
# EPSRC Service Level Agreement with STFC for Computational Science Support

FY 2019/20 Annual Report

(*Covering the period 1 April 2019 – 31 March 2020*)

**May 2020**

# Table of Contents

# Background

The Scientific Computing Department (SCD) of the Science and Technology Facilities Council (STFC) provides computational science support for a number of scientific communities funded by EPSRC, and organised in Collaborative Computational Projects (CCPs) and High End Computing (HEC) Consortia. This programme of work is carried out by staff at the Daresbury and Rutherford Appleton Laboratories under a Service Level Agreement (**SLA**) with EPSRC, and its main objectives are:

- Develop and maintain codes and methods to keep them internationally relevant for current and evolving scientific challenges and hardware capabilities
- Widen participation in the exploitation of methods and codes through training and scientific collaboration
- Support collaboration and coordination of the various communities to broaden and strengthen the UK-based research activities aligned with EPSRC's goals
- Provide career paths and professional development opportunities for computational scientists and engineers focused on method and software development
- Widen engagement with the broader UK and international communities engaged in developing methods and software for computational science and engineering

The nature of the support provided is tailored to the needs of the communities and can include:

- **Development of theory, algorithms, and software**: This is a key element of support for many current projects, resulting in long-term, continued expansion and updating of the software programs. It may include the consolidation of existing codes into a more sustainable community software package
- **User support and training**: This includes individual support and training as well as help to organise and conduct events such as workshops, summer schools and study weekends. Support for Centres of Doctoral Training is also offered
- **Outreach and promotion of computational science and engineering activities**: facilitate the exchange of expertise, and tools, reaching out to new communities, including experimentalists and industry, nationally and world-wide
- **Collaboration on scientific projects**: working together with scientists in the communities to advance scientific research and help nurture projects and develop new opportunities
- **Porting, optimisation, and benchmarking** on local and national computing platforms including High Performance Computing (HPC) and new architectures: evaluation of new hardware and software technologies
- **Maintenance, distribution**, license management, dissemination and demonstration of software
- **Management of scientific data**: This includes activities such as, for example, the development of visualisation and workflow management tools, database and archiving, and verification and validation activities
- **Co-ordinate and nurture existing and new communities**, from practical tasks such as organising community meetings, to representing the communities in strategic activities in the UK Research Councils and abroad.

**CoSeC**, the Computational Science Centre for Research Communities, brings together these activities with those in support of CCP4 (partly funded by a BBSRC grant), CCP-EM (funded by an MRC grant), and CCP-WSI (funded by an EPSRC grant): https://www.scd.stfc.ac.uk/Pages/CoSeC.aspx

The communities currently supported are summarised in the table below:

| Project | Title | Project Chair | CoSeC Project manager |
|---|---|---|---|
| CCP5 | The Computer Simulation of Condensed Phases. | Prof Paola Carbone | Dr Alin Elena |
| CCP9 | Computational Electronic Structure of Condensed Matter | Prof Stewart Clark | Dr Leon Petit |
| CCP-Mag | CCP on Computational Magnetism | Prof Julie Staunton | Dr Kun Cao |

| | | | |
|---|---|---|---|
| CCP-NC | NMR Crystallography | Dr Paul Hodgkinson | Dr Simone Sturniolo |
| CCPQ | Quantum dynamics in Atomic Molecular and Optical Physics | Prof Graham Worth | Dr Martin Plummer |
| CCP-Plasma | The Plasma-CCP Network | Prof Tony Arber | Dr Joseph Parker |
| CCPi | Tomographic Imaging | Prof Philip Withers | Dr Edoardo Pasca |
| CCP PET-MR | Computational Collaborative Project in Synergistic PET-MR Reconstruction | Prof Kris Thielemans | Dr Evgueni Ovtchinnikov |
| CCP-BioSim | Biomolecular simulation at the life sciences interface | Prof Adrian Mulholland | Dr Tom Keal |
| CCP-WSI+ | Wave structure interaction + | Prof Deboah Greaves | Dr Stephen Longshaw |
| CCP Turbulence | Turbulence | Prof Sylvain Laizet | Dr David Emerson |
| CCP-NTH | Nuclear thermal Hydraulics | Prof Shuisheng Hu | Dr David Emerson |
| CCP-QC | Quantum computing | Prof Viv Kendon | Dr Alin Elena |
| MCC | UK Materials Chemistry Consortium | Prof Richard Catlow | Dr Tom Keal |
| HEC-BioSim | High-End Computing Consortium in biomolecular simulation | Dr Syma Khalid | Dr James Gebbie |
| UKCP | United Kingdom Car-Parrinello Consortium | Prof Matt Probert | Dr Dominik Jochym |
| HEC-Plasma | Plasma High-end Computing Consortium | Prof Tony Arber | Dr Joseph Parker |
| UKCOMES | UK Consortium on Mesoscale Engineering Sciences | Prof Kai Luo | Dr Michael Seaton |
| UK-AMOR | UK Atomic, Molecular and Optical physics R-matrix Consortium | Prof Jonathan Tennyson | Dr Martin Plummer |
| UKTC | UK Turbulence Consortium | Dr Sylvain Laizet | Dr David Emerson |
| UKCTRF | UK Consortium on Turbulent Reacting Flows | Prof Nilanjan Chakraborty | Dr David Emerson |

Following the 2019 EPSRC CCP call the following CCPs were not renewed: CCP5, CCP-Mag, CCPQ and CCP-Plasma. Four new CCPs were funded and can be seen highlighted in yellow in the table above and below: CCP-WSI+, CCP Turbulence, CCP-NTH and CCP-QC

A brief description of each community can be found in the community-specific sections below. More information is available at http://www.ccp.ac.uk/ for the CCPs and at https://www.epsrc.ac.uk/research/facilities/hpc/access/highendcomputingconsortia/ for the HEC consortia. The current level of support awarded to the communities is as follows. Please note that this table shows the original level of support awarded following the EPSRC call in 2015 and the current level of support (in brackets) following the most recent EPSRC CCP call in 2019.

| Community | Core support per project (FTEs per annum) |
|---|---|
| CCP5 | 3.4 (0.0 from 31 March 2020) |
| CCP9 | 2.6 (3.0 from 1 April 2020) |
| CCP-Mag | 0.8 (0.0 from 31 March 2020) |
| CCP-NC | 1.4 |
| CCPQ | 2.0 (0.0 from 31 March 2020) |
| CCP-Plasma | 0.8 (0.0 from 31 March 2020) |
| CCPi | 1.3 (1.5 from 1 April 2020) |
| CCP-PET/MR | 1.25 (1.8 from 1 April 2020) |
| CCP-BioSim | 1.25 (1.95 from 1 Aprl 2020) |
| CCP-WSI+ | 2.0 (from 1 April 2020) |
| CCP Turbulence | 1.75 (from 1 April 2020) |
| CCP-NTH | 1.0 (from 1 April 2020) |
| CCP-QC | 0.6 (from 1 April 2020) |
| Materials Chemistry | 2.5 (2.0 from 31 Oct 2018) |
| HEC-BioSim | 1.0 (0.8 from 31 Oct 2018) |
| UKCP | 1.0 |
| HEC Plasma | 0.2 |
| UK COMES | 1.0 (0.6 from 31 May 2018) |
| UK-AMOR | 0.2 (from 1 April 2018) |
| UKTC | 0.3 (from 30 June 2018) |
| UKCTRF | 0.13 (from 07 Jan 2019) |

## *CoSeC Project Office*

The objectives of the Project Office are to provide coordination through the CCP Steering Panel and programme oversight through periodic reports, and input to reviews, to EPSRC and the SLA Steering Committee, support for the organisation of networking and training events, and manage the progress of the programme staff, L&D and liaise with the communities supported, in particular to help increase the impact of the programme.

### Summary Report (1 April 2019 – 31 March 2020)

During the reporting period the Project Office continued to work with Technopolis to complete the commissioned SLA impact report. The report is now final and published. We also assisted with the preparation of submissions to the recent CCP call as required.

The inaugraul CoSeC Impact Award was launched, with a reasonable number of nominations received, and a winner for the 2020 award was selected unanimously by an external judging panel. The aim of launching this new award is three-fold: it is a means of recognising the work of researchers early in their careers who have been, or continue to be, supported by CoSeC; it is a means of raising awareness of the CoSeC communities, and finally, it is a means of acquiring evidence of the impact of CoSeC and the communities it supports on science, society and the economy.

### Outline Plans 2020/21 (1 April 2020 – 31 March 2021)

*(Detailed plans and milestones are included in the Technical Appendix for reference)*

Moving forward into 2020/21 the project office will look to implement any changes required as a result of the latest CCP call, including welcoming four new CCPs. The 2021 CoSeC Impact Award will be

launched towards the end of 2020. As we approach the end of the current five year SLA cycle we will also begin to prepare any renewal documentation required by EPSRC as part of that process.

## *CCP5 – Computer Simulation of Condensed Phases*

CCP5 is the Collaborative Computational Project for computer simulation of condensed phase materials at length scales spanning from atomistic to mesoscopic levels. Founded more than 35 years ago, CCP5 has promoted the involvement of UK scientists in collaborative research achieved via software and methodology development, training, networking and outreach. It provides support for all UK scientists engaged in developing, applying and exploiting computer simulation methods for condensed matter systems. CCP5 has over 330 UK members and over 330 international members, which comprise research active academic faculty staff in 35 different UK universities and at least 18 other UK industrial, charitable or government organisations. A distinctive feature of CCP5 is its successful strategy of developing and disseminating new codes and methods for all kinds of materials problems. These include solid-state materials, polymers, colloidal solutions, liquids and mixtures, liquid crystals, surfaces and interfaces, homogeneous and heterogeneous catalysts, mineral, bio-mineral, organic and bio-molecular systems.

Our core software support covers numerical energy minimisation, classical molecular dynamics and Monte Carlo simulation, ranging from atomistic to multi-scale molecular systems. An increasing effort is exerted to tackle major challenges in cutting edge parallel simulations, linking atomistic and higher level models with first principles (quantum), spanning longer time- and length-scales by means of coarse-graining and mesoscale modelling so as to provide reliable multi-scale simulation protocols.

CCP5 major software and methodology support includes five active projects which together account for over 4,000 active licence holders worldwide in 80 countries. DL_POLY is a general purpose, classical, particle dynamics program. DL_MESO is a general purpose Dissipative Particle Dynamics program. DL_MONTE is a general purpose particle Monte Carlo program. ChemShell is an advanced command line environment with tools and methods for modelling materials systems simultaneously in classical and quantum terms. DL_FIELD is a cheminformatics program for conversion of materials structures from XYZ/PDB description to structure and force-field model files suitable for input into DL_POLY, DL_MESO and DL_MONTE.

CCP5 also provides funding for undergraduate student bursaries, workshop and conference funding and international visitor tours in the UK as well as an extensive range of training events including the annual CCP5 summer school.

### Summary Report (1 April 2019 – 31 March 2020)

DL_POLY 4.09.03 was released in August 2019, this is a bug release, solving few critical issues related to domain decomposition heuristic. One minor update for DL_MESO was released in June 2019, which fixed some identified bugs in the Widom insertion utility for calculating excess chemical potentials from DPD simulations and added OpenMP multithreading to that utility to speed up calculations allowing users to do more science per unit of time. DL_Monte has got an automated testing framework which will greatly benefit the developers and improve the quality of code for users - for developers the code is automatically testing saving time and ensuring correctness and for users there is better trust of their results

DL_POLY 4.10 is in work with new features in testing phase and a release date provisional for November 2019. The new version has the advantages of pure modularisation, thread safe data by design and separation of data and computation, as described in a previous report. Further software developments happened in parallel for addition of new features with Dr Alex Buccheri at the University of Bristol, Dr Oliver Dicks and Aaron Diver at QMUL and Dr Jacob Wilkins at University of Oxford. Ivan

Scivetti's work on Empirical Valence bond implementation and Forward Flux Sampling continues. Prof Simone Meloni University of Ferarra and Dr Marco Tortora, Sapienza Rome visited Daresbury Laboratory this summer and restrained molecular dynamics and a library of collective variables was implemented. All these additions will greatly expand the scope of the DL_POLY code for the benefit of the users. Users will be able to do new types of science, eg access dftb formalism from dlpoly, or monitor coordination of atoms live during the simulation.

A proof of concept jupyter notebook running in cloud was demonstrated this summer, eliminating the initial setup barrier for novice users. The instance runs temporary at jupyter.elena.space, with plan to move it to a more permanent place at training.ccp5.ac.uk

A new python package for DL_POLY was released dlpoly-py in collaboration with Dr Jacob Wilkins from University of Oxford, https://gitlab.com/drFaustroll/dlpoly-py this allows users to devise complex workflows from python scripts and mix with other python modules, like python-ase and mdanalysis for pre and post processeing of inputs and ouputs. The impact of this has two aspects: easy automation of complex tasks, like changing temperature for the simulation or the statistical ensemble; reproducibility of a computer experiment, in that a user will be able to encode all the manual steps in a python script together with any pre- and post- processing steps.

Cross-fertilisation between CCP5 and other CCPs continued this summer mainly via the CCP5 Summer School which took place in Durham over two weeks in July, where various CoSeC members lectured, Alin Elena, CCP5, Fortran and Practicals, Michael Seaton, Mesoscale Simulations, Sarah Fegan, Bio-molecular simulations and Barry Searle, First principle methods. This cross-fertilisation makes other communities aware of what ccp5 is doing and the other way round and sometimes starts new collaborations. for the summer school, for example where we interact with CCP9/UKCOMES/CCPBiosim we show case their strengths to the students and the students benefit from having expert practitioners in the field teaching them

The following events were sponsored by CCP5 and some of them directly supported by CoSeC members: Molecules at Surfaces: What do we really know? 12-13 September 2019, University of Liverpool, Physics by the Lake, August 5-16 2019, University of Stirling, Strathclyde Adsoprtion Summer School, 18-21 June 2019, University of Strathclyde, Glasgow, CP2K Users and Developers Meeting, 14th of June 2019, Imperial College London, Scientific Application of Quantum Computing Meeting,14 May 2019 at the LSE Bankside House, DL_POLY Developers Meeting, QMUL, 12-13 November 2019,Two temperature molecular dynamics, Daresbury, 9th of March 2020; Department of Materials Postdoctoral and Fellow Symposium Imperial College, London 16th of December 2019, New Horizons 2020, York, 10th January 2020.

The CCP5 AGM conference organised at London School of Economics, 16th -18th of September 2019, was attended by over 80 participants, from UK and overseas, from academia and industry. The event involved a selection of very high quality international speakers. The CCP5 AGM for 2020 will be organised at University of Huddersfield by Marco Molinari and Alin Elena. This will mark 40th anniversary of CCP5.

CCP5 Summer Bursaries for 2019 were awarded to a number of five students. Reports will be available on the website once all are received. This programme will continue in 2020 also.

Two new members were elected to the CCP5 committee, Dr Flor Siperstein from University of Manchester and Prof Fernanda Duarte from University of Oxford. Effort is ongoing on keeping the website up to date and new outreach and impact sections were added. A new chair was elected Dr Paola Carbone from University of Manchester. Part of the period Prof Neil Allan, Dr Paola Carbone and

Dr Alin Elena spent on writing the renewal grant for CCP5. A comprehensive member survey was undertaken.

Dr Ivan Scivetti had two invited talks at Universidad Autonoma de Madrid, in May 2019, one invited talk at Molecules at surface, University of Liverpool and an oral presentation at AGM 2019. Dr Michael Seaton had an invited talk at the HPC 2019 conference in Borovets, Bulgaria in September 2019. Dr Andrey Brukhno had two oral presentations at CECAM workshop in Liverpool (July 2019) and CCP5 AGM in London (Sept 2019). Dr John Purton gave a talk at University of Bath Symposium on Energy Materials

Various CoSeC members were involved in supporting PhD and Master students over the summer. Students involved were from QMUL, Cardiff University, Imperial College, Sheffield Hallam and University of Huddersfield. Daniele Visco from Imperial College was awarded best poster prize at IMSE symposium.

The renewal process for CCP5 was not successful this will create a lot of problems for the community and for the software maintenance programme. Discussions within community are undergoing to address the problem. Dr Paola Carbone the new chair will report more.

We had a few postponed sponsored events due to covid, with the most important by far the 2020 CCP5 summer school which was postponed to next July and will be merged with the 2021 edition extending the number of participants to 100+. The rest of the postponed events were: 4th Manchester Conference on Multiscale modelling March 30-April 1 2020, Introduction to Modern Fortran, April 15-17 King's College London, Data Enabled Atomistic Modelling, with other CCPs, June 11-12, Cosener House, Oxfordshire, Physics by the Lake, University of Stirling, August 3-14 2020, IWOM2020: Thomas Young Centre International Workshop on Organic Materials - Charge Transfer and Photo-induced Processes; 17th to 19th June 2020, London, UK, Hermes 2020 Summer school, 24-28 July 2020. With the exception of the Multiscale conference none of these events implied a financial penalty.

### Outline Plans 2020/21 (1 April 2020 – 31 March 2021)

*(Detailed plans and milestones are included in the Technical Appendix for reference)*

CCP5 was not renewed as part of the 2019 EPSRC call "Collaborative Computational Projects: Community Building, Networking and Core Computational Science Support".

## *CCP9 – Electronic Structure of Solids*

The Collaborative Computational Project on Computational Electronic Structure of Condensed Matter (CCP9) brings together UK researchers in the field of ab-initio electronic structure calculations of condensed matter systems. The field includes the study of metals, semiconductors, magnets, and superconductors from atomistic quantum mechanical calculations. The activities of CCP9 encompass such areas as magneto-electronics, photonics, nanotechnology, high-temperature superconductors, novel wide band gap semiconductors, and magneto-calorics. CCP9 provides a network that connects UK research groups in electronic structure, facilitates UK participation in the larger European Ψk Network, and supports a number of cutting edge computational codes.

### Summary Report (1 April 2019 – 31 March 2020)

Questaal is the main UK code for going beyond DFT using many body perturbation theory (GW, BSE). It provides the theoretical methodology to treat a variety of materials problems, i.e. their electronic, magnetic and spectral properties, with unprecedented accuracy and reliability. It's important it continues

to be developed to extend its functionalities, such as combining DMFT with GW, and improve its performance.

The QUESTAAL results for the international Delta codes validation and verification effort have been uploaded to the website https://molmod.ugent.be/deltacodesdft by Jerome Jackson. This benchmarking exercise is an important community lead initiative to demonstrate the mutual consistency of the different electronic structure codes as well as their accuracy and reliability. Jerome Jackson is corresponding author on invited QUESTAAL review article for the special issue of Computer Physics Communications, marking its 50th anniversary. An abridged version of this article was furthermore published as Psi-k highlight in March 2020.

The 3rd Daresbury QUESTAAL hands-on course was organized 13-17 May 2019 at Daresbury Laboratory. Thirty participants from both the UK and further afield took part. The course was organized in collaboration with the code developers from King's College. Lectures by invited speakers as well as tutorials were given, introducing the students to the capabilities of the code, including latest DMFT and BSE developments to deal with spectroscopies.

Ongoing theory-experiment collaboration with Warwick University and Ames laboratory (US), with the goal to probe the interactions that govern the rare-earth/transition-metal compounds, providing important guidelines for future quantitative modelling of technologically useful materials such as are used for permanent magnets and magnetic cooling. Magneto caloric effects currently finds specialty application for cooling to ultra-low temperatures, permanent magnets find application in for example hard-disk drives , MRI scanners, and electricity generators among others. Leon Petit is the first author on an article investigating the competition of different magnetic interactions in Gd-intermetallic alloys, and a further paper highlighting the electronic mechanism behind the giant non-hysteretic magnetocalric effect in Eu2In has been submitted.

CECAM/Psi-k co-funding was secured for the SPR-KKR hands-on course in collaboration with CCP-Mag and colleagues from MLU Munich, and the 4th Daresbury QUESTAAL school in collaboration with the code developers at King's College. The hands-on courses will teach a new generation of materials scientist how to use our electronic structure codes.

During the reporting period, a number of successful proposals have been submitted: Jerome Jackson is co-investigator on a successful EPSRC New Investigator Award proposal by Tom Ostler of Sheffield Hallam entitled "Fast Magnetisation Dynamics". Leon Petit is co-investigator on a successful EPSRC proposal entitled "Emergent Magnetism and Spin Interactions in Metallo-Molecular Interfaces". The aim is to develop the tunable and enhanced properties such as magnetoresistive and gated magnetism in these compounds.

Work has been ongoing to implement OpenMP parallelization in the CRYSTAL code. The OpenMP parallelisation of bi-electronic forces was finalized in November 2019, the OpenMP parallelisation of DFT forces was completed in March 2020. An important aspect of this work has been the code modernisation necessary to support OpenMP which also improves the code maintainability for the developers. Barry Searle joined the CRYSTAL developer meeting on two occasions in Turin as a result of which the work on two-electron forces was merged with the main code in January 2020. The DFT forces and any additional work should be merged in June 2020 followed by a code freeze prior to the next release of CRYSTAL early in 2021.

A CCP9 working group meeting was organized in London on the 27th of August to discuss the CCP9 renewal proposal with the members of the community, and based on their suggestions the Daresbury team collaborated with Stewart Clark in putting together the proposal that was submitted to EPSRC in early October. The proposal was successful, being allocated 3 FTE working on Community driven

projects during the 2020-2025 period, as well as full support for the entire range of proposed networking events. The CCP9 CoSeC team visited Stewart Clark at the end of January to discuss the projects lined up for the two first years of the new CCP9 period.

Psi-k: CCP9 provides support for the European electronic structure network Psi-k in the form of finance administration and editing the Psi-k scientific highlight. Four highlights were edited during the reporting period, and submitted to the mailing-list which now reaches more than 5000 people.

**Problems, issues and difficulties encountered during reporting period:** As noted earlier, a number of events had to be postponed due to CoVid-19. We are waiting for feedback from EPSRC to hear whether we will be allowed to use funds from the current CCP9 period to pay for these delayed events even if they take place after November 2020.

### Outline Plans 2020/21 (1 April 2020 – 31 March 2021)

*(Detailed plans and milestones are included in the Technical Appendix for reference)*

The CCP9 CoSeC activities of 20/21 will focus on working with the Wannier90 library, developing the spin susceptibilities in MBPT for the QUESTAAL code, and studying the electronic structure of correlated interface materials. Our work will focus on understanding the different behaviour of correlated materials at surfaces/interfaces, compared to bulk, with catalysts being a possible application. The network support will among others include organizing and tutoring a number of hands-on courses for CASTEP (widening participation by going outside the usual theory community, and engage with potential experimental and industrial users of our codes), KKR, QUESTAAL, and CRYSTAL.

## *CCPmag – Computational Multiscale Magnetism*

Magnetism and its microscopic understanding are of high importance in a number of vital technologies, starting from the energy sector, such as permanent magnets in the generators, to computer technology, for instance magnetic RAM, hard drive technology. Modelling magnetic properties can speed up the process of developing novel materials for these applications.

The UK has mature communities in magnetic modelling at various length scales, as well as a strong experimental community, mainly around the ISIS neutron scattering facilities at Rutherford Appleton Laboratory (RAL). This embraces simulations on different length scales, starting from ab initio calculations of magnetic properties (based on the full quantum mechanical description), over atomistic modelling, representing magnetic materials by spin models, up to micromagnetics, utilising finite element techniques to describe larger systems or whole devices.

The Collaborative Computational Project on Computational Magnetism was established in 2015 to bring researchers from these communities together and make the expertise of ab initio electronic structure codes, partially developed within the CCP9 community) available to them.

### Summary Report (1 April 2019 – 31 March 2020)

Kun Cao is working on the implementation of the calculation of magnetic exchange constants in the full-potential plane wave pseudopotential method (Quantum Espresso) code. The calculated J's can then act as input for atomistic/micromagnetic simulations or as guidelines for experimentalists to fit their data. It would benefit both the community of spin dynamics simulation and inelastic neutron scattering.

Within the ISIS project, Kun Cao has finished calculations of dynamic spin susceptibility of Fe along experimentally measured high symmetry lines in the BZ using his implementation based on Quantum Espresso. In order to get a quantitative comparison between theory and the experimental data, Kun

Cao has now also completed additional calculations along selected off-symmetry lines to take into account the experimental resolution of wave vectors. This paves the way to create an implementation to allow routine calculations of spin fluctuations for both theorists and experimentalists, which would act as a standard tool to study spin fluctuations by the community.

The collaboration between Kun Cao and Nick Brookes from ESRF on a RIXS study of magnetic excitations in Fe and Ni has produced a manuscript which has been submitted to PRB. We have prepared a reply to the referee report and revised our manuscript correspondingly. We are about to resubmit the paper. Kun Cao is collaborating with Prof. Lixin He (USTC, China) on magnetic phase diagram of quasi-two dimensional Fe3GeTe2. We have written up a manuscript and are about to submit it.

### Outline Plans 2020/21 (1 April 2020 – 31 March 2021)

*(Detailed plans and milestones are included in the Technical Appendix for reference)*

CCPmag was not renewed as part of the 2019 EPSRC call "Collaborative Computational Projects: Community Building, Networking and Core Computational Science Support".

## *CCPNC – NMR Crystallography*

Nuclear Magnetic Resonance (NMR) is a useful technique to determine chemical structure, especially in compounds of which it is hard to produce single crystals big enough for diffraction techniques, as commonly found in organic molecules. NMR Crystallography is the technique of using quantum-mechanical simulations to predict NMR spectra to a high degree of precision, and combining this with experiment to open new ways of exploring structure in not yet understood crystals, such as new pharmaceuticals. CCP-NC has the objective of disseminating and promoting this approach throughout the experimental NMR community in the UK and worldwide.

### Summary Report (1 April 2019 – 31 March 2020)

The Soprano library has been maintained and upgraded further, reaching version 0.8.7.1. As the library is now mostly mature, this has consisted mostly of ordinary maintenance and bugfixes, plus the inclusion of a new powder averaging method for NMR spectra (SHREWD). In addition, the ability to identify special Wyckoff positions in periodic systems has been added, a new, more convenient interface for NMR tensor data has been created, and work on tutorials in the documentation has started.

The Soprano library is used to manipulate files out of NMR calculations and random structure searching calculations. It replaces and expands on the old MagresPython library and is already in use especially in the St. Andrews group. It's also a key element of our work with the muon spectroscopy community as it provides e.g. random generation, treatment of periodicity, and so on. It is also part of the back-end for the CCP-NC magres database.

The CCP-NC database website has been finally released and is available at https://www.ccpnc.ac.uk/database/. The last stretch of development has required some additional functionality. Search functionality has been expanded and empowered with new fields. In addition, in order to support the broader needs of the user base, the possibility has been introduced for users to choose the license they want to apply to their data upon submitting it. Support for the Creative Commons CC-by license has thus been added. Work is undergoing to agree on a final schema for the database entries themselves, and to refactor some parts of the code that need it.

The new regularized SCAN exchange correlation functional (rSCAN) has been developed and fully implemented in CASTEP in time for the release of v20.0. Papers have been published documenting the

development process, benchmarking the efficacy of the functional, and highlighting the modifications to the original SCAN meta-GGA functional made in order to enhance its performance and allow it to be numerically stable when generating pseudopotentials across the periodic table. The new functional will allow for more accurate prediction of band gaps and NMR parameters, thus enhancing the usefulness of CASTEP for the community.

The library tools developed together with the muon community under the name of “pymuon-suite” have been further expanded. Existing functionality has undergone extensive redesign to provide a more unified and easy to use interface. The capability to generate statistical samplings of configurations starting from the results of phonon calculations in order to average observable quantities with quantum effects in the harmonic approximation is of particular interest for the NMR community as well. In addition, new functionality has been developed to compute the dipolar field distribution at a given site in presence of a random distribution of dipoles, which especially in magnetic materials has useful applications for NMR too.

Work has been done on developing a method for the averaging of NMR frequencies and couplings, both in an infinite speed limit and in dynamical conditions. These are relevant because CASTEP produces static results, but in real life, systems are often dynamic due to vibrations but also more complex processes like rotations of CH3 groups. The idea here is to have a series of tools to do the post-processing step that allows people working especially with organic crystals to get simulated spectra closer to their actual experimental results by incorporating dynamical effects. This is still mostly in development. There's also an old software called EXPRESS that was used for similar purposes that's now discontinued, and one of the goals stated in the CCP-NC proposal for renewal was to rewrite it or replace it with something else.This has resulted in various scripts and a paper that has been submitted to Physical Review E for review.

## Outline Plans 2020/21 (1 April 2020 – 31 March 2021)

*(Detailed plans and milestones are included in the Technical Appendix for reference)*

In 2020, the CCP-NC will renew its efforts in supporting the solid state NMR community while branching out in new specialised branches of NMR crystallography, specifically concerning disordered materials. Existing tools will be maintained and expanded upon in order to better satisfy these needs. In particular, the library Soprano will be fully merged with MagresPython and will be expanded with more functionality focused around the generation and analysis of disordered structures, such as defective or amorphous materials. In addition, MagresView will be updated to keep up with modern standards of web design and expanded to support these new interests. The development of MagresView 2.0 will come back into focus after having been put on hold due to the reduction in available FTEs during the 2019-2020 period.

# *CCPQ – Quantum Dynamics in Atomic Molecular and Optical Physics*

The remit of CCPQ is to develop code for quantum dynamics simulations, solving the time-dependent Schrödinger equation for both heavy (nuclei) and light (electrons) particles. It has 3 main community codes covering different areas; 1) TNT: A package coding Tensor Network Theory for coherent many-body nuclear dynamics, 2) R-Matrix Suite: A set of programs for electron (positron) -atom and -molecule scattering, (ultrafast) laser pulse interactions and related problems. The package is also currently being adapted for ultracold molecule-molecule (atom-atom) interactions, bound states and resonances (RMAT_REACT), mainly through HEC UK-AMOR but with CCPQ CoSeC support and 3) Quantics: A package based on the MCTDH algorithm for molecular quantum dynamics.

The R-Matrix suite is a set of programs for electron (positron) -atom and -molecule scattering, (ultrafast) laser pulse interactions and related problems. The scientific goals are to: study attosecond physics in

atomic and molecular systems; compute fusion-relevant atomic data; study electron-molecule collisions with relevance to biological radiation damage, fusion energy and technology.

The TNT suite considers strongly correlated dynamics of nuclei in low temperature solids and other supercooled systems. The Quantics package studies reactive molecular collisions (chemical reactions) in full ab initio detail. Apart from general scientific understanding of the structures and reactions, these packages are also used to simulate the behaviour of the interacting components of quantum computers.

The collaboration has also had related code development in the field of coherent dynamics in the CCE (Cluster Correlation Expansion) code and in the field of anti-matter scattering related to the R-Matrix work. It also supports codes and research in low-energy antimatter interactions (positron and antihydrogen collisions with atoms and molecules), with input in this area from Aberdeen and Belfast, the latter employing many-body-theory diagrammatic techniques.

## Summary Report (1 April 2019 – 31 March 2020)

CoSeC CCPQ support work for 2019-20 followed planned milestones with some changes due to circumstances and new milestones added. Martin Plummer took part in management meetings for UK_AMOR, the R-matrix HEC, and he and Andrew Sunderland attended the CCPQ Software Flagship Project code workshop in June, which reached out to diverse R-matrix communities and beyond to other scientists and industrial applications. Martin organized, with Ben Dudson (York) and Colin Roach (CCFE) of CCP/HEC-PLASMA, a new 3 day workshop to present state of the art use and exploitation of ab initio atomic, molecular and optical physics (AMO) data in the plasma community with applications to nuclear fusion, industrial plasmas, low temperature and astrophysical plasmas. The meeting, with around 20 speakers, was to take place at CCFE on 1-3 April 2020 as a major step to bring the two CCP communities and their experimental and industrial counterparts together, to explore new calculations and how AMO data can be targeted for use in larger scale simulations and experimental diagnostics. Postponed due to COVID-90, it should take place in 2020-2021. Martin attended ECAM's Extended Software Development Workshop on Quantum Dynamics (Durham, July 2019) with Software Engineering Group (SEG) members Alan Kyffin and Gemma Poulter who devised and presented a course to 30 chemists and physicists on SEG's 'ANVIL' continuous integration and testing (CIT) tool for version control, sustainability, cross platform porting and validation/verification of large community code packages.

Martin presented posters at the International Conference on Photonic, Electronic and Atomic Collisions (ICPEAC 2019, Deauville) and its satellite POSMOL (positronic interactions and electron molecule collisions, Belgrade), on double-continuum R-matrix theory, on the molecular R-matrix code package UKRMol+ and Andrew's memory optimization work, and antimatter theory of Ps- formation in He anti-H collisions (coordinate transformations in multidimensional integrals). These boosted CoSeC's international recognition and generated interest and discussions.

Exploitation of the CoSeC double-continuum work and codes is part of a successful EPSRC proposal submitted by QUB with co-Is at UCL and Daresbury Laboratory (Martin), including a request for SEG sustainability support, on double ionization processes in laser atom physics combining the ab initio Belfast RMT code with an analytic method (the Coulomb quantum orbit strong field approximation). The 3-year project will begin in May 2021. The research will shed light on photoabsorption processes, yielding deeper understanding of fundamental atomic-scale light-matter interactions. This will have substantial impact on experimental techniques such as laser-induced electron diffraction and high-harmonic spectroscopy. The team were also awarded a CECAM international 3-day workshop for July 20. The workshop, Quantum Battles in Attoscience, designed to attract diverse communities of approaches to attosecond physics for atoms and molecules, will now take place in July 2021, with a virtual preliminary workshop at the start of July 2020. CoSeC support also assisted the PI and steering

panel in the 2019 CCP Call for the renewal of CCPQ and developed new work packages for core support. Extremely disappointingly, this proposal was not successful, despite its quality and ambition. The CCPQ community will continue to interact informally and via its website until new funding is found, and through the UK-AMOR consortium.

Coding and scientific development continued. Andrew completed and uploaded his memory optimization work to the UKRMol+ GitLab repository, enabling much larger electron collision calculations with detailed structure and larger molecules, of interest in industry and medicine, to be made. The UKRMol+ developers asked him to extend this further over a greater range of possible calculations with broader parallelism. This took place in the final quarter of the year to be completed in April 2020. Alan completed transfer of the full TNT codebase and repository to GitLab with TNT code manager M Lubasch (Oxford). They also produced a complete Python interface for non-specialist users. Andrew completed work with the RMT developers on optimising code adapted for GPU BLAS, allowing faster, more responsive calculations. Martin added further functionality to the RMAT_REACT adaptation of DL's 'PFARM' code for UCL's ultracold chemistry simulations, along with optimization and user-interface work on the double-continuum code and antimatter work. He helped with a tribute meeting for Professor PG Burke organised by Paul Durham (again, postponed).

## Outline Plans 2020/21 (1 April 2020 – 31 March 2021)

*(Detailed plans and milestones are included in the Technical Appendix for reference)*

CCPQ was not renewed as part of the 2019 EPSRC call "Collaborative Computational Projects: Community Building, Networking and Core Computational Science Support".

The original plans for CCPQ in 2020-2021 assumed that CCPQ would be funded. As it was not funded (a grave disappointment given the quality of the proposal), bulk CoSeC funding was withdrawn from CCPQ. There is a small amount of residual carry-over funding left for administration, organisation of postponed meetings, back-up as required of AQuA-DIP (or UK-AMOR) support including continued one-to-one training of related PhD students and early career PDRAs, ongoing and new proposals for funding and shorter-term follow-up of existing milestones during 2020-2021.

# *CCP-Plasma – HEC-Plasma Physics*

CCP-Plasma includes researchers from UK universities, the Culham Centre for Fusion Energy (CCFE), the Rutherford Appleton Laboratory (RAL) and AWE. The study of plasma physics covers a huge range of scales and applications. It is core to the development of laboratory experiments such as fusion power, new light sources and the next generation of particle accelerators. On the largest scales it is fundamental to our understanding of astrophysics.

CCP-Plasma was established in 2007 with the aim of pooling the collective expertise across these disparate subjects and developing core plasma physics simulation codes, and training packages, for UK science. CCP-Plasma supports three core codes EPOCH, BOUT++ and GS2. We also have one Flagship project for the development of a radiation-hydrodynamics ALE code for laser plasmas (Odin). 1FTE of CoSeC core support is split evenly between the codes GS2 and BOUT++, which both focus on modelling plasma in magnetic confinement fusion devices. This CoSeC support is funded 20% from the HEC-Plasma Consortium until 28/5/2023.

The Plasma High-End Computing (HEC) Consortium is an EPSRC funded collaboration of UK plasma physics researchers. The Plasma HEC supports research in the simulation of plasmas, including those plasmas of relevance to magnetic confinement fusion, laser-plasma interactions, and laser fusion energy. The software development includes a commitment to optimizing key codes, with input from EPCC, RAL and Warwick Computer Science, and developing new physics packages as required to maintain the UK's role in laser-plasma physics and all approaches to fusion for energy. The consortium

supports meetings and software development. It also manages a block allocation of national super-computer time to be used for plasma physics research.

## Plasma Summary Report (1 April 2019 – 31 March 2020)

The project receives 1FTE of support from Joseph Parker, divided equally between the codes GS2 and BOUT++. Work in this reporting period has been delayed, owing to 5 months paternity leave taken by Joseph Parker.

Minimizing plasma turbulence in tokamaks (nuclear fusion devices) is vital for achieving fusion, but simulations of fusion plasmas are computationally very expensive. GS2 is a plasma turbulence code which has been developed since the 1990's and is already highly optimized; however further optimizations are required for simulations to include all physical effects and to resolve necessary space or time scales.

In collaboration with OeRC, we have optimized GS2's inherently load-unbalanced implicit field solve step which is the current scaling bottleneck. We have distinguished between communication required for the calculation, and communication that is only required for certain diagnostics, removing unnecessary communication. We have implemented a strategy to allow some regions of the domain to replicate small computations (rather than work-share) when the cost of sending small amounts of data outweighs the gain from parallelization. Our new approach is over x2 faster than the old approach at O(3000) cores for typical problems, and continues near-ideal scaling beyond O(10000) cores, while the old approach stopped scaling at O(6000) cores.

Realistic simulations of future fusion power plants requires that plasma modelling (such as that provided by BOUT++) includes the modelling of atomic and molecular interactions. Codes and databases exist for the latter, but none are adequately coupled to plasma modelling codes. We have assessed the available atomic/molecular codes and also discussed with colleagues at Culham the possibility of developing a new code. We are now involved in a EUROfusion project to develop a new plasma code which would include atomic/molecular effects. Our assessment of the atomic/molecular codes has fed into the design specification for this European code. We anticipate that molecular effects will be incorporated into BOUT++ as a prototype for this new code, and have therefore delayed the code coupling aspect of this work to coordinate with the European project.

Simulations of the tokomak edge require the inversion of elliptic operators to obtain the electrostatic potential from the vorticity and density. Work on operator inversion in BOUT++ in the last reporting cycle allowed CCFE to undertake realistic, fully 3D simulations of turbulence in tokomk plasmas. However, the inversion remains the performance and scaling bottleneck. This is because the 3D solver is based on iterations in two dimensions, with a direct solver used in the remaining dimension. The direct solver requires all-to-all communication, and therefore does not scale. We have developed the theory for a novel iterative solver for the remaining dimension that should scale to high core counts. We have implemented this and are currently profiling and optimizing. Improved scalability will allow realistic simulations of tokomak turbulence to utilize more cores, significantly reducing time-to-solution and allowing code users to perform more high-fidelity simulations.

## Outline Plans 2020/21 (1 April 2020 – 31 March 2021)

*(Detailed plans and milestones are included in the Technical Appendix for reference)*

When Archer 2 begins service in 2020, it is vital for our users that the transition from Archer is seemless. We will port our codes, perform benchmarking, provide user guidance and make first-order optimizations. Our subsequent work on GS2 will focus on increased scalabilty. At present, GS2 distributes the 5D arrays containing particle distribution functions, but smaller 3D arrays containing electromagnetic fields are replicated on each MPI rank. This replication is unnecessary for computation

and increasingly wastes memory on larger real-space grids. Scalability is increasingly hampered by communications overhead to update EM fields on all ranks. We will refactor the code to remove the need for each rank to store EM fields. To fully exploit the parallelization of the electromagnetic fields, we will also reimplement GS2's diagnostics module to use parallel I/O.

BOUT++ work will also focus on efficiency and scalability, as detailed simulations in full tokomak geometry remain challenging. We will optimize current implementations of the elliptic and parabolic operator inversion routines for calculating potentials, and for the preconditioning of fully implicit schemes. These routines are the current scaling bottlenecks. This work naturally leads on to work packages for subsequent years on improving parallelisation strategies, and merging loops to improve cache efficiency and scaling.

## *CCPi – Tomographic Imaging*

The CCPi was established in 2012 to support the emerging UK tomography community with a toolbox of algorithms to increase the quality and level of information that can be extracted by computed tomography. In 2015 four major parts were defined to enable modular development of CCPi codes: pre-processing techniques for image calibration and noise reduction, reconstruction techniques to create a 3D volume data set from projections, segmentation and quantification techniques that can extract relevant objective values, and a software framework to enable exploitation within a wide range of existing commercial and open source software environments.

The size of this community has continued to grow with many academic groups around the UK taking up tomographic imaging and purchasing new lab based x-ray CT scanners. In 2012 there was an estimated 50,000 CT imaging sources around the world. The size of our community has arisen from ~250 in 2013 to just over 400 in 2018.

Our focus aims to bring together the UK imaging community in maximising the return on investment in imaging software through developing, maintaining, sustaining and prompting the CCPi core imaging toolbox. Primarily we concentrate on advancing the state-of-the-play of the field through: 1) reducing the barrier to access in multi-modality image analysis algorithms, be the data coming from mid-range facilities hosted by the universities or from large scale synchrotron or neutron facilities; 2) improving the accessibility and distribution of the codes; and 3) establishing a national multidisciplinary image analysis focal point for the multidisciplinary community comprising of algorithm developers, material scientists, instrument manufacturers, and instrument scientists.

The CCPi core staff actively involves in three major annual imaging events in the country, each having 50+ attendees. These include an X-Ray user group symposium (ToScA) managed by the National History Museum (NHM) and Royal Microscopical Society (RMS), a technical forum supported by RCaH and DLS; and a "dimensional XCT" conference supported by NPL that is leading to formal BSI/ISO standards.

### Summary Report (1 April 2019 – 31 March 2020)

The team is focusing on the development of the Core Imaging Library (CIL), a toolbox of algorithms and tools for tomographic imaging. This includes algorithms for pre-processing, reconstruction, quantification, segmentation, and visualisation. This is the first attempt in the UK to improve the take-up of iterative reconstruction algorithms by the community. This challenge is being achieved by: simplifying the software installation process; improving documentation; releasing sample code and demos; as well as making group specific readers and writers available. During the reporting period there have been two CIL releases version (19.07, 19.10) adding functionality. With CIL we aim at reach more

scientists/users by simplifying the installation and distribution and by providing a simple Python interface to a variety of codes and methods that have been developed within CCPi

Current development in tomography beamlines in synchrotrons, neutron sources and lab machines is focussed on 4D/5D tomography. This includes 3D imaging of a sample while changing its status (e.g. its temperature) and multi-channel (or spectral) imaging, where sources of different energies can be selectively acquired. This means that there is a push for faster acquisition and energy selective imaging which inherently reduce the signal-to-noise ratio, SNR. For instance, the ISIS/IMAT beamline can resolve up to 2300 different energies of the impinging neutrons, dramatically reducing the statistics of the collected data per energy with regard to the “white beam” acquisition. In X ray imaging the dose delivered to the sample may change its characteristics. That’s why it’s necessary to develop methods that are resilient to low SNR data acquisition, which is something the users are going to be interested in. Also, there’s a trend in acquiring dynamical data. In this setting we expect to be imaging the sample multiple times; in this case the acquired data will suffer from low SNR. Again methods that are resilient to low SNR are necessary

The main additions to the CIL are addressing the challenges introduced by low SNR data, and spectral (multichannel) tomography. The CIL includes an implementation of a new reconstruction framework for iterative methods; ‘the regularisation toolkit’ for de-noising and regularisation during iterative reconstruction; and ‘TomoPhantom’ a software for generation of synthetic phantoms for CT to be used as benchmark of reconstruction algorithms.

In the reporting period efforts have been spent to enable the use of the CIL reconstruction framework with the CCP PET/MR software SIRF. This allows the use of novel reconstruction algorithms by the PET/MR community. The development with different community is a large push for producing better codes. The PET/MR community, focussed on medical imaging, is developing a similar framework to that of one part of the CIL. However, the CIL has a specific package for optimisation which PET/MR lacks. The PET/MR has clearly benefitted from the fact of being able to use the CIL to perform reconstructions that were not yet available to their community. The CIL, and therefore the CCPi community, benefitted of the interaction with PET/MR as we could readily utilise coding conventions, software structures that had been already discussed and thoroughly tested. Additionally, by trying to be open to a different community, we have been able to produce a code that is more versatile than it was initially designed for: not only can we apply our methods to data from Xray/neutron, but also from PET and MRI.

In preparation for the co-organised Symposium on Synergistic Reconstruction (3-6 Nov Chester) the team participated to two CCPPETMR hackathons and organised 1 hackathon and a training course on the CIL. The symposium consisting in a 2 days scientific session followed by a software training (SIRF and CIL) attracted a participation of 100 and 50 people, for the scientific and training sessions respectively.

In December we ran a user orientated hackathon with participation from ISIS, DLS i14, UoM this included an interactive training session on the CIL, and allowed us to assess their requirements and begin to broaden our user base. This was followed by a hackathon in March focussing on the flagship development on multispectral tomographic reconstruction with the developments well received by IMAT.

In the reporting period we have also co-organised a series of events: in June, the XCT-fringe meeting and 2-day Turing summer school at RAL and the fourth annual workshop on Advances in X-Ray Imaging (UoM, 3Dimgination, DLS) at DLS; in September ToScA. However these have since been cancelled/postponed due to the pandemic.

The core support effort (1.56 FTE) was provided by Edoardo Pasca, Gemma Fardell with external contribution by Martin Turner (University of Manchester, SCD Visiting Scientist).

**Problems, issues and difficulties encountered during reporting period:** Future community events including training days and hackathons have been postponed due to covid-19.

## Outline Plans 2020/21 (1 April 2020 – 31 March 2021)

*(Detailed plans and milestones are included in the Technical Appendix for reference)*

During the years 2017/2019 a large effort has been put in the new development of the iterative reconstruction framework, also in collaboration with CCPi flagship. We expect a period of lower development activity, more focused on deployment at university labs and facilities like DLS, ISIS/IMAT and CLF/EPAC. Also, we will focus on optimisation of the current code base of the DVC tool and distribution to wider community. Regarding new development, we will focus on investigation of novel methods based on ML/AI for computed tomography.

# *CCP PET-MR - Positron Emission Tomography (PET) and Magnetic Resonance (MR) Imaging*

For medical imaging, the UK is a globally leading country. It has the highest number of Positron Emission Tomography and Magnetic Resonance (PET-MR) medical imaging machines per capita in the world, evenly spread throughout the country. The CCP PET-MR project established in 2015 aims at bringing together the best of the UK's PET-MR imaging expertise to capitalise on the investment in this area. New research shows that the use of MRI intermediate results can improve PET imaging quality and vice versa, and latest scanners can acquire MR and PET data simultaneously. Our CCP is dedicated to exploiting exciting new capabilities that the synergy of MR and PET imaging can deliver. The main deliverable of the project is an open source PET-MR reconstruction software framework we named SIRF (Synergistic Image Reconstruction Framework). SIRF is simple enough in use for educational and research purposes, thus reducing the "barrier for entry" for new contributors to PET-MR imaging research and development, and at the same time powerful enough to process real scanner data.

STFC CoSeC support for this CCP currently focusses on developing the SIRF code base that provides an easy-to-use script-language (Python and Matlab) environment built around existing open source reconstruction software. This includes maintaining network, website, running workshops and training courses, on top of the software engineering effort that contributes to SIRF development, testing, deployment and documentation.

## Summary Report (1 April 2019 – 31 March 2020)

Our work during the reported period progressed according to the job plan: software development and engineering efforts, adding content to our website www.ccppetmr.ac.uk, maintaining our steadily growing mailing lists (we now have 101 members on the CCP-PETMR announcement list, 34 on the developers and 72 on the users lists), organising working group and executive meetings, organising a series of well-attended Developers Days to present and discuss progress in our software framework development, and organising the third and fourth CCP PETMR Hackathons (also attended by members of CCPi).

In the reported period we published Release 2.0 of our Open source software suite SIRF on 24 May 2019, Release 2.1 on 20 Nov 2019. These releases are based on the PET reconstruction package STIR (Software for Tomographic Image Reconstruction) and the MR reconstruction package Gadgetron and are capable of processing measured data from the Siemens PET-MR scanner, with work-in-

progress to support the GE PET-MR system. Major new advances achieved in SIRF 2.0 and 2.1 are the addition of the image registration capability (work led by the CCP PETMR Flagship project), a crucial step in enabling synergy between PET and MR modalities, and integration with the CCPi Core Imaging Library (CIL). CIL contains advanced regularisation techniques and recent algorithms for optimisation of non-smooth objective functions. CIL largely adapted SIRF terminology and (Python) class structure, with some aliases now introduced into SIRF. This joint effort of CCP PETMR and CCPi was facilitated by joint CoSeC staff between the 2 projects.

Major highlights of this reporting period were our next three CCP PET-MR Hackathons. Hackathon 3 was held at University of Hull 24-26 July 2019, and it was attended by 10 developers from KCL, UCL, Leeds, Hull and STFC, including CCPi developers. The hackathon was organised around 4 themes: (i) try to use the SIRF in an HPC cluster SCARF, (ii) image registration, (iii) denoising images, and (iv) GPU projector in STIR and Gadgetron GPU gadgets. Participants were divided in the 5 groups (two groups for (iv)), with occasional cross-checks and CoSeC staff and the PI floating between groups. We had good interaction with CCPi, and the feedback from the participants was excellent. Hackathon 4 was held at University of Bath 23-24 Sep 2019, and it was attended by 10 developers from UCL, Bath, Manchester, Leeds, NPL and STFC, including 2 CoSeC staff. The Hackathon started with a half-day training for new SIRF users, and then proceeded to working on 3 themes: (i) try to reconstruct PET dataset with CIL implementation of FISTA algorithm, (ii) finalise adding CIL in SIRF-SuperBuild, (iii) incorporate Hybrid Kernel algorithm from STIR into SIRF, and (iv) get basic synergistic functionality up and running and prepare demos for the training in November. On 27-29 Jan 2020, we held our 5th Hackathon, hosted by STFC at Coseners House, Abingdon. This Hackathon, attended by 9 researchers from STFC, UCL, MLU (Germany) and PTB (Germany), was focused on Motion Correction Image Reconstruction. We had to abandon our 6th Hackathon, planned for 16-18 Mar 2020, because of the COVID-19 pandemic.

Another highlight was the one day PET-MRI School for students and early stage researchers at PSMR 2019, 8th Conference on PET/MR and SPECT/MR, 15–17 April 2019, Munich, Germany (28 participants) with a hands-on PET-MR software training session using SIRF.

On 3-6 Nov 2019 we held Synergistic Reconstruction Symposium, organized jointly by CCPPETMR and CCPi, followed by 2-day software hands-on training session introducing our Synergistic Image Reconstruction Framework. The Symposium was attended by about 100 people, and the training by 40.

CoSeC staff has led the submission of a journal paper on SIRF: E. Ovtchinnikov, R. Brown, et al., SIRF: Synergistic Image Reconstruction Framework, published in the special 50th anniversary issue of Computer Physics Communications in Dec 2019.

Finally, our major achievement is the extension (and increase) of support from EPSRC for further 5 years, our CCP being renamed CCP on Synergistic Reconstruction in Biomedical Imaging to reflect widening of the scope of investigation and development.

Our development of SIRF provides PET-MR community - for the first time ever - with self-contained user-friendly software environment that facilitates the development of synergistic PET-MR imaging methods and testing them on real data. The ensuing acceleration of the development of efficient reconstruction algorithms will eventually translate into higher accuracy of scanners images, helping to achieve better quality and longevity of life for people suffering from cancer dementia and other serious illnesses.

## Outline Plans 2020/21 (1 April 2020 – 31 March 2021)

*(Detailed plans and milestones are included in the Technical Appendix for reference)*

In 2020/21 we plan to continue our software development effort aiming at joint motion estimation, list mode reconstruction without conversion to sinograms, time-of-flight reconstruction and motion-guided reconstruction. We plan to have some of these features in SIRF Release 3.0, to be published in the last quarter of 2020.

We will continue our engagement with the CCP PET-MR (now CCP SyneRBI) community by maintaining our website and mailing lists, organizing meetings, developers' days and other events. We will continue to organize and support training courses and developers' workshops and assist in new proposal writing. Embedding within Institute of Nuclear Medicine at UCL Hospital for one or more days a week will be maintained, and similar arrangements with KCL will be sought.

## *CCPBioSim - Biomolecular Simulation at the Life Sciences Interface*

CCPBioSim is the Collaborative Computational Project in biomolecular simulation at the life sciences interface, bringing together chemists, physicists and chemical engineers as well as researchers from all branches of "molecule-oriented" biochemistry and biology. Simulations help to analyse how enzymes catalyse biochemical reactions, and how proteins adopt their functional structures e.g. within cell membranes. They contribute to the design of drugs and catalysts, and in understanding the molecular basis of disease. Our aim is to involve experimentalists and computational specialists in this work, sharing the belief that the best science can be done when theory and experiment are closely integrated. CCPBioSim engages with early career researchers and non-experts through the provision of tutorials and workshops enabling them to become proficient and productive users of biomolecular simulation techniques. We are also actively engaged in developing new advanced methods, which in future will be used by our community to deliver new and exciting science.

From April 2020 onwards CCPBioSim is supported by 1.95 FTE of SLA core effort. Of this 1.5 FTE is allocated to core software development, including 1.0 FTE towards a portfolio of short, community-driven software projects and 0.5 FTE for longer-term maintenance of CCPBioSim software and flagship projects. 0.25 FTE supports coordination of SLA effort, website maintenance, mailing list management and networking activities, while the remaining 0.2 FTE is for infrastructure support including our cloud-based training platform. For further information please see www.ccpbiosim.ac.uk.

### Summary Report (1 April 2019 – 31 March 2020)

The cloud-based infrastructure for hosting workshop material in Docker containers has been developed with James Gebbie-Rayet (HECBioSim). Our new training can be accessed through our website (www.ccpbiosim.ac.uk/training). This online resource allows members of our community to access our training material at any time from any place giving greater flexibility to the training options available.

We supported the May training week in Bristol and three workshops in early September with registrations and the cloud-based training infrastructure. We also provided support with the online practical session for a QM/MM workshop in Columbia on 12 September 2019. Sarah Fegan along with Richard Henchman (University of Manchester) and Anđela Šarić (UCL) taught the advanced course in biomolecular simulations as part of the CCP5 Summer School. This course (16-18 July 2019) was attended by 20 students.

Sarah assisted with the organisation for the CCPBioSim Annual Conference held in Bristol 4-6 September 2019. Sarah also led the organisation of the 4th Manchester Multiscale conference (30 March – 1 April 2020) while Tom Keal was on shared parental leave, and Tom rejoined the organising committee in November. This conference series is jointly run by CCPBioSim and CCP5 and this time

was also sponsored by the CECAM Daresbury Node. There were nine confirmed invited speakers: Maria Fyta (Stuttgart, Germany), Frauke Gräter (Heidelberg, Germany), Syma Khalid (Southampton, UK), Horst Lillig (Greifswald, Germany), Céline Merlet (Toulouse, France), Irina Paci (Victoria, Canada), Lars Pastewka (Freiburg, Germany), Anđela Šarić (UCL, UK), and Rebecca Wade (Heidelberg, Germany). The conference had a very successful registration period, being fully booked with almost 100 registrations. Unfortunately, the conference had to be cancelled in the light of the coronavirus situation, but we would hope to rerun it at a later date if finances permit. Tom also organised an associated ChemShell training workshop to be held at the University of Manchester straight after the conference (1-2 April 2020), and again was in high demand with the capacity raised on two occasions up to 40 attendees. Unfortunately it too had to be cancelled, but will be run again as soon as circumstances permit.

The multiscale modelling of cytochrome P450s project continues under the supervision of Tom with Kakali Sen and Ya-Wen Hsiao working on the atomistic and QM/MM parts of the workflow and Sarah working on the coarse-grained simulations. This work builds on the original CCPBioSim flagship project protocol and is investigating potentially clinically important drug-drug interactions in P450 systems involving warfarin and dapsone. Initial CG simulations and atomistic MD on the warfarin, dapsone and combined dapsone/warfarin P450 systems have been carried out and we expect the project to result in a series of publications before it ends in October 2020.

Work started on the joint project with CCP4/CCP-EM on the conformations of gylcosylated proteins. A dimer of N-glycosylated beta-D-glucosidases (PDB ID: 5FJI) was chosen as a first system. Atomistic MD simulations have been setup using CHARMM-GUI and the Charmm36 forcefield and the initial simulations used the GROMACS program. Plans have been made to look at running simulations with OpenMM and integrating with existing CCP-EM tools and workflows, which will allow the project work to be run more efficiently and help similar projects in future.

Investigation into the integration of Lomap with FESetup revealed that the way in which FESetup handles input files will have to be restructured. This is a bit more complex than originally expected, but should lead to an improved underlying structure to FESetup which will make adding additional free energy functionality more straightforward for the benefit of the community. The FESetup documentation has been made available online and an FESetup email list is in place to provide user support. Links for documentation and user support can be found at www.ccpbiosim.ac.uk/software.

## Outline Plans 2020/21 (1 April 2020 – 31 March 2021)

*(Detailed plans and milestones are included in the Technical Appendix for reference)*

Following renewal of CCPBioSim, work will begin on the first short project involving meshing the EMDB for mesoscale simulations in collaboration with Sarah Harris and CCP-EM. A second short project will begin in November.

Ongoing maintenance of FESetup will include an upgrade to Python 3. Work on the multiscale P450 and the gylcosylated proteins projects will continue until the end of the CCP cycle in October with publications expected.

All training events are currently on hold due to the coronavirus situation, but will start up again as soon as circumstances permit. These will include a Py-ChemShell training workshop in Manchester, the regular Bristol training week, the advanced course at the CCP5 Summer School and other workshops to be decided. Our self-guided online training remains available for all members of the community.

# *MCC – Materials Chemistry Consortium*

The Materials Chemistry Consortium exploits high end computing in a broad programme of work modelling and predicting the structures, properties and reactivities of materials. The consortium is a broadly based but coherent grouping comprising 36 university groups, with the emphasis on modelling at the atomic and molecular level but with growing links to models at larger length and time scales. Founded in 1994, the current scientific programme is built around seven related themes: catalysis, energy storage and generation, surface and interfacial phenomena, nano- and defect structures, soft matter, biomaterials, environmental materials. The Consortium has an active programme of code support, development and optimisation, tapping into the ecosystem of UK based software development initiatives including CoSeC.

CoSeC supports the consortium across the range of techniques used by its members, embracing both force-field methods employing static and dynamical simulation methodologies and electronic structure methods with a strong emphasis in recent years on Density Functional Theory (DFT) techniques employing both periodic boundary conditions and embedded cluster implementations. The four main codes supported by CoSeC are: DL_POLY, DL_FIELD, ChemShell, and CRYSTAL.

## Summary Report (1 April 2019 – 31 March 2020)

**ChemShell/DL-FIND:** John Purton has written a prototype interface to DL_MONTE functionality in ChemShell, allowing Monte Carlo calculations to be run within the ChemShell framework. The DL_MONTE toolkit, dlmontepython, has been adapted to allow ChemShell to make use of its class library, enabling combined workflows in either ChemShell or dlmontepython. Future work will target complex workflows involving Monte Carlo stages and abstracting the Monte Carlo simulation driver so that other energy evaluators can be used. John has also developed an interoperability mechanism for ionic material forcefields for DL_POLY, DL_MONTE and GULP. A small fraction of MCC time remains allocated to Tom Keal to supervise the ChemShell work package in the SAINT flagship project, which continues to progress according to plans. John Purton has taken over delivery of the work package, which in this period has included implementation of the Chemical Reaction Dynamics method of Meyer et al and testing of the mechanisms with an empirical potential using GULP. Work is ongoing to integrate the QM package CP2K to perform full QM/MM calculations of metallic systems. As part of this work John wrote a simple (NVE) python-based MD code within ChemShell for testing purposes which may be extended in future. Both Tcl-ChemShell and Py-ChemShell continue to be maintained on ARCHER and THOMAS for use by the MCC community. Altogether, these developments increase the range of materials that users can study and methods they can use in their own QM/MM simulations.

**DL_POLY/DL_FIELD**: The allocated effort during this period was rescaled, with all DL_FIELD work put on hold until April 2020. The re-scoped effort was used to supply supervision by Ilian Todorov over the DL_POLY_4.10 refactoring exercise. The refactoring process has led to the creation of an extensive testing suite of different model systems, testing consistently feature after feature as available in the manual as well as testing the limits of a number of concepts in the package. As an outcome of this exercise a number of issues have been filed in the project GitLab repository and addressed over the reporting period. These have led to a large set of improvements (including a Py-ChemShell driver), updates and bugfixes, some of which were carried out in collaboration with Kostya Trachenko's team at QMUL. The outcome of this effort was first prelease in the summer with a full update 4.09.3 planned for November 2019. The DL_POLY and DL_FIELD projects were showcased at the training, provided at the international summer school at Antofagasta (Chile) in May. The school attracted over 30 researchers from 6 Latin American states. This effort has substantially increased the robustness of the code which will result in more reliable simulations for users

**SAINT:** Barry Searle is assisting the core SAINT development team with the generation of symmetrised surface models from bulk unit cells within the backend software. The creation and display of surface models is now working and test calculations of bulk and surface structures with the backend software have been successful.

**CRYSTAL:** Ian Bush has started work on the distributed memory parallelisation calculation of properties that require large amounts of memory, such as those that depend upon the eigenvectors of the system, or require a three dimensional grid in a large unit cell. Initial versions of projected densities of states, band structures and representations of the charge density and electrostatic potential on a grid have been completed, and these will form part of the next release of CRYSTAL. Ian has started to develop benchmarks to compare the performance of periodic ab initio codes used by the MCC. The intention is that these will be for systems of scientific relevance for the consortium. The initial targets will be to compare CRYSTAL and CP2K, and VASP and CASTEP. Discussions have been held with consortium members and code authors to start scoping what is required, and Matthew Watkins at Lincoln has been visited. The parallelisation work will increase the parallel scaling performance of the software, allowing users to target larger systems in future.

In January 2020 MCC and CoSeC developed a proposal to establish a working group for exascale materials modelling under the ExCALIBUR strategic priorities fund call, involving Ian Bush, Alin Elena and Tom Keal. The proposal was funded and preparatory work has begun with a focus on organising an online kick-off meeting in May. 0.5 FTE will be assigned to ExCALIBUR for the duration of the project (Ian and Tom at 0.2 FTE and Alin at 0.1 FTE).

## Outline Plans 2020/21 (1 April 2020 – 31 March 2021)

*(Detailed plans and milestones are included in the Technical Appendix for reference)*

**ChemShell** development will continue to focus on the new Py-ChemShell code. Py-ChemShell developments under the SLA will include preparing recent developments for a new release (formerly work supported by CCP5) and an interface to MolPro to enable the advanced QM/MM embedding methods and calculations on multiple excited states. ChemShell will continue to be maintained on ARCHER and Tier 2 systems such as UCL's Thomas facility.

**DL_POLY:** The supervision, support and development will continue with a sole focus on the refactored version. The need for all-to-all coulombic calculation by ChemShell will be addressed via systolic loops. The same strategy will be adopted for frozen-frozen exclusion interaction to remove the approximate scheme adopted at the moment due memory overload on large processor counts.

**DL_FIELD:** Improve interoperability of force field data structures will significantly improve the ease of setting up force field models. This will remove inherent force field-dependant restrictions on format-specific user's configuration files. Such efforts require code restructuring and changes to some data structures in the force field libraries. The DL_FIELD development work will also concentrate on implementations and rigorous testing of the bio-inorganic model setup workflows, including code refactoring, in order to ensure software robustness to handle a wide range of mixed component models and conditions.

New force field schemes will be implemented to broaden the capability of DL_FIELD, as well as demonstrate the capability of DL_FIELD/DL_POLY by managing large-scale simulations of graphene-cellulose nanocomposites for lightweight electronic devices. Such models cannot directly implement in some traditional bio-packages such as AMBER.

**ExCALIBUR**: We shall form part of the steering panel for this grant, so being involved in organising the meetings between the interested partners, and taking the lead on scaling out algorithms and applications to higher core counts. We shall also investigate the possibility of interfacing CRYSTAL with

the python interface that is available for CASTEP, so aiding interoperability between the 2 codes. We will also investigate complex workflow blueprints on exascale systems, with a prototype design demonstration involving Py-ChemShell, and benchmark high performance I/O approaches.

## *UKCP – UK Car-Parrinello Consortium*

The United Kingdom Car-Parrinello Consortium (UKCP) is a group of researchers across the UK who develop `first principles' quantum mechanical techniques for studying atomistic systems and apply them to a very wide variety of systems. The UKCP consortium is one of the longest-running High-End Computing Consortia in the UK, and has been funded almost continuously by EPSRC since the 1990s. The 1 FTE of core support is currently focused on supporting the CASTEP code, one of the UK flagship first principles codes.

### Summary Report (1 April 2019 – 31 March 2020)

Two major new pieces of functionality have been added to CASTEP during this period, to be made generally available in the academic release of version 20.1. Albert Bartok-Partay (AB-P) has delivered new code machinery for meta-GGA functionals, which allow for more accurate calculation of materials properties in a framework comparable in computational expense to the most widely used approximations in the field. In a first of its kind development this was done consistently with the generation of pseudopotentials – a central part to any plane-wave code, such as CASTEP, that allows accurate yet computationally tractable treatment of electrons close to atomic nuclei. In the process, AB-P discovered numerical and mathematical pathologies in the existing state-of-the-art models, leading him to develop a new, more robust model for the approach. These methodologies and comparative results have been published, the latter in the high-profile Physical Review B in collaboration with other world leading theorists.

The other major new functionality is the implementation of second-order non-linear optical coefficients by Dominik Jochym (DBJ). This will allow for the modelling of non-linear optical properties such as Second Harmonic Generation (used for example in blue laser technology). The approach takes advantage of the unique implementation of CASTEP's "curvature operator" that accelerates the rate of numerical convergence and improves the time-to-science over implementations in other plane-wave codes.

In March 2019 it was announced at the first CASTEP User Group Meeting that the source code for CASTEP will be made available free-of-charge for academic use worldwide. DBJ has worked closely with the CASTEP Developer Group (CDG) and Cambridge Enterprise (CE) to draft and produce updated license terms. STFC (as part of UKRI) have entered an agreement with CE to administer distribution of academic CASTEP licenses and the source code. As a result, DBJ's efforts have been redirected to the development and deployment of suitable license and source code distribution mechanisms that comply with the legal terms. A "soft-launch" of the new licensing was done to coincide with the academic release of CASTEP 19.1 and the workshops in August. This change in academic licensing for CASTEP has been enthusiastically received, with over 200 license requests from groups outside of the UK. The anticipated roll out of STFC's implementation of the Elucid licensing platform has been delayed to financial year 2020/21. The additional administrative burden of license management on DBJ has been noted and a plan put in place to address this in the next reporting period. Workplans for 2020/21 have been adjusted accordingly.

**Problems, issues and difficulties encountered during reporting period:** License management for CASTEP was intended to be low-impact on UKCP support, with the expectation of the Elucid license platform to be in place by January 2020. This is yet to come into production and DBJ continues to process license requests by email. Elucid is "soon" to be in place and minimal change should be needed

to the prototype workflow developed in October 2019. A data entry project has been offered to Georgia Lomas to reduce the administrative workload on DBJ in order to focus on an in-house distribution platform for CASTEP.

### Outline Plans 2020/21 (1 April 2020 – 31 March 2021)

*(Detailed plans and milestones are included in the Technical Appendix for reference)*

A reassessment of priorities has led to a modified set of activities for the 2019/20 and 2020/21 period, particularly in the wake of the worldwide academic release of CASTEP.

In direct support of the UKCP and wider CASTEP community, the annual CASTEP software release management, teaching workshop and code developer's workshop will be carried out. Maintenance of the software development and testing platform will allow for the continued growth of the CASTEP package and the software standards required and expected by the commercial partners and academic community. A custom software distribution web site that complies with CASTEP's new licensing needs will be launched. The ongoing development and maintenance of this platform will be carried out. It is expected that the approximately 300 license holding research groups world-wide will continue to grow in number and will make use of this platform to download the CASTEP package.

Work in collaboration with Jonathan Yates (Oxford) on an interface between CASTEP and YAMBO will enable accurate simulations of electronic excitations to be performed. Support will be provided to Matt Probert (York) and Scott Donaldson (UKCP Intern) in releasing a novel multi-objective Genetic Algorithm tool alongside CASTEP. This technique allows for structural searches based not only on energetics, but also on desirable materials properties.

For new software developments a long-awaited implementation of "van der Waals DFT" will be prototyped within CASTEP. This major new functionality promises to improve the underlying approximations used in CASTEP, with particular impact in the field of pharmaceuticals and organic electronics.

Note that with regards to workshops, on-line attendance options are being considered as circumstances require.

## *UKCOMES - UK Consortium on Mesoscale Engineering Sciences*

The United Kingdom Consortium On Mesoscale Engineering Sciences (UKCOMES) – founded in 2013 – is a group of researchers across the UK who develop and apply mesoscopic modelling techniques to explore systems of scientific and industrial interest at scales between atomistic and continuum-based levels. Several modelling techniques are applied in this consortium, but the most frequently used and studied is the Lattice Boltzmann Equation (LBE) method, a particle-based statistical technique capable of modelling fluid flows with complex geometries and interactions between multiple fluids and phases.

The 0.6 FTE of core support per year is focussed on developing DL_MESO, the consortium's community code for LBE simulations, by adding new functionality and optimising for various computing architectures. Both activities allow for a wider range of systems to be modelled with available computing resources, including the UK's national supercomputer ARCHER. Development, porting and optimisation support is also provided for other LBE codes used by the community, particularly MPLB and HemeLB, which are respectively based on structured mesh libraries and sparse system geometries

### Summary Report (1 April 2019 – 31 March 2020)

The work has been proceeding mostly as planned with support from Michael Seaton and Jianping Meng. Michael Seaton claimed less from October 2019 until March 2020 due to the need to use other available funding with a corresponding reduction in code development effort.

The UKCOMES community's aim is to model fluid systems of scientific and engineering interest. This requires development of both mesoscale modelling techniques to incorporate some finer detail at larger scales and codes that can apply those techniques to large (industrial-scale) problems.

Jianping Meng leads the developments on the EPSRC-funded "High-Level Mesoscale Modelling System" (HiLEMMS) project, where a set of abstractions have been proposed and released publicly (https://gitlab.com/jpmeng/hilemms). HiLeMMS is a coding abstraction for Lattice Boltzmann that allows coupling of this method to structured grid libraries, with the aim of automated code generation for various computer architectures. This would enable researchers operating at the mesoscale to carry out efficient large-scale calculations on any given computer system. Input parameters can now be read from a JSON file for both sequential and parallel simulations. HiLeMMS identifies aspects of a required Lattice Boltzmann simulation (e.g. grid size, number of fluids) by means of metadata, which is used in the generation of a resulting simulation code for a specific problem. JSON is a compact file format that allows metadata to be specified and read in efficiently.

Changes have been continuously made in the MPLB LBE code to implement the HiLeMMS definition, and the code has also been released publicly in GitHub under the BSD licence (https://github.com/jpmeng/MPLB). The MPLB code is based on the Oxford Parallel library for Structured-mesh solvers (OPS), which can run on a wide range of hardware (e.g., GPUs) via the support of various threading models. MPI can also be used to connect multiple accelerators, thus enabling large-scale numerical simulations for the UKCOMES community.

An alternative method of applying Lishchuk immiscible fluid/fluid interactions was implemented in DL_MESO by Michael Seaton. This implementation exploits the sparseness of fluids at any given lattice point and only requires a small number of fluid lattices and additional book-keeping to keep track of larger numbers of fluids, making simulations of large numbers of immiscible drops or fluid-filled vesicles (representing e.g. red blood cells) more efficient or even feasible. Fluid-filled vesicles that do not combine together could represent red blood cells (RBCs): being able to model these would allow blood flows with explicit RBCs to be studied and extend the application of mesoscale modelling techniques to medical problems. (Another code supported by UKCOMES, HemeLB, uses lattice Boltzmann to model blood flows in arteries, although this does not explicitly model RBCs.)

In summary the overall aim of UKCOMES is to produce simulations on a scale that represent real-life physical systems, and for these simulations to run at computing and time scales that are competitive with experimental time-scales. Simulations at these scales can not only offer greater insight about a particular system, but they can also be a viable substitute for some real-life experiments, saving resources when developing new products or chemical formulations.

**Problems, issues and difficulties encountered during reporting period:** Michael Seaton had to reduce support for UKCOMES between October 2019 and March 2020 due to a requirement to use funding for an Ada Lovelace Centre project at that time: this has led to postponement of a number of milestones for DL_MESO, which will now be completed in 2020/2021.

## Outline Plans 2020/21 (1 April 2020 – 31 March 2021)

*(Detailed plans and milestones are included in the Technical Appendix for reference)*

The Lishchuk immiscible fluid model in DL_MESO's LBE code will be adapted to give non-spherical drops as fluid-filled vesicles; this can be used to model a wider range of fluid/fluid suspensions, including representation of red blood cells for haemodynamics studies.

Multi-block technique is important for reducing the memory consumption where there is complex geometry layout in the problem, e.g., backstep and T-junction etc. Based on the general infrastructure provided by the OPS library, interfaces will be designed according to the specific requirement of the lattice Boltzmann method. These interfaces will be then implemented in the MPLB code and work together with the JSON configuration file. Afterwards, the technique will be utilised for testing LES models using typical problems, e.g., backstep flows.

## *HEC Plasma Physics*

HECPlasma and CCPPlasma resources have been combined into a single workplan. Please see the report above under CCPPlasma.

## *HECBioSim*

HECBioSim exists to bring High-End Computing for biomolecular simulation to a wider community, including users from industry and experimental bioscientists, and to engage physical and computer scientists in biological applications. The Consortium works closely with CCPBioSim.

HECBioSim is supported by 1.0 FTE of SLA core effort, which provides support for scientists applying for time on ARCHER, primarily through maintenance of the HECBioSim web portal. It includes help on preparation of applications e.g. with the HECtime resource calculator, and on reporting the outcomes of approved projects. The SLA post also works on a variety of codes for biomolecular simulation and analysis appropriate to High End Computing. For further information please see www.hecbiosim.ac.uk.

### Summary Report (1 April 2019 – 30 September 2019)

A significant portion of time has been dedicated to delivery of the new CCP/HECBioSim cloud based training system. Due to procurement issues in the previous reporting year most of the work on getting this project delivered fell into this reporting year. Initial work focused on producing a set of training courses that had all of the courses stored within a single Docker container, this was due to a very narrow time window between procurement of cloud time and the CCPBioSim training week workshops. After this event ran, it was clear that large monolithic Docker containers caused technical issues with resourcing and we needed a different approach if we were to run courses online in future. Work focused on re-engineering this whole platform to fit the micro-services approach that is more common in applications deployed into the cloud.

To accomplish this micro-services approach, a kubernetes cluster was prepared with a deployment of Traefik which acts as an ingress controller and routing system. Helm charts were prepared for each training course to provision a JupyterHub instance for each workshop container and deployed onto the kubernetes cluster. The result being each workshop course being split into its own elastically scalable environment, thus eliminating the technical issues with resourcing seen with the previous version.

This means that resources scale on a per course per user basis allowing both the online mode and instructor led modes of delivery without any extra infrastructure. This training platform represents a step change in how the BioSim consortia can deliver their training programmes. It removes the awkward technical setup of running courses in university computer clusters and allows new content to be delivered quickly. It provides for more advanced instructor led courses to be delivered as beginner courses can now be done online as a prerequisite.

To complete the cloud training project delivery, a subdomain of the CCPBioSim website was routed into the cloud ingress controller, and a new highly visual training webpage was completed to advertise and link into the training courses.

Longbow is developed under CoSeC in collaboration with Nottingham University. Longbow functions to seamlessly offload simulations to powerful super computers in a manner that appears to the user as the simulation had run on their own PC. In our community, users fit broadly into two groups as far as HPC is concerned. The expert users and those that have never used any big machines, and there is nothing in between.

For the experts, longbow enables them to rapidly submit large volumes of simulations without spending long tedious hours making submit scripts and transferring files. This can all be done automatically with a single command. For the beginner users, and these include non-linux folks and experimentalists dabbling with simulation. These folks do not understand the commandline tools or which ones to use for what, however they do know how to run basic md simulations on their own computer. Longbow lowers the technical barriers to using larger HPC machines by abstracting this all away.

This means that experienced groups spend more time doing science and not mundane and tedious computing tasks leading to higher productivity. Whilst for the beginners it enables them to submit simulations to large scale facilities such as tier 0, 1, 2 or local HPC without having to learn heaps about each machine. The tool is primarily focused on these two groups of users and either lowering technical barriers to access or by accelerating the amount of work they can do.

Longbow handles all file transfer, scheduling, monitoring and is capable of doing complex multi-machine submission patterns. Currently Longbow has been downloaded 7,934 times split between our website and the command-line installer pip.

Version 1.5.3, originally listed for delivery in Q2 2019, delivery was pushed to the next work year due to the prioritisation of other tasks, new issues being reported by users in the meantime and the delays on delivery of ARCHER2 as support for this new machine was due in this version. Although this is the case, work on this release has been progressing with currently 13 of the 16 tasks that define this release version having been completed. The vast majority of improvements made to the code so far are fixing user reported crashes or annoyances, such as a bug when used with the latest version of python 3.7, a critical crash when working with non-supported 3rd party programs, an error when installing Longbow into Docker containers and some issues with logging. In addition to this, the documenting of the Longbow software library has been scripted to auto-generate and build alongside the other auto-generating user and developer documentation and moved under version control. It is anticipated that these fixes and improvements will enhance user experience and improve the adoption of Longbow in fields outside of the remit of biosimulation. It is intended that Longbow after this major release, will take the development priority of a mature software package and thus will be updated less frequently and with only minor patches released.

Work supporting consortium users with ARCHER and JADE has mainly been focused on addressing system problems with the adoption of JADE. Benchmarks have been run for both ARCHER and JADE with the current versions of the MD codes and the figures on our website updated. On JADE a number of issues with software have been resolved, this has mainly been recompilation of software packages that were broken during system package refreshing but there were a number of issues with the way some users were running their simulations. To address this, a new guidance webpage has been added to the HECBioSim website with instructions on how to make use of the GPUs on JADE for each of the MD packages we support.

Two projects to compile the latest codes and get benchmark data on UK Tier 2 facilities Thomas and ISAMBARD were completed. These projects were focused on getting performance data for comparison to ARCHER to inform the community and EPSRC on the level of interim support across various tier 2 facilities would be needed during the ARCHER to ARCHER2 changeover. Work on both facilities progressed smoothly, some codes and versions were already present on Thomas so work there was mainly tweaking the benchmarks suite to its architecture and running the suite. On ISAMBARD, work required some experimentation to get codes compiled as this is an ARM64 architecture. Once compilation was complete, again the benchmark suite was tweaked to the machine architecture and performance figures collected and available to the community, EPSRC and the computing centers that run the services.

Other work on the HECBioSim website so far has included patching the server and content management framework to the latest version. A new publications page has been created and a small number of papers have been listed, this is anticipated to grow over time as we collect historical data and update with new publications. This will help us to feed in more detailed information into reports but to also publish more widely the success stories arising from the support of the HECBioSim consortium.

A new course has been written for the Longbow job submission tool. This course was delivered as an instructor led course at the CCPBioSim training week in May to 22 delegates. It was written to be compatible with the new JupyterHub cloud based training system and is now available as one of the training modules that consortium members can complete online. Delegates at the training course reported that they found the course useful, they would like to use Longbow as part of their research and that they would like to see more training in the use of HPC in the life sciences.

A project in collaboration with Charlie Laughton to interrogate the communities existing training catalogue and to bring it in line with the new cloud based training approach was completed. This involved auditing previously made material and firstly placing the material into repositories on github. Then splitting the material up into sensible modules and building Docker containers for each module and pushing to our central repository. These containers were then pulled into the cloud based training environment and made available for workshops and online training.

## Outline Plans 2020/21 (1 April 2020 – 31 March 2021)

*(Detailed plans and milestones are included in the Technical Appendix for reference)*

The 2020-2021 year will see the arrival of a number of computing facilities from ARCHER 2 to a number of new UK tier 2 facilities. Plans in this year will include tasks around supporting access to and supporting these facilities with maintaining a high standard of software support where HECBioSim has a partnership.

The benchmark suite will have to be tested and potentially updated on some of the new architectures, scientific codes will be compiled and benchmarked on these HPC machines. Some new benchmarking regimes will be developed and a UK wide outlook on HPC along with guides and best practice information will be compiled into one single resource.

Work will be done maintaining the HECBioSim website, and also on the BioSim cloud based training infrastructure, as these are web facing resources, work has to be done to maintain the information presented and to the security aspects of running such services. Some tasks moved from the previous work year due to high demand from new tasks and more projects are currently in the scoping stage at the current time.

# *UK-AMOR*

The UK-AMOR consortium exists to develop and exploit the UK Atomic, Molecular and Optical Physics R-matrix codes on ARCHER. The R-Matrix suite is a set of programs for electron (positron) -atom and -molecule scattering, (ultrafast) laser pulse interactions and related problems. UK-AMOR's scientific goals are to: study attosecond physics in atomic and molecular systems; compute fusion-relevant atomic data; study electron-molecule collisions with relevance to biological radiation damage, fusion energy and technology. The consortium allows high-end use of the CCPQ codes RMATRIX I, PRMAT (RMATRXII, PFARM), UKRMol(+) and RMAT_REACT, and RMT. Major code development projects of UK-AMOR include the development of a novel R-matrix treatment of ultracold chemistry, RMAT_REACT, which will be used to study key systems (atom-atom and molecule-molecule) in parallel with experiments being performed in the UK and abroad, and algorithmic improvements made to existing codes to improve functionality and scalability.

The consortium is the focus for UK and international activity for users of the UK-AMOR codes suite. It is committed to: making its codes widely available; interacting closely with CCPQ; engaging with industrial partners and international collaborators through workshops and other activities; broadening the user community for the AMOR methodology and codes; providing training for the next generation of computational scientists in the UK.

UK_AMOR has received some administrative CoSeC support from CCPQ funding (M Plummer) and an average of 0.2FTE per year direct CoSEC support for code developments.

## Summary Report (1 April 2019 – 31 March 2020)

CoSeC support given to UK-AMOR in this period was a mixture of direct support by Andrew Sunderland, and Martin Plummer's support from CCPQ SLA/CoSeC funding as described in the CCPQ report. Direct support for UK-AMOR, at an average of 0.2FTE per year, concentrated on preparations to put a reliable molecule-molecule bound state code into the RMAT_REACT interface. Andrew (a computational scientist with some AMO experience) attended the CCPQ Flagship project R-MADAM workshop on R-matrix codes and studied the theory of bound state location, along with code used previously by Martin for electron-atom work which showed fewer signs of numerical instability than the existing electron molecule R-matrix code. The result of this study is that the RMAT_REACT bound state code will be adapted using stability techniques from the atomic code module (with minor modifications for heavy particles). The new code should be more accurate and efficient than it would be if we made simple adjustments to the original electron-molecule bound state module, and will better serve general users by being more stable, producing accurate results that less experienced users can rely on, and increased computational efficiency both in time and energy.

The original aim was to begin to put this study into practice in the latter part of the reporting year. The very disappointing results of the CCP renewal call, with CCPQ not funded, affected the plans. Andrew was asked to further extend his very successful CCPQ CoSeC memory optimization work on the UK-AMOR code UKRMol+ (described in the CCPQ report) and complete this before CCPQ CoSeC support ended. Martin took over direct UK-AMOR support from March, with an increased effort level now planned for 2020-2021 to give an average of 0.3FTE for the year (0.1FTE carried over from 2018-19). He has begun adapting his RMAT_REACT work to interface to the bound state code and will be consolidating the bound state code production, which will speed-up development of a full production code. UK-AMOR itself continues to flourish, with full use of its ARCHER time by members and travel grants awarded to present, for example, at various satellites and the main meeting of the International Conference on Photonic, Electronic and Atomic collisions (ICPEAC) in July 2019.

## Outline Plans 2020/21 (1 April 2020 – 31 March 2021)

*(Detailed plans and milestones are included in the Technical Appendix for reference)*

CoSeC Support for UK-AMOR was originally planned as a mixture of coding development for the RMAT_REACT project by Andy Sunderland and more general software sustainability work from the Software Engineering Group, but due to the reduced allocation of an average of 0.2FTE per year, the RMAT_REACT work was given precedence. Following the extremely disappointing result of the CCP renewal call, with CCPQ not being funded despite the quality of the proposal, the UK-AMOR CoSeC support is now the only community core support apart from some small residual carried-over CCPQ support. MP has taken over the support, 0.3FTE in 2020-2021, and will be following up AGS's UK-AMOR work of 2019-20 by continuing to develop the RMAT_REACT bound state code for ultracold molecular collisions, subject to any decisions made by UK-AMOR's management committee during the year. He will extend the interface he already developed for the collision code PFARM to allow use of techniques in a numerically stable bound state code developed previously for electron collisions. Unavoidable peculiarities (localized instabilities, controlled using targeted quadruple precision where needed) in the bound state code under construction will be investigated fully. His UK-AMOR management committee duties and CoSeC reporting duties are now funded through CoSeC's UK-AMOR support. He will work with UK-AMOR's management committee to try to secure new funding, such as through ARCHER-2 eCSE calls and by supporting general grant applications (for example, the CCPQ/UK-AMOR code RMT is the subject of a new EPSRC 3 year physics grant beginning in May 2020, with MP as a co-I and UK-AMOR co-I Dr A Brown the PI).

## *UKTC - The UK Turbulence Consortium*

Understanding, predicting and controlling turbulent flows is of central importance and a limiting factor to a vast range of industries: naval, aeronautical, automotive, power generation, process, pharmaceutical, meteorological and environmental. For example, energy-efficient fluid mixing is a major limiting factor in the chemical industry which is the second largest industrial consumer of energy worldwide, but also in the pharmaceutical and process industries. How can one mix best with as little power loss as possible? The cost of pumping oil and gas through pipelines is directly proportional to the frictional losses due to turbulence. Polymer additives can modify the turbulence in these pipelines to reduce frictional losses and speed up the flow by one or two orders of magnitude with the same power input. How can one maximise this effect which remains poorly understood? Moving vehicles often generate noise, and in the case of airplanes, it can have a severe detrimental impact close to airports. What kind of airbrake could generate enough drag with as little aeroacoustic losses as possible? Worldwide ocean shipping consumes about 2.1 billion barrels of oil per year while the airline industry uses 1.5 billion barrels per year. How can turbulent boundary layers be controlled to significantly reduce friction drag which will simultaneously reduce costs and emissions? Simulating and understanding turbulent flows to answer the previous questions is one of the most challenging problems in science. Many of the environmental and energy-related issues we face today cannot possibly be tackled without a better understanding of turbulence. The overarching objective of the UK Turbulence Consortium (UKTC) is to facilitate world-class turbulence research using national High-End Computing (HEC) resources. This involves performing numerical experiments with turbulence-resolving computational approaches. Such simulations are ab initio calculations based on first principles and produce data to answer basic questions regarding the physics and modelling of turbulent flows found across a range of engineering, physiological and geophysical applications. The consortium serves as a forum to communicate research and HEC expertise within the UK turbulence community, and to help UK science remain internationally leading in this aspect of HEC-based research.

### Summary Report (1 April 2019 – 31 March 2020)

During this reporting period we have extended the benchmark study of the Taylor-Green vortex (TGV) test case to the turbulent flow regime using Code_Saturne up to Re=10000. Code_Saturne was used

to compute the TGV with a range of grid sizes and time steps at different Reynolds numbers. At Re = 1600, the agreement between Code_Saturne and the high-order spectral element simulation is known to be excellent apart from minor dependency at the peak of the integral of enstrophy. Tests up to Re = 6400 worked well with grid independent solutions being achieved. However, it was found that the finer mesh size and smaller time step needed to achieve results comparable to higher-order DNS codes, particularly at high Reynolds number, proved too expensive with large data storage and CPU time being required due to its second-order nature. We continue to look for a suitable high Reynolds number benchmark test case suitable for testing and evaluating the UKTC code base.

The feasibility of extracting the Lagrangian particle tracking module from Code_Saturne was investigated. Although the subroutines in the Lagrangian particle tracking module are written in C, there is a Fortran wrapper available which will allow the subroutines to be incorporated into Fortran or C codes in the community. Code_Saturne can track particles by post-processing if the concentration of particles is low, i.e. one-way coupling between the particle and fluid. Particle size and density are required as inputs. It can also track particles during the simulation, i.e. full coupling between the particle and fluid. However, the particle tracking module is currently available only for RANS calculations.

Preliminary work on in-situ visualisation has started to investigate the potential of ParaView - Catalyst on ARCHER which could help to substantial reduce the size of the output files for Code_Saturne. In turn, if successful this strategy could be rolled out to other UKTC codes..

### Outline Plans 2020/21 (1 April 2020 – 31 March 2021)

*(Detailed plans and milestones are included in the Technical Appendix for reference)*

OpenSBLI currently uses traditional no-slip boundary conditions whereas it is know that high-speed flows can involve both slip and temperature jump. Using kinetic theory through the moment method, we will incorporate high-order components to test and validate impact on results.

As we perform simulations on large-scale facilities, such as ARCHER2, the need for in situ visualisation becomes a key consideration. Previous work considered Code_Saturne as an exemplar to understand how we can devlop an in situ capability in other UKTC codes. We will now explore what is done in nek5000 to get a better understanding of best practice in delivering this capability.

## *UKCTRF - UK Consortium on Turbulent Reacting Flows*

The UKCTRF will offer a step change in the fundamental understanding of turbulent reacting flows using advanced numerical simulations utilising developments in HPC, and translate physical insights into high-fidelity models for engineering simulations through collaborative and complementary research.

### Summary Report (1 April 2019 – 31 March 2020)

Over the past twelve months work has focused on further developing the HAMISH code, which is a new CFD solver for turbulent reacting flows using adaptive mesh refinement (AMR). In the reporting period Jian Fang's work focused on having a stable HAMISH code for researchers in Cambridge University and Newcastle University. This has allowed researchers to begin testing and validating the initial developments. He also finished converting the HAMISH code from f77 to f90 format for better compatability with future hardware developments. This work allowed the introduction of bit operation-based Morton code algorithms for dynamic partitioning of the code and introduce dynamic memory allocation to minimise memory use, particularly using AMR.

### Outline Plans 2020/21 (1 April 2020 – 31 March 2021)

*(Detailed plans and milestones are included in the Technical Appendix for reference)*

Work will continue on the optimisation of the HAMISH code whilst also working on the development of the capability of simulating combustion flows and the immersed boundary method for the code.

## *Software Outlook*

Software Outlook focuses on software technologies that are vitally important to the development and optimisation of the world-leading scientific software produced by the CCPs. This includes evaluation of new programming techniques that are essential for the timely and cost-effective exploitation of current and near-future High Performance Computing systems, demonstrating how specific software technologies can be applied to existing applications, and providing guidance on the principles of best practice within software engineering.

### Summary Report (1 April 2019 – 31 March 2020)

During this reporting period, Software Outlook has been working on the hybrid programming and maths/physics libraries work packages.

Hybrid Programming: Using an initial benchmark to compare the porting process and the code performance prior to comparison with more complex codes, benchmarking runs have been performed and a technical report has been written. Unfortunately, staff illness meant that work on the more complex codes was delayed. This was further hampered by staff availability but the comparison of frameworks restarted in November 2019 and we have been using a CCPi code to compare the different frameworks, which has been a good example for showing how data layout can significantly hamper code performance. A Technical Report should be available within the next few months. In August, Jony Castagna (STFC Hartree Centre) represented Software Outlook at Sheffield University's First GPU Hackathon as one of the mentors. A total of 45 people covering a large range of scientific applications (including developers of CCP codes such as CASTEP) attended the highly successful hackathon with further details available at https://rse.shef.ac.uk/blog/2019-10-10-gpu-hackathon/. The importance of this work is highlighted by some attendees reporting that they achieved a 200 times speed-up of their code. This will enable users to either run many more simulations in an allotted time period or run more sophisticated simulations in that time period.

Maths/Physics Libraries - FFT Solvers: Due to staff illness, this project was delayed. However, the three reports for the C/C++, Fortran and GPU FFT library benchmarks have been completed. At the time of writing, one has been converted to a RAL Tech Report and the others are in the process. Currently, FFTW is widely used amongst the CCPs and HECs but our benchmarks have revealed that the MKL library can give significant execution time savings and, in the case of the MPI versions, is more reliable. By decreasing the execution time of the code, more simulations can be performed in the same time period or more complex simulations could be performed.

Maths/Physics Libraries - Linear Algebra: Staff illness and staffing availability have meant that the FFT work was prioritised ahead of the linear solvers work and the majority of this work package was moved to the 2020-21 plan. However, work was done with Joseph Parker (CCP-PLASMA) to analyse a proposed method for solving linear systems and understand why Joseph's method sometimes converges but sometimes diverges from the solution. Use of the correct linear algebra solver (method and software) can dramatically reduce execution times of scientific codes or reduce memory usage, which will allow the user to run more simulations, increase the accuracy of the simulation or run more sophisticated simulations.

As mentioned above, Software Outlook has had problems due to staff illness and availability of staff. Since July, Software Outlook has not been fully staffed but we will be back up to full strength in April 2020, when the Code Coupling work will also start.

### Outline Plans 2020/21 (1 April 2020 – 31 March 2021)

*(Detailed plans and milestones are included in the Technical Appendix for reference)*

Linear Solvers: Physics and mathematics problems lie at the heart of most simulation codes and, hence, there are normally parts of the code where external physics and maths libraries can be used. These parts of the code can quickly become a bottleneck if the wrong library or method is used. For example, in BOUT++ (CCP-Plasma), the code is using time-steps that are much smaller than they need to be because the convergence of the underlying iterative linear algebra method is so poor. By improving the convergence through more effective preconditioners, this will also enable the use of a larger time-step with the potential of dramatically improving the execution times and energy consumption values for these computational simulations. As part of Software Outlook's 2019/20 work, benchmarking of linear algebra libraries for use in examples from BOUT++, EPOCH and other CCP codes will have been started. In 2020/21, Software Outlook will complete this work.

Code Coupling at Scale: Since the initial code coupling work was performed by Software Outlook, code coupling tools have evolved along with their documentation. Historically, considerable effort has been given to in-house development of coupling libraries etc. This work package will aim to assess/demonstrate the practical applicability of existing code coupling tools (PLE, MUI, etc.) in coupling two independent codes. It also allow us to provide recommendations for, and improvements to, available code coupling tools to maximise their value across a range of HPC applications.

Code Containerisation: Code containerisation using technologies such as Docker is starting to become a popular activity and be used by CCPs. There are now a number of containerisation options available and it is important that the correct chose is made and CCPs avoid potential lock-in. We are aware of CCPs that have had both positive and negative experiences with containers and we will work to understand the differences. Software Outlook will review current containerisation options and promote the best to the CCPs/HECs.

Best Practices in Software Engineering: One of the roles of Software Outlook is to promote best practice in software engineering methodologies. Identifying what makes "high-quality software for researchers" will be part of this process and we will help organise a workshop with the SSI in 2020. As part of this work package, we will identify existing best practice guidance and promote it to CCP and HEC developers. Additionally, with the help of our Hartree Centre colleagues, we will develop a series of articles on best practice including scalability testing, roofline modelling, performance portability and continuous integration.

## Resource Planning and Yearly Costing

This section summarises the approach to resource allocation and yearly costing.

**Level of funding and resource for the current SLA**

During the current SLA cycle, the total funding agreed corresponds to 113 FTE over the 5 years. This figure includes the effect of the £200k drop in funding from 2018/19 until the end of the SLA imposed by EPSRC's own budget constraints and made known to STFC at the time of the SLA renewal in 2016. In addition to delivering the full 113 FTEs of effort for the current five year SLA cycle, we have agreed that during this cycle we will make up for a shortfall in effort delivered that arose for several projects during the previous SLA cycle (up to 2015/16), as indicated in the table below. The table also reflects

the outcome of the EPSRC call for HEC consortia in 2017, ie an increase of the number of consortia supported by the SLA from 5 to 8, plus the outcome of the EPSRC call for CCPs in 2019.

**Planned Resource Profiles**

In this current planning process, for each project we allow the amount of effort to vary from year to year and to be re-adjusted yearly as long as each project recovers its full support over the five years of the SLA cycle.

The following table shows the predicted spend profiling for the five year cycle. For the yearly allocation, the figure in brakets shows the allocation before the HEC renewal in 2017 and the CCP renewal in 2019, with the figure not in brackets showing the revised figure after the renewals. For 2016/17, 2017/18, 2018/19 and 2019/20 two figures are included – in **black**, the planned effort used to calculate the annual cost of the CoSeC programme, and in **red** the actual effort provided. For the current year the figures in blue are the predicted effort for 2020/21. This table now also includes the four new CCPs that were approved following the 2019 EPSRC CCP call (highlighted in yellow).

| | Agreed Carry Over | Yearly Alloc | 2016-17 | 2017-18 | 2018-19 | 2019-20 | 2020-21 | Original Total | Total after HEC/CCP Renewals |
|---|---|---|---|---|---|---|---|---|---|
| Project Office | 0.00 | (2.25)<br>1.50 | 2.25<br>**2.23** | 2.25<br>**2.23** | 1.45<br>**1.79** | 1.50<br>**1.60** | 0.90 | **11.25** | **8.64** |
| CCP5 | 0.42 | (3.20)<br>0.00 | 2.82<br>**1.95** | 3.56<br>**3.19** | 3.26<br>**3.68** | 3.51<br>**4.42** | 0.00 | **16.42** | **13.22** |
| CCP9 | 0.83 | (2.40)<br>3.00 | 2.71<br>**2.64** | 2.70<br>**2.28** | 2.07<br>**2.70** | 2.21<br>**2.72** | 3.00 | **12.83** | **13.43** |
| CCPmag | 0.47 | (0.74)<br>0.00 | 0.83<br>**0.64** | 0.40<br>**0.70** | 0.76<br>**1.24** | 0.72<br>**0.51** | 0.00 | **4.17** | **3.43** |
| CCPNC | 1.80 | (1.30)<br>1.40 | 1.50<br>**1.62** | 1.30<br>**2.00** | 0.70<br>**1.79** | 0.64<br>**0.43** | 0.80 | **8.30** | **8.40** |
| CCPQ | 0.00 | (1.86)<br>0.00 | 1.86<br>**1.69** | 1.86<br>**1.80** | 2.00<br>**1.92** | 1.83<br>**1.94** | 0.12 | **9.30** | **7.44** |
| CCP Plasma | 0.48 | (0.75)<br>0.00 | 1.00<br>**0.79** | 0.80<br>**0.79** | 0.80<br>**0.85** | 0.32<br>**0.57** | 0.52 | **4.23** | **3.48** |
| CCPi | 0.33 | (1.20)<br>1.50 | 1.45<br>**0.82** | 1.50<br>**1.01** | 1.02<br>**1.13** | 1.07<br>**1.56** | 2.00 | **6.33** | **6.63** |
| CCPPetMR | 0.33 | (1.20)<br>1.80 | 1.15<br>**1.09** | 1.50<br>**1.53** | 1.40<br>**1.44** | 1.30<br>**0.98** | 2.00 | **6.33** | **6.93** |
| CCPBioSim | 0.00 | (1.20)<br>1.95 | 1.16<br>**1.19** | 1.20<br>**0.79** | 1.20<br>**1.23** | 1.40<br>**1.43** | 1.70 | **6.00** | **6.75** |
| CCP-QC | 0.00 | (0.00)<br>0.60 | 0.00<br>**0.00** | 0.00<br>**0.00** | 0.00<br>**0.00** | 0.00<br>**0.00** | 0.60 | **0.00** | **0.60** |
| CCP-WSI+ | 0.00 | (0.00)<br>2.00 | 0.00<br>**0.00** | 0.00<br>**0.00** | 0.00<br>**0.00** | 0.00<br>**0.00** | 2.00 | **0.00** | **2.00** |
| CCP Turbulence | 0.00 | (0.00)<br>1.75 | 0.00<br>**0.00** | 0.00<br>**0.00** | 0.00<br>**0.00** | 0.00<br>**0.00** | 1.75 | **0.00** | **1.75** |
| CCP NTH | 0.00 | (0.00)<br>1.00 | 0.00<br>**0.00** | 0.00<br>**0.00** | 0.00<br>**0.00** | 0.00<br>**0.00** | 1.00 | **0.00** | **1.00** |
| MCC | 0.00 | (2.50)<br>2.00 | 2.50<br>**2.53** | 2.50<br>**2.32** | 2.15<br>**1.95** | 1.45<br>**0.87** | 2.10 | **12.50** | **11.15** |
| UKCP | 0.42 | (1.00)<br>1.00 | 1.42<br>**1.36** | 1.00<br>**0.97** | 1.00<br>**0.98** | 0.48<br>**0.86** | 0.80 | **5.42** | **5.42** |

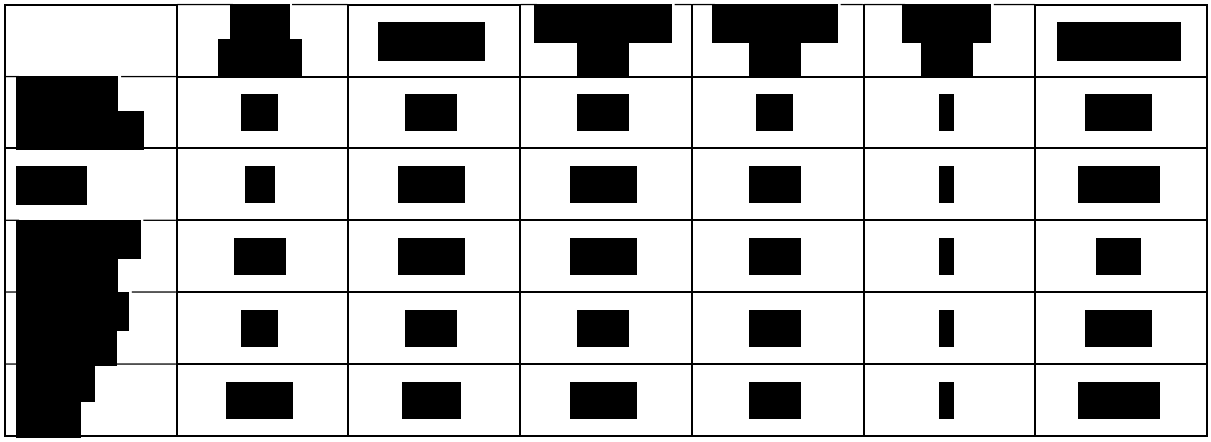
| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| UK-COMES | 0.00 | (1.00)<br>0.60 | 1.00<br>**0.93** | 1.00<br>**1.16** | 0.60<br>**0.75** | 0.60<br>**0.41** | 0.60 | **5.00** | **3.87** |
| HEC Plasma | 0.00 | (0.20)<br>0.20 | 0.20<br>**0.20** | 0.20<br>**0.20** | 0.20<br>**0.21** | 0.20<br>**0.14** | 0.20 | **1.00** | **1.00** |
| HECBioSim | 0.50 | (1.00)<br>0.80 | 1.00<br>**0.98** | 1.00<br>**1.01** | 1.00<br>**1.02** | 0.50<br>**1.01** | 1.00 | **5.50** | **5.00** |
| UKAMOR | 0.00 | (0.00)<br>0.20 | 0.00<br>**0.00** | 0.00<br>**0.00** | 0.20<br>**0.09** | 0.30<br>**0.20** | 0.31 | **0.00** | **0.60** |
| UK Turbulence | 0.00 | (0.00)<br>0.40 | 0.00<br>**0.00** | 0.00<br>**0.00** | 0.30<br>**0.10** | 0.40<br>**0.41** | 0.40 | **0.00** | **1.10** |
| UKCTRF | 0.00 | (0.00)<br>0.50 | 0.00<br>**0.00** | 0.00<br>**0.00** | 0.13<br>**0.14** | 0.50<br>**0.54** | 0.50 | **0.00** | **1.13** |
| Software Outlook | 0.10 | (2.00)<br>1.50 | 2.00<br>**2.31** | 2.00<br>**1.64** | 1.50<br>**1.54** | 1.40<br>**1.13** | 1.90 | **10.10** | **8.60** |
| **FTE TOTALS** | **5.68** | **(23.80)<br>23.70** | 24.85<br>**22.97** | 24.77<br>**23.62** | 21.74<br>**24.55** | 20.33<br>**21.73** | **24.20** | **124.68** | **121.57** |

**Planned vs Actual Resource**

While the resource planning is now done over a 5-year period, the costing of the programme needs to be calculated on a yearly basis based on the planned effort. Deviations between planned and actual effort are calculated and used to reprofile the effort and costing for future years. For 2016/17, 2017/18, 2018/19 and 2019/20, the difference between delivered and planned effort for each project has been calculated and redistributed over future years so that by the end of the SLA cycle each project will spend the amount of FTEs awarded for this SLA plus any historical underspend that a project may have incurred.

**2019/20 Planned Resource and Costing**

Of the planned 23.99 FTEs for 2019/20, 20.33 FTEs have been costed via the SLA, while 3.66 FTEs will be provided based on funding received previously (i.e. carry-over from previous SLA or adjustment due to deviation between planned and actual effort). The detailed cost breakdown for 2019/20 for the 20.33 FTEs is as follows:



Please note that recurrent expenditure includes:

- Travel and Subsistence – this includes T&S for our staff attending management/ project meetings. Includes T&S for visitors.
- Consumables – includes items such as telephone charges, training, photo repro, registration fees and workshop costs.
- HW/SW – this includes capital purchase of desktop/ laptop systems, software maintenance and applications packages.

- Participation in Supercomputing, production of annual report and general support activities around the SLA including the web.
- Hardware and software maintenance costs.
- Computing Infrastructure covers maintenance of servers, software licenses and printing consumables. Staff effort supports hardware and software maintenance and systems management of desktop systems and fileservers and support of visualisation and grid infrastructure.

**CoSeC Starters and Leavers**

New members of staff joining the CoSeC programme included:

- Ya-Wen Hsiao (CCPBioSim)
- Ian Bush (MCC)
- Jony Castagna (Software Outlook)
- Jack Taylor (Software Outlook)

**Size of Communities Supported**

The following table shows the community size of each of the supported CCPs and HECs in terms of the number of members of each community.

| CCP/HEC (Main supported code) | Start Date | Community Size (Members) | Allocation (FTEs per year) |
|---|---|---|---|
| **CCP5** (DL_POLY etc.) | 1980 | 1050 | 3.20 |
| **CCP9** (Questaal) | 1981 | 450 | 2.40 |
| **CCPPlasma** (GS2 / BOUT++) | 2007 | 150 (with HECPlasma) | 0.75 |
| **CCP-NC** (MagresView) | 2011 | 60 | 1.30 |
| **CCPQ** (R-Matrix, TNT, Quantics) | 2011 | 150 | 1.86 |
| **CCPBioSim** (FESetup) | 2011 | 402 (with HECBioSim) | 1.20 |
| **CCPi** (CCPi CIL) | 2012 | 380 | 1.20 |
| **CCP-Mag** (KKR) | 2015 | 44 | 0.74 |
| **CCP PETMR** (SIRF) | 2015 | 80 | 1.20 |
| **UKCP** (CASTEP) | 1990 | 150 | 1.00 |
| **MCC** (CRYSTAL, Chemshell) | 1994 | 464 | 2.00 |
| **HECBioSim** (Longbow) | 2013 | 402 (with CCP BioSim) | 0.80 |
| **UKCOMES** (DL_MESO) | 2013 | 150 | 0.60 |
| **HEC-Plasma** (GS2, BOUT++) | 2013 | 150 (with CCP Plasma) | 0.20 |
| **UK-AMOR** (R-Matrix) | 2018 | 40 | 0.20 |
| **UKTC** (Code_Saturne) | 2018 | 47 | 0.40 |
| **UKCTRF** (SENGA+) | 2019 | 47 | 0.50 |

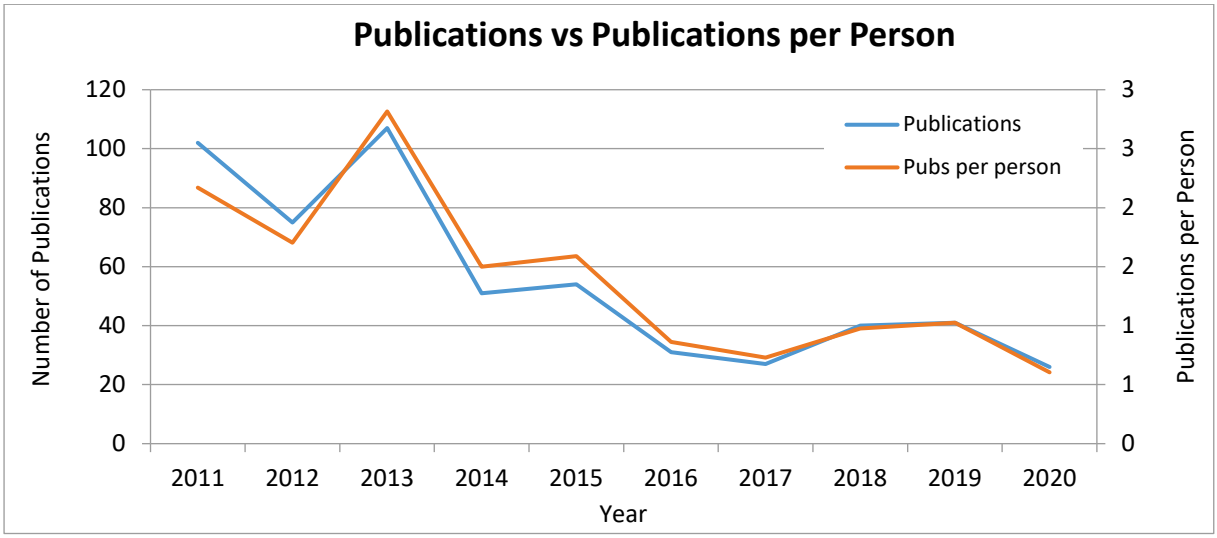
# Metrics

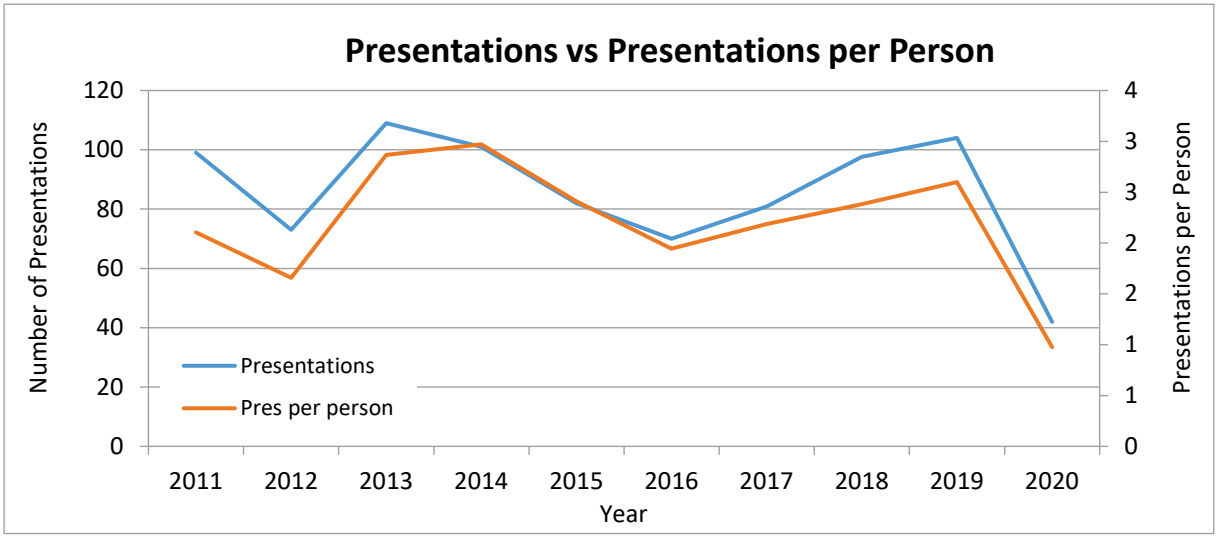
The metrics currently used for this programme are defined as:

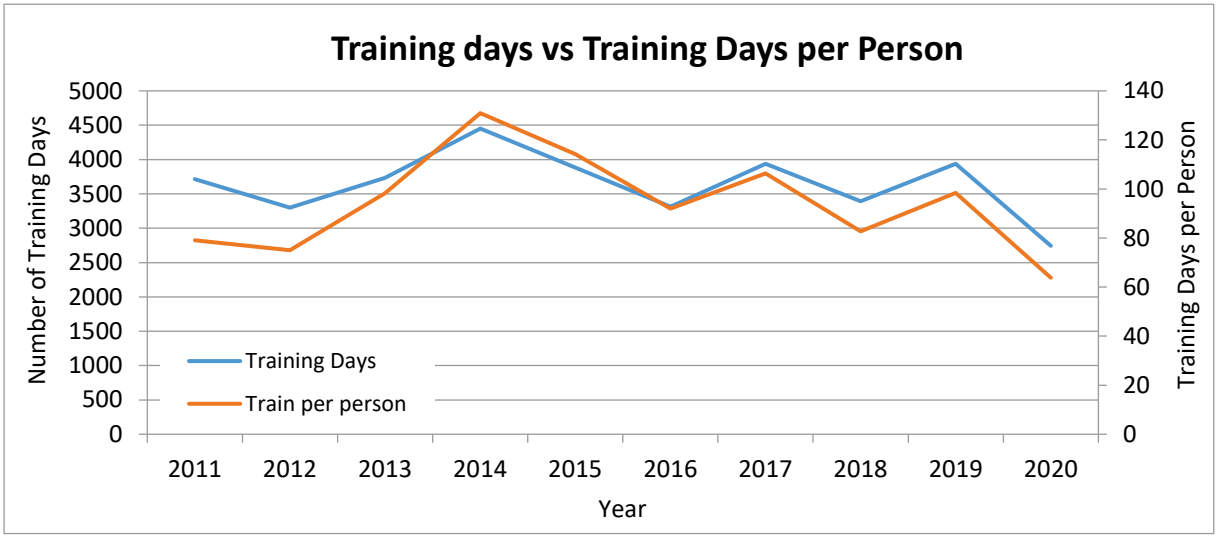
- Number of citations in peer-reviewed journals of a publication about software supported by SLA-funded staff. Please note that not all software packages we support have a citeable publication.
- Number of training days delivered by SLA-funded staff. This metric measures outputs, ie how many people were trained and over how many days.
- Number of publications in peer-reviewed journals authored by SLA funded staff.
- Number of scientific/technical presentations at external events delivered by SLA-funded staff.

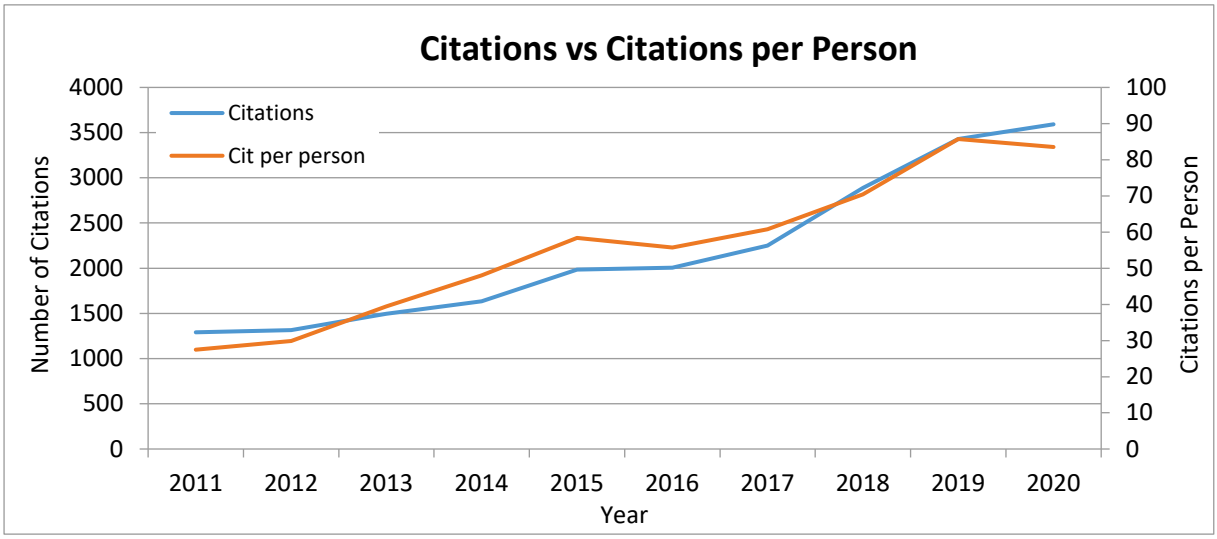
The metrics are per financial year, with the exception of the citations metric, which is for calendar year (for ease of data collection). Please note that for an interim report the data collection for the current year is partial for all metrics except the citations.

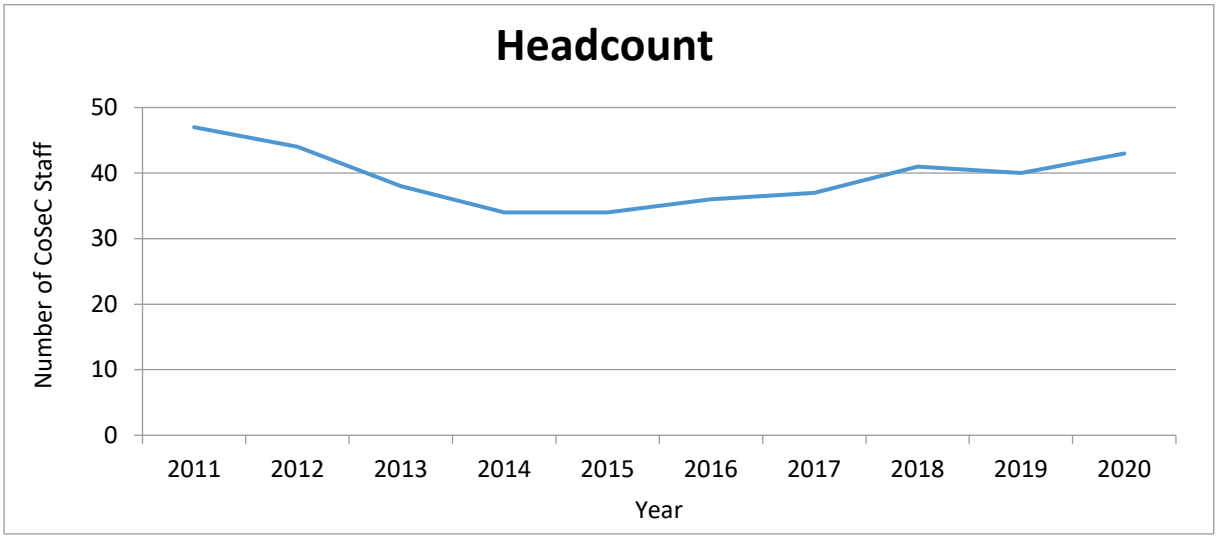
In the metrics graphs below, the blue line refers to the total number while the orange line refers to the number per scientific/technical staff member. This is useful because metrics are included for any person irrespective of whether they are partially or fully funded by the SLA. For information, a graph of the scientific/technical staff headcount has also been included.






*Please note: these metrics are being double checked for the meeting on 10 June 2020.*

**Metrics breakdown by area**

| | FTE | Publications | Presentations | Training Days | Citations |
|---|---|---|---|---|---|
| CCP5 (including DL_ codes) | 3.20 | 4 | 5 | 417.25 | 468 |
| CCP9 | 2.40 | 3 | 3 | 550 | 220 |
| CCPmag | 0.74 | 0 | 0 | 0 | 7 |
| CCP-NC | 1.30 | 4.5 | 4.5 | 163 | 1018 |
| CCPQ | 1.86 | 1 | 3 | 118 | 100 |
| CCP Plasma/HEC Plasma | 0.95 | 0 | 1 | 14 | 80 |
| CCPi | 1.20 | 0 | 2 | 60 | 0 |
| CCP PET-MR | 1.20 | 1 | 1 | 118 | 0 |
| CCP BioSim/HEC BioSim | 2.20 | 1 | 3 | 170.25 | 11 |
| MCC (including ChemShell, CRYSTAL) | 2.50 | 3 | 7 | 345.25 | 201 |
| UKCP (including CASTEP) | 1.00 | 2.5 | 2.5 | 217 | 1472 |
| UKCOMES | 1.00 | 1 | 5 | 337.25 | 15 |
| UK-AMOR * | 0.20 | 0 | 0 | 0 | 0 |
| UKTC | 0.40 | 1 | 1 | 0 | 0 |
| UKCTRF | 0.50 | 1 | 2 | 0 | 0 |
| Software Outlook | 2.00 | 3 | 2 | 234 | 0 |
| **Totals** | | **26** | **42** | **2744** | **3592** |

*NOTE: fractional figures arise due to cross-project activities where the metric is shared between multiple projects.*

** Given the limited specific UK-AMOR FTE available all metrics related to CoSEC support for UK-AMOR for this reporting year are included under the CCPQ CoSeC support report. This will change in 2020-2021.*