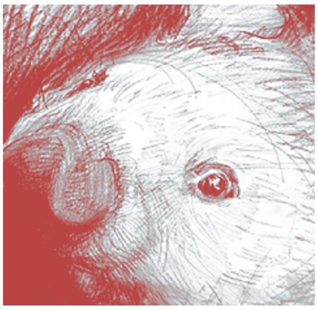

WOMBAT 2024

Workshop on Optimisation, Metric Bounds, Approximation and Transversality

4th - 6th December

Discipline of Buisness Analytics, The University of Sydney

Program and Abstracts

THE UNIVERSITY OF SYDNEY

We would like to acknowledge and pay respect to the traditional owners of the land on which we meet; the Gadigal people of the Eora Nation. It is upon their ancestral lands that the University of Sydney (Camperdown Campus) is built.

Workshop Information

We warmly welcome participants to the Workshop on Optimisation, Metric Bounds, Approximation and Transversality (WOMBAT) in 2024. This year, the workshop is hosted by the Discipline of Business Analytics at the University of Sydney.

Venue

The workshop will be held in the Abercrombie Building H70 at the University of Sydney Business School, corner of Abercrombie St and Codrington St, Darlington NSW 2006 (Google Maps link).

- All talks will be held in **Case Study Lecture Theatre 2140, Level 2 of H70**. The room has HDMI and USD-C ports to connect to projectors, or you can use a USB drive to load apresentation onto the room's desktop computer.
- The welcome reception on Wednesday 4th December will be held in **The Refectory, Level 5 of H70**, starting at **5:30pm**.

Catering

Morning tea and lunch are provided on all days. Afternoon tea is provided only on Wednesday 4th and Thursday 5th December.

Tourism information

The nearest shopping and dining strips to the University of Sydney are King St (Newton) and Glebe PointRd (Glebe). It is a short bus trip from the University to Haymarket/Chinatown or the city.

- Sydney tourism
- Visit New South Wales
- Sydney weather forecast
- Transport NSW (public transport information)
- Sydney Airport transport – the easiest way from the airport is to catch the T8 train to the city to Central-Station, then change to a T1/T2/T4/T9 train to Redfern (1 stop) or catch a bus (423/426/428/430)from Pitt St to City Rd (University of Sydney stop).

All Sydney public transport allows you to tap on/off with your credit card or digital wallet (Transport NSW payment information). The TripView Lite app is an excellent (free) way to plan trips on publictransport.

Acknowledgements

We gratefully acknowledge the administrative and financial support from the Discipline of Business Analytics at the University of Sydney. We would also like to thank our financial sponsor Gurobi for their generosity in supporting this event. Finally, we thank the University of Sydney Business School for providing rooms to host the workshop.

Organising committee

Li Chen (The University of Sydney)
Nam Ho-Nguyen (The University of Sydney)
Dmytro Matsypura (The University of Sydney)

Program — Wednesday 4th December 2024

09:00–10:00	**SUMMER SCHOOL** **Fred Roosta** – Newton-Type Methods: Exploring the Interplay Between Outer and Inner Iterations Part I
MORNING TEA 10:00–10:30	
10:30–12:00	**Eddie Anderson** – Stochastic optimisation when the distribution changes over time: a Wasserstein approach **Qiuzhuang Sun** – Optimal stopping under imperfect condition monitoring for non-Markovian systems **Tomas Lagos** – The Online Shortest Path Problem: Learning Travel Times Using a Multi-Armed Bandit Framework
LUNCH BREAK 12:00–13:30	
13:30–15:00	**Lindon Roberts** – Model Construction for Convex-Constrained Derivative-Free Optimization **Felipe Atenas** – Weakly convex optimization: methods of descent and epsilon-subdifferentials **Hongzhi Liao** – Finite Convergence of Circumcentered-Reflection Method on Closed Polyhedral Cones in Euclidean Spaces
AFTERNOON TEA 15:00–15:30	
15:30–17:00	**Kerry He** – Interior Point Methods for Quantum Relative Entropy Programming **Alexander Lim** – Faithful-Newton Methods **Oscar Smee** – Inexact Newton-type Methods for Optimisation with Nonnegativity Constraints
WELCOME RECEPTION 17:30–19:30 **The Refectory – H70 Level 5**	

Program — Thursday 5th December 2024

09:00–10:00	**SUMMER SCHOOL** **Fred Roosta** – Newton-Type Methods: Exploring the Interplay Between Outer and Inner Iterations Part II
MORNING TEA 10:00–10:30	
10:30–12:00	**Scott B Lindstrom** – Betting on duality: curious extrapolations and the Douglas–Rachford ADMM. **Queenie Yingkun Huang** – Piecewise sum-of-squares-convex moment optimisation via SDPs **Tan Pham** – A proximal splitting algorithm for generalized DC programming with applications in signal recovery
LUNCH BREAK 12:00–13:30	
13:30–15:00	**Hoa Bui** – On cutting plane algorithms for nonlinear binary optimization **Yuen-Man (Mandy) Pun** – Online Non-Stationary Quasar-Convex Optimization **Johnson Phosavanh** – Minimizing the number of late jobs and total late work with step-learning
AFTERNOON TEA 15:00–15:30	
15:30–17:00	**KEYNOTE TALK** **Melvyn Sim** – Estimation and Prediction Procedures for Unified Robust Decision Models

Program — Friday 6th December 2024

09:00–10:00	**SUMMER SCHOOL** **Fred Roosta** – Newton-Type Methods: Exploring the Interplay Between Outer and Inner Iterations Part III
MORNING TEA 10:00–10:30	
10:30–11:30	**Mareike Dressler** – Separability of Hermitian Tensors **Anthony Wirth** – Correlation Clustering Nuggets
LUNCH AND WORKSHOP CLOSE 11:30	

Abstracts

Keynote Talk — Thursday 5th December, 15:30–17:00

Estimation and Prediction Procedures for Unified Robust Decision Models

Professor Melvyn Sim

Over the past two decades, robust optimization techniques have efficiently addressed decision problems under uncertainty, offering high assurance of feasibility without being overly conservative. However, research on estimating parameters for these robust decision models from data has been lacking. In this paper, we focus on a unified framework for robust decision models that integrate robust optimization and robust satisficing paradigms. In particular, we identify two layers of uncertainty: outcome uncertainty, involving deviations from specified inputs based on historical data, and estimation uncertainty, addressing deviations from latent input vectors, such as the unobservable true mean. We introduce estimation and prediction procedures tailored to reduce conservativeness while enhancing feasibility within this unified robust decision framework. The concept of minimum volume confidence set is introduced, which minimizes the size of the outcome confidence set while considering the likelihood of infeasibility, thereby improving the precision of uncertainty characterization. This method also accommodates asymmetric uncertainty by adjusting the confidence set accordingly. Additionally, our framework incorporates an affine predictive model that leverages side information to improve input vector predictions, seamlessly integrating into robust decision models. Our method has been implemented in the algebraic modelling toolbox, RSOME, facilitating practical applications.

Summer School — Daily, 09:00–10:00

Newton-Type Methods: Exploring the Interplay Between Outer and Inner Iterations

Professor Fred Roosta

Compared to first-order optimisation algorithms, Newton-type methods offer significant theoretical and empirical advantages, including enhanced resilience to ill-conditioning, reduced sensitivity to hyperparameter uncertainty, superior local convergence, improved communication efficiency in distributed applications, and affine invariance. However, unlike the simplicity of first-order methods, second-order algorithms often involve non-trivial subproblems that must be solved at each iteration. Although the appropriate resolution of these subproblems is crucial to the effectiveness of the optimisation algorithm, they are often treated as an afterthought. This series of lectures will explore various Newton-type methods, focusing on the role of subproblem solvers. It will demonstrate how effectively leveraging suitable solvers can eliminate unnecessary safeguards and assumptions, resulting in simpler algorithms and analyses. We will cover a range of optimisation settings, including convex and non-convex problems, as well as unconstrained and constrained scenarios, along with both shared memory and distributed environments.

Wednesday 4th December, 10:30–12:00

Stochastic optimisation when the distribution changes over time: a Wasserstein approach

Eddie Anderson, e.anderson@imperial.ac.uk

In many practical situations we observe data drawn from a distribution that varies unpredictably over time. We explore the structure of the resulting stochastic optimisation problem when, at each time period, the change in distribution takes place in a way that gives equal likelihood to any distribution at the same Wasserstein distance from the distribution at the previous time period.

Optimal stopping under imperfect condition monitoring for non-Markovian systems

Qiuzhuang Sun, qiuzhuang.sun@sydney.edu.au

Using mission abort as an example, we study the optimal stopping problem for non-Markovian systems. While most on-demand mission-critical systems are engineered to be reliable to support critical tasks, occasional failures may still occur during missions. To increase system survivability, aborting the mission before an imminent failure is a common practice. We consider optimal mission abort for a system whose deterioration follows a general three-state (normal, defective, failed) semi-Markov chain. The failure is assumed self-revealed, while the healthy and defective states have to be predicted from imperfect condition monitoring data. Due to the non-Markovian process dynamics, optimal mission abort for this partially observable system is an intractable stopping problem. For a tractable solution, we introduce a novel tool of Erlang mixtures to approximate non-exponential sojourn times in the semi-Markov chain. This allows us to approximate the original process by a surrogate continuous-time Markov chain whose optimal control policy can be solved through a partially observable Markov decision process (POMDP). We show that the POMDP optimal policies converge almost surely to the optimal abort decision rules when the Erlang rate parameter diverges. This implies that the expected cost by adopting the POMDP solution converges to the optimal expected cost. Next, we provide comprehensive structural results on the optimal policy of the surrogate POMDP. We develop a modified point-based value iteration algorithm based on the results to numerically solve the surrogate POMDP. We further consider mission abort in a multi-task setting where a system executes several tasks consecutively before a thorough inspection. Through a case study on an unmanned aerial vehicle, we demonstrate the capability of real-time implementation of our model, even when the condition-monitoring signals are generated with high frequency.

The Online Shortest Path Problem: Learning Travel Times Using a Multi-Armed Bandit Framework
Tomas Lagos, tomasignacio.lagos@sydney.edu.au
In the age of e-commerce, logistics companies often operate within extensive road networks without accurate knowledge of travel times for their specific fleet of vehicles. Moreover, millions of dollars are spent on routing services that fail to accurately capture the unique characteristics of the drivers and vehicles of the companies. In this work, we address the challenge faced by a logistic operator with limited travel time information, aiming to find the optimal expected shortest path between origin-destination pairs. We model this problem as an online shortest path problem, common to many lastmile routing settings; given a graph whose arcsí travel times are stochastic and follow an unknown distribution, the objective is to find a vehicle route of minimum travel time from an origin to a destination. The planner progressively collects travel condition data as drivers complete their routes. Inspired by the combinatorial multiarmed bandit and kriging literature, we propose three methods with distinct features to effectively learn the optimal shortest path, highlighting the practical advantages of incorporating spatial correlation in the learning process. Our approach balances exploration (improving estimates for unexplored arcs) and exploitation (executing the minimum expected time path) using the Thompson sampling algorithm. In each iteration, our algorithm executes the path that minimizes the expected travel time based on data from a posterior distribution of the speeds of the arcs. We conduct a computational study comprising two settings: a set of four artificial instances and a real-life case study. The case study uses empirical data of taxis in the 17-km-radius area of the center of Beijing, encompassing Beijingís ì5th Ring Road.î In both settings, our algorithms demonstrate efficient and effective balancing of the exploration-exploitation trade-off.

Wednesday 4th December, 13:30–15:00

Model Construction for Convex-Constrained Derivative-Free Optimization
Lindon Roberts, lindon.roberts@sydney.edu.au
In model-based derivative-free optimisation algorithms, black-box functions are typically approximated using polynomial interpolation models. Most existing model-based DFO methods for constrained optimisation assume the ability to construct sufficiently accurate interpolation models, but this is not always achievable when sampling only feasible points. In this talk, I will outline a new approximation theory for linear and quadratic interpolation models in the presence of convex constraints.

Weakly convex optimization: methods of descent and epsilon-subdifferentials
Felipe Atenas, felipe.atenas@unimelb.edu.au
As modern optimization models become increasingly more complex, alternatives to convexity become a necessity. A benign form of nonconvexity is weak convexity, which captures several machine learning applications. In this talk, we discuss weakly convex minimization from a theoretical and algorithmic point of view. We present a framework for analyzing convergence of a class of descent methods, including methods of gradient and proximal type, as well as bundle methods. We establish linear rates of convergence to critical points under a subdifferential-based error bound on the distance to these points, akin to the popular Kurdyka-?ojasiewicz inequality. Furthermore, we introduce an enlarged subdifferential for weakly convex functions and analyze variational principles. Using this novel object, we explore numerical consequences for algorithmic patterns of epsilon-subgradient descent to find critical points of weakly convex functions.

Finite Convergence of Circumcentered-Reflection Method on Closed Polyhedral Cones in Euclidean Spaces
Hongzhi Liao, hongzhi.liao@unsw.edu.au
The Circumcentered-Reflection Method (CRM) is a relatively new projection method for solving convex feasibility problems. It has been proved that CRM enjoys better convergence performance than classic projection methods, such as the DouglasñRachford and alternating projection method. In this work, our first main theo-

rem shows CRM can find a feasible point for the intersection of two closed convex cones in $\mathbb{R}^2$ from any initial guess in the Euclidean plane. Then we apply this theorem to two polyhedral sets in $\mathbb{R}^2$ and two wedge-like sets in $\mathbb{R}^n$ and prove CRM can converge to a point in the intersections from any initial point. In addition, we propose a new technique, the sphere-centered reflection method, to show CRM can find a feasible point in finitely many iterations for the intersection of two proper polyhedral cones in $\mathbb{R}^3$ if the initial guess is in an outer cone of the intersection, which means every point except the origin of the intersection is an interior point of the outer cone. Finally, we raise an example to show that the finite convergence may fail for the intersection of two proper polyhedral cones from $\mathbb{R}^3$ if the initial guess is outside of the aforementioned outer cone.

Wednesday 4th December, 15:30–17:00

Interior Point Methods for Quantum Relative Entropy Programming

Kerry He, Kerry.He1@monash.edu
Quantum relative entropy programs are convex optimization problems which minimize a linear functional over an affine section of the epigraph of the quantum relative entropy function. Recently, the self-concordance of a natural barrier function was proved for this set. This has opened up the opportunity to use interior-point methods for nonsymmetric cone programs to solve these optimization problems. We present an efficient implementation of primal-dual interior point methods to solve quantum relative entropy programs, and show how common structures arising from applications in quantum information theory can be exploited to improve the efficiency of these methods.

Faithful-Newton Methods

Alexander Lim, alexander.lim@uq.edu.au
One of the most computational expensive operations in Newtonís method is the evaluation of a Newtonís step, which involves the computation of the inverse of the Hessian matrix. In large dimension problems, this operation becomes too expensive to carry out. Therefore, in many algorithms, like Newton-CG, Newton-CR, Newton-MR etc., they avoid computing the inverse of the Hessian, but using iterative solvers, like CG, CR or MINRES, to approximate a Newtonís step instead. However, this raises the questions ìhow do we judge the quality of an approximated Newtonís step?î, ìis solving a step very close to a Newtonís step a good idea?î. In an attempt to answer these questions from the computational cost perspective, we developed a variant of Newtonís method, called Faithful-Newton.

Inexact Newton-type Methods for Optimisation with Nonnegativity Constraints

Oscar Smee, o.smee@uq.edu.au
We consider solving large scale nonconvex optimisation problems with nonnegativity constraints. Such problems arise frequently in machine learning, such as nonnegative least-squares, nonnegative matrix factorisation, as well as problems with sparsity-inducing regularisation. In such settings, first-order methods, despite their simplicity, can be prohibitively slow on ill-conditioned problems or become trapped near saddle regions, while most second-order alternatives involve non-trivially challenging subproblems. The two-metric projection framework, initially proposed by Bertsekas (1982), alleviates these issues and achieves the best of both worlds by combining projected gradient steps at the boundary of the feasible region with Newton steps in the interior in such a way that feasibility can be maintained by simple projection onto the nonnegative orthant. We develop extensions of the two-metric projection framework, which by inexactly solving the subproblems as well as employing non-positive curvature directions, are suitable for large scale and nonconvex settings. We obtain state-of-the-art convergence rates for various classes of non-convex problems and demonstrate competitive practical performance on a variety of problems.

Thursday 5th December, 10:30–12:00

Betting on duality: curious extrapolations and the Douglas–Rachford ADMM.

Scott B Lindstrom, scott.lindstrom@curtin.edu.au
I will look at some ways that the duality-like relationship connecting alternating direction method of multipliers (ADMM) and the Douglas–Rachford method is used for extrapolation schemes.

Piecewise sum-of-squares-convex moment optimisation via SDPs

Queenie Yingkun Huang, yingkun.huang@unsw.edu.au
In this talk, we present exact Semi-Definite Program (SDP) reformulations for a class of moment optimisation problems involving piecewise Sums-of-Squares (SOS) convex functions and projected spectrahedral

support sets. These moment problems cover real-world applications such as the newsvendor and revenue maximisation problems with higher-order moments. Generally, solving moment problems is computationally intractable because evaluating a multi-dimensional integral for the expectations and searching through the infinite-dimensional space of probability distributions is numerically hard.

Our approach involves establishing an SOS representation for the non-negativity of a piecewise SOS-convex function on a projected spectrahedron and employing conic program duality. We show that both the optimal value and an optimal probability measure of the original moment problem can be found by solving a single SDP, which can be solved efficiently using commonly available software.

This talk is based on: Huang, Q.Y., Jeyakumar, V. and Li, G., 2024. Piecewise SOS-convex moment optimization and applications via exact semi-definite programs. EURO Journal on Computational Optimization, p.100094. https://doi.org/10.1016/j.ejco.2024.100094

A proximal splitting algorithm for generalized DC programming with applications in signal recovery

Tan Pham, pntan.iac@gmail.com

The difference-of-convex (DC) program is a crucial approach to nonconvex optimization problems due to its structure, which encompasses a wide range of practical applications. In this paper, we aim to tackle a generalized class of DC programs, where the objective function is formed by summing a possibly nonsmooth nonconvex function and a differentiable nonconvex function with Lipschitz continuous gradient, and then subtracting a nonsmooth continuous convex function. We develop a proximal splitting algorithm that utilizes proximal evaluation for the concave part and DouglasñRachford splitting for the remaining components. The algorithm guarantees subsequential convergence to a stationary point of the problem model. Under the widely used Kurdykañ?ojasiewicz property, we establish global convergence of the full sequence of iterates and derive convergence rates for both the iterates and the objective function values, without assuming the concave part is differentiable. The performance of the proposed algorithm is tested on signal recovery problems with a nonconvex regularization term and exhibits competitive results compared to notable algorithms in the literature on both synthetic data and real-world data.

Thursday 5th December, 13:30–15:00

On cutting plane algorithms for nonlinear binary optimization

Hoa Bui, hoa.bui@curtin.edu.au

The cutting plane method is known to converge for concave discrete maximization problems. This talk introduces a new condition for convergence that is weaker than concavity, allowing for extensions to generalized concave cases such as quasiconcavity and pseudoconcavity. More significantly, the new convergence condition enables the use of smaller penalty values in the standard concave reformulation technique for general nonconcave problems, yielding stronger cuts and faster convergence. In concave cases, this condition can even be applied to "de-concavify" concave problems to enhance convergence speed. Numerical results show that exploiting the new convergence condition can yield substantial computational improvements for complex, large-scale discrete optimization problems with up to one thousand variables.

Online Non-Stationary Quasar-Convex Optimization

Yuen-Man (Mandy) Pun, yuenman.pun@anu.edu.au

This work studies the time-varying online stochastic optimization problems with quasar-convex losses, which has recently been shown to have applications in identification of linear dynamical systems and generalized linear models. We apply online gradient descent and establish regret bounds in terms of cumulative path variation and cumulative gradient variance. We then show that these results can be applied to problems including generalized linear models, phase retrieval, and tomographic reconstruction. Numerical results are then presented to support our theoretical findings.

Minimizing the number of late jobs and total late work with step-learning

Johnson Phosavanh, johnson.phosavanh@sydney.edu.au

We study single-machine scheduling problems with step-learning, where an improvement in processing time is experienced if a job is started at, or after, a job-dependent learning-date. We consider minimizing two functions: the number of late jobs and the total late work, and we show that when at least a common due-date or common learning-date is assumed, the problem is NP-hard in the ordinary sense; however, when both are arbitrary, the problem becomes strongly NP-hard. For each of the problems where at least one of the dates is assumed to be common, we analyze the structure of an optimal job schedule with and without idle time and propose pseudo-polynomial time dynamic programming algorithms. We also show that under some assumptions, the problem is equivalent to minimizing the makespan, and we provide a more efficient algorithm than the existing literature. Lastly, we show that our analysis can also be applied to the case of step-deterioration, where instead, the processing times of jobs increase at a given date.

Friday 5th December, 10:30–11:30

Separability of Hermitian Tensors

Mareike Dressler, m.dressler@unsw.edu.au

Hermitian tensors naturally generalize Hermitian matrices, yet they exhibit distinct properties. A Hermitian tensor is separable if it admits a Hermitian decomposition with only positive coefficients, i.e. it can be expressed as a sum of rank-1 positive semidefinite (psd) Hermitian tensors. In this talk, I show how to detect separability of Hermitian tensors. This problem is equivalent to the long-standing quantum separability problem in quantum physics, which seeks to determine whether a given quantum state is entangled - a question that traces back to the foundational work of Einstein and Schrödinger (1935). We formulate the separability problem as a truncated moment problem and then provide a semidefinite relaxation algorithm to solve it. I conclude with outlining an alternative approach based on psd decompositions of separable Hermitian tensors and general tensor decomposition methods, which proves to be more computationally efficient for detecting separability in tensors of larger sizes.

Correlation Clustering Nuggets

Anthony Wirth, anthony.wirth@sydney.edu.au

Correlation Clustering is an important problem in the constrained clustering family. Given (inconsistent) advice about which pairs of items should be co-clustered and which should be separated, the aim is to produce a clustering that observes as much of the advice as possible. An NP-hard problem, it has a rich collection of algorithms, as well as strong connections to ranking, consensus clustering, maximizing modularity, and other cut-style network problems. Correlation clustering arose from questions about entity resolution in natural-language and database processing. In this presentation, I describe several of my and my colleagues' forays into correlation clustering.

List of Participants

Sara Abtahi
Eddie Anderson
Felipe Atenas
Hoa Bui
Li Chen
Mareike Dressler
Andrew Eberhard
Joe Gurr
Kerry He
Nam Ho-Nguyen
Queenie Yingkun Huang
Jeya Jeyakumar
Tomas Lagos
Jessica Leung
Hongzhi Liao
Alexander Lim
Scott B Lindstrom
Fangyu Liu
Pasindu Marasinghe
Dmytro Matsypura
Michael Nefiodovas
Tan Pham
Johnson Phosavanh
Yuen-Man (Mandy) Pun
Md Hifzur Rahaman
Lindon Roberts
Fred Roosta
Melvyn Sim
Oscar Smee
Qiuzhuang Sun
Yiwei Wu
Qi (Wenky) Wang
Anthony Wirth
Wei Zhang