M.Sc. in Mathematical Modelling and Scientific Computing Dissertation Projects

February 2025

Contents

1 Industrial Projects

Descriptions of industrial projects have been omitted from the web version of the projects booklet. Please contact Kathryn Gillow, gillowk@maths.ox.ac.uk, if you believe you should have access to these.

2 Biological, Medical and Physical Application Projects

2.1 Understanding the Emergence of Oscillations in Biological Pattern Formation

Supervisor: Prof. Mohit Dalwadi
Contact: dalwadi@maths.ox.ac.uk

The size, shape and patterning of animals varies greatly during development, and can be controlled by chemical signalling molecules called 'morphogens' that generate patterns of cell fates by activating different genes in a concentration-dependent manner [1]. Recent work has shown that biologically-observed pattern scaling with system size can be achieved via an 'expansion-repression' mechanism [2]. We have found that this patterning mechanism is susceptible to a Hopf bifurcation when the system size increases beyond a critical length, resulting in limit cycles in concentration-space (Figure 1). Since oscillations will severely hinder the robustness of biological pattern formation, we hypothesise that this critical length corresponds to a fundamental limit on the sizes of tissues that can be patterned via this mechanism. However, the mechanistic cause of these oscillations is currently not well understood.

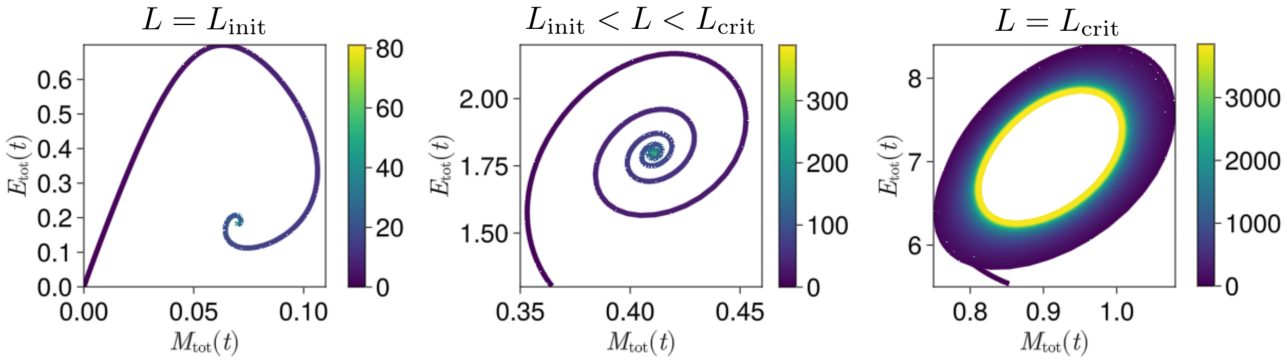

Figure 1: Limit cycles in concentration-space can be observed in simulations when the system size L is increased from an initial length L_{init} to the critical length L_{crit} (colour indicates simulation time).

In this project, you will analyse two coupled nonlinear partial differential equations (PDEs) that model the expansion-repression mechanism in developing tissues as a reaction-diffusion system. This will involve technical asymptotic and nonlinear dynamical systems analysis, in combination with numerically solving the PDEs using, for example, the method of lines. The main, and more technically challenging, aspect of the project will involve analysing the Hopf bifurcations in the system where spontaneous oscillations appear. This will require non-standard linear stability analyses of heterogeneous steady states, involving the technique of matched asymptotic expansions to systematically derive and solve the appropriate equations in different regions of space. If time permits, weakly nonlinear analyses could also be carried out to determine how and why this bifurcation transitions from supercritical to subcritical, and how this should be interpreted biologically.

This project aims to improve our understanding of this oscillation phenomenon by using applied mathematics to determine the relationships between key length and time scales within the system. A successful project would provide a mathematical theory and biological interpretation of the mechanism that generates the observed Hopf bifurcations for this system.

A key prerequisite for this project is excellent technical skills in asymptotic analysis. This project

involves working with Lewis Mosby and Zena Hadjivasiliou from the Francis Crick Institute.

References

[1] L. S. Mosby, A. E. Bowen, and Z. Hadjivasiliou. Morphogens in the evolution of size, shape and patterning. *Development*, 151(18):dev202412, 2024.

[2] D. Ben-Zvi, G. Pyrowolakis, N. Barkai, and B.-Z. Shilo. Expansion-repression mechanism for scaling the Dpp activation gradient in Drosophila wing imaginal discs. *Curr. Biol.*, 21(16):1391–1396, 2011.

2.2 Dynamic Robustness of Cell Polarity in Early Biological Development

Supervisor: Prof. Mohit Dalwadi
Contact: dalwadi@maths.ox.ac.uk

The robust establishment of cell polarity is crucial for early animal development. The PAR signalling pathway plays a central role in the onset of this asymmetry in the model organism *Caenorhabditis elegans* (*C. elegans*) as well as other organisms, including humans, and has a highly conserved function in organizing cell polarity [1]. Experiments show that intracellular signalling cues for polarity vary spatio-temporally over the *C. elegans* cell cycle and across cell divisions. Ultimately polarization must also be robust to changes in initial state and remain capable of reorientation and polarity reversal in response to competing cues. Steady state analysis of mathematical models of this process suggest several steady states are possible [2]. However, the actual system is continually varying spatio-temporally, and recent experiments suggest that despite these dynamic transitions the system invariably reaches the correct states in healthy development (see Figure 2). It is thought that the spatio-temporal variations contribute to the robustness of the outcomes, but the precise mechanisms underlying this are currently unknown.

In this project, you will analyse coupled nonlinear partial differential equations (PDEs) that model the PAR signalling pathway in early *C. elegans* development as a reaction-diffusion system. In particular, this will involve analysing dynamic transitions across bifurcations in the equations, and will involve technical asymptotic and nonlinear dynamical systems analysis, in combination with numerical solutions for PDEs. The analysis will involve techniques such as matched asymptotic expansions and weakly nonlinear analyses. Since the PAR network is conserved across organisms, it will be of particular interest to understand how the relatively fast cell cycle of *C. elegans* affects the dynamic robustness generated and whether slower cell cycles require different robustness mechanisms.

A successful project would provide a mathematical theory and biological interpretation of the mechanism that generates the observed robustness in the system. More broadly, if we can gain mechanistic insight into how dynamic effects in the PAR pathway contribute to the robustness of establishing cell polarity, we can contribute to genetic and pharmacological treatments for the many disorders associated with its dysregulation.

A key prerequisite for this project is excellent technical skills in asymptotic analysis. This project involves working with Nathan Goehring and KangBo Ng from the Francis Crick Institute.

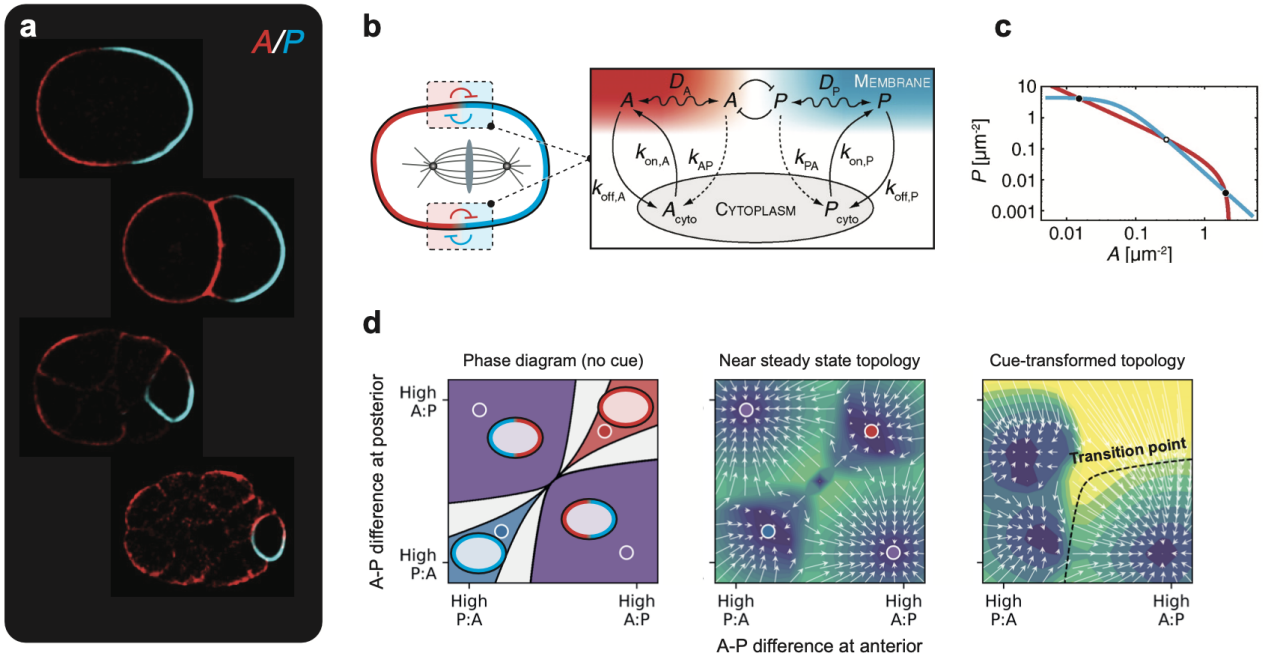

Figure 2:

References

[1] Nathan W. Goehring, Philipp Khuc Trong, Justin S. Bois, Debanjan Chowdhury, Ernesto M. Nicola, Anthony A. Hyman, and Stephan W. Grill. Polarization of PAR proteins by advective triggering of a pattern-forming system. *Science*, 334(6059):1137–1141, 2011.

[2] Philipp Khuc Trong, Ernesto M. Nicola, Nathan W. Goehring, K. Vijay Kumar and Stephan W. Grill. Parameter-space topology of models for cell polarity. *New Journal of Physics*, 16(6):065009, 2014.

2.3 Phase Behavior of Biomolecules

Supervisors: Dr Giulia Celora and Prof. Andreas Münch
Contact: celora@maths.ox.ac.uk and muench@maths.ox.ac.uk

Biomolecular condensates are dynamic, liquid-like compartments within cells composed of macromolecules such as proteins and RNAs. Recent experimental findings indicate that these structures form through liquid-liquid phase separation (LLPS), akin to the way oil droplets spontaneously emerge in water. This discovery has provided new insights into intracellular organization in both healthy and diseased states. Notably, disruptions in LLPS have been linked to conditions such as neurodegeneration and cancer. While the oil droplet analogy is useful, it falls short in explaining why only certain biomolecules undergo phase separation within cells and how this process is regulated by intracellular and external environmental factors. This highlights the need to refine existing theories of phase separation to shade light on the function and control of biomolecular condensates within cells.

This project aims to investigate the electro-chemical aspects of protein liquid-liquid phase separation (LLPS), specifically considering how protonation and deprotonation reactions influence their net charge variability. Gaining insights into this process will be crucial for addressing open questions in cell biology, particularly in understanding how cells adapt to stressful environmental conditions.

We have recently developed a theory for studying the electro-chemical regulation of LLPS of polyelectrolytes (such as RNA) [1], which are polymers that have residues of a single type (either positive or negative charge). We would like to extend our theory to LLPS of polyampholytes (such as proteins) — polymers that contain both positively charged and negatively charged residues.

The first goal would be to familiarize yourself with the phase field theory developed in [1, 2] and the Maxwell construction approach to compute phase separated equilibria [1]. The project might follow different directions depending on the student's interests:

1. Use numerical computation to compute the phase diagram of weak polyampholytes using Maxwell constructions.
2. Derive approximate analytical solutions in specific regimes for which explicit phase-separated solution can be computed.
3. Extend the work to include spatial heterogeneity and develop numerical algorithm to solve for localised solutions.

Various extensions are possible, for example extension of the model to look at the dynamics of phase separation in simple geometries.

References

[1] G. L. Celora, R. Blossey, A. Münch, and B. Wagner. Counterion-controlled phase equilibria in a charge-regulated polymer solution. *J. Chem. Phys.*, 159(18):184902, 14 November 2023. https://doi.org/10.1063/5.0169610.

[2] Y. Avni, R. Podgornik, and D. Andelman. Critical behavior of charge-regulated macro-ions. *J. Chem. Phys.* 153(2):024901, 14 July 2020. https://doi.org/10.1063/5.0011623.

2.4 Visco-Elastic Phase Separation

Supervisor: Prof. Andreas Münch
Contact: muench@maths.ox.ac.uk

Much like oil droplets in a salad dressing or the "eyes" on a fatty soup, it an everyday experience that poorly miscible liquids demix — or phase separate — depending on their material properties and environmental factors. Recently, the importance of *liquid-liquid phase separation* for the formation of membraneless organelles in cell biology has been discovered, and has spawned a large and rapidly growing body of research on this phenomenon and its role for cell biology as well as in pathologies such as neuro-degenerative diseases. It is important to understand when phase separation arises, but also how it evolves i.e. the dynamics of phase separation.

This project will focus on the case of liquid-liquid phase separation that takes into account that one of the constituents are often large molecules such as proteins, and the other is much more mobile small

molecules (usually a solvent) and study, the commonly called, *viscoelastic phase separation*. Insights from this problem will be fundamental to a range of applications and open problems in cell biology [1].

Models for this phenomenon are found for example in [2] We have recently developed a theory for the long-time evolution of the phase-separated regions that we would like to test numerically.

The task in this project could be, for example, to

(a) Implement a numerical algorithm to solve the phase field model in [2], in particular (9) (FEniCS templates/code snippets are available);

(b) Derive approximate analytic solution for simple geometries, building on our existing theory, and comparing these with the numerical results.

The first goal would be to verify and discuss the theory developed for these viscoelastic phase separation models. Various extensions are possible, for example modification of the models, and consideration of other stages of the phase separation process and other geometries.

References

[1] H. Tanaka. Viscoelastic Phase Separation in Biological Cells. *Communications Physics*, 5(1):1–12, 2022.

[2] D. Zhou, P. Zhang, and W. E. Modified Models of Polymer Phase Separation. *Physical Review E*, 73(6):061801, 2006.

2.5 Dynamical Systems and Network Modelling of Stratosphere-Troposphere Coupling

Supervisors: Prof. Irene Moroz, Prof. Renaud Lambiotte, and Dr Scott Osprey
Contact: moroz@maths.ox.ac.uk, lambiotte@maths.ox.ac.uk, and Scott.Osprey@physics.ox.ac.uk

A weakening of the winter Arctic stratospheric polar vortex can propagate downward into the lower troposphere, with a significant impact on the predictability of surface weather. In most cases, couplings are identified through the occurrence of Sudden Stratospheric Warnings. This can occur when geopotential height crosses a value in the stratosphere and the polar vortex starts to break down.

An earlier investigation [1] used Marwan's recurrence methods and network analysis on a coupled 6-D troposphere-stratosphere model, due to Lorenz [2].

The current project aims to extend the original analysis to include more realistic coupled stratosphere-troposhere dynamical models, such as shallow water equations.

References

[1] Susheel Adusumilli. *Stratosphere-Troposphere Interactions Diagnosed through Complex Network Analysis*. M.Sc. in Mathematical Modelling and Scientific Computing dissertation, Mathematical Institute, University of Oxford, 2015.

[2] Edward N. Lorenz. Irregularity: a fundamental property of the atmosphere. *Tellus A*, 36(2):98–110, 1984.

[3] Norbert Marwan, M. Carmen Romano, Marco Thiel, and Jürgen Kurths. Recurrence plots for the analysis of complex systems. *Physics Reports*, 438(5):237–329, 2007.

2.6 Kuramoto Systems with Multiple Phases

Supervisor: Prof. Peter Grindrod
Contact: grindrod@maths.ox.ac.uk

Kuramoto systems usually couple-up units that are simple oscillators, with a single phase variable each, all on a lattice... but what if each unit has multiple phase variables? How will we analyse the resulting patterns or categorise such, and how can we visualise them? These systems are simple models for human cognition.

2.7 Modelling of a Network of Interacting Spiking Neurons

Supervisors: Dr Alexandra Holzinger and Prof. Maria Bruna
Contact: holzinger@maths.ox.ac.uk and bruna@maths.ox.ac.uk

Background and problem statement: This dissertation project covers a mathematical research area at the interface between partial differential equations and stochastic interacting particle systems. Motivated by statistical physics, mean-field limits have become a powerful tool in the derivation of Partial Differential Equations (PDEs). Due to their fruitful applications in physics, biology, neuroscience and artificial intelligence, this mathematical concept became highly influential. The core strength of mean-field limits is the dimension reduction from a large system of equations to an effective, partial differential equation representing the macroscopic level [1]. In this project, we will use this framework to model a system of interacting neurons of different types [2], which could represent variations in a population of neurons.

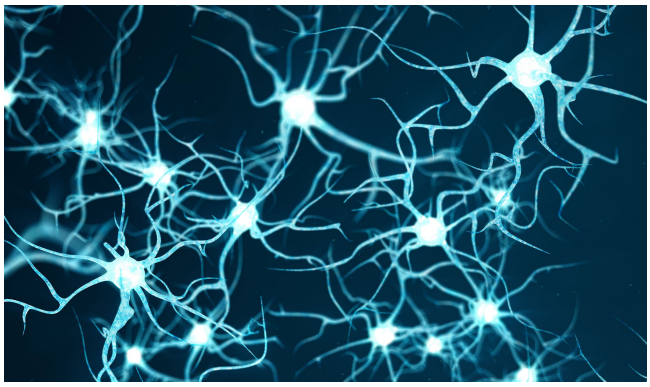

Figure 3: SEM image of a neural network (Paul de Koninck, Laval University).

Description of the planned approach and the techniques needed: The starting point will be a system of ordinary (stochastic) differential equations representing networks of Hodgkin–Huxley or FitzHugh–Nagumo neurons. The student will investigate their behaviour via simulations and identify changes depending on model parameters. They will then consider the associated mean-field PDE model and again use simulations to compare the macroscopic behaviour to the stochastic system as the number of neurons N is varied. Finally, it is interesting to consider the role of fluctuations, which are neglected in the mean-field PDE model, and how these can be accurately taken into account in an effective model setting. Useful background and/or special topics are:

- Stochastic processes, e.g., Stochastic Modelling of Biological Processes.
- Mathematical Physiology.

Reasonable expected outcome of the project: Development of a new PDE model for two species of neurons from a microscopic system. Validation of the derived model using numerical simulations. Review of the rigorous proof in [2] and possibly extend it to different neuron types. Characterise the effect of fluctuations in two-species models via simulations or analytical tools.

References

[1] H. Tanaka and M. Hitsuda. Central limit theorem for a simple diffusion model of interacting particles. *Hiroshima Mathematical Journal*, 11(2):415–423, 1981.

[2] J. Baladron, D. Fasoli, O. Faugeras, and J. Touboul. Mean-field description and propagation of chaos in networks of Hodgkin-Huxley and FitzHugh-Nagumo neurons. *The Journal of Mathematical Neuroscience*, 2:1–50, 2012.

[3] B. Duchet, C. Bick, and Á. Byrne. Mean-field approximations with adaptive coupling for networks with spike-timing-dependent plasticity. *Neural Computation*, 35(9):1481–1528, 2023.

2.8 Modelling Subglacial Discharge from Beneath Marine-Terminating Glaciers

Supervisor: Prof. Ian Hewitt
Contact: hewitt@maths.ox.ac.uk

Marine-terminating glaciers in Greenland and Antarctica discharge ice into the ocean. Their dynamics control the mass balance of these large ice sheets, and are therefore a fundamental factor influencing sea-level change. They are also the location at which a large amount of meltwater flows from beneath the ice sheet into the ocean. This project aims to use a mathematical model to investigate what is going on in these hard-to-probe regions.

The project would build on recent modelling work that has considered the flow of cold fresh water from beneath the ice sheet into the ocean, and how this interacts with the ocean water, which — due to its larger density — may intrude beneath the ice sheet as a gravity current [1]. Since the ocean water is generally warmer, this may have an important influence on the melting and stability of the ice sheet [2].

Depending on the student's interest, the specific line of enquiry could include modifying the model to:

1. consider the role of seasonally varying subglacial discharge
2. consider the role of the tides which, like a river estuary, cause daily variations in the outflow, possibly incorporating the tidally-induced uplift of the ice.
3. consider sediment transport in the subglacial water flow, perhaps following ideas from Ref. [3].

There is potential to collaborate with colleagues at the Geological Survey of Denmark and Greenland, who are planning to take measurements from a marine-terminating glacier in Greenland later this year.

This project would require interest in mathematical modelling and fluid mechanics. It will require developing numerical methods and/or asymptotic methods to solve systems of PDEs.

References

[1] E. A. Wilson, A. J. Wells, I. J. Hewitt, and C. Cenedese. The dynamics of a subglacial salt wedge. *Journal of Fluid Mechanics*, 895, 2020. https://doi.org/10.1017/jfm.2020.308

[2] A. Bradley and I. J. Hewitt. Tipping point in ice-sheet grounding-zone melting due to ocean water intrusion. *Nature Geoscience*, 17:631–637, 2024. https://doi.org/10.1038/s41561-024-01465-7

[3] I. J. Hewitt and T. T. Creyts. A model for the formation of eskers. *Geophys. Res. Lett.*, 46(12): 6673–6680, 2019. https://doi.org/10.1029/2019GL082304

2.9 Modelling the Zoonotic Capabilities of Mpox and other Poxviruses from Viral Genomic Traits

Supervisor: Dr Jasmina Panovska-Griffiths
Contact: jasmina.panovska-griffiths@queens.ox.ac.uk
Collaborator: Dr Liam Brierly, University of Glasgow

Mpox virus re-emerged to cause a globally significant outbreak in 2022 and has continued to cause national outbreaks in Central Africa into 2024. Although mpox is firmly recognised as an epidemic threat, differences in infection propensity may exist between its distinct clades (clade IIb causing the global outbreak, but clade I circulating in central Africa).

Furthermore, several other animal viruses in the *O*rthopox genus and *Poxviridae* family are also known to occasionally cause zoonoses [6] and many further viruses are suspected — but yet not determined — to have potential to infect humans.

It is possible to identify signals of host suitability in compositional biases adapted toward/against use of particular coding patterns or 'motifs' in virus genome sequences [1]. These motifs can be used as features to train machine learning models that can then successfully estimate which virus species can cause zoonotic infection [2, 4]. Similarly, bioinformatic approaches have begun to predict animal hosts origins of orthopoxviruses [5]. However, to understand which orthopoxviruses might have zoonotic potential at the level of clade and subspecies, more targeted models are necessary.

This project will compare the genomic and proteomic characteristics in poxvirus genome sequences from human infections versus other vertebrates to identify "pre-adaptations" in animal viruses that predispose to zoonotic spillover.

The project will start by extracting viral sequences from NCBI Virus, which currently hosts over 2700 whole genome sequences from 28 orthopoxvirus species within 15 host species. The student will learn how to programmatically query, filter, and process sequences via open-source tools in a programming language such as R [3].

The student will calculate measures of genome composition and functional protein chemistry already known to be likely predictors of zoonotic capability, e.g., Conjoint Triad, Composition-Transition-Distribution, or Pseudo-amino acid composition.

The student will then train machine learning classification algorithm(s), (e.g., random forests, support vector machines) to can classify human versus non-human hosts from sequence features. The resulting system can then be tested to make predictions on zoonotic capability of out-of-sample viruses.

Expected outcomes:

1. Data analysis via processing data on genomic motifs of orthopoxviruses.
2. Developing machine learning models to predict zoonotic transmission from orthopoxviruses.
3. Develop a ranked model system to predict likely zoonoses among currently known viruses circulating in animals.
4. Journal publication of the results.

Prerequisite MMSC courses: Mathematical Methods; Perturbation Methods; Optimisation for Data Science; Theories of Deep Learning.

Desirable courses: Mathematical Modelling; Continuous Optimisation; Optimal Control; Stochastic Modelling of Biological Processes.

References

[1] Simon A. Babayan, Richard J. Orton, and Daniel G. Streicker. Predicting Reservoir Hosts and Arthropod Vectors from Evolutionary Signatures in RNA Virus Genomes. *Science* 362(6414):577–80, 2018. https://doi.org/10.1126/science.aap9072

[2] Jakub M. Bartoszewicz, Anja Seidel, and Bernhard Y. Renard. Interpretable Detection of Novel Human Viruses from Genome Sequencing Data. *NAR Genomics and Bioinformatics*, 3(1):lqab004, 2021. https://doi.org/10.1093/nargab/lqab004

[3] Liam Brierley. Using open-access tools (taxize, rentrez) to find coronaviruses, their genetic sequences, and their hosts. *rOpenSci blog*, 2020. https://ropensci.org/blog/2020/11/10/coronaviruses-and-hosts

[4] Nardus Mollentze, Simon A. Babayan, and Daniel G. Streicker. Identifying and Prioritizing Potential Human-Infecting Viruses from Their Genome Sequences. *PLOS Biology*, 19(9): e3001390, 2021. https://doi.org/10.1371/journal.pbio.3001390

[5] Katie K. Tseng, Heather Koehler, Daniel J. Becker, Rory Gibb, Colin J. Carlson, Maria del Pilar Fernandez, and Stephanie N. Seifert. Viral Genomic Features Predict Orthopoxvirus Reservoir Hosts. *bioRxiv*, 2023. https://doi.org/10.1101/2023.10.26.564211

[6] Mark E. J. Woolhouse, Liam Brierley, Chris McCaffery, and Sam Lycett. Assessing the Epidemic Potential of RNA and DNA Viruses. *Emerging Infectious Diseases*, 22(12):2037–44. https://doi.org/10.3201/eid2212.160123

3 Numerical Analysis Projects

3.1 Numerical Linear Algebra or Approximation Theory

Supervisor: Prof. Yuji Nakatsukasa
Contact: nakatsukasa@maths.ox.ac.uk

I would be happy to supervise MMSC dissertations on topics in (randomised) numerical linear algebra (NLA) or Approximation Theory (Approx).

Depending on your preference, the focus can be theoretical, computational/algorithmic, or application oriented. Potential projects include

- NLA: Explore the decay of singular values for matrices consisting of parameter-dependent vectors (e.g. [7]).
- NLA: Topics in randomised algorithms in numerical linear algebra, e.g. for least-squares problems [4], low-rank approximation [5], or trace estimation [3].
- Approx: Exploring the use of the AAA algorithm for rational approximation, particularly in the context of model order reduction [1, 2].
- Approx: Explore the phenomenon of overfitting [6] from the perspective of Lebesgue constants [8].

References

[1] A. C. Antoulas, C. A. Beattie, and S. Güğercin. *Interpolatory methods for model reduction*. SIAM, 2020.

[2] P. Benner, S. Gugercin, and K. Willcox. A survey of projection-based model reduction methods for parametric dynamical systems. *SIAM Rev.*, 57(4):483–531, 2015.

[3] A. Cortinovis and D. Kressner. On randomized trace estimates for indefinite matrices with an application to determinants. *Found. Comput. Math.*, 22(3):875–903, 2022.

[4] E. N. Epperly, M. Meier, and Y. Nakatsukasa. Fast randomized least-squares solvers can be just as accurate and stable as classical direct solvers. *arXiv preprint:* https://arxiv.org/abs/2406.03468, 2024.

[5] N. Halko, P.-G. Martinsson, and J. A. Tropp. Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions. *SIAM Rev.*, 53(2):217–288, 2011.

[6] T. Hastie, A. Montanari, S. Rosset, and R. J. Tibshirani. Surprises in high-dimensional ridgeless least squares interpolation. *Ann. Stat.*, 50(2):949–986, 2022.

[7] D. Kressner and C. Tobler. Low-rank tensor Krylov subspace methods for parametrized linear systems. *SIAM J. Matrix Anal. Appl.*, 32(4):1288–1316, 2011.

[8] L. N. Trefethen. *Approximation Theory and Approximation Practice, Extended Edition*. SIAM, 2019.

3.2 Numerical Linear Algebra

Supervisor: Dr Nathaniel Pritchard
Contact: pritchard@maths.ox.ac.uk

Description of Proposal: I am happy to supervise MMSC disserations on topic relating to Randomized Numerical Linear Algebra. Depending on what you prefer, the project can be more theoretical, computational, or application focused. I have included two potential projects, both of which will require some coding. The projects are

1. **Cimmino's Algorithm (joint with Prof. Yuji Nakasukasa):** The Randomized Kaczmarz algorithm has been extensively studied because of it's nice geometric interpretations and linear convergence [1]. In practice, it often suffers from a lack of parallelizability and slow linear convergence. In this project, we will focus on a related method known as Cimmino's algorithm, which has an equally intuitive geometric motivation and is easily parallelizable, but has a convergence behavior that can be worse than Randomized Kaczmarz [2]. Specifically, this project aims to examine different properties of Cimmino's algorithm to understand its performance in comparison to Randomized Kaczmarz.

2. **Randomized CUR:** SVD has been long-used in data analysis to for low-rank approximations. These low-rank models are then used to accelerate model prediction, cluster data points, or compress images [3]. A related approach is known as CUR which aims to find rows and columns of the data that best represent the data as a whole. Determining the subset that best represents the matrix is a combinatorially hard problem, but randomized approaches have been able to find close to optimal solutions [4, 5]. Unfortunately, the randomized approaches can produce different CUR approximations with different runs of the algorithm, which can be disadvantageous for some applications. This project will examine the variability in the different randomized CUR approaches, considering aspects like the consistency of index selection, and the spectral properties of the CUR approximation residual.

References

[1] Thomas Strohmer and Roman Vershynin. A randomized Kaczmarz algorithm with exponential convergence. *Journal of Fourier Analysis and Applications* 15(2):262–278, 2009.

[2] Magherita Guida and Carlo Sbordone. The Reflection Method for the Numerical Solution of Linear Systems. *SIAM Review* 65(4):1137–1151, 2023.

[3] Michael W. Mahoney and Petros Drineas. CUR matrix decompositions for improved data analysis. *Proceedings of the National Academy of Sciences* 106(3):697–702, 2009.

[4] Alexander Osinsky. Close to optimal column approximations with a single SVD. *arXiv preprint*, https://arxiv.org/abs/2308.09068, 2023.

[5] Yijun Dong and Per-Gunnar Martinsson. Simpler is better: a comparative study of randomized pivoting algorithms for CUR and interpolative decompositions. *Advances in Computational Mathematics* 49(4):66, 2023.

3.3 Second-Order Optimization Methods for Scientific Machine Learning

Supervisors: Prof. Coralia Cartis and Dr Sadok Jerad
Contact: cartis@maths.ox.ac.uk and jerad@maths.ox.ac.uk

Leveraging the past success of deep learning in addressing complex statistical problems, [3] introduced a novel class of deep learning architectures known as Physical-Informed Neural Networks (PINNs). These innovative methods represent a promising alternative to traditional numerical Partial Differential Equation (PDE) solvers for both forward and inverse problems. PINNs offer significant advantages, including faster computation of approximate solutions compared to standard mesh discretization techniques.

However, when high-precision solutions are required, PINNs can suffer from ill-conditioning. Standard adaptive gradient methods optimizers, such as Adam, may not be sufficient in these scenarios. Recent research by [4] suggests that quasi-Newton methods and second-order methods can yield better results in such cases. Since PINNs' optimization problems suffer from both non-convexity and ill-conditioning, second-order methods emerge as a possible candidate to address these issues. The goal of this master dissertation is to empirically test the performance of various second-order methods (trust-region, linesearch, BFGS and others) and to explore some new variants that are more suitable in the stochastic case [2, 1].

First, a comprehensive review of existing optimization techniques used in PINNs will be conducted, focusing on their strengths and limitations. Experiments will be performed to compare the performance of second-order methods with standard adaptive gradient methods such as Adam. The project aims to investigate the potential advantage of stochastic second order optimization methods over standard adaptive gradient methods in the context of PINNs and to develop scalable stochastic second order algorithms. The results are expected to contribute to the development of more efficient and accurate deep learning solutions for complex PDE problems and to a better understanding of randomized Newton methods. Core courses in nonlinear optimization and numerical linear algebra will be very useful. For the PINNs part of the project, prior knowledge of deep learning is required. In addition, for the theoretical analysis, some knowledge of probability will be important for the investigation of stochastic algorithms.

References

[1] Coralia Cartis, Jaroslav Fowkes, and Zhen Shao. Randomised subspace methods for non-convex optimization, with applications to nonlinear least-squares. *arXiv preprint:* https://arxiv.org/abs/2211.09873, 2022.

[2] Serge Gratton, Sadok Jerad, and Philippe L. Toint. A stochastic objective-function-free adaptive regularization method with optimal complexity. *arXiv preprint:* https://arxiv.org/abs/2407.08018, 2024.

[3] M. Raissi, P. Perdikaris, and G. E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019.

[4] Pratik Rathore, Weimu Lei, Zachary Frangella, Lu Lu, and Madeleine Udell. Challenges in training pinns: a loss landscape perspective. In *Proceedings of the 41st International Conference on Machine Learning*, ICML'24, 2024.

3.4 Continuous and Discrete Perspectives on Optimization Algorithms

Supervisor: Prof. Coralia Cartis
Contact: cartis@maths.ox.ac.uk

There is a close connection between optimization algorithms and dynamical systems: the iterates' trajectory can be viewed as a discretised underlying ODE/dynamical system. This perspective has given nice connections between the fields of control, dynamical systems and ODEs and optimization; see papers of Attouch et al [1, 2] or online. In this project, I am keen that we look in particular at applying such (discretised) continuous perspectives to global optimization, and looking at (discretised) trajectories of deterministic and stochastic dynamical systems (such as Langevin). Alternatively, one may look into second order methods from dynamical system perspective. This project will NOT address accelerated first-order methods.

References

[1] H. Attouch and J. Fadili. From the Ravine method to the Nesterov method and vice versa: a dynamical system perspective. *arXiv preprint:* https://arxiv.org/abs/2201.11643, 2022.

[2] H. Attouch, Z. Chbani, J. Fadili, and H. Riahi. First-order optimization algorithms via inertial systems with Hessian driven damping. *Mathematical Programming*, 193(1):113–155, 2022.

[3] B. Gidas. Global optimization via the Langevin equation. In: *24th IEEE Conference on Decision and Control, Fort Lauderdale, FL, USA*, pp. 774-778, 1985. https://doi.org/10.1109/CDC.1985.268602

[4] B. Gidas. The Langevin Equation as a Global Minimization Algorithm. In: E. Bienenstock, F. F. Soulié, and G. Weisbuch(eds) *Disordered Systems and Biological Organization*. NATO ASI Series, vol 20. Springer, Berlin, Heidelberg, 1986. https://doi.org/10.1007/978-3-642-82657-3_31

[5] Pan Xu, Jinghui Chen, Difan Zou, and Quanquan Gu. Global Convergence of Langevin Dynamics Based Algorithms for Nonconvex Optimization. Part of *Advances in Neural Information Processing Systems 31*, (NeurIPS 2018), 2018.

[6] Y. Huang, M. Herty, D. Kalise, and N. Kantas. Fast and robust consensus-based optimization via optimal feedback control. *arXiv preprint:* https://arxiv.org/abs/2411.03051, 2024.

3.5 Finding Multiple Solutions of Optimization Problems

Supervisor: Prof. Coralia Cartis
Contact: cartis@maths.ox.ac.uk

This project looks at using local optimization algorithms to find multiple local/global solutions of optimization problems; in particular it would use deflation techniques to do so, see Farrell et al [2]. An extension of the latter work for optimization was already undertaken in Riley et al [1]. Here, we

are looking at either improving the latter by addition of trust region of regularization or at extending to derivative free optimization or for fixed point systems.

References

[1] Alban Bloor Riley, Marcus Webb, and Michael L. Baker. Deflation Techniques for Finding Multiple Local Minima of a Nonlinear Least Squares Problem. *arXiv preprint:* https://arxiv.org/abs/2409.14438, 2024.

[2] P. E. Farrell, A. Birkisson, and S. W. Funke. Deflation Techniques for Finding Distinct Solutions of Nonlinear Partial Differential Equations. *SIAM J. Sci. Comp.*, 37:A2026–A2045, 2015. https://doi.org/10.1137/140984798

3.6 Multilevel Monte Carlo Projects

Supervisor: Prof. Mike Giles
Contact: gilesm@maths.ox.ac.uk

Here are some outline ideas for computational projects. MLMC refers to Multilevel Monte Carlo, for which I have many references available from https://people.maths.ox.ac.uk/gilesm/mlmc.html and https://people.maths.ox.ac.uk/gilesm/mlmc_community.html.

These projects are all best suited to students with solid programming skills in either Python or C/C++. I can assist any students who are interested in improving their C/C++ skills by using OpenMP for multi-threaded parallelisation on CPUs, or CUDA for many-threaded paralleisation on GPUs.

Broad avenues of investigation are:

- MLMC for stochastic PDEs, possibly based on [1] and possibly also involving HPC implementations using CUDA (NVIDIA GPUs) or OpenMP (CPUs).
- MLMC with an HPC focus on reduced precision arithmetic using CUDA (NVIDIA GPUs) or OpenMP (CPUs), based on [2].
- MLMC for randomised Numerical Linear Algebra, possibly jointly supervised by Prof Yuji Nakatsukasa.
- A topic in High Performance Computing, using OpenMP on CPUs or CUDA on GPUs, perhaps as part of a project with another supervisor. I have some specific ideas above for CUDA projects, but they would need someone with very strong programming skills who is interested in immersing themselves in the challenges of parallel programming on GPUs.

References

[1] Federico Cornalba and Julian Fischer. Multilevel Monte Carlo Methods for the Dean–Kawasaki Equation from Fluctuating Hydrodynamics. *SIAM Journal on Numerical Analysis*, 63(1):262–287, 2025.

[2] Mike Giles and Oliver Sheridan-Methven. Analysis of nested multilevel Monte Carlo using approximate Normal random variables *arXiv preprint:* https://arxiv.org/abs/2102.08164, 2021.

3.7 Hybrid Solvers for Linear Inverse Problems with an Unmatched Backprojector

Supervisor: Dr Malena Sabaté Landman
Contact: sabatelandma@maths.ox.ac.uk

Inverse problems — or the reconstruction of hidden objects or parameters from possibly noisy indirect measurements — naturally appear in a wide variety of applications, e.g. in medicine (computed tomography, CT), industry (non-destructive testing) and science (geophysics). We consider linear discrete inverse problems which can be of very large-scale and whose reconstructions are typically very sensitive to perturbations in the measurements. Using Krylov subspace methods is a very natural choice of solver in this case, since they are efficient; matrix-free, and have regularizing properties. However, in some cases, the matrix-vector products with the projector A and the backprojector B are constructed separately such that $B \neq A^T$ exactly. This is also known as the unmatched backprojector problem and it is ubiquitous in tomography applications where it appears as a by-product of the efficient implementations of the corresponding matrix-vector products in hardware.

In this project, you will study a class of Krylov subspace solvers, called AB/BA-GMRES [1]; which were originally developed to mitigate the effect of the unmatched backprojector problem. When assuming noise in the measurements, these iterative projection methods only produce meaningful reconstructions if they are stopped early (the noise is amplified after that). However, this produces reconstructions that are very sensitive to the choice of stopping criterion. An alternative would be to use an ℓ_2, a.k.a Tikhonov, regularization in the projected problems. We call this a 'hybrid' method [2].

By the end of this project, you should have developed a hybrid version of AB/BA-GMRES, as well as a simple implementation in MATLAB to perform comparisons to other standard Krylov solvers for CT applications. You should have also gained a theoretical understanding of the convergence of this new method.

References

[1] P. C. Hansen, K. Hayami, and K. Morikuni. GMRES methods for tomographic reconstruction with an unmatched back projector. *Journal of Computational and Applied Mathematics*, 413:377–427, 2022.

[2] J. Chung and S. Gazzola. Computational Methods for Large-Scale Inverse Problems: A Survey on Hybrid Projection Methods. *SIAM Review*, 66:205–284, 2024.

3.8 Parallel Inner-Product Free Krylov Solvers

Supervisors: Dr Malena Sabaté Landman and Prof. Mike Giles
Contact: sabatelandma@maths.ox.ac.uk and gilesm@maths.ox.ac.uk

Inverse problems — or the reconstruction of hidden objects or parameters from possibly noisy indirect measurements — naturally appear in a wide variety of applications, e.g. in medicine (computed tomography, CT), industry (non-destructive testing) and science (geophysics). We consider linear discrete inverse problems which can be of very large-scale and whose reconstructions are typically very

sensitive to perturbations in the measurements. Using Krylov subspace methods is a very natural choice of solver in this case, since they are efficient; matrix-free, and have regularizing properties. Moreover, the development and commercialization of modern computer architectures, has sparked a renewed interest in methods exploiting low precision arithmetic and parallelization. In this case, avoiding inner products might be beneficial since they can cause undesired early stopping of the algorithms due to reorthogonalizing leading to a significant loss of information in low precision; and require global communication.

In this project, you will study two solvers based on the Hessenberg method: the changing minimal residual Hessenberg method (CMRH) [1, 2] and the least-squares LU method (LSLU) [3]; and explore new efficient parallel implementations. Note that for CMRH, which was initially developed as a solver for linear systems, some parallel implementations already exist, see, e.g. [4]. However, these assume that the system is dense, and are focused on making matrix-vector-products and swap operations more efficient. In this project, we will assume that matrix-vector are cheap, for example, because the system matrix is sparse or structured. Since LSLU is a new method, no parallel implementations exist.

By the end of this project, you should have developed efficient parallel versions of CMRH and LSLU. You will also work towards coding parallel implementations to assess the performance improvements of the new methods compared to conventional Krylov solvers in the context of linear large-scale inverse problems.

References

[1] H. Sadok. CMRH: A new method for solving nonsymmetric linear systems based on the Hessenberg reduction algorithm. *Numerical Algorithms*, 20:485–501, 1999.

[2] A. N. Brown, J. G. Nagy, and M. Sabaté Landman. H-CMRH: An Inner Product Free Hybrid Krylov Method for Large-Scale Inverse Problems. *SIMAX*, 46:232–255, 2024.

[3] A. N. Brown, J. Chung, J. G. Nagy, and M. Sabaté Landman. Inner Product Free Krylov Methods for Large-Scale Inverse Problems. *arXiv preprint*, https://arxiv.org/abs/2409.05239, 2024.

[4] Sébastien Duminil, Mohammed Heyouni, Philippe Marion, and Hassane Sadok. Algorithms for the CMRH method for dense linear systems. *Numerical Algorithms*, 71:383–394, 2016.

3.9 Automated Parallel Goal-Based Adaptivity

Supervisor: Prof. Patrick Farrell
Contact: farrellp@maths.ox.ac.uk

Background and problem statement: When solving a mathematical problem numerically, it is crucial that the accuracy of the approximation be quantified. Scientists, engineers, and policymakers use numerical solutions for design and decisions; can they rely on those approximations?

Since the 1990s, a mathematical theory of *goal-oriented error estimation and adaptivity* for partial differential equations (PDEs) has been developed [1]. If the variational statement of our PDE problem is

$$u \in V : \quad F(u; v) = 0 \quad \text{for all } v \in V,$$

and our discrete approximation is

$$u_h \in V_h : \quad F(u_h; v_h) = 0 \quad \text{for all } v_h \in V_h,$$

for a discrete approximation space V_h, the aim of goal-based error estimation is to compute η such that

$$\eta \approx |J(u) - J(u_h)|$$

for a goal functional $J : V \to \mathbb{R}$. The aim of goal-based adaptivity is to efficiently design an adapted mesh with associated discrete function space V_h such that

$$\eta < \varepsilon$$

for a target approximation error $\varepsilon > 0$. As a concrete example, F might represent the Navier–Stokes equations, u the velocity, pressure, and temperature of the fluid, and J represent the drag coefficient of a candidate wing design.

Heretofore goal-based adaptivity has been applied on a case-by-case basis. The framework requires tailoring to each individual situation, limiting its applicability for scientists and engineers. In this project we aim to automate the application of goal-based adaptivity in the finite element computing environment Firedrake [2].

Planned approach: Our initial steps will follow the work conducted by Rognes and Logg in 2013 [3]. This approach automated some of the components of goal-based adaptivity (strong residual estimation, adjoint derivation), but only worked in serial, and did not have the tools to efficiently approximate the adjoint solutions required for an effective implementation. Our approach will work natively in parallel, and we will investigate new ideas on cheap adjoint approximations that might improve the computational efficiency of the adaptive procedure.

Once this is done, if time permits, we will investigate extensions not considered by Rognes and Logg [3]: to time-dependent problems, to eigenproblems, and to variational inequalities.

Desired outcomes: One output of the MSc research project will be a software contribution to the Firedrake system making goal-based adaptivity available to non-expert users. An aspirational outcome would be to apply goal-based adaptivity to three-dimensional simulations of convection in a nuclear fusion reactor.

Prerequisites: This project depends on a good working knowledge of *Finite Element Methods for PDEs* and *Python in Scientific Computing*.

References

[1] R. Becker and R. Rannacher. An optimal control approach to *a posteriori* error estimation in finite element methods. *Acta Numerica*, 10:1–102, 2001. https://doi.org/10.1017/S0962492901000010

[2] F. Rathgeber, D. A. Ham, L. Mitchell, M. Lange, F. Luporini, A. T. T. Mcrae, G.-T. Bercea, G. R. Markall, and P. H. J. Kelly. Firedrake: automating the finite element method by composing abstractions. *ACM Transactions on Mathematical Software*, 43(3):1–27, 2016. https://doi.org/10.1145/2998441

[3] M. E. Rognes and A. Logg. Automated goal-oriented error control I: Stationary variational problems. *SIAM Journal on Scientific Computing*, 35(3):C173–C193, 2013. https://doi.org/10.1137/10081962X

3.10 Adaptive Multigrid Solvers for Electromagnetism

Supervisor: Prof. Patrick Farrell
Contact: farrellp@maths.ox.ac.uk

Background and problem statement: Adaptive discretisations aim to compute an approximation $u_h \in V_h$ of the solution of a partial differential equation (PDE) $u \in V$ that achieve a target error bound $\varepsilon > 0$:

$$\|u_h - u\| \lesssim \varepsilon.$$

To do this, they first solve on a coarse mesh, estimate the local contributions to the error, and refine those areas of the mesh that contribute the most error. The process is then iterated until the error bound is satisfied.

Adaptive procedures thus naturally build hierarchies of meshes, and require the repeated solution of the same PDE. They naturally lend themselves to multigrid solvers, which exploit this hierarchy of meshes to accelerate the solution of the PDE on the finest grid to hand.

The finite element computing framework Firedrake [1] has excellent support for multigrid solvers [2]. As of very recent work, it also has excellent support for adaptive discretisations using Netgen [3, 4]. In this project we will combine the two, allowing multigrid solvers to be used on mesh hierarchies built from adaptivity. We will then employ the developed technology to investigate potential combined adaptive strategies and solvers for some model PDEs arising in electromagnetism.

Planned approach: Our first steps will implement the work of Miraci et al. [5], who efficiently combine adaptivity and multigrid to develop fast and p-robust solvers for the Poisson equation (p being the polynomial degree). However, the software developed will apply to much more general problems, allowing us to branch out to PDEs not yet considered in the literature. A particular target will be to develop p-robust adaptive multigrid solvers for the canonical problems of curl and div:

$$u \in H(\mathrm{curl}) : (u, v) + (\mathrm{curl}\, u, \mathrm{curl}\, v) = (f, v) \text{ for all } v \in H(\mathrm{curl}),$$
$$u \in H(\mathrm{div}) : (u, v) + (\mathrm{div}\, u, \mathrm{div}\, v) = (f, v) \text{ for all } v \in H(\mathrm{div}).$$

These PDEs are important in the fast solution of Maxwell's equations arising in electromagnetism.

Desired outcomes: One output of the MSc research project will be a software contribution to the Firedrake system making possible the combination of adaptivity and multigrid. If we achieve good results on adaptive multigrid solvers for the canonical problems of curl and div, they will be publishable in a good numerical analysis or scientific computing journal.

Prerequisites: This project depends on a good working knowledge of *Finite Element Methods for PDEs*, *Numerical Linear Algebra*, and *Python in Scientific Computing*.

References

[1] F. Rathgeber, D. A. Ham, L. Mitchell, M. Lange, F. Luporini, A. T. T. Mcrae, G.-T. Bercea, G. R. Markall, and P. H. J. Kelly. Firedrake: automating the finite element method by composing abstractions. *ACM Transactions on Mathematical Software*, 43(3):1–27, 2016. https://doi.org/10.1145/2998441

[2] Robert C. Kirby and Lawrence Mitchell. Solver composition across the PDE/linear algebra barrier. *SIAM Journal on Scientific Computing*, 40(1):C76–C98, 2018. https://doi.org/10.1137/17M1133208

[3] J. Schöberl. NETGEN: An advancing front 2D/3D-mesh generator based on abstract rules. *Computing and Visualization in Science*, 1(1):41–52, 1997. https://doi.org/10.1007/s007910050004

[4] Jack Betteridge, Patrick E. Farrell, Matthias Hochsteger, Christopher Lackner, Joachim Schöberl, Stefano Zampini, and Umberto Zerbinati. ngsPETSc: A coupling between NETGEN/NGSolve and PETSc. *Journal of Open Source Software*, 9(104):7359, 2024.

[5] A. Miraçi, J. Papež, and M. Vohralík. A multilevel algebraic error estimator and the corresponding iterative solver with p-robust behavior. *SIAM Journal on Numerical Analysis*, 58(5):2856–2884, 2020. https://doi.org/10.1137/19m1275929

3.11 Spectral Methods for Planetary Dynamics and Melting

Supervisors: Prof. Patrick Farrell and Prof. Richard Katz
Contact: farrellp@maths.ox.ac.uk and richard.katz@earth.ox.ac.uk

Background and problem statement: Melting of planetary materials (rock, ice) takes place at the boundaries of mineral grains. It leads to a network of interconnected pores that enable the melt to segregate from the residual solids. These grain-scale processes of melting and melt percolation are fundamental to the large-scale evolution of planetary interiors. Segregation of melt is coupled to the large-scale, creeping (viscous) flow of the polycrystalline solid phase — indeed, large-scale convective flow of the solid phase can drive and organise melting. Averaging over the grain scale enables derivation of a continuum theory for the dynamics of rock with a local volume-fraction of melt $\phi(\boldsymbol{x}, t)$, based on mass and momentum conservation for interpenetrating liquid (ℓ) and solid (s) phases at zero Reynolds number,

$$\begin{aligned} \partial_t \phi + \boldsymbol{\nabla} \cdot [\phi \boldsymbol{v}_\ell] &= \Gamma/\rho_\ell, \\ \partial_t (1-\phi) + \boldsymbol{\nabla} \cdot [(1-\phi)\boldsymbol{v}_s] &= -\Gamma/\rho_s, \\ \phi\left(\boldsymbol{v}_\ell - \boldsymbol{v}_s\right) &= -M_\phi\left(\boldsymbol{\nabla} P - \rho_\ell \boldsymbol{g}\right), \\ \boldsymbol{\nabla} \cdot \overline{\boldsymbol{\sigma}} + \overline{\rho}\boldsymbol{g} &= 0. \end{aligned}$$

Here, $\boldsymbol{v}_j$ is a phase velocity, ρ_j is a phase density, Γ is the melting rate, M_ϕ is the mobility of the pore liquid (a nonlinear function of ϕ), P is liquid pressure, and $\boldsymbol{g}$ is gravitational acceleration. An overline represents a phase-averaged quantity, e.g., the stress, $\overline{\boldsymbol{\sigma}} = \phi\boldsymbol{\sigma}_\ell + (1-\phi)\boldsymbol{\sigma}_s$. Constitutive laws provide closure for the stress tensors in terms of the pressure and phase velocities.

Known solutions to these nonlinear governing equations are characterised by emergence of small length- and time-scales. The problem is to efficiently obtain stable, robust numerical solutions under these challenging conditions.

Planned approach: This project will explore the potential to analyse the above system of equations using spectral decompositions. It will access spectral methods via the Dedalus software framework (https://dedalus-project.org/). This framework is actively developed to solve systems of elliptic, hypobolic and parabolic PDEs, and to model multiphysical flows. Despite its broad applicability, it has never been used for two-phase flows, where the solid and liquid velocities are defined throughout space. This project will use Dedalus to solve the system of equations above in configurations inspired by the internal dynamics of planetary bodies (including Earth). It may address issues associated with two-phase mass conservation, internal boundaries separating zero and non-zero ϕ, nonlinear constitutive laws, coupling of additional physics (granular, thermal, etc), localisation and the emergence of small length scales.

Desired outcomes: The governing PDEs admit analytical solutions describing solitary waves and the eigenfunctions of a linearised problem; benchmarking the numerical solutions against these is an initial target for research. Further work would aim to assess local and global conservation of mass, and to explore non-linear evolution of a mechanical instability to shear bands. An aspirational outcome is the modelling of Rayleigh-Bénard convection of the solid phase with melting, melt segregation, and freezing (direct relevance to Europa's convecting ice shell). These outcomes would lay the groundwork for use of Dedalus in applications to planetary bodies.

3.12 High-Order-in-Time Particle Method for the Landau Equation

Supervisor: Dr Andrea Medaglia
Contact: medaglia@maths.ox.ac.uk

Kinetic equations play an important role in describing different types of phenomena involving large numbers of interacting particles or agents evolving over time. These models have been adopted effectively in various fields of research, ranging from classical rarefied gas dynamics and plasmas, to semiconductors and new dynamics in socio-economic, biological, life and computational sciences. In particular, plasma models are studied at the kinetic level by the Landau-Fokker-Planck equation and have gained a lot of interest due to the important applications related to carbon-neutral fusion reactors. In this project, the candidate will explore higher-order-in-time particle methods for the Landau equation, with particular attention on the relation between the computational cost and the accuracy achieved.

Indeed, amongst the various numerical methods, particle-type methods are widely used. In [1] a gradient flow perspective of the collisional operator has been proposed, built upon an analogy with the heat equation and the 2-Wasserstein distance. This reformulation, together with a regularization of the entropy functional, leads to a fully coupled system of ODEs for the particles. This method has been effectively adopted in subsequent works in the inhomogeneous case, and for the space homogeneous multispecies Landau equation. In the literature, only first-order-in-time, both explicit and implicit, methods have been proposed. However, for a possible application to the fully inhomogeneous

multispecies equation, a careful study of the time integrator is required. In fact, in the presence of multispecies or the coupling with the electromagnetic field, one can only have numerical conservation of the main invariant quantities up to Δt, being Δt the time discretization. Furthermore, given the complexity of the full problem, the relationship between computational cost and accuracy (in time) must be carefully studied.

Based on the existing literature, the aim of this project is to investigate high-order-in-time methods of the deterministic particle scheme, starting from the space homogeneous Landau equation as a gradient flow. The first step is to replicate the results obtained in the seminal work [1]. Then, the candidate will implement carefully higher-order methods, like explicit Runge-Kutta or implicit Crank-Nicolson, and study the computational complexity of the new scheme and possible application of optimization techniques like random batch methods.

References

[1] J. A. Carrillo, J. Hu, L. Wang, and J. Wu. A particle method for the homogeneous Landau equation. *J. Comput. Phys.*, 7:100066, 2020.

4 Data Science

4.1 Deep Learning

Supervisor: Prof. Jared Tanner
Contact: tanner@maths.ox.ac.uk

Most modern deep learning architectures are now built on transformer blocks which are a combination of an attention mechanism and a fully connected layer. These structures are under intense study due to their use in the most effective text, image, and video generation tools. This project will allow students to select from distinct topics such as the below: sketching attention and value matrices for improved computational efficiency, investigating the presence and properties of attention sinks, or the spectral gap associated with softmax attention and the accompanying rank-collapse.

Central to these projects is the attention and transformer mechanism which a student should be aware of either from attending Theories of Deep Learning or analogous. The main secondary tool will be numerical linear algebra, such as eigenvalue and singular value analysis and techniques from randomised numerical linear algebra. Lastly, for some themes, the project may include pruning and low-rank plus sparse approximation of matrices with the most interesting in this case being structured sparse models. These project will include training small networks in order to investigate the effectiveness of the proposed method, but won't use large models such as used in commercial methods.

The student will become very familiar with the attention and transformer module and accompanying modules such as layer-norm. Beyond this what will be achieved depends greatly on the particular theme chosen. For instance, an investigation of the spectral gap of softmax attention layers could include random matrix theory of correlated matrices. Alternatively, for structured pruning one could consider measures and algorithms for projecting matrices to structured sparse models and how such models interact with attention sinks. Lastly, if randomised numerical linear algebra methods are considered then different having structures can be explored and linked to the attention sink structure observed in modern networks.

Students interested in a project within the scope of the above theme are welcome to contact Prof. Jared Tanner discuss suitability and structure of a specific focused topic.

References

[1] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention Is All You Need. In *Proceedings of the 31st Conference on Neural Information Processing Systems*, (NIPS 2017), 2017.

[2] Yihe Dong, Jean-Baptiste Cordonnier, and Andreas Loukas. Attention is not all you need: pure attention loses rank doubly exponentially with depth. In Proceedings of the 38th International Conference on Machine Learning, PMLR 139:2793–2803, 2021.

[3] Insu Han, Rajesh Jayaram, Amin Karbasi, Vahab Mirrokni, David P. Woodruff, and Amir Zandieh. HyperAttention: Long-context Attention in Near-Linear Time. *arXiv preprint:* https://arxiv.org/abs/2310.05869, 2023.

[4] Alireza Naderi, Thiziri Nait Saada, and Jared Tanner. Mind the Gap: a Spectral Analysis of Rank Collapse and Signal Propagation in Attention Layers. *arXiv preprint:* https://arxiv.org/abs/2410.07799, 2024.

[5] Guangxuan Xiao, Yuandong Tian, Beidi Chen, Song Han, and Mike Lewis. Efficient Streaming Language Models with Attention Sinks. *arXiv preprint:* https://arxiv.org/abs/2309.17453, 2023.

[6] Xiangming Gu, Tianyu Pang, Chao Du, Qian Liu, Fengzhuo Zhang, Cunxiao Du, Ye Wang, and Min Lin. When Attention Sink Emerges in Language Models: An Empirical View. *arXiv preprint:* https://arxiv.org/abs/2410.10781, 2024.

[7] Tri Dao, Beidi Chen, Kaizhao Liang, Jiaming Yang, Zhao Song, Atri Rudra, and Christopher Ré. Pixelated Butterfly: Simple and Efficient Sparse training for Neural Network Models. *arXiv preprint:* https://arxiv.org/abs/2112.00029, 2021.

4.2 How do Properties of Deep Neural Networks Promote (or not) Robustness?

Supervisors: Dr Vicky Kouni and Prof. Jared Tanner
Contact: kouni@maths.ox.ac.uk and tanner@maths.ox.ac.uk

Background and problem statement: Several classes of deep neural networks (DNNs) exhibit interesting properties. For instance, ReLU-type DNNs can have an inherent sparsifying effect [1], i.e., represent data with few nonzero entries, while residual-type DNNs can enhance data processing, by incorporating knowledge of prior data structure [2]. Nevertheless, it is quite unclear how these properties ripple out to the case where the DNNs' input data are contaminated with adversarial noise, i.e., noise that is carefully created to be indistinguishable [3].

Planned approach: This project will explore the robustness, i.e., resilience to adversarial noise, of assorted DNN architectures, in terms of the DNNs' properties. Particularly, a cornerstone in adversarial noise and related robustness pertains to selecting the exact method generating the noise, and test it for a given architecture. Robustness of models is expected to vary, but to what extent and what specific architectural properties affect the robustness is an exciting working avenue. The project's direction will be mainly on the experimental part, with many interesting applications, e.g., on audio and image data, while still requiring mathematical thinking and reasoning, with knowledge in one or more of the following fields: probabilities, optimization, data science, linear algebra.

Expected outcomes: Our aim for this project lies primarily in the development of a robustness bench- mark, which will coarsely pertain to *"the "x" architecture, because of its "y" properties, is more robust to the "z" adversarial noise".*

References

[1] Ilan Price, Nicholas Daultry Ball, Samuel C. H. Lam, Adam C. Jones, and Jared Tanner. Deep Neural Network Initialization with Sparsity Inducing Activations. *arXiv preprint:* https://arxiv.org/abs/2402.16184, 2024.

[2] Vasiliki Kouni, Georgios Paraskevopoulos, Holger Rauhut, and George C. Alexandropoulos. ADMM-DAD Net: A deep unfolding network for analysis compressed sensing. In *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 1506–1510, 2022.

[3] Peiyu Xiong, Michael Tegegn, Jaskeerat Singh Sarin, Shubhraneel Pal, and Julia Rubin. It Is All About Data: A Survey on the Effects of Data on Adversarial Robustness. In *ACM Computing Surveys*, 56(7):1–41, 2024.

4.3 Non-Markovian Models for Interacting Systems

Supervisor: Prof. Renaud Lambiotte
Contact: lambiotte@maths.ox.ac.uk

Background and problem statement: Networks offer a powerful framework for representing the structure and dynamics of complex systems across different scales. In these models, direct interactions between components are depicted as edges, while indirect interactions are captured through paths and walks. A common assumption is that system flows follow sequences of independent transitions — as a Markov process. However, real-world path data often exhibit higher-order dependencies, necessitating more advanced modelling approaches.

Description of the approach planned and the techniques needed: To address this issue, researchers have introduced network models with memory, often under the form of a higher-order Markov model. Integrating Markov models of varying orders within a multi-order framework significantly enhances next-element prediction in real-world systems. However, enforcing a single fixed Markov order across an entire system imposes a rigid constraint, as real-world dependencies often extend across multiple orders. The uneven distribution of observations in real data can lead such models to overfit certain regions while underfitting others simultaneously. A fixed-order modelling is also often impractical in practice, as the state space explodes with higher-order models.

What you'd hope to achieve: The purpose of this project will be to explore the possibility to constructing compact network representations from path data, balancing accuracy and model complexity.

References

[1] Rohit Sahasrabuddhe, Renaud Lambiotte, and Martin Rosvall. Concise network models of memory dynamics reveal explainable patterns in path data. *arXiv preprint:* https://arxiv.org/abs/2501.08302, 2025.

[2] Renaud Lambiotte, Martin Rosvall, and Ingo Scholtes. From networks to optimal higher-order models of complex systems. *Nature Physics*, 15(4):313–320, 2019.