Mini workshop on Optimization	
Date: 11 Dec 2025 **Venue: Block S17, Room 04-06** **Address:** Department of Mathematics, 10 Lower Kent Ridge Road, Singapore 119076	

0930-1000	**Stephen Wright, University of Wisconsin-Madison**
	TBA
1000-1030	**Shaoning Han, NUS**
	Analysis of A Class of Heaviside Composite Minimization Problems
	The minimization of nonclosed functions is a difficult topic that has been minimally studied. Among such functions is a Heaviside composite function that is the composition of a Heaviside function with a possibly nonsmooth multivariate function. Unifying a statistical estimation problem with hierarchical selection of variables and a sample average approximation of composite chance constrained stochastic programs, a Heaviside composite optimization problem is one whose objective and constraints are defined by sums of possibly nonlinear multiples of such composite functions. In this talk, we first present difficulties of tackling such problems using existing optimization techniques. Then we establish stationarity conditions for the problem, and introduce a sufficient condition, termed local convex-like property, under which the proposed stationary point is a local minimizer. Finally, we briefly discuss solution strategies via lifting and reformulation techniques.
1030-1100	**Coffee break**
1100-1130	**Guanyi Wang, NUS**
	Fast Presolving Framework For Sparsity Constrained Convex Programming: Screening-Based Cutting Plane Generation and Selection
	Screening is a widely utilized presolving method for Mixed-Integer Programming (MIP). It aims to certify a priori whether one or multiple specific binary variables can be fixed to optimal values based on solutions from convex relaxations. This paper studies the challenge of solving Sparsity-constrained (strongly) Convex Programming (SCP) and proposes the Screening-based Cut Presolving Framework (SCPF). SCPF contains two parts: a Screening-based Cut Generation (SCG) rule and a Screening-based Cut Selection (SCS) method. We demonstrate that the SCG provides superior screening ability compared to existing safe screening rules, and achieves a finer balance between screening effectiveness and computational overhead. Furthermore, we provide theoretical guarantees for the SCS method to ensure the selection of cuts with high screening ability. Extensive numerical experiments validate these theoretical findings and indicate that the proposed framework significantly outperforms state-of-the-art methods. Notably, our SCPF yields a 2x to 3x reduction in total running time across challenging scenarios, including high-dimensional synthetic datasets, complex real-world instances, and sparse identification of nonlinear dynamics.
1130-1200	**Meixia Lin, SUTD**
	Low rank convex clustering for matrix-valued observations
	Common clustering methods, such as \$k\$-means and convex clustering, group similar vector-valued observations into clusters. However, with the increasing prevalence of matrix-valued observations, which often exhibit low rank characteristics, there is a growing need for specialized clustering techniques for these data types. In this paper, we propose a low rank convex clustering model tailored for matrix-valued observations. Our approach extends the convex clustering model originally designed for vector-valued data to classify matrix-valued

	observations. Additionally, it serves as a convex relaxation of the low rank k-means method proposed by Z. Lyu, and D. Xia (J. R. Stat. Soc. B, to appear). Theoretically, we establish exact cluster recovery for finite samples and asymptotic cluster recovery as the sample size approaches infinity. We also give a finite sample bound on prediction error in terms of centroid estimation, and further establish the prediction consistency. To make the model practically useful, we develop an efficient double-loop algorithm for solving it. Extensive numerical experiments are conducted to show the effectiveness of our proposed model.
1200-1330	**Lunch break**
1330-1400	**Yong sheng Soh, NUS**
	Symmetries and Semidefinite Programming
	In this talk I will discuss data processing problems that involve some notion of symmetry. Prominent examples include graph alignment, image registration, cryo-EM and group invariant regularizer learning. A core object in these problems is the convex hull of the orbit of a function acted on by the group action that characterizes the symmetry of interest. In particular, we argue that tractable descriptions of such sets (or its outer approximations) translate to tractable algorithms for solving these associated data processing problem. I will focus on the case where the group is SO(2) and SO(3), in which case the associated convex hull admits spectrahedral descriptions. I will discuss harmonic analytical aspects as well as numerical implementations associated with these methods.
1400-1430	**Antonios Varvitsiotis, SUTD**
	Concave/Monotone Games via Sum-of-Squares Hierarchies
	Concavity and its refinements underpin tractability in multiplayer games, where players independently choose actions to maximize their own payoffs which depend on other players' actions. In *concave* games, where players' strategy sets are compact and convex, and their payoffs are concave in their own actions, strong guarantees follow: Nash equilibria always exist and decentralized algorithms converge to equilibria. If the game is furthermore *monotone*, an even stronger guarantee holds: Nash equilibria are unique under strictness assumptions. However, we show that *certifying* concavity or monotonicity is NP-hard, already for games where utilities are multivariate polynomials and compact, convex basic semialgebraic strategy sets, an expressive class that captures extensive-form games with imperfect recall. On the positive side, we develop two hierarchies of sum-of-squares programs that certify concavity and monotonicity of a given game, and each level of the hierarchies can be solved in polynomial time. We show that almost all concave/monotone games are certified at some finite level of the hierarchies. Subsequently, we introduce the classes of SOS-concave/monotone games, which globally approximate concave/monotone games, and show that for any given game we can compute the closest SOS-concave/monotone game in polynomial time. Finally, we apply our techniques to canonical examples of extensive-form games with imperfect recall.
1430-1500	**Di Hou, NUS**
	A Low-rank Augmented Lagrangian Method for Polyhedral–SDP and Moment-SOS Relaxations of Polynomial Optimization
	Polynomial optimization problems (POPs) can be reformulated as geometric convex conic programs, as shown by Kim, Kojima, and Toh (SIOPT 30:1251–1273, 2020), though such formulations remain NP-hard. In this work, we prove that several well-known relaxations can be unified under a common polyhedral–SDP framework, which arises by approximating the intractable cone by tractable intersections of polyhedral cones with the positive semidefinite matrix cone. Although effective in providing tight lower bounds, these relaxations become computationally expensive as the number of variables and constraints grows at the rate of $\Omega(n^{2\tau})$ with the relaxation order τ. To address this challenge, we propose RiNNAL-POP, a low-rank augmented Lagrangian method (ALM) tailored to solve large-scale polyhedral–SDP relaxations of POPs. To efficiently handle the $\Omega(n^{2\tau})$ nonnegativity and consistency constraints, we design a tailored projection scheme whose computational cost scales linearly with the number of variables. In addition, we identify a hidden facial structure in the polyhedral--SDP relaxation, which enables us to eliminate a large number of linear constraints by restricting the matrix variable to affine subspaces corresponding to exposed faces of the semidefinite cone. The latter enables us to efficiently solve the factorized ALM subproblems over the affine subspaces. At

	each ALM iteration, we additionally carry out a single projected gradient step with respect to the original matrix variable to automatically adjust the rank and escape from spurious local minima when necessary. We also extend our RiNNAL-POP algorithmic framework to solve moment-SOS relaxations of POPs. Extensive numerical experiments on various benchmark problems demonstrate the robustness and efficiency of RiNNAL-POP in solving large-scale polyhedral–SDP relaxations.
1500-1530	**Coffee Break**
1530-1600	**Hemant Tyagi, NTU**
	Joint learning of a network of linear dynamical systems via total variation penalization
	We consider the problem of joint estimation of the parameters of \$m\$ linear dynamical systems, given access to single realizations of their respective trajectories, each of length \$T\$. The linear systems are assumed to reside on the nodes of an undirected and connected graph \$G = ([m], E)\$, and the system matrices are assumed to either vary smoothly or exhibit small number of ``jumps'' across the edges. We consider a total variation penalized least-squares estimator and derive non-asymptotic bounds on the mean squared error (MSE) which hold with high probability. In particular, the bounds imply for certain choices of well connected \$G\$ that the MSE goes to zero as \$m\$ increases, even when \$T\$ is constant. Joint work with Claire Donnat and Olga Klopp.
1600-1630	**Xin Tong, NUS**
	Stochastic Gradient Descent with Adaptive Data
	Stochastic gradient descent (SGD) is a powerful optimization technique that is particularly useful in online learning scenarios. Its convergence analysis/effectiveness is relatively well understood under the assumption that the data samples are independent and identically distributed (iid). However, applying online learning to policy optimization problems in operations research involves a distinct challenge: the policy changes the environment and thereby affects the data used to update the policy. The adaptively generated data stream involves samples that are non-stationary, no longer independent from each other, and affected by previous decisions. The influence of previous decisions on the data-generating environment introduces estimation bias in the gradients, which presents a potential source of instability for online learning not present in the iid case. In this paper, we introduce simple criteria for the adaptively generated data stream to guarantee the convergence of SGD. We show that the convergence speed of SGD with adaptive data is largely similar to the classical iid setting, as long as the mixing time of the policy induced dynamics is factored in. Our Lyapunov-function analysis allows one to translate existing stability analysis of systems studied in operations research into convergence rates for SGD, and we demonstrate this for queuing and inventory management problems. We also showcase how our result can be applied to study an actor-critic policy gradient algorithm.
1630-1700	
1800-2000	**Dinner, by invitation**