chemistry

## Programme

HPC Symposium, Wednesday 6<sup>th</sup> June 2018, in 3WN2.1

09:30 [Registration and refreshments](#)

10:00 [Welcome](#)

*Prof Steve Parker, Chair of the HPC Management Group*

10:15 [Session One](#)

*Chair: Prof Jonathan Dawes, Director Institute of Mathematical Innovation*

10:15 [Keynote 1: The GPU Data Centre for AI](#)

*Dr Timothy Lanfear, NVIDIA*

10:50 [Keynote 2: An AI system for radiological reporting and prioritisation](#)

*Prof Giovanni Montana, University Of Warwick*

11:15 [Delegation photograph](#)

11:20 [Break and poster viewing – 3WN Foyer](#)

11:45 [Session two](#)

*Chair: Dr James Hook, Mathematical Sciences*

11:45 [Automatic Recognition and Classification of Hip Fractures](#)

*Dr Ellen Murphy, Institute for Mathematical Innovation*

12:00 [Distributed Deep Reinforcement Learning for Autonomous Control](#)

*Gordon Rennie, Computer Science*

12:15 [Supercharging the materials discovery process with machine learning](#)

*Daniel Davies, Chemistry*

12:30 [Session three](#)

*Chair: Dr Michael Carley, Mechanical Engineering*

12:30 [Quick fire 3-minute poster talks \(see page 4\)](#)

13:00 [Lunch – 3WN Foyer](#)

14:00 [Session four](#)

*Chair: Dr Katharine Fraser, Mechanical Engineering*

14:00 [Keynote 3: Small holes - big impact: A molecular view of adsorption](#)

*Prof Tina Düren, Chemical Engineering*

14:30 [Molecular dynamics simulations of biological membranes and membrane proteins](#)

*Vitoria Oakes, Chemistry*

14:45 [Crystal-Size Framework Flexibility in MOFs](#)

*Megan Thompson, Chemical Engineering*

- 15:00 Molecular Modelling of Hydrogen Storage in Cellulose  
*Megan Stalker, Chemistry*
- 15:15 Using MB (Mega Bytes) to visually analyse MB (Mega Bases)  
*Catherine-Axa Wilkins, Biology and Biochemistry*
- 15:30 Break and poster viewing – 3WN Foyer
- 16:00 Session four  
Chair: *Dr Russell Bradford, Computer Science*
- 16:00 A Parallel Fast Multipole Method Implementation for Electrostatic Interactions  
*William Saunders, Mathematical Sciences*
- 16:15 Parallel implementation of the Time-Evolving Block Decimation algorithm  
*Paul Secular, Physics*
- 16:30 Computational methods for locating phase transitions in solids and fluids  
*Tom Underwood, Physics*
- 16:45 Keynote 4: Tales from Supercomputing  
*Jess Jones, Cray UK Ltd*
- 17:00 Presentation of prizes
- 17:10 Pizza reception in 3WN Foyer
- 18:00 Close of meeting

## Posters

Dedicated session for researchers to give a three-minute introduction to themselves and their poster.

1. SLATE-based solvers for semi-implicit hybridised DG method in fluid dynamics  
*Jack Betteridge, Mathematical Sciences*
2. Streamlining the Hydrodynamic Modelling of OceanWave Energy Converters  
*J A Bridgwater Court, Mechanical Engineering*
3. Modelling the design spaces of construction kits  
*J. A. Gopsill, Mechanical Engineering*
4. The importance of the membrane in mediating the mechanism of Monoamine Oxidase B  
*Hannah Jones, Biology and Biochemistry*
5. Force Field Validation for Light Gas Adsorption in Metal-Organic Frameworks  
*Mike Kallo, Chemical Engineering*
6. pylj: an open-source python library for teaching molecular simulation  
*Andrew R. McCluskey, Chemistry*
7. Modelling Proteins as a Mechanical Network  
*Thomas McManus, Physics*
8. Hazardous Compounds at the Soil-Water Interface  
*Leyorla Ohene-Yeboah, Chemistry*
9. Computational Modelling of Defects in Battery Materials  
*Alex Squires, Chemistry*
10. How well do stratospheric reanalyses reproduce satellite temperature data?  
*Corwin Wright, Electrical Engineering*

## List of attendees

First Name	Surname	Department / Organisation
Stefano	Angioni	Computing Services
Eliot	Ayache	Physics
Jonathan	Barnard	Chemistry
Jack	Betteridge	Mathematical Sciences
Russell	Bradford	Computer Science
James	Bridgwater Court	Mechanical Engineering
Michael	Carley	Mechanical Engineering
James	Cave	Physics
Steven	Chapman	Computing Services
Antony	Cleave	ClusterVision
Matthew	Cook	Physics
Andrew	Cookson	Mechanical Engineering
Daniel	Davies	Chemistry
Jonathan	Dawes	Institute for Mathematical Innovation
Jess	Dillon	Chemistry
Sergey	Dolgov	Mathematical Sciences
Carmen	Domene	Chemistry
Gaël	Donval	Chemical Engineering
Tina	Düren	Chemical Engineering
Anastasia	Efthymiadou	Physics
Joseph	Flitcroft	Chemistry
Katharine	Fraser	Mechanical Engineering
Dibyajyoti	Ghosh	Physics
Kiran	Gill	Chemical Engineering
Richie	Gill	Mechanical Engineering
Edward	Gokmen	Computer Science
James	Gopsill	Mechanical Engineering
James	Grant	Computing Services
Surani	Gunasekera	Physics
Claire	Hobday	Chemical Engineering
James	Hook	Mathematical Sciences
Roger	Jardine	Computing Services
Hannah	Jones	Biology and Biochemistry
Jess	Jones	Cray UK Ltd
Mike	Kallo	Chemical Engineering
Surendra	Kaushik	Independent
Nathaniel	Kelly	Mechanical Engineering
Timothy	Lanfear	NVIDIA
Matthew	Lennox	Chemical Engineering

First Name	Surname	Department / Organisation
Ivi	Martins de Carvalho	Chemical Engineering
Roshan	Mathew	Computing Services
Andrew	McCluskey	Chemistry
Thomas	McManus	Physics
Federica Angela	Mevoli	Architecture and Civil Engineering
Giovanni	Montana	University of Warwick
Eike	Mueller	Mathematical Sciences
Ellen	Murphy	Institute for Mathematical Innovation
Victoria	Oakes	Chemistry
Leyorla	Ohene-Yeboah	Chemistry
Anuradha	Pallipurath	Chemistry
Steve	Parker	Chemistry
Owen	Pembery	Mathematical Sciences
Laura	Redfern	ClusterVision
Gordon	Rennie	Computer Science
Andrew	Rhead	Mechanical Engineering
Francesco	Rizzo	Mechanical Engineering
Will	Saunders	Mathematical Sciences
Paul	Secular	Physics
Michael	Shapiro	Biology and Biochemistry
Jonathan	Skelton	Chemistry
Alexander	Smith	Physics
Megan	Stalker	Chemistry
Ben	Thomas	Mechanical Engineering
Matt	Thomason	Cray UK Ltd
Megan	Thompson	Chemical Engineering
Ian	Thompson	Physics
Tom	Underwood	Physics
Robert	Watson	Electronic and Electrical Engineering
Catherine-Axa	Wilkins	Biology and Biochemistry
Ian	Williams	Chemistry
Corwin	Wright	Electronic and Electrical Engineering
Artem	Yurlov	Mathematical Sciences

## Abstracts

## Keynotes

---

### The GPU Data Centre for AI

*Dr Timothy Lanfear, NVIDIA*

The on-going AI revolution is an exciting time in technology history. All sorts of new applications are emerging with enormous computational and data processing demands. NVIDIA has identified autonomous vehicles, smart cities and healthcare as key markets to focus on. The GPU is an excellent platform for AI computation, but new demands are being placed on the data centre infrastructure. In this presentation we will look at how the IT needs of the emerging AI industry can be met.

---

### An AI system for radiological reporting and prioritisation

*Prof Giovanni Montana, University of Warwick*

X-rays are commonly performed imaging tests that use small amounts of radiation to produce pictures of the organs, tissues, and bones of the body. X-rays of the chest are used to detect abnormalities or diseases of the airways, blood vessels, bones, heart, and lungs. In this work we present a stochastic attention-based model that is capable of learning what regions within a chest X-ray scan should be visually explored in order to conclude that the scan contains a specific radiological abnormality.

---

### Small holes - big impact: A molecular view of adsorption

*Tina Düren, Chemical Engineering*

Adsorption is a process that relies on the interaction of guest molecules with the surface of porous solids which have nanometre size pores and surface areas of 1000 m<sup>2</sup> / g or more. Applications include e.g. energy storage, carbon capture, separations in the petrochemical industry, drug delivery and sensing. In this presentation I will give an overview of how we (mis)use Balena for molecular simulations not only allow to predict the macroscopic performance of these materials for industrial applications but more importantly to gain detailed picture on the molecular level which is not easily accessible by experimental methods. The simulations allow to establish property / performance relationships for particular applications, provide insight into adsorption mechanisms and allow testing materials before they have been synthesised all of which ultimately helps to design new materials with improved performance. I will introduce the molecular world of adsorption showing fascinating and often surprising phenomena when, for example, gates open or solids start to breathe in the presence of guest molecules.

---

### Tales from Supercomputing

*Jess Jones, Cray UK Ltd*

A look at the work that goes on behind the scenes during the procurement and installation of a large supercomputer, from the point of view of the vendor. We look at the Met Office's latest supercomputer, and give a stratospheric overview of what went on during its procurement, from conception to the point where it passed acceptance testing.

## Talks

---

### Supercharging the materials discovery process with machine learning

*Daniel W. Davies (Chemistry), Keith T. Butler, Aron Walsh*

d.w.davies@bath.ac.uk

The design of new compounds for use in renewable energy applications such as solar cell devices is crucial to the sustainable scalability of these technologies. Thanks to a recent and ongoing paradigm shift towards data-led research in materials science, we are now able to relate compositions of chemical elements to certain target properties with models built using machine learning techniques. I will show how we have trained one such model on computational chemistry calculations carried out using HPC systems. We have applied this model as a filter to a space of 1.1 million hypothetical oxygen-containing compositions, in order to isolate those which show promise as light-absorbing semiconductors for use in solar energy applications.

To calculate their properties more accurately, it is necessary to predict what arrangement a given composition of elements will adopt in 3D space. This step of "crystal structure prediction" is usually carried out using extremely expensive evolutionary algorithms and can take many weeks to run for a single composition, even using highly parallelised code. I will show how we can sidestep this bottleneck by employing yet another model built from materials data. This model can be applied in a matter of minutes on a desktop workstation, and by scaling this approach up and running it in parallel on the Balena HPC system, we have been able to suggest over 400,000 crystal structures in just two days.

**Keywords** - Computational Chemistry, Energy Materials, Machine Learning, Data Mining

---

### Automatic recognition and classification of hip fractures

*Ellen Murphy (Institute for Mathematical Innovation), C.L. Gregson, O.A. Von Arx, M.R. Whitehouse, C.J. Budd, H.S. Gill*

e.murphy@bath.ac.uk

Hip fractures are a major cause of death and disability for older people and are one of the costliest treatments for the NHS (£1Billion annually). Given the high mortality rates (9% within 30 days; 30% within 1 year), any improvement in classification, and hence treatment, will have significant benefits. Hip fractures are classified using a standard classification system. However, there is wide variation between hospitals regarding who performs the classification. As classification influences the chosen treatment, this can affect patient outcomes. We are developing a machine learning based method to automatically classify hip fractures using X-rays, with the eventual aim of standardising classification across the NHS. There are two stages in our proposed classification process: automatically recognise the hip joints in the X-ray image; and then classify the hip fracture given this selected region. Using a GPU and caffe, we have trained a fully convolutional network to automatically locate the hip joint, scoring an intersection-over-union above 0.8 (>0.5 is correct) for 93% of the test set. We are currently training a convolutional neural network to automatically classify the type of fracture, using a dataset of X-ray images that have been hand-labelled by a panel of experts. Details of our methodology and results of our study will be given.

**Keywords** - Neural networks, machine learning, medical imaging, GPUs

---



---

## Molecular dynamics simulations of biological membranes and membrane proteins

*Victoria Oakes (Department of Chemistry), Carmen Domene*  
v.oakes@bath.ac.uk

Cell membranes are lipid bilayers that cover the surface of cells, separating the cell interior from the surrounding environment. Membrane proteins, which are embedded in these assemblies, are responsible for regulating the transport of materials and signals into and out of the cell. Recently, there have been significant advancements in structure determination of membrane proteins, allowing them to be studied on an atomic level. In my research, we use molecular dynamics to simulate structures in a model cell environment in order to understand their behaviour in situ, and directly observe a wide range of biological phenomenon such as drug binding. Membrane proteins are the targets of over 50% of the known small molecule drugs; therefore, understanding how they act on a molecular level is crucial to the rational, affordable design of improved pharmaceuticals.

**Keywords** - Computational Chemistry, molecular dynamics, ion channels, membrane lipids

---

## Distributed Deep Reinforcement Learning for Autonomous Control

*Gordon Rennie (Department of Computer Science), Özgür Şimşek*  
grr28@bath.ac.uk

Reinforcement learning is a type of machine learning in which agents learn through repeated interactions with an environment, with potential application to a variety of real-world control problems. In recent years advances have been made by augmenting reinforcement learning algorithms with deep neural networks, allowing agents to equal and surpass human expert level performance in complex games such as Go, chess, and various video games. To reduce training time for complex tasks, distributed implementations of reinforcement learning algorithms have been developed which allow many agents to interact in parallel to generate large training data sets. This talk will introduce the fundamentals of deep reinforcement learning, outline distributed architectures for training agents at scale, and describe an upcoming thesis project using distributed deep reinforcement learning on Balena to teach an agent to autonomously control aircraft in a range of autopilot tasks.

**Keywords** - reinforcement learning, deep learning, autonomous control, UAV

---

## A Parallel Fast Multipole Method Implementation for Electrostatic Interactions

*William Saunders (Mathematical Sciences), Eike Müller, James Grant*  
w.r.saunders@bath.ac.uk

Electrostatic interactions occur between charged particles. Including these interactions in molecular simulation codes is important to correctly model the physical properties of many materials. However, the cost of naive implementations of electrostatic interactions grows with the square of the number of particles, making realistic simulations unfeasible. We present an introduction to the Fast Multipole Method (FMM) which is an existing technique to compute electrostatic interactions with a computational complexity proportional to number of charged particles. We present a parallel FMM implementation within our framework for performance portable atomistic simulations and present results that compare performance with alternative algorithms.

**Keywords** - Fast Multipole Method, Electrostatics, MPI

---

## Parallel implementation of the Time-Evolving Block Decimation algorithm

*Paul Secular (Physics), Stephen Clark*

p.m.secular@bath.ac.uk

A common problem in condensed matter physics is to understand the dynamics of quantum lattice models. Such models are often described by local Hamiltonians, that is Hamiltonians with short-ranged interactions. Guifré Vidal's Time-Evolving Block Decimation (TEBD) algorithm allows one to simulate the time-evolution of quantum states on 1D lattices under such Hamiltonians. Like the Density Matrix Renormalisation Group (DMRG) algorithm, which is used for finding ground states, TEBD uses a tensor network representation of quantum states known as a matrix product state (MPS) or tensor train. In this talk I report on the development of an open-source implementation of the TEBD algorithm employing hybrid MPI /OpenMP parallelism and present preliminary results.

**Keywords** - Quantum Mechanics, Tensor networks, TEBD, DMRG, MPI, OpenMP

---

## Molecular Modelling of Hydrogen Storage in Cellulose

*Megan R. Stalker (Chemistry), Steve C. Parker, James Grant, Tim J. Mays*

m.stalker@bath.ac.uk

Hydrogen is a renewable and clean burning fuel which can provide on-demand power, making it a strong candidate to become an integral part of the future energy landscape. Due to the low volumetric density of hydrogen gas, the storage of hydrogen remains challenging. Porous polymers provide a lightweight alternative to traditional crystalline hydrogen storage materials. Cellulose is the most abundant organic polymer in the biosphere, and its biodegradability makes it more sustainable than conventional polymers. To the best of our knowledge, cellulose is yet to be investigated as a hydrogen storage candidate.

The project aims to apply molecular modelling to evaluate the feasibility of cellulose as a candidate for hydrogen storage. Firstly, a computational workflow was developed to model a range of cellulose structures. Grand canonical Monte Carlo methods were used for hydrogen sorption simulations in the various cellulose structures. Finally, the transport behaviour of hydrogen molecules within the various cellulose structures was investigated using molecular dynamics.

Cellulose structure and pore size were found to influence the hydrogen storage capacity. Slit pore model cellulose structures displayed higher hydrogen density than was observed in vacuum, indicating favourable sorption. Reasonable mobility of hydrogen was also observed in such structures, indicating sorption and desorption of hydrogen from cellulose is also possible. Further work has been undertaken to improve the accuracy of the models of cellulose structures.

**Keywords** - Computational Chemistry, Cellulose, Hydrogen storage, Sustainability

---

## Crystal-Size Framework Flexibility in MOFs

*Megan Thompson (Chemical Engineering), Claire Hobday, Tina Düren*

m.j.thompson@bath.ac.uk

Metal-organic frameworks often exhibit framework flexibility upon interaction with guest molecules due to weak coordination bonds between the nodes and ligands. Such MOFs spark interest for sensory and separation applications, but the transformation between the open and closed states is not well understood.

ZIF-8 and DUT-8 are flexible metal-organic frameworks that show different gate opening behaviour depending on the particle size. To get a better understanding of these experimental observations, the work presented uses a combination of density functional theory and grand-canonical Monte Carlo simulations to model N<sub>2</sub> adsorption in micro and nano-sized MOF particles.

**Keywords** - Molecular simulation, Metal-Organic Frameworks, Adsorption

---

---

## Computational methods for locating phase transitions in solids and fluids

*Tom L. Underwood (Physics)*

t.l.underwood@bath.ac.uk

It is curious that at certain temperatures the properties of a substance (e.g. water) change dramatically. For example, upon heating a solid through its melting point it becomes a liquid; and upon heating a liquid through its boiling point it becomes a gas. Such a dramatic change in a substance's properties is known as a *phase transition*, and the ability to accurately predict - from theory alone - the temperatures at which phase transitions occur would be invaluable to many fields (e.g. physics, chemistry, materials science). Unfortunately making such predictions is extremely challenging. Progress has been made on two fronts: improvements in computer power; and the development of computational methods which enable the problem to be solved more efficiently (thus allowing one to 'make the most' of the computer power at their disposal).

Here we consider computational methods. It turns out that the problem of pinpointing the location of a phase transition amounts to the problem of 'learning' the probability distribution over some quantity which characterises the phase of the substance in question (e.g. solid, liquid or gas). A plethora of methods exist to do this, though some are more suitable for certain problems than others. We will elucidate the general approach which underpins most of these methods. We will then present the results of recent calculations which show some of these methods in action, with a focus on the problem of determining the positions of solid-solid and liquid-gas phase transitions.

**Keywords** - statistical mechanics, Monte Carlo, free energy, computational chemistry

---

## Using MB (Mega Bytes) to visually analyse MB (Mega Bases)

*Catherine-Axa Wilkins (Biology and Biochemistry), James Doughty, Rod Scott*

c.wilkins@bath.ac.uk

Biologists may require the high-end graphics cards available on Balena to view large and complex models of biological molecules or to critically evaluate large genomic datasets. Here, we report on the use of Balena's visualization capability to load several large next generation DNA sequenced, aligned and annotated datasets using IGV (Integrated Genome Viewer). The purpose of this was to determine if potential and significant differences exist at a sequence level in two groups of DNA datasets. This could only have been possible if screening of these large files was carried out simultaneously by drawing on the core capability of Balena's visualization nodes.

**Keywords** - Visualisation, large DNA datasets, Balena capability

## Posters

---

### SLATE-based solvers for semi-implicit hybridised DG methods in fluid dynamics

*Jack Betteridge (Mathematical Sciences), Eike Müller, Ivan Graham and Thomas Gibson, Lawrence Mitchell (Imperial College London)*

j.d.betteridge@Bath.ac.uk

Many research problems in science and engineering rely on the solution of partial differential equations (PDEs) on a large scale and at a high resolution; one example is numerical weather prediction (NWP). This requires the use of modern high performance computing (HPC) technology and corresponding modern discretisation techniques. To make optimal use of the computational hardware, those methods have to be implemented using the latest software packages.

The development of competitive semi-implicit time-stepping schemes for Discontinuous Galerkin (DG) methods would help to popularise their use in operational code. Our recent research has focused on speeding up the numerical solution of the shallow water equations (related to the Navier-Stokes equations for NWP), by using a semi-implicit time-stepping scheme with our own efficient DG solvers. Our code currently uses the Firedrake software framework, which is capable of generating high performance parallel code for these DG solvers. The new SLATE language, which is built into Firedrake, allows the easy implementation of hybridized DG methods. With the help of our collaborators at Imperial College London, we were able to improve our existing code and add further functionality to the Firedrake code base. I will present some of our latest results on the development of solvers and bespoke preconditioners for the shallow water equations.

**Keywords** - Partial Differential Equations, Discontinuous Galerkin, High-order, Numerical Weather Prediction, Shallow Water Equations, Firedrake

---

### Streamlining the Hydrodynamic Modelling of Ocean Wave Energy Converters

*J A Bridgwater Court (Mechanical Engineering), A J Hillis, A R Plummer*

j.a.bridgwater.court@bath.ac.uk

Ocean surface waves represent a largely untapped source of renewable energy that is nowhere near the levels of technical maturity of resources such as solar and wind due to the cost of obtaining reliable data for use in design development. Developing an ocean wave energy device requires detailed modelling of the interaction between the waves and the wave energy converter (WEC) body, coupled with that between WEC and controlled power take-off circuit. Unfortunately it is generally the case that for any geometry and control strategy of practical interest this coupled interaction is too nonlinear to be accurately modelled using computationally inexpensive methods, while the more accurate approaches - computational fluid dynamics (CFD) and smoothed particle hydrodynamics (SPH) - are too burdensome to be viable, meaning device developers need to pay large sums to inform their work using physical laboratory and real-sea tests. This research aims to find when in the device development process the simpler, more linearised models can be used through comparison of these approaches with detailed CFD simulations, developed in OpenFOAM R and coupled with a representative power take-off model incorporating control, in order to provide confidence in the numerical modelling of WECs and thereby delay the point at which developers need to pay for expensive wave tank testing and sea trials until these are truly necessary, as well as highlighting modelling approaches that bridge the gap between the rapid, low-fidelity linearised models and the slow, high-fidelity CFD and SPH methods.

**Keywords** - Computational Fluid Dynamics, MPI, OpenFOAM, WSI, Wave Energy

---

---

## Modelling the design spaces of construction kits

*J. A. Gopsill (Mechanical Engineering)*

j.a.gopsill@bath.ac.uk

Construction Kits (CKs) are a fundamental tool for the generation and communication of ideas. First appearing as a means to support children's development through the building of models, they have since been used to support the learning of languages and physical environments. The application of CKs to support New Product Development (NPD) is also well established where they are used extensively to support the creation and sharing of ideas. They have been particularly successful in supporting configuration design tasks, such as city, town, office and manufacturing facilities planning, where stakeholders from multiple disciplines come together to design. Whilst their ubiquity across society is apparent, it is surprising to see that there is little research on how a CK represents and constrains the design space. The manner in which a CK constrains the design space may have profound effects on the solutions generated by individuals as well as supporting the development of their problem solving skills. This poster presents the development of code that utilises High-Performance Computing (HPC) to compute the design spaces of CKs so that a better understanding of this phenomena can be developed.

**Keywords** - Construction Kits, Combinatorics, Python, MPI, Engineering Design

---

## The importance of the membrane in mediating the mechanism of Monoamine Oxidase B

*Hannah Jones (Biology and Biochemistry), Rory Crean, Marc van der Kamp, Christopher Pudney*

h.b.l.jones@bath.ac.uk

Human monoamine oxidase B (MAO-B) catalyses the oxidation of amines in the brain, and is inhibited for the treatment of both Parkinson's disease and depression. Despite its clinical importance, its catalytic mechanism remains unclear. Experimental evidence indicating half-site reactivity and the involvement of the membrane environment in MAO-B were further investigated using molecular dynamic (MD) simulations. Our MD simulations identify a hitherto un-recognized entrance for substrate binding, membrane modulated substrate access, a half-site reactive binding site, and a largely inaccessible active site. The study will inform new routes and targets for rational drug design based on a principle of dynamic structural ensembles.

**Keywords** - Computational biochemistry, molecular dynamic simulations, GPUs

---

## Force Field Validation for Light Gas Adsorption in Metal-Organic Frameworks

*Mike Kallo (Chemical Engineering), Ping Yee Tan, Matthew Lennox*

m.kallo@bath.ac.uk

Carbon dioxide is a gas participating in global warming effects. It is released in the atmosphere by our petroleum based economy, having a harmful impact on the environment. Metal organic frameworks (MOFs) have shown promise for carbon capture processes and the use of molecular simulation to identify high-performance MOFs and explain fundamental adsorption phenomena has become more and more common. Simulating the system under study representatively, by considering all the affecting parameters and methodology limitations, is essential. In this project we study computationally carbon dioxide adsorption in several MOFs, by using grand canonical Monte Carlo (GCMC) techniques. Several different, commonly used force fields are implemented in order to describe the adsorption phenomena and simulation validity is evaluated via comparison with experimental data. Different MOFs are tested on their capacity of adsorbing CO<sub>2</sub> molecules on a range of pressures at ambient temperature.

We will show that while DREIDING is able to provide quantitative and qualitative agreement with experimental isotherms in many cases, the predictive power of common GCMC approaches is system dependent. Energy calculations revealed that the Van der Waals carbon dioxide-metal organic framework interactions are significantly more important than the rest of the interactions for the system. Comparison of energies predicted from GCMC calculations with available quantum chemical simulation data reveals that even simple force fields such as DREIDING are able to predict CO<sub>2</sub>-MOF interactions with reasonable accuracy. In our present work we study a 2D thin film MOF system, CuBDC for CO<sub>2</sub>/methane separation process since Metal Organic Framework Nanosheets (MONs) demonstrate a high selectivity adsorption of carbon dioxide over methane.

**Keywords** - Simulation, CO<sub>2</sub>, adsorption, GCMC, force field, thin films, MONs, MOFs

---

### [pylj: an open-source python library for teaching molecular simulation](#)

*Andrew R. McCluskey (Chemistry), Benjamin J. Morgan, Karen J. Edler, Stephen C. Parker*  
a.r.mccluskey@bath.ac.uk

pylj is an open-source Python library designed to facilitate the interaction between students, both undergraduate and postgraduate, and molecular simulation. Designed around the Jupyter notebook framework it is easy to implement in either classroom or computational laboratory. Additionally, due to the open-source, and documented, nature of the code it is easy for educators to add unique, custom extensions.

pylj is capable of performing the molecular dynamics simulation of a 2D argon system, under both NVE and NVT ensembles using the Velocity-Verlet integrator. Furthermore, a series of sampling classes exist, and it is easy to design a custom sampling class. We will present an example exercise that can be undertaken using pylj, where the ideal, and non-ideal nature of argon gas at different pressures is investigated.

pylj can be tested online, running on an interactive Jupyter notebook, or can be run locally by installation via pip or directly from source. For more information about pylj please visit <http://pythoninchemistry.org/pylj>.

**Keywords** - Teaching, Python, Jupyter, Molecular Simulation

---

### [Modelling Proteins as a Mechanical Network](#)

*Thomas McManus (Physics), Stephen Wells, Alison Walker*  
t.j.mcmanus@bath.ac.uk

In the field of protein bioinformatics, it has long been clear that understanding a protein's structure is key to understanding its function and biological role. In more recent years it has also been made apparent that the global length scale changes to a structure as it performs its function, can be just as informative as the atomic level information. Elastic network modelling is a proven method used to find the low frequency motions of a protein structure, predicting these global conformational pathways. Our improved implementation of FRODA (Framework Rigidity Optimised Dynamic Algorithm), along with a new encompassing package of tools for analysing the chemical structure, SafetyFIRST, finds and explores these pathways with minimal computational requirement, permitting  $\mu$ s-ms simulations of large complexes on a desktop machine. We present the outline of this method here-in, along with an algorithm for completely *in silico* protein design, capitalizing on the speed of this method to reduce experimental waste, both timewise and financially. Furthermore, it has been shown that the pairing of these tools with relatively short simulations of more accurate 'golden standard' methods, such as Molecular Dynamics, allows not only fast but also highly accurate explorations of these motions, by use of FRODA and SafetyFIRST as an effective 'shortcut' for starting configurations.

Moving forward the efficiency of this tool will be used to both: unlock the behaviour of high variance systems - namely the antibodies present in human serum and their antigen binding motions, focusing on HIV protease at first - and use HPC parallelism to explore structures on size and time scales not yet conceivable with atomic level resolution.

**Keywords** - Computational Biophysics, Efficient Simulation, Bioinformatics, Protein Structure, FRODA, SafetyFIRST, Elastic Modelling, In Silico Protein Design

---

## Hazardous Compounds at the Soil-Water Interface

*Leyorla Ohene-Yeboah (Chemistry), Jonathan Skelton, James Grant, Steve Parker*

[l.a.ohene-yeboah@bath.ac.uk](mailto:l.a.ohene-yeboah@bath.ac.uk)

The accumulation and persistence of hazardous compounds (HCs) in surface and ground water as well as living organisms have emerged as an adverse effect of human anthropogenic behaviour. A variety of HCs, from emerging contaminants found in pharmaceutical residues and personal care products to household chemicals, biocides/pesticides and manufacturing wastes have been found to persist in the environment. Some of these compounds have been shown to cause adverse effects in aquatic organism as well as promote increased risk of developing thyroid disorders, tumours and diabetes[1].

In this study, we aim to build on the foundations of our understanding of how HCs interact with the environment by applying atomistic simulation methods to determine the physicochemical factors controlling the distribution of pollutants and their metabolites in aqueous and terrestrial environments, and then apply this to identifying sustainable ways of control their transport. Initial studies centred on HCs adsorption on two model clay surfaces; sodium montmorillonite and pyrophyllite. Simulations were performed using a combination of dispersion corrected DFT and classical MD methods, to calculate the binding free energy and identify favourable sites for HC adsorption on clays in vacuum and water. Our initial investigation focused on 10 molecules including dibenzodioxins and two poly-chlorinated derivatives, chloro- and hexachloro-benzene, emerging contaminants amphetamine and two other derivatives and MDA and its analogue MDMA. These compounds were chosen to detect the links between the chemical properties of these HCs and their possible fate and effects in the environment.

[1] *Water Res.*, 2014, 72, 3-27.

**Keywords** - Computational Chemistry, Environmental Chemistry, GPUs, Emerging Contaminants, Persistent organic pollutants

---

## Computational Modelling of Defects in Battery Materials

*Alex Squires (Chemistry), Alison B. Walker, Benjamin J. Morgan*

[ags49@bath.ac.uk](mailto:ags49@bath.ac.uk)

To explore the properties of lithium ion batteries (LIBs), the defect chemistry of their constituent materials must be considered; however, as LIBs function via a mechanism of chemical extremes, it is not simple to model them without resorting to computationally expensive methods. Such an approach is required as Li is accommodated and transported by generating or annihilating point defects. It is defect concentrations that give insight into thermodynamic activities, both theoretical capacity and cell voltage can be expressed through the study of point defects. Kinetic properties also relate strongly to defect chemistry, as the vast majority of potential charge carriers (Li and electrons) in the system are of no interest; their energies are such that they are inaccessible to the processes we are interested in describing. Defects form the system charge carriers, we must understand their chemistry to understand battery system kinetics.

**Keywords** - Computational Chemistry

---

---

## How well do stratospheric reanalyses reproduce satellite temperature data?

*Corwin Wright (Elec. Eng.), Neil Hindley*

c.wright@bath.ac.uk

Atmospheric reanalyses are weather and climate models which are widely used in the atmospheric sciences as proxies for the state of the atmosphere in the recent past. This is particularly the case in the stratosphere, where historical observations are sparse. Here, we oversample stratospheric temperature data from six modern reanalyses to produce synthetic satellite observations, which we directly compare to retrieved temperatures from the COSMIC, HIRDLS and SABER instruments and to brightness temperatures from the AIRS instrument for the period 2003- 2012. The highest correlations are seen at high latitudes and the lowest in the sub-tropics, but root-mean-square (RMS) differences are highest (10K or greater) in high-latitude winter. At all latitudes, correlations decrease and RMS differences increase with increasing height. Differences at high altitudes become especially large during disrupted periods such as the post-sudden stratospheric warming recovery phase, where zonal-mean differences can be as high as 18K between different datasets. Finally, we show evidence that the full-input reanalyses are more tightly correlated with each other than with observations, even observations which they assimilate. Analysis of co-located observations indicates that the reanalyses may be over-tuned to match their comparators. If so, this presents significant implications for future reanalysis development.

**Keywords** - atmospheric dynamics, weather forecasting, reanalysis, satellites, stratosphere



## Acknowledgements

First and foremost, we would like to thank all our speakers, poster presenters and session chairs for facilitating the insightful and thought-provoking discussions during the day.

Thank you to the prize committee for judging the contributions and awarding the prizes, and to ClusterVision for providing one of the prizes.

Special thanks go to the members of Organising Committee and HPC Team for their invaluable input during the planning stages and in coordinating the event:

Gaël Donval  
James Grant  
Jonathan Skelton

Steven Chapman  
Roshan Mathew

Finally, thank you to you, the attendees, for your participation and for truly making the Symposium a real success.

For more information about the University's HPC service visit: [go.bath.ac.uk/hpc](http://go.bath.ac.uk/hpc)

You can also follow us on Twitter: @BathHPC

Notes



