# Table of contents

<table>
<thead>
<tr>
<th>Name</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>About the School</td>
<td>3</td>
</tr>
<tr>
<td>Professor Howard Bondell</td>
<td>4</td>
</tr>
<tr>
<td>Professor Kostya Borovkov</td>
<td>5</td>
</tr>
<tr>
<td>Dr Douglas Brumley</td>
<td>6</td>
</tr>
<tr>
<td>Associate Professor Alysson M Costa</td>
<td>7</td>
</tr>
<tr>
<td>Associate Professor Diarmuid Crowley</td>
<td>8</td>
</tr>
<tr>
<td>Professor Jan de Gier</td>
<td>9</td>
</tr>
<tr>
<td>Professor Aurore Delaigle</td>
<td>10</td>
</tr>
<tr>
<td>Associate Professor Sue Finch</td>
<td>11</td>
</tr>
<tr>
<td>Associate Professor Jennifer Flegg</td>
<td>12</td>
</tr>
<tr>
<td>Professor Peter Forrester</td>
<td>13</td>
</tr>
<tr>
<td>Associate Professor Nora Ganter</td>
<td>14</td>
</tr>
<tr>
<td>Dr Xi Geng</td>
<td>15</td>
</tr>
<tr>
<td>Dr Alexandru Ghitza</td>
<td>16</td>
</tr>
<tr>
<td>Dr Mingming Gong</td>
<td>17</td>
</tr>
<tr>
<td>Dr Hailong Guo</td>
<td>18</td>
</tr>
<tr>
<td>Dr Jack Hall</td>
<td>19</td>
</tr>
<tr>
<td>Dr Sophie Hautphenne</td>
<td>20</td>
</tr>
<tr>
<td>Professor Mark Holmes</td>
<td>21</td>
</tr>
<tr>
<td>Dr Wei Huang</td>
<td>22</td>
</tr>
<tr>
<td>Dr Mario Kieburg</td>
<td>23</td>
</tr>
<tr>
<td>Dr Johanna Knapp</td>
<td>24</td>
</tr>
<tr>
<td>Dr Pavel Krupskiy</td>
<td>25</td>
</tr>
<tr>
<td>Associate Professor Kim-Anh Lê Cao</td>
<td>26</td>
</tr>
<tr>
<td>Dr Dennis Leung</td>
<td>27</td>
</tr>
<tr>
<td>Professor Malwina Luczak</td>
<td>28</td>
</tr>
<tr>
<td>Dr James Mark Osborne</td>
<td>29</td>
</tr>
<tr>
<td>Professor James McCaw</td>
<td>30</td>
</tr>
<tr>
<td>Dr Daniel Murfet</td>
<td>31</td>
</tr>
<tr>
<td>Professor Aleks Owczarek</td>
<td>32</td>
</tr>
<tr>
<td>Dr Liuhua Peng</td>
<td>33</td>
</tr>
<tr>
<td>Associate Professor Guoqi Qian</td>
<td>34</td>
</tr>
<tr>
<td>Dr Thomas Quella</td>
<td>35</td>
</tr>
<tr>
<td>Associate Professor Charl Ras</td>
<td>36</td>
</tr>
<tr>
<td>Dr Lawrence Reeves</td>
<td>37</td>
</tr>
<tr>
<td>Associate Professor David Ridout</td>
<td>38</td>
</tr>
<tr>
<td>Associate Professor Nathan Ross</td>
<td>39</td>
</tr>
<tr>
<td>Dr Volker Schlue</td>
<td>40</td>
</tr>
<tr>
<td>Associate Professor Karim Seghouane</td>
<td>41</td>
</tr>
<tr>
<td>Professor Kate Smith-Miles</td>
<td>42</td>
</tr>
<tr>
<td>Dr Matthew Tam</td>
<td>43</td>
</tr>
<tr>
<td>Professor Peter Taylor</td>
<td>44</td>
</tr>
<tr>
<td>Professor Kari Vilonen</td>
<td>45</td>
</tr>
<tr>
<td>Dr Chenyan Wu</td>
<td>46</td>
</tr>
<tr>
<td>Dr Binzhou Xia</td>
<td>47</td>
</tr>
<tr>
<td>Dr Ting Xue</td>
<td>48</td>
</tr>
<tr>
<td>Dr Yaping Yang</td>
<td>49</td>
</tr>
<tr>
<td>Dr Lele (Joyce) Zhang</td>
<td>50</td>
</tr>
<tr>
<td>Dr Gufang Zhao</td>
<td>51</td>
</tr>
<tr>
<td>Professor Paul Zinn-Justin</td>
<td>52</td>
</tr>
</tbody>
</table>
About the School

The number of academic staff, including postdoctoral research fellows continues to grow. Several
members of staff are Fellows of the Australian Academy of Science and have been awarded
prestigious Fellowships and Awards.

The School offers a wide range of subjects to undergraduate and postgraduate students and
runs a Mathematics and Statistics Learning Centre that looks after the large number of
undergraduate students.

The School hosts two ARC Centres of Excellence, an ARC Industrial Transformation
Training Centre, and has several ARC Laureate, Future and DECRA Fellows.

In the pursuit of academic excellence, the School is a partner in the
collaborative research centres: Melbourne Integrative Genomics
and the Melbourne Centre for Data Science and is home to
the Statistical Consulting Centre that engages with
industry to provide high quality statistical
consulting.
John W. Tukey reportedly told a colleague, “the best thing about being a statistician is that you get to play in everyone’s backyard.” Needless to say, this precise sentiment is even more applicable in the world today than when it was first uttered some 50 years ago. It is truly an exciting time to be involved in the development of statistical and machine learning methodology in today’s realm of data science. I am passionate about using my mathematical and statistical expertise to develop methods for applications across diverse areas.

With the influx of data now available in both the public and private sectors, large and complex data sets have become the norm rather than the exception. Advances in data science are essential to extract meaningful information from this data for purposes of understanding and decision making.

**Statistical modelling, uncertainty quantification, and decision-making with complex data**

We develop innovative statistical methodology that is both interpretable and theoretically justified. By incorporating principled statistical approaches together with modern machine learning methods, our work enhances the discovery of knowledge by developing tools to analyse these complex datasets, while accounting for the uncertainty inherent in data.

**Bayesian inference and optimisation of physical and mechanistic systems**

Modern Bayesian inference methods provide tools to handle complex data from both a computational and modelling perspective. These approaches are also ideally suited to incorporate physical and mechanistic models to be reconciled with observable data for purposes of optimisation and decision-making. Our work develops methods to efficiently handle applications with large numbers of input variables and enables the discovery of the relevant features that separate the signal from the noise.

**Application-focused statistical and machine learning methods**

In many applications, the collection and format of the data does not fit nicely into an existing framework. Along with colleagues at the Melbourne Centre for Data Science, we tailor statistical and machine learning methods specifically for a targeted problem. Some examples of issues that arise include handling missing data that may not be missing at random, streaming and high frequency data, and complex sampling schemes such as repeated measures or longitudinal studies. Incorporating these complexities into flexible machine learning approaches allows for the removal of biases and produces appropriate uncertainty quantification. We have successfully worked to enable research advances for a variety of disciplines, including applications to network security, energy forecasting, climate change and public health.
For me, there are several motivations to do research. One is the overwhelming feeling one has when making a scientific discovery — albeit a small one. Another is that I mostly work on mathematical problems that have direct relevance to real-life systems, and solving them can bring about positive change. Yet another is that doing research makes me a better lecturer and supervisor for our students, helping me to inspire them as well.

**Boundary crossing by random processes**

Evaluating boundary crossing probabilities for random processes is a fundamental problem in many areas of probability theory and its applications, with some of the most eminent problems coming from collective risk theory, statistics, change-point detection and mathematical finance. Among the central topics in some other important areas in probability theory and its applications (such as extreme value theory, dam/storage problems and queueing and telecommunication theory), we often also find special cases of boundary crossing problems. Despite the large amount of work done in this area, there are many interesting open problems of which solutions can be of substantial importance for potential applications.

**Dynamic point processes with preferential attachment mechanisms**

A dynamic version of the Neyman contagious point process can be used for modelling the spatial dynamics of biological populations, including species invasion scenarios. The dynamics of such models have not been fully explored yet — there are open problems on more detailed descriptions of the evolution of the points population, both in distributional terms and in terms of its almost sure behaviour.
The Brumley lab utilises mathematics, microfluidics and microscopy to study a range of dynamic processes in biology, including bacterial motility, symbioses, and nutrient cycling and flows around coral reefs.

Microscale fluid flows in biology

The ways in which microorganisms interact with one another and with their surroundings underpins their collective functioning in the environment. Biogeochemical cycling in the ocean, infection of humans by pathogens, and microbial symbioses in coral reefs are all driven by the dynamics of microorganisms.

The Brumley lab has developed a range of new microfluidic tools, high-speed imaging techniques and mathematical frameworks for investigating the motion of microorganisms at the single-cell level, and is using these to predict ecosystem-level outcomes.
My mission is to identify and model decision-making problems in different areas of applications, with a focus on issues that will contribute to the well-being of the population and allow resources available to be used in an efficient and equitable way. I am to extend the limits of available solution methods to allow for the solution of even larger and more realistic models.

Decision-making problems are found in the most diverse areas of application: how can you design trams and trains schedules to better serve the public? When to release water from reservoirs in order to improve the health of endangered species in a river system? When to schedule elective surgeries to maximise the service provided? How to better route delivery trucks for faster and more efficient delivery of goods? Where to install vaccination centres to better deliver service to the population? How to schedule high school lectures to better fit the needs of students and teachers? Which crops should be planted to maximise yields while minimising impact on natural resources?

In Operations research, we tackle these (and many other) questions with an approach that relies on converting the problem to a mathematical model and then solving this model with specialised solution algorithms that can obtain the best solution for realistic models with tens of thousands of variables. Modelling and solving go hand to hand, and the research questions are often related to the obtention of the best models for the algorithms available and, conversely, on the extension of the available algorithms to better solve the models developed.
To understand the structure of high-dimensional manifolds through their classification and the mappings between them to deepen our understanding of the nature of space.

Topology of manifolds

The idea of a space, a collection of locations related in a global structure, is a powerful metaphor. Abstractly conceived, spaces need not be three dimensional but can be multi-dimensional. Topology is the field of mathematics which studies space, and manifolds are a fundamental class of space which arise as solutions to complicated equations in many variables.

I study the structure of multi-dimensional manifolds and the relationships between them; in other words, the possible structures of high-dimensional spaces.

Surfaces are classified by genus: the 2-sphere (genus 0), the torus (genus 1).

A blackboard from a talk I gave in Edinburgh in 2013.

jdgie@unimelb.edu.au
findanexpert.unimelb.edu.au/profile/9355-jan-de-gier
I aim to unravel and understand the underlying mathematical structures and patterns that give rise to emerging macroscopic behaviour in physical and biological systems.

**Solvable lattice models**

Solvable lattice models provide useful mathematical frameworks for modeling real world phenomena. Examples of solvable lattice models are quantum spin chains as systems to implement quantum computation, models for metals, cold atoms and superconductivity. Another example is exclusion processes as general transport models in biology, physics and engineering, and more general stochastic processes to model phenomena that involve some element of randomness, such as interface or domain growth. A third calls of models are random tilings as models for quasicrystals and crystal melting.

The study of solvable lattice models uses a variety of techniques in pure mathematics and mathematical physics, ranging from algebraic concepts such as the Yang-Baxter equation, Hecke algebras and quantum groups, to analytic methods such as complex analysis and elliptic curves. Due to this wide variety of methods, the study of solvable lattice models often produces unexpected links between different areas of research.

**Multivariable polynomials**

I study connections between enumerative combinatorics and statistical mechanics on the one hand, and (symmetric) polynomials and representation theory on the other. Schur and Macdonald polynomials are important classes of polynomials and we have developed a new framework to study and generalise such polynomials using solvable models and matrix product methods. Multi-variable polynomials may be used in problems involving many degrees of freedom, or data points. They are computationally expensive and complex to deal with in a brute force manner, but smart underlying mathematical structures and patterns can unravel these complexities.
My goal in research is to focus on new frontier research problems of statistics that genuinely require a solution. I am dedicated to developing flexible methods that do not rely on stringent or unrealistic conditions, and ensuring their validity by developing rigorous theoretical arguments that explore their limit of application.

Nonparametric techniques for data with measurement errors

A lot of my research has been devoted to developing statistical techniques for analysing data observed with measurement errors. Often, those data make standard statistical methods fail, and the challenge is to incorporate some correction. Because of important applications to a variety of problems, measurement errors have become a very active and popular topic of research in statistics.

A large part of my own research in the area originates from the need to analyse data from nutrition. Dietary data are used to estimate the distribution of usual (that is, average long-term) intake of various nutrients and food groups, monitor such intakes over time and relate individual usual intakes to health outcome measures. My major achievements in the area include developing highly effective and practical nonparametric statistical techniques, which often involve converting abstract, theoretical concepts into completely realistic and fully applicable methodology.

Nonparametric techniques for group testing data

Another of my research interests concerns the statistical analysis of group testing data. There, the primary goal is to analyse data from infectious disease studies, particularly the presence/absence of an infectious disease as detected from a blood or a urine test.

In large infectious disease studies, it is sometimes not possible to test individuals separately — only the result of tests on groups is available (a technique that has been used in the COVID-19 pandemic). This is referred to as group testing data. My research in the area so far has consisted of developing nonparametric techniques for estimating the influence of various biological, environmental or other factors on the incidence of the disease. The task is challenging because when the data are grouped as in these problems, we lose a lot of information about individual disease status. Moreover, the blood tests are usually imperfect, adding an extra layer of uncertainty to the data.

Functional data analysis

Another of my research interests concerns the analysis of functional data, ie data that are in the form of curves. For example, growth curves of children and yearly rainfall or temperature curves are functional data. One of my research directions has been to develop dimension reduction procedures that can keep as much of the information from the data as possible when the goal is to classify and cluster the individuals into groups that share similarities. Another challenge of functional data is that they are often observed sparsely: instead of observing the entire curves of interest, we may be able to observe the curves only at few points. I have developed several new procedures for analysing functional data in the case where only fragments of curves have been observed.
In collaborating with other researchers and providing them with statistical support, I aim to enhance the quality of the design, methodology and implementation of data-based research, provide robust data analytics and support effective, transparent communication of the findings of empirical research. My research in statistical literacy, data visualisation and statistical consulting and communication supports these goals.

Enhancing the application of data analytics in real-world practice

Reports of research findings and claims based on data pervade our everyday lives through a wide range of media channels. The societal value of such research and claims depends in part on the quality of the communication from the producers of the research, and on the capacity of potential consumers of the research to understand the findings.

My research aims to enhance the quality of data-based research outputs and communication, with a focus on data visualisation. It also aims to improve statistical literacy and statistical pedagogy in order to promote statistical literacy and critical thinking about quantitative information more broadly.
I am an applied mathematician, which means that I use mathematical tools to help answer real-world problems. In particular, I apply mathematics and statistics to problems in biology. As an example, I develop predictive statistical models in space and time for the level of antimalarial drug resistance. These predictive maps fill in the gaps where no information is available on drug resistance and have been used by health agencies to develop new policies about where and when certain drugs are appropriate to use.

Mathematical biology and medicine

My research focuses on using mathematics and statistics to answer questions in biology and medicine. In particular, I develop mathematical models in areas such as wound healing, tumour growth and infectious disease epidemiology.
My research in random matrix theory enjoys international recognition, with my research monograph ‘Log-gases and random matrices’ being cited over 1,400 times. This gives me the opportunity to play a leadership role in the advancement of studies in random matrix theory in the Asia Oceania region, where it is an emerging research topic.

My primary research field is random matrix theory. This is a broad research topic with links to quantum physics, number theory, disordered and complex systems, communications engineering, and neural networks, among other applications. My main tools are methods of special functions in one and several variables, which facilitate analytic computations.

Dip-ramp-plateau is a signature of many-body quantum chaos. This duality formula facilitates analytic computations in random matrix theory.
I have a passion for involving students in research-oriented activities from day one of their education at university. Much of this is a community-building exercise, and we are very excited to have just started offering a student-organised seminar in which the students get to vote on the topics and do the presentations (currently on neural networks, previously on open key encryption). I view research and teaching as a unit and endeavour to create a research training environment that offers our students, from first year undergraduates to PhDs, the same level of opportunity that I enjoyed in my education overseas. The vision is to create a pan-Australian PhD program in mathematics with a strong cohort component, an emphasis on communication skills, and to create more high-level offerings for interested undergraduate students.

Algebraic topology studies the geometry of elastic objects. This young discipline has seen a remarkable journey over the past two decades with profound applications to machine learning, sensor fields, computational neuroscience and more. The breakthrough here is that data science is moving away from thinking of data as discrete points and moving towards understanding data as continuous (or algebraic) structures, giving a much more holistic view and demanding a completely new toolset — so new in fact that the applications have not all found their ways into the textbooks yet.

My own research is on the foundations of mathematics, but through student training and conversations with colleagues I interact indirectly with these applications rather frequently.

For instance, a former MSc student of mine is now designing efficient fiber optic networks, while another is working as an analyst in a renewable energy startup. Others are pursuing PhDs or postdocs overseas.

The Fano plane is at the heart of many famous phenomena, from Bott periodicity in pure mathematics to the triality symmetry in physics and the existence of error-correcting codes, used to send images from space.

Mathematical ideas are created: the power of pen and chalk.
My research and academic path are driven by the intrinsic elegance of mathematics. Since I was a PhD student, I took the ‘pursuit of deeper truth’ as the fundamental goal of my life. The essential reason for my research and ultimate objective is to have a fundamental understanding about the deeper intrinsic mechanism that drives the various phenomena that we observe mathematically and physically. I choose specific areas because they fit my taste for mathematics and involve a broad range of deep mathematics, which is something I particularly appreciate.

I am a mathematician working in the interplay between analysis and probability. My core research areas are stochastic analysis and rough path theory. Broadly speaking, stochastic analysis is the differential calculus of stochastic processes and the analysis of stochastic differential equations (SDEs). One of the core problems in this area is related to the study of fine quantitative properties of SDE solutions. Rough path theory is an analytic theory of integration and differential equations with respect to highly irregular paths, providing robust analytic tools for studying problems in stochastic analysis. My research requires and emphasises the use of a wide range of mathematical tools from analysis, geometry and algebra apart from core probabilistic methods.
As number theory becomes increasingly technical and requires highly specialised knowledge from many branches of pure mathematics, we are witnessing a ghettoisation phenomenon where experts in different aspects of the same object of study are unable to communicate about their methods and results. My aim is to help reconnect these areas in order to facilitate the exchange of expertise and lower the barrier to entry for newcomers.

Modular forms and friends

The earliest known instance of a number-theoretic result is on a clay tablet from around 1800 BCE. A lot has happened in the intervening 4000 years, much of it motivated by the desire to resolve difficult problems such as Fermat’s Last Theorem, or understand the subtleties of the distribution of prime numbers. In the process, we gradually moved from concrete questions about whole numbers to increasingly abstract constructions that model the surprisingly subtle behaviours of integers and their generalisations. One such construction is that of modular forms, a type of function that has a lot of symmetries and encodes deep arithmetic information. I study the structure of the spaces of modular forms from a variety of points of view. A wide web of conjectures relates them to geometric objects such as elliptic curves (well-known for their cryptographic applications) and algebraic objects such as Galois representations, and this study takes me and my students across large swaths of the contemporary pure mathematics landscape. Computational tools for number theory: numerical experimentation has always played a part in the discovery process in number theory; Gauss used to count prime numbers in intervals in his spare time, eventually obtaining the list of all primes up to three million. Nowadays, we have computers that can do the tedious work for us, but the subtleties of arithmetic behaviour and the abstract nature of the tools we use to study it lead to the need for efficient algorithms for experimentation in number theory. I develop, implement, and use such algorithms for computing with modular forms, elliptic curves, and their generalisations. I believe in providing equal and unrestricted access to these tools to anyone who would use them and make them available as free and open-source software (for instance within SageMath or related systems).
My research aims to promote causal revolution of data-centric intelligence and science — constructing machines that can communicate in the language of cause and effect and answer ‘why’ questions by inferring from data.

My research interests lie at the intersection of machine learning and causal reasoning. I develop causal structure discovery algorithms from various kinds of observational data. Meanwhile, I explore causal reasoning methods and insights to tackle challenging problems in machine learning, such as transferability, generalisability and robustness. I am also interested in machine learning and reasoning problems arising from areas such as computer vision, bioinformatics and financial analysis.
My primary goal is to develop and analyse efficient and high-accuracy numerical methods for solving differential equations from different areas like physics, chemistry, and material science.

**Numerical solution of differential equations**
Differential equations are the mathematical language to describe the real world. We shall use mathematical knowledge and computer tools to simulate and understand the real world.

**Modelling and simulation of topological materials**
Topological materials are a class of quantum materials whose properties are preserved under topological transformations. We shall analyze the topological properties of topological materials from a mathematical viewpoint. Based on mathematical analysis, our research will develop state-of-the-art numerical methods for simulating the topological materials. We also propose to use the shape optimization approach to design topological materials.

**Deep learning for partial differential equations**
One of the main challenges of classical numerical methods like finite difference methods and finite element methods is the curse of dimensionality. We shall use machine learning and artificial intelligence to develop cutting-edge computational methods for solving high-dimensional problems from real applications.
An overarching goal of mine is to gain a better understanding of the algebra underlying the sophisticated geometries that arise in the classification problems that are pervasive in mathematics and its applications to physics.

Algebraic geometry has a long history. Its origins lie in the geometry of the set of solutions to systems of polynomial equations in several variables. Although simple-sounding, algebraic geometries are sufficiently rich to describe problems that arise in diverse areas, such as theoretical physics and cosmology, biology, statistics, quantum computing, cryptography, machine learning, robotics, and control theory. Being algebraic, the geometry can be computed by hand or using software packages.

Classification is fundamental to science. Well-known examples of classifications are: taxonomy in biology, the periodic table in chemistry and the elementary particles in physics. In mathematics we not only classify objects, but also study the very notion of classification itself. From this abstract point of view, the easiest type of classification is a simple list. We can also classify objects that can vary in a continuous fashion — like equations, whose coefficients we can vary. The study of this type of classification is the subject of moduli theory. Moduli theory encourages us to think about a single object of interest by analysing the geometry of a moduli space parameterising all similar objects. For example, how does the nature of an equation change if we adjust one or more of its coefficients just a bit?
My goal is to produce research in applied probability that can be used by practitioners (mainly biologists). Therefore, my research in applied probability involves the development of algorithmic and computational techniques, as well as statistical methods to fit stochastic models to various types of data. I am working in close collaboration with conservation biologists on the management of endangered species (including the Chatham Island black robins and Lord Howe currawongs) and hope to produce elements of environmental benefit to society.

My work is motivated by real-world problems arising in population biology, ecology, and epidemiology. My research focuses primarily on branching processes, which are a flexible class of probabilistic models well suited to the study of populations that evolve randomly over time.

**Analysing the probability of extinction of complex populations**

The survival or extinction of a population is subject to random events. The establishment of a population depends on several features, including the reproductive fitness of individuals and environmental factors. I am characterising the probability of extinction of various models which incorporate a number of these features. I am motivated by the practical use of these models, so my research also involves developing efficient computational methods.

**Stochastic models applied to conservation biology**

I am particularly interested in developing Markov models for endangered populations. One example is the Chatham island black robin, which was saved from the brink of extinction in the early 1980s. I am developing statistical methods to estimate model parameters, and I am deriving techniques that enable us to use the fitted models to analyse population viability and to design conservation strategies.
I initially chose to work in the area of Probability Theory because I found it fascinating, even though I didn’t understand it. In my research, I continue to look for problems that are easy to state but are counterintuitive or hard to understand/solve. Probability is a fertile breeding ground for such problems.

I study various types of random processes, where one typically describes some local rules and is interested in describing the global behaviour of a system. For example, in a grid-like city at each street intersection either (with probability $p$) put signs pointing north and east or (with probability $1-p$) put signs pointing south and west. This is an example of a random medium defined via local rules. We may now ask various global questions, for example, starting from the middle of the city, which points can you get to by following signs? More generally, I study various kinds of random processes such as reinforcement processes, random walks, random media (the above is an example of a random medium), interacting particles and queues.

This picture illustrates the probability of losing (being first eliminated) in a simple 3-player poker game when your opponents start with $x$ and $y$ respectively, and you start with $\$1000-x-y$. 
In modern society, we have so much information to deal with and so many decisions to make every day. I hope my research can help to develop more advanced data analysis techniques. By incorporating them into machine learning and AI technologies, we can one day have powerful enough AI to assist everyone in daily information gathering and decision making.

Causal inference

Functional data are the data in the form of curves. For example, the data in the Berkley Growth Study, which consists of the heights of 93 individuals measured from their birth to eighteen years old, are functional data. The functional data analysis involves a full scale of statistical problems such as hypothesis testing, regression and classification. One challenge of the functional data analysis is that the data are usually observed partially — namely, each individual is only observable on a subset of the domain of interest. We have worked on the case where the random observable sub-domain is assumed to be independent of the data value. There, we established identifiability conditions under which the covariance function of the data can be recovered consistently and proposed a consistent estimator for it. Using this technique, we can recover the unobserved parts of the functional data.

Partially observed functional data

Most studies in the health and social sciences are motivated by causal rather than associative questions. For example, what fraction of past crimes could have been avoided by a given policy? What is the efficacy of a given drug in a given population? These questions could be answered straightforwardly if the potential outcome of any value of the cause is available while all the confounders are under control. However, in practice, such potential outcomes are usually unobservable. We aim to tackle this challenge using advanced techniques such as moment equation balancing and develop efficient methodologies to conduct causal inferences from different data types. The problem becomes even more complicated when we are given multivariate (dynamic) data, and the goal is to learn the causal structure among the variables from the data. This is called the causal discovery problem in Machine Learning. We aim to establish a framework to solve the problem using differential equations.
I am studying problems in Random Matrix Theory as it is exciting to see many different fields and interdisciplinary groups working on similar problems, only considering them from a different point of view. I like to work with mathematical rigour but also need to see the applications, which are a perfect motivational motor. The rich mathematical structures and the combination of tools which do not look as if they would be related are a creative challenge, which is an incentive in itself. My credo is: 'throw everything in one pot, stir it a bit and see what the outcome is'.

Random matrices and their applications

Random Matrix Theory (RMT) is the theory where matrices are sampled randomly. It is generally a statistical and analytical mathematical tool which applications cover physical systems such as condensed matter theory, quantum chaos, quantum field theories and quantum information, as well as topics such as in engineering (telecommunications, information theory and machine learning) and statistics (time series analysis and stochastic processes). RMT is also very rich in several other mathematical tools comprising not only probability theory and linear algebra as the name suggests but is also intimately connected to differential geometry, graded algebras such as supersymmetry, harmonic analysis due to the relation to matrix groups, number theory and many more. The focus of my research lies in the mathematical and theoretical development of the tools and the understanding of RMT. Moreover, I also have several research projects on its applications such as quantum information, strongly interacting quantum field theory, time series analysis and telecommunications.

The Mexican Hat Potential is a standard example of spontaneous symmetry breaking. The latter is the reason why random matrices can be taken as an approximation for Quantum Field Theories in the low-energy limit.
I am interested in understanding the fundamental laws of nature. Through my research, I hope to contribute to a better understanding of how the fundamental forces can be unified within one theoretical framework: string theory.

The physical mathematics of extra dimensions

My main area of research is string theory. String theory predicts the existence of extra dimensions, and the mathematical properties of the extra dimensions determine the physics that comes out of string theory. A prominent class of examples is called Calabi-Yau spaces. My research is concerned with the construction and analysis of Calabi-Yaus in the context of string theory, where I am applying standard techniques from physics to uncover mathematical properties of these spaces.

This approach is sometimes called 'physical mathematics' and has led to exciting new results both in physics and mathematics. Among those are highly non-trivial connections between different Calabi-Yaus, known as string dualities.
I am working on the development of very flexible models for high-dimensional data with complex dependence structures that standard models cannot capture. The proposed models can improve the fit in the tails of a multivariate distribution, and these models therefore can improve the forecasts of extreme events, such as severe floods or extreme pollution levels.

**Multivariate dependence modelling**

My research interests include the development of flexible models for multivariate data with complex dependence structures and fast inference methods for these models. My research is focused on modelling multivariate extremes (data that arise when very unusual or extreme events are observed, such as large negative returns on a stock market, severe flood or drought events), spatial data (data that are measured at different locations, such as pollution levels or weather data) and copula modelling. Copulas are functions used to combine univariate marginals into multivariate distributions. These functions can be used to construct flexible multivariate distributions, then used to model high-dimensional data with complex dependence structures. Potential applications include modelling financial returns, forecasting extreme weather events, and analysis of image data (such as fMRI data).

- Time-varying dependence in stock returns from the consumer staples sector.
- Interpolation maps for the average wind speed in Netherlands.
Our research focuses on the development of computational methods, their applications in areas informed by biology, and the training of the new generation of computational biologists and data analysts. Our area of expertise is in the integration of biological ‘omics data (transcriptomics, proteomics, metabolomics etc, as well as microbiome, metagenomics, single cell transcriptomics and multi-omics) with multivariate and dimension reduction methodologies, selection of features of biomarkers in large biological data sets, and R software development. Our group provides critical collaborative expertise to biologists, bioinformaticians, statisticians and clinicians and welcomes budding data analysts.

Data integration methods using multivariate dimension reduction methods

We develop dimension reduction techniques to integrate the information from cutting-edge sequencing technologies that give rise to ‘omics data (transcriptomics: the study of all transcripts; proteomics: the study of all proteins; metabolomics: the study of all metabolites; metagenomics: the study of all bacteria). The challenge we are facing is the large number of features or omics molecules (hundreds of thousands) compared to the number of samples (less than 50 individuals) and the difference in scale of the data we analyse. We are currently expanding our methods for bulk experiments to single cell assays.

Statistical methods for microbiome studies

There are major statistical and computational challenges in analysing microbial communities that hinder the potential of microbiome research. We are interested in characterising and understanding microbiome-host interactions. Some of our methods’ developments aim at addressing batch effects in microbiome experiments and analyse scarce temporal sampling in time-course studies. We analyse microbiome datasets from our collaborators for a wide range of studies, including investigating the role of gut and oral microbiome in spondyloarthropathy diseases, the development of intestinal or salivary microbiota in toddlers and infants, and investigating the gut-brain crosstalk in Huntington’s disease.

Development of the mixOmics R toolkit package (www.mixOmics.org)

mixOmics is one of the few R packages dedicated to the integration of multiple ‘omics data (19 novel methodologies implemented so far, amongst which 13 were developed by our lab) and with an increasing uptake from the research community. The package is in the five per cent top downloads in Bioconductor. Programming developments are ongoing for efficient programming for large-scale studies and the development of novel multivariate methods. We routinely teach multi-day workshops.

Examples of methods, each designed to answer a specific question, and graphical outputs to understand the results.
Hypothesis testing and the dual task of computing confidence intervals arguably constitute the most rigorous paradigm in statistics for uncertainty assessment, at least from the frequentist point of view. Within this paradigm, I conduct both theoretical and methodological research to develop new analytical techniques that are suitable for modern scientific applications where data can be ‘big’ in different senses, eg tens of thousands of measurements for different genetic markers or millions of voxel measurements for fMRI data.

Multiple testing for false discovery control

To make data-driven scientific discoveries with good reproducibility, researchers often have to routinely test a multitude of hypotheses simultaneously based on experimental data. For example, one may want to identify the potential few SNPs associated with a disease from many of them in a genetic dataset. My current research seeks to expand and refine the current methodologies on multiple comparisons in practical applications. This development is naturally propelled by the surge in demand to tease out, in a principled manner, the ‘gems’ from the ‘haystacks’ of data, such as biological data in omics scale, imaging data in millions of voxels, and many more, found in fields where large-scale datasets are generated.

High-dimensional inference

Sometimes, even a single hypothesis can be ‘big’; this is true when the underlying dimension of the associated test is ‘high’. For example, how would we calibrate the critical value if we were to test whether p random variables are mutually independent when the number of samples (n) we have is small compared to p? Classical asymptotic results assuming p (the underlying dimension) is fixed will give a bad approximation. Therefore, one needs to develop a new suite of high-dimensional asymptotic results to tackle problems of this kind.
I aim to study near-critical phenomena in epidemic models, both to deepen our understanding of specific simple models and to gain insights into the behaviour of a wide range of more complex models of epidemics.

Emergence and elimination of a disease involves the process of infectious transitions and recoveries being pushed across a critical threshold. For example, a pathogen mutation can increase the transmission rate and make a previously ‘subcritical’ disease (ie not infectious enough to cause a large outbreak) into a ‘supercritical’ one where a large outbreak may occur. Equally, an infection control measure such as vaccination may decrease the transmission to below the threshold, and the disease may even be completely eliminated from the population. The passage across the critical threshold constitutes a phase transition, and there is an associated critical window of parameter values, which is certainly of theoretical and, at least in principle, of practical interest. However, phase transitions can often be intricate, with delicate parameter tuning required to describe them; there has yet been very little rigorous work in the literature regarding near-critical phenomena in epidemics.

I work in probability theory and applications, with a particular emphasis on Markov processes, random graphs and applications to the spread of epidemics on networks.

Markov processes can be used as models for large complex random systems arising in various applications, including biology (eg epidemics), computer science (eg allocation models and random graphs), and statistical mechanics (eg the Ising model). For a Markov process it is sometimes possible to show that over a finite time it remains close to a deterministic process solving a differential equation; in my work I have developed new methods to show that in suitable cases the approximation holds long-term and in equilibrium. An important topic underlying large swathes of probability theory and applications is the study of the speed of convergence to equilibrium of Markov chains. My contributions to this area involve developing new techniques for analysing the speed of convergence and applying them to prove results for specific chains.

My work in random graphs has focused on phase transitions, ie instances where a small change in one model parameter leads to a dramatic change in qualitative behaviour. I have determined the precise nature and location of the phase transition for a number of important models and developed new methods that can be used for other problems in random graphs.

Theoretical analysis of stochastic models of population growth and the spread of epidemics in populations has been the focus of much of my recent work. Traditionally, biological processes have mainly been studied via differential equations and their exact or approximate solutions, but in recent years there has been increased interest in probabilistic models, which are often able to capture more accurately the true behaviour of such processes. Surprisingly, the long-term behaviour of even some very simple-looking stochastic epidemic models is still far from completely understood. Thus there has been a growing need for new mathematical approaches specifically tailored to such models, many of which have features similar to those in random graphs and statistical mechanics.
I am trying to develop mathematical and computational methods and tools to help unravel the complex nonlinear interactions controlling organ development, homeostasis and disease.

**Multicellular systems biology**

I combine mathematical modelling, numerical methods and computational simulation to understand multicellular processes — in particular, how cellular and subcellular interactions manifest at the tissue and organ level and how perturbations to these systems (mutations and drug interventions) influence the organs’ form and function.

Example applications. **Left:** multicellular simulation of colorectal crypt. **Centre:** flow in a deforming vessel bifurcation demonstrating bloodflow tissue interaction. **Right:** multicellular simulation of C elegans germline.
My research goal is to advance our understanding of infectious diseases by applying mathematical and statistical theory. Through the application of rigorous analytical methods to data, my research will enable society to more effectively respond to emerging and re-emerging diseases, advancing public health.

How do infectious pathogens like influenza, SARS-CoV-2 and malaria spread through the community? How do we most effectively respond to dangerous new pathogens? Can we protect ourselves and society by learning more about how the body responds to infection? All of these questions can be addressed through the development and application of mathematical models of infection and transmission. My research team uses mathematics, computational modelling and statistics to learn more about infectious diseases, how they spread and how society can most effectively respond. We develop and analyse models and apply them to data across multiple scales. We study how pathogens replicate within the body and how the immune system and drugs inhibit that replication. We also study how pathogens spread from human to human and how public health responses — such as hygiene recommendations, physical distancing and vaccination — modify transmission.

The team develop new mathematical models for these biological processes and study their dynamics. Our research is deeply cross-disciplinary, engaging colleagues from clinical and biomedical fields through to psychology, epidemiology and public health. We contribute to fundamental knowledge in the life and mathematical sciences. Furthermore, we make direct contributions to public health policy and response measures for major infectious disease threats such as COVID-19 and malaria.


Model fit and validation for human volunteers infected with malaria. Our mathematical models of malaria infection capture the aspects of infection dynamics, enabling their use to predict how infectious humans are, enabling new approaches to malaria control (Cao P et al., eLife 2019).
To understand singularities and their many manifestations in the world.

Algebraic geometry
I work in singularity theory on foundational aspects of the theory of hypersurface singularities and their associated categories. My work touches on aspects of mathematical physics, and I have often collaborated with researchers in topological field theory and string theory on problems in algebraic geometry arising from those subjects.

Mathematical logic
I have worked in the area of proof theory, on linear logic and its semantics, on connections between logic and algebra and between logic and the theory of computation.

Theory of deep learning
I am working with colleagues in the Melbourne Deep Learning Group on a new approach which aims to lay the foundations for a mathematical theory of deep learning based on algebraic geometry and statistical physics.
Lattice Polymers, also known as lattice random walks, are a rich set of models describing the universal geometric and topological behaviour of long chain polymers which are ubiquitous in the world around us, from paints and plastics to the DNA in our cells. These models provide key insights into the behaviour of such polymers, but they also have deep connections to a variety of other models in mathematical physics, combinatorics and probability theory. Their study allows us to uncover both new physical phenomena and new mathematical structures.

My area of expertise is mathematical statistical mechanics and, in particular, the area of phase transitions and critical phenomena of model polymer systems, namely lattice walk models, which lies within the discipline of mathematical physics. My work endeavours to uncover the universal geometric and topological features of long chain molecules such as DNA in a variety of generic conditions. The models I study arise naturally in discrete mathematics and combinatorics and in stochastic processes as well as in theoretical physics. The two main topics of interest are:

1. The exact solution of interacting directed walk systems, and
2. Numerical analysis — both Monte Carlo computer simulation and exact enumeration techniques of lattice walks.

Two directed walks in a strip.

Self-avoiding walk as a model polymer.
My research lies in many areas in statistics, especially in Statistical theory. In particular, I am interested in extreme value theory with applications in economics and finance, distributed and online updating statistical inference for big data, high-dimensional data analysis, bootstrap, nonparametric inference, variable selection and model selection, random forest, and hypothesis testing in time series.

**Extreme value theory with applications in economics and finance**

In probability and statistics, the behaviour of a distribution function in the tail region has been extensively investigated with the advancement of extreme value theory. Extreme value theory basically restricts the tail region of the distribution function to resemble a limited class of functions that can be fitted to the tail. The extreme value index plays a key role and serves as a basis for statistical applications of extreme value theory, thus statistical inference for the extreme value index and related topics are of great interest and have been broadly studied in the literature of statistics. In practice, extreme value theory combined with statistical inference offers a wide variety of applications in many scientific areas such as hydrology, environmental research and meteorology, geology, insurance, finance, and economics. In my research, we utilize extreme value theory to study extreme events and tail behaviour of datasets from many different perspectives. Our theoretical results are applied to many practical problems in economics and finance including risk management and investment evaluation.

**Distributed and online updating statistical inference for big data**

Massive data with rapidly increasing size are encountered in many scientific fields with a need for new statistical analysis. When the size of the data is huge, the implementations of statistical inference can be computationally expensive and slow. The issue is even worse when the data are stored in different locations, and we need to consider the computational cost of data communication and transformation. In my research, we aim to study the distributed inference for big data by proposing novel algorithms and investigating their theoretical properties.

**Bootstrap and related areas**

Bootstrap is widely used for accessing uncertainty of estimates and constructing confidence intervals since it is easy to implement and is second-order accurate under regularity conditions. Different versions of bootstrap have been implemented and studied in many areas of Statistics. In my research, we study or propose different bootstrap algorithms applied to different types of datasets and different setups, and their theoretical consistency is built.
Much of my research involves the development of foundational graph-theoretical tools, models and algorithms for solving optimisation problems from telecommunications, VLSI microchip design, wireless sensor networks and underground mining networks.

We undertake wide-ranging state-of-the-art research in statistics and data science, including spatiotemporal statistics, statistical machine learning, stochastic modelling, computational statistics, and statistical methods for climatology, epigenomics, and geo-hazard research. Examples are:

- Spatiotemporal modelling and extreme-event forecasting on big nonstationary geo-hazards data
- Spatiotemporal statistical methods for climate change research, including for seasonal tropical cyclone modelling and prediction, and for continental temperature and precipitation profiling
- Regression, clustering and classification for big and complex data/text mining and machine learning, including regression clustering,
- Support vector machines and ensemble of decision trees
- Statistical genomic and epigenomic association studies using cloud and clustered supercomputing
- Efficient computing algorithms for model selection in high-dimensional model space
- Decision and information theoretic methods for model selection
- Markov chain Monte Carlo including Gibbs sampler, EM algorithm and Bayesian computing
- Time series analysis, including change-points and dependence structure identification
- Generalised linear models and composite likelihood for categorical, longitudinal and panel data, and
- Statistical modelling and inference for capture-recapture data in ecological studies.

I develop innovative quantitative methods and real-world applications to address statistical challenges encountered in a broad range of science, engineering, industry and business fields.

We undertake wide-ranging state-of-the-art research in statistics and data science, including spatiotemporal statistics, statistical machine learning, stochastic modelling, computational statistics, and statistical methods for climatology, epigenomics, and geo-hazard research. Examples are:

- Spatiotemporal modelling and extreme-event forecasting on big nonstationary geo-hazards data
- Spatiotemporal statistical methods for climate change research, including for seasonal tropical cyclone modelling and prediction, and for continental temperature and precipitation profiling
- Regression, clustering and classification for big and complex data/text mining and machine learning, including regression clustering,
- Support vector machines and ensemble of decision trees
- Statistical genomic and epigenomic association studies using cloud and clustered supercomputing
- Efficient computing algorithms for model selection in high-dimensional model space
- Decision and information theoretic methods for model selection
- Markov chain Monte Carlo including Gibbs sampler, EM algorithm and Bayesian computing
- Time series analysis, including change-points and dependence structure identification
- Generalised linear models and composite likelihood for categorical, longitudinal and panel data, and
- Statistical modelling and inference for capture-recapture data in ecological studies.
Since I was a child, I was fascinated by nature and natural phenomena. Over time this fascination only grew with seeing more and more phenomena and becoming more proficient in the mathematical tools to describe them. At some point, this also sparked my interest in the intrinsic beauty of mathematical theories. Since then, both passions evolved hand in hand as I attempt to make my modest contribution to advancing the knowledge of humankind.

**Topological quantum matter**

On the level of individual particles the fundamental constituents of matter (such as electrons) and their interactions (such as the Coulomb force) are well understood. However, in the presence of a large number of particles, the mutual interplay between interactions may lead to unexpected emergent phenomena, such as high-temperature superconductivity. Another phenomenon is associated with anyonic quasi-particles that can be manipulated to form braids and knots. These constitute a crucial ingredient of fault-tolerant topological quantum computation.

In my research, I am striving for a theoretical and mathematical description of such phenomena. This includes aspects of classification (What phenomena are out there?), characterisation (How to recognise and identify them if they are encountered?) and realisation (in theoretical toy models but also with experiments in mind).

Besides relying on purely theoretical and mathematical reasoning usually based on symmetry and entanglement considerations, the investigations also partly involve numerical simulations using tensor networks, neural nets or quantum computers.

**Critical phenomena and statistical learning**

Roughly speaking, critical phenomena occur in systems which cannot decide which of two or more possibilities with fundamentally different properties they should realise. Critical phenomena are omnipresent in physics, but they also play a crucial role in applied mathematics, probability theory, statistical learning or even more remote areas such as sociology or finance.

Criticality goes hand in hand with large fluctuations and as a consequence, quantities of interest frequently diverge at a critical point. These divergencies can be described by power laws featuring so-called critical exponents. These critical exponents are usually insensitive to the microscopic details of a system, and hence permit us to make quite general and very useful statements about the behaviour of whole classes of systems. For example, in statistical learning the critical exponents are linked to the behavior of the generalisation error, the error in predicting outcomes associated with unseen data. And surprisingly the results turn out to be relatively insensitive to the precise architecture of the network.

In my research on critical phenomena, I am making use of a range of methods such as conformal field theory (CFT), symmetries and renormalisation group ideas. Beyond the discussion of individual systems of interest, a strong focus of my research is on conceptual aspects and general structural properties of general critical phenomena.
Much of my research involves the development of foundational graph-theoretical tools, models and algorithms for solving optimisation problems from telecommunications, VLSI microchip design, wireless sensor networks and underground mining networks.

My primary research interest is in fundamental algorithmics for combinatorial optimisation. I have extensive experience in network optimisation, particularly in the design and analysis of exact algorithms.
The study of symmetry, as modelled by group theory, uses geometric and topological techniques. Finitely generated groups can be approached as geometric objects, and this point of view has been extremely fruitful over the past twenty-five years. To a finitely generated group, one can associate a metric space, the Cayley graph, from which one then aims to deduce as much algebraic information as possible.

Symmetry is omnipresent in both our everyday lives and in the abstract world of pure mathematics. My interest in this research field is motivated by the desire to understand and analyse the mathematical abstraction of symmetry.
I study the mathematical structures that underlie my favourite physical theories. These days that’s mostly two-dimensional conformal field theories, especially the logarithmic ones, though I have dabbled in the past in integrability. On the mathematical side, I study the representation theory of vertex operator algebras and superalgebras. Sometimes I even prove theorems. But, pretty much anything that involves cool math is fine with me.

One of my long-term interests is the class of conformal field theories known as the fractional-level Wess–Zumino–Witten models and the W-algebras obtained from them by quantum hamiltonian reduction. These are generally logarithmic, meaning that the corresponding vertex operator algebras have nonsemisimple module categories. This sounds hard (and it is), but there’s a lot of cool 21st century mathematics tied up in this game, from ‘relaxed’ versions of highest-weight modules to generalisations of modular tensor categories and exotic new types of modular forms.

Physicists have also recently renewed their interest in this area because of newly found dualities connecting these types of conformal field theories to higher-dimensional supersymmetric gauge theories.

If you recognise this as the roots of sl(3) and the weight of its fundamental representations, then you’ll know that I like to work on things related to Lie algebras.

If you recognise these as extended Kac tables for the logarithmic minimal models of central charge c=-2 and c=0, then you’ll know that I’m a big fan of conformal field theory.
The main goal of my research is to understand probability models that frequently form the basis of likelihoods in statistical procedures. Such an understanding is necessary to quantify when a procedure ‘works’, meaning that its output is of an appropriate precision given the quantity and quality of the input data. I hope that my research helps guide practitioners to effective methods of analysis.

My research is at the interface of probability and statistics and addresses problems that are fundamental to ‘big data’ applications. Here, the term ‘big’ refers not only to datasets of unprecedented size, leading not only to impossible computing scenarios using standard methods of analysis, but also to ones that have novel structures for which standard methods of analysis don’t yet exist. Such datasets with the latter property are now routinely generated in many applications, for example, genetic data or network-structured data such as social or contact networks.

A main theme of my research is to understand probability models used for such data arising from real-world applications, with the ultimate goal of guiding practitioners to appropriate methods of analysis. My work has focused on understanding the structure of a number of random network and genealogical models, many times in concert with developing a powerful tool in modern probability called Stein’s method. In particular, I have adapted the method to high and infinite-dimensional data settings of the kind that are now routinely encountered.

Genealogies

Consider the following problem from population genetics. A sample of size n is drawn from a population of size N, and the genes are sequenced for the purpose of inferring population parameters, for example, the past size of a viral population. For fixed sample size n, as the population size N tends to infinity the sampling distribution of types converges to the Ewen’s Sampling Formula (ESF). Thus, it is standard practice to use the ESF likelihoods for inference. However, in modern genetics studies, faster and cheaper sequencing has made sample sizes large relative to population size, so it is no longer appropriate to think of n as ‘fixed’ and N as ‘infinity’. How large can n be relative to N so that the ESF is a good approximation to the true sampling distribution? A consequence of my work is a partial answer to this question which gives an explicit upper bound in the error between probabilities under the finite-N sampling distribution and the ESF, which is mathematically rigorous and meaningful in its application.

Networks

A standard approach to understanding the structure of a given network, such as a social or contact network, is to fit a parametric network model to the data. Properties of the network can be read from those of the model with the inferred parameters. For example, when modelling the spread of a disease in a community, it is important to first carefully model the structure of the underlying contact network. This will determine, for instance, if there are a few people with many contacts (super-spreaders) or if most people have close to the average number of contacts.

A popular parametric network model in sociology is the Exponential random graph model (ERGM). While ERGMs are widely used, their properties are not well understood. Some recent mathematical work suggests that in ‘high temperature’ parameter regimes ERGMs behave similarly to much simpler and well-understood Erdős-Rényi random networks (ERRNs). Some of my recent work derives explicit bounds on the difference between average quantities in ERGMs and ERRNs — here again the result is mathematically rigorous and can be used by practitioners to determine if their model is not appropriate.
I am driven by the explanation of many natural phenomena in mathematical terms. This is especially true in general relativity, where mathematical prediction often precedes experimental discovery. Indeed, gravitational waves were predicted 100 years before they were detected. I am aspiring to advance the mathematical theory needed to understand the many predictions of general relativity theory.

My research interests are in partial differential equations, geometry and analysis, and I am particularly interested in mathematical problems that arise in general relativity theory. These questions pertain to the study of gravitational waves, black holes and models of the expanding universe. My primary research aims to understand the evolution of expanding black hole spacetimes and to analyse the nonlinear nature of gravitational waves in the presence of a positive cosmological constant.
From sensors to information and decision, we develop everything in between. The type of sensors, their complexity and the nature of the phenomenon they are used to monitor are constantly evolving and changing. Extraction of valuable information for analysis and decision making from the data provided by these sensors requires the development of new and complex statistical methods and the associated algorithms. These statistical methods and algorithms differ from problem to problem and task to task. Their development is a challenge which is always different, and the result can have a direct impact in areas such as health, education, environment and defence.

Statistical signal and image processing

Images and video can, for example, be found in the aerospace, cinema, environment and health industries. Signals, on the other hand, are found in the music, defence, communication and seismic industries. Given such information support, my research aims to develop machine learning and statistical methods and the associated algorithms to extract the carried information for analysing this type of data or decision making.

Examples of tasks can be the detection of a signal hidden in noise, recognition of a specific or unknown object, identification of a person, analysis of motion in a video, compression, anomaly detection and border security. The methods and algorithms that are developed can be used for performing tasks from a single signal or image, a batch set of signals or images, a mixture of signals and images, streaming signals and images or a set of distributed sensors, each delivering a separate piece of information.

Associate Professor
Karim Seghouane

Statistics
Signal processing
Image processing
Medical imaging
Physiological signal analysis

abd-krim.seghouane@unimelb.edu.au
findanexpert.unimelb.edu.au/profile/470194-karim-seghouane
I am driven to use my mathematical knowledge to tackle significant problems impacting society. I enjoy interdisciplinary collaboration and have worked with colleagues in many fields, including climate science, stem cell biology, robotics, and finance, to name just a few. One of the most significant societal challenges in the last decade or so has been the rise in popularity of algorithms to help make decisions. The vexing question of whether we can and should trust algorithms has provided the impetus for my Instance Space Methodology, to provide researchers and industry with the methodologies and tools to stress-test algorithms and identify their strengths and weaknesses. Using mathematical and statistical techniques, we can reveal and understand the conditions under which an algorithm is reliable and can be trusted. I hope this work will have far-reaching consequences, not just for improving research practice, but for real-world deployment of algorithms, particularly to address the growing concerns around algorithmic bias.

**Stress-testing algorithms for establishing trust**

Whether we trust an algorithm depends critically on how we test it. The choice of test examples, or instances, used to test an algorithm should not be arbitrary, random, or cherry-picked to make an algorithm look successful. Honest assessment of an algorithm’s strengths and weaknesses relies on being able to prove that the test instances are comprehensive and possess important properties, such as diversity, difficulty, discrimination, and lack of bias.

We have developed a powerful methodology known as Instance Space Analysis, using mathematical and statistical techniques, for visualising the whole instance space, and where an algorithm has been tested, where it should be tested with new test instances, and how the strengths and weaknesses of various algorithms can be understood in terms of instance properties. The methodology is broadly applicable, and has been applied to many problem domains in optimisation, machine learning, image processing, and other fields. Online tools are available at matilda.unimelb.edu.au to support researchers performing Instance Space Analysis for their problem domains, and stress testing their algorithms.

**Industrial optimisation**

Optimisation problems found in industry typically fall into two categories: model-based (where the goal is mathematically defined in terms of the decision variables) or model-free (aka black-box optimisation, where there is no mathematical statement of how the decisions affect the solution quality, only a limited set of past experiments that have been done to understand the cause-effect relationships between decisions and outcomes). We work with many industry partners in the ARC Training Centre for Optimisation Technologies, Integrated Methodologies, and Applications (OPTIMA) to tackle large-scale industrial optimisation problems of both types. Challenges include handling difficult constraints and multiple objectives, handling multi-fidelity sources of information, integrating predictive models with optimisation to manage uncertainty, decomposing problems into smaller components to handle large scale problems, and constructing real-world like test instances to rigorously stress test developed algorithms to ensure they are robust and fit for real-world deployment.

---

**Melbourne Laureate Professor**

**Kate Smith-Miles**

*Optimisation*

*Operations research*

*Machine learning*

*Data science*

*Algorithm testing*
Despite their ubiquitous use, many software packages routinely rely on a number of algorithms from mathematical optimisation which are not properly understood. My work aims to develop the mathematical theory required to rigorously justify the use of such algorithms and thereby ensure the integrity of the decision tools they produce.

Optimisation is the field of mathematics concerned with the problem of selecting a ‘best element’ from a set of permissible alternatives. Typically, these problems are expressed in terms of minimising or maximising the quantity of interest subject to constraints on the decisions which affect said quantity. My particular research interest is the development of computational methods for solving optimisation problems.
I’m passionate about applied probability.

Systems that are driven by stochastic or random effects are ubiquitous. It is crucial for society’s wellbeing, in economic and other terms, that we understand such stochastic systems and learn to control and optimise their operation. My research advances the mathematical theory of stochastic modelling at a deep level, connects with practitioners in science, economics, social science and engineering and reaches out to the public with the goal of educating people about stochastic phenomena.
My goal is to understand certain basic questions regarding representation theory, most of which are motivated by the Langlands program. In its most general incarnation, the Langlands program can be regarded as a kind of grand unification scheme involving several areas of mathematics. One immediate goal is to bring certain parts of this program involving real groups to conclusion.

Representation theory, algebraic geometry and the Langlands program

My research involves the study of representation theory, often done with the aid of geometric methods. Much of my research involves real groups: algebraic objects which describe the fundamental symmetries occurring in nature. They are important both in number theory and in the physical sciences, for example, quantum mechanics and particle physics. My work aims to address deep fundamental questions about these symmetries, and my research is motivated and guided by the Langlands program. This program predicts deep connections among different areas of mathematics as well as unexpected dualities. The study of such dualities is central in my work, which involves a variety of techniques. The geometric techniques are often aided by analysis which involve understanding the structure of a system of partial differential equations. A novel feature is the suggestion that the theory of mixed Hodge modules governs the representation theory of real groups. Most recently, I have been working on theory of character sheaves in the context of graded Lie algebras. This work is connected to many other areas of representation theory, and the (at the moment still distant) goal is to apply it to representation theory of p-adic groups.
I love thinking about and striving to understand beautiful structures in elusive concepts. I would like to build upon them and make new discoveries. These studies may provide powerful new tools to other branches of mathematics and physics where understanding and manipulation of structures are key. The influence propagates to other disciplines, which in turn may have real-world applications that none of us that help create them can foresee.

My research area is number theory. I study automorphic forms and automorphic representations. Their invariants, such as L-functions, are mirrored in the representations of Galois groups. This is an aspect of the Langlands Program. I focus on the study of the behaviour of automorphic representations under theta correspondence and more generally under endoscopic lift/descent. We may glean important information on L-functions from this investigation. In particular, we may contribute to the resolution of conjectures such as the Riemann Hypothesis, which concerns the Riemann zeta function that the L-functions generalise. I also study Abelian varieties and their moduli spaces; my research aims to construct explicit Abelian varieties which produce L-functions with desirable vanishing properties at the critical point. The invariants attached to Abelian varieties, such as L-function and rank of groups of rational points, are the objects that appear in the Birch-Swinnerton-Dyer conjecture.

**Theorem (Jiang-Wu, corollary to which gives “≥”)**

\[ F^{\text{out}}(g) = \dim Y + 2r \Rightarrow \]

\[ \int_{\mathcal{F}} \mathcal{R}_{\mathbb{Z}} = \mathcal{E}(g, s, f_{X(G)}) \mathcal{F}_{X_1, Y}(g, 1, f_{\phi}) d\mathcal{g} \neq 0 \]

with \( s_0 = \frac{3}{2}(\dim X - (\dim Y + 2r) + 2) \). Here

- \( J = \begin{cases} \text{Sp}_{2n+2} & \text{if } r \text{ even;} \\ \text{Jacobi group with Levi } & \text{Sp}_{2n+2-2r-1} \text{ if } r \text{ odd.} \\ \end{cases} \)

- \( J \in \text{Sp}(X_1) = \text{Sp}_{2n+2} \).

**Galois Orbits in A-tree**

**Definition**

For a set of vertices define the centre to be the central edge/vertex on any longest path that connects two vertices.

**Proposition**

The centre associated to the set \( \{ \mathcal{P} | x \in \text{Gal}(\mathbb{F}/\mathbb{F}) \} \) is fixed under \( \text{Gal}(F'/F) \).
While there are always concrete math problems that I’m working on, the ultimate goal is always to reveal the surprising aspect of the mathematical object. This can be the pattern, analog, generalisation or link between seemingly unrelated branches. It is the intrinsic beauty of pure math that makes me appreciate it, look for it and try to touch it.

Symmetries of Cayley digraphs

Algebraic graph theory is an important branch of mathematics and Cayley digraphs are fundamental objects of study in this field. Given a group and a subset of the group one can construct a digraph, called a Cayley digraph, by regarding the elements of the group as vertices and drawing an arc between two vertices whenever their quotient lies in the subset.

Since introduced by Arthur Cayley in 1878, the study of Cayley digraphs has been at the intersection of group theory and graph theory, and is now a central topic in algebraic graph theory and geometric group theory. The aim of research in this area is to characterise and construct Cayley digraphs with different levels of symmetries, using tools from combinatorics and permutation group theory.
My current research aims to establish a Springer theory in the setting of graded Lie algebras. The theory provides a framework that connects many central objects in representation theory, such as Hecke algebras, affine Springer fibres and Hessenberg varieties. As such, a long-term goal is to apply our theory to gain insight into representation theory in these more complex contexts, as well as geometry of fundamental objects in algebraic geometry and number theory.

Algebraic groups and Springer theory

My research is concerned with geometric aspects of representation theory, a mathematical study of the basic symmetries that occur in nature. Representation theory is a central area of mathematics which facilitates connections among different areas and has many applications to sciences such as quantum physics and quantum chemistry. It studies groups, the algebraic structures which encode symmetries, via their representations: linear algebra data that is easier to understand. A fundamental role in representation theory of finite groups is played by the Springer theory and character sheaves. My research has focused on developing Springer theory and the theory of character sheaves in more general settings. Our theory is expected to have applications in representations of real groups and p-adic groups which encode more complicated symmetries, as well as in number theory and algebraic geometry.
I study the relation of Morava E-theories (oriented cohomology theories in topology) and the character formulas built using Kazhdan–Lusztig polynomials in representation theory. I am interested in this project since it not only provides new insight into classical constructions, but also produces new objects. It also uncovers the profound connection between two fields of mathematics.

My primary research area is geometric representation theory. The representation theory studies the basic symmetries of mathematics and physics. The geometric representation theory is the study of the problems and objects of representation theory by discovering new geometric tools and insights and has close and deep connections to many branches of mathematics. I study the quantum groups which are deformations of the basic symmetries and their representations which can be realised as cohomology of certain moduli spaces of sheaves arising from mathematical physics.
I would like to use my knowledge and skills to solve real-life problems, such as last-mile delivery, staff rostering, public transport scheduling, traffic flow management and signal optimisation, and see the models, methods and solutions that we develop implemented in practice and benefit people.

Our research applies techniques in Operations Research to model and optimise real-life problems arising in areas such as transportation, city logistics and health care systems. Insights obtained from our research will inform stakeholders, including companies and government agencies, to make smarter decisions and achieve better performance.

**City logistics**
Growing freight demand from e-Commerce combined with low utilisation of freight vehicles by independent operators has considerable effects on urban traffic congestion. Our group has been working on a platform that is used for modelling an open, shared and integrated urban freight network and designing viable and effective solutions that reduce the total network costs and benefit all stakeholders. The platform can estimate the effects of the solutions on the financial performance of shippers and carriers and the impacts on traffic congestion. It helps understand the factors that motivate the stakeholders to participate in such an integrated network in practice.

**Traffic flow modelling and control**
Sophisticated traffic signal controllers are essential for achieving good performance of urban transportation networks. Our group develops a stochastic microsimulation model for studying the fundamental relationship between traffic flow and density and identifying the factors that determine a network’s performance, which include traffic demand’s spatial and temporal patterns, traffic signal settings and road lengths. The model is also used for devising traffic control strategies, including green-wave and perimeter control, and evaluating their performance in mitigating traffic congestion.
To push further knowledge in pure mathematics.

I am interested in representation theory and related algebraic geometry and topology. In particular, I am interested in various types of quantum symmetries often of mathematicophysical nature arising from geometry. The methods I use involve cohomology theories, derived category and moduli spaces.
I'm interested in the mathematical aspects of fundamental theories of physics and how they can in turn shed new light on problems of a purely mathematical nature.

My main area of research is the interaction between quantum integrable systems (a topic from mathematical physics) and pure mathematics, in particular, algebraic combinatorics and enumerative geometry. I also have an interest in computational methods in algebraic geometry and combinatorics, and in computer algebra systems.