Discrete Mathematics and Data Science Research Team Seminars

All seminars are held at 15:10-16:00 in Room M/1.02, Senghennydd Road, Cardiff unless stated otherwise.

Programme Coordinators:

Dr Iskander Aliev (MATHS), Dr P. Corcoran (COMPS) and Dr A. Gagarin(MATHS/COMPS)

Date	Speaker	Seminar
22 June 2017	Dr Varun Kanade (Oxford University)	From which world is your graph? Discovering statistical structure from links is a fundamental problem in the analysis of social networks. Choosing a misspecified model, or equivalently, an incorrect inference algorithm will result in an invalid analysis or even falsely uncover patterns that are in fact artifacts of the model. This work focuses on unifying two of the most widely used link-formation models: the stochastic blockmodel (SBM) and the small world (or latent space) model (SWM).

Date	Speaker	Seminar
		Integrating techniques from kernel learning, spectral graph theory, and nonlinear dimensionality reduction, we develop the first statistically sound polynomial-time algorithm to discover latent patterns in sparse graphs for both models. When the network comes from an SBM, the algorithm outputs a block structure. When it is from an SWM, the algorithm outputs estimates of each node's latent position. We report experiments performed on a twitter dataset collected between October and November 2016. Based on joint work with Cheng Li, Zhenming Liu and Felix Wong.
1 June 2017	Dr Peter Richtarik (University of Edinburgh)	Stochastic reformulations of linear systems and efficient randomized algorithms

Date	Speaker	Seminar
		Please note different time and venue: 15:00 in WX3.07 (Queens Buildings). We propose a new paradigm for solving linear systems with a very large number of equations. In our paradigm, the system is first reformulated into a stochastic problem, and then solved with a suitable (typically randomized) algorithm. Our stochastic reformulation is flexible as it depends on a user-defined parameter in the form of a distribution defining an ensemble of random matrices. The choice of the distribution directly influences the "condition number" of the reformulation, which leads to the novel concept of "randomized preconditioning". We give necessary and sufficient conditions for the reformulation to be exact, i.e., for the

Date	Speaker	Seminar
		solution set of the stochastic problem to be identical to the solution set of the linear system. We also show that the reformulation can be equivalently seen as a stochastic optimization problem, stochastically preconditioned linear system, stochastic fixed-point problem and as a probabilistic intersection problem. For instance, the condition number of the reformulation is equal to the condition number of the stochastically preconditioned linear system, and to the condition number of associated with the Hessian of the objective function appearing the stochastic optimization reformulation.
		Further, we propose and analyze basic, parallel and accelerated stochastic algorithms for solving the reformulated problem, with linear

Date	Speaker	Seminar
		convergence rates. The methods have natural and sometimes surprising interpretations from the viewpoint of each of the four reformulations.For instance, the methods can be interpreted as basic, parallel and accelerated variants of stochastic gradient descent, stochastic Newton descent, stochastic projection method and stochastic fixed-point method. The complexity of the basic variants scales linearly with the condition number of the reformulation, while the accelerated variants scale with the square root of the condition number. Moreover, all our methods lend themselves to a natural dual interpretation as "stochastic subspace ascent" methods, a novel class of optimization algorithms not analyzed before. Stochastic dual

Date	Speaker	Seminar
		coordinate ascent and stochastic dual Newton ascent arise in special cases. We prove global linear convergence of all our algorithms. Further, we highlight a close connection to recent algorithmic developments in machine learning through casting the problem as an instance of the Empirical Risk Minimization problem in a new regime not studied before.

The above development can be extended to matrix inversion. In particular, we develop and analyze a broad family of stochastic/randomized algorithms for inverting a matrix, with specialized variants maintaining symmetry and/or positive definiteness of the iterates. All methods in the family converge

Date	Speaker	Seminar
		globally and linearly, with explicit rates. In special cases, we obtain stochastic block variants of several quasi-Newton updates, including bad Broyden (BB), good Broyden (GB), Powell-symmetric-Broyden (PSB), Davidon-Fletcher-Powell (DFP) and Broyden-Fletcher-Goldfarb-Shanno (BFGS). Ours are the first stochastic versions of these updates shown to converge to an inverse of a fixed matrix. Through a dual viewpoint we uncover a fundamental link between quasi-Newton updates and approximate inverse preconditioning. Further, we develop an adaptive variant of randomized block BFGS, where we modify the distribution underlying the stochasticity of the method throughout the iterative process to achieve faster convergence. Further,

Date	Speaker	Seminar
		for rectangular and non-invertible matrices, variants of our methods can be shown to converge to the Moore-Penrose pseudoinverse.
28 April 2017	Richard Booth (School of Computer Science and Informatics, Cardiff University)	Quantifying distance between argumentation graph-labellings An argumentation framework can be seen as a graph that expresses, in an abstract way, the conflicting information of an underlying logical knowledge base. The nodes of the graph correspond to arguments, and the edges correspond to attacks between arguments. This conflicting information often allows for the presence of more than one possible reasonable position which one can take, with each position corresponding to a particular labelling of the graph that labels arguments "accepted",

Date	Speaker	Seminar
		"rejected" or "undecided". A relevant question, therefore, is how much these positions differ from each other. In this talk, we will examine the issue of how to define meaningful measures of distance between the (complete) labellings of a given argumentation framework. We provide concrete distance measures based on argument-wise label difference, as well as based on the notion of *issues*, and examine their properties.
24 March 2017	Dr Frank Langbein (School of Computer Science and Informatics, Cardiff University)	Robust Control of Quantum Spin Networks Networks of spin-1/2 particles form basic components of quantum devices for networking, simulation and computing. Information stored in spin states can propagate through such networks without any charge transport in a wave-like manner,

Date	Speaker	Seminar
		dispersing and refocusing over the network. We specifically study single excitation dynamics in spin rings and chains. The resulting traffic shows surprising non-classical behaviour, such as the presence of an anti-core, a node with minimal traffic instead of the congestion core in traditional networks. Optimal control of quantum systems can be utilised to direct this traffic. In particular shaping the energy landscape via static bias fields, instead of the usual dynamic modulation or switching of control fields, yields highly efficient, simple control schemes.

Due to the limited degrees of freedom, finding such controls via optimisation is more difficult than for dynamic control. Nevertheless, high-fidelity results can be |

Date	Speaker	Seminar
		achieved. Some of the resulting controls show the astounding property of being optimally robust at optimal performance, as demonstrated by log-sensitivity results and the rank correlation between fidelity and the mu of the structured singular value analysis.
		For practical implementation of such control schemes, network topologies and couplings must be identified from measurements of specific devices to build a model suitable to find the necessary controls. For this we present an approach for discriminating between different network structures and learning model parameters. Overall this holds the promise to achieve robust, scalable quantum devices, such as routers and memories, complex spin dynamics

Date	Speaker	Seminar
		simulators, towards computing devices.
17 March 2017	Andrei Gagarin (Cardiff)	The bondage number of graphs on topological surfaces and Teschner's conjecture
		The domination number of a graph is the smallest number of its vertices which have all the other vertices in their neighborhoods. The bondage number of a graph is the smallest number of its edges whose removal results in a graph having a larger domination number. In a sense, the bondage number measures integrity and reliability of the smallest dominating sets with respect to edge removals in graphs, which may correspond, e.g., to link failures in communication networks. The decision problem for the bondage number is known to be NP-hard.

Date	Speaker	Seminar
		We provide constant upper bounds for the bondage number with respect to embeddability of graphs on topological surfaces, and improve upper bounds for the bondage number in terms of the maximum vertex degree and the orientable and non-orientable graph genera. Also, we present stronger upper bounds for graphs with no triangles and graphs with the number of vertices larger than a certain threshold in terms of graph genera. This settles Teschner's Conjecture in the affirmative for almost all graphs. As an auxiliary result, we show tight lower bounds for the number of vertices of graphs 2-cell embeddable on topological surfaces of a given genus. (Joint work with Vadim Zverovich, University of

Date	Speaker	Seminar
		the West of England, Bristol, UK).
17 February 2017	Prof Owen Jones (Cardiff)	The cross-entropy method for discrete optimisation

The cross-entropy method is a versatile heuristic tool for solving difficult estimation and optimization problems, based on Kullback–Leibler (or cross-entropy) minimization. It has been successfully applied to combinatorial optimisation problems such as the travelling salesman problem, the quadratic assignment problem, and the max-cut problem. It can also be used for noisy problems such as the buffer allocation problem.

In this talk I will give an introduction to the Cross-Entropy method, and then present some theoretical results that show it can find an optimal solution with

Date	Speaker	Seminar
		probability arbitrarily close to 1, and give conditions under which an optimal solution is generated eventually with probability 1.
25 November 2016	Dr. Shoaib Jameel (School of Computer Science and Informatics, Cardiff University)	Towards embeddings for information retrieval

We present two recent models to generate embeddings. In the first model, we concentrate on generating conceptual subspaces. Conceptual spaces are geometric representations of conceptual knowledge in which entities correspond to points, natural properties correspond to convex regions, and the dimensions of the space correspond to salient features. While conceptual spaces enable elegant models of various cognitive phenomena, the lack of automated methods for constructing such representations have so |

Date	Speaker	Seminar
		far limited their application in artificial intelligence. To address this issue, we propose a method which learns a vector-space embedding of entities from Wikipedia and constrains this embedding such that entities of the same semantic type are located in some lower-dimensional subspace. In the second part of the talk, we also propose a new word embedding model, inspired by GloVe, which is formulated as a feasible least squares optimization problem. In contrast to existing models, we explicitly represent the uncertainty about the exact definition of each word vector. To this end, we estimate the error that results from using noisy co-occurrence counts in the formulation of the model, and we model the imprecision that results from including

Date	Speaker	Seminar
		uninformative context words. Our experimental results demonstrate that these models compare favourably with existing models.
1 November 2016	Dr. Ender Ozcan(School of Computer Science, University of Nottingham)	Improving Heuristic Optimisation via Tensor Analysis

The advanced optimisation techniques underpin many of the automated decision support systems, producing "high quality", or the "best" solutions to the challenging real-world problems. Nevertheless, the complexity of such techniques has increased along with the complexity of problems to the point where their configuration and/or control is in itself a significant challenge. There is a growing number of studies in addressing this issue with the usage of various data science techniques. In this talk, |

Date	Speaker	Seminar
		we illustrate the potential of tensor analysis as an advanced machine learning technique for improving the performance of several heuristic optimisation methods on selected case studies, including nurse rostering.
28 October 2016	Dr. Yu-Kun Lai(Cardiff University, School of Computer Science and Informatics)	Analysis and Processing of 3D Shape Collections

With the availability of cheap 3D cameras and accessible 3D modelling packages, 3D geometric shapes have proliferated in recent years. Processing potentially large-scale shape collections effectively provides both opportunities and challenges. My talk will overview some of our recent effort in this research direction. We develop geometric modelling techniques that greatly benefit from shape collections for more meaning and |

Date	Speaker	Seminar
		intuitive interactive shape editing, shape interpolation and morphing editing. We also take on the challenge by developing techniques to allow interactive exploration of large-scale heterogeneous shape repositories using efficient shape analysis and visualisation.
30 September 2016	Prof. Anatoly Zhigljavsky (Cardiff)	Linear regression analysis with large number of predictors
14 September 2016	Prof. Igor Pak (UCLA)	How to prove Steinitz's theorem Steinitz's theorem is a classical but very remarkable result characterising graphs of convex polytopes in R^3. In this talk, I will first survey several known proofs, and present one that is especially simple. I will then discuss the quantitative version and recent advances in this

Date	Speaker	Seminar
		direction. Joint work with Stedman Wilson.