IITB-Monash Research Academy

Indian Institute of Technology Bombay | MONASH University

An Indian-Australian research partnership

Project Title:	Scalable Algorithms for Structured Optimization Problems	
Project Number	**IMURA1096**	
Monash Main Supervisor (Name, Email Id, Phone)	Dr James Saunderson, james.saunderson@monash.edu	*Full name, Email*
Monash Co-supervisor(s) (Name, Email Id, Phone)		
Monash Head of Dept/Centre (Name,Email)	Prof. Scott Tyo, scott.tyo@monash.edu	*Full name, email*
Monash Department:	Electrical and Computer Systems Engineering	
Monash ADGR (Name,Email)	Prof Timothy Scott, timothy.scott@monash.edu	*Full name, email*
IITB Main Supervisor (Name, Email Id, Phone)	Vishnu Narayanan, vishnu@iitb.ac.in	*Full name, Email*
IITB Co-supervisor(s) (Name, Email Id, Phone)		*Full name, Email*
IITB Head of Dept (Name, Email, Phone)	Jayendran Venkateswaran, jayendran@iitb.ac.in	*Full name, email*
IITB Department:	Industrial Engineering and Operations Research	

Research Clusters:

Research Themes:

Highlight which of the Academy's CLUSTERS this project will address? *(Please nominate JUST* ***one.*** *For more information, see www.iitbmonash.org)*		**Highlight which of the Academy's Theme(s) this project will address?** *(Feel free to nominate more than one. For more information, see www.iitbmonash.org)*	
1	*Material Science/Engineering (including Nano, Metallurgy)*	***1***	*Artificial Intelligence and Advanced Computational Modelling*
2	*Energy, Green Chem, Chemistry, Catalysis, Reaction Eng*	2	*Circular Economy*
3	*Math, CFD, Modelling, Manufacturing*	3	*Clean Energy*
4	*CSE, IT, Optimisation, Data, Sensors, Systems, Signal Processing, Control*	4	*Health Sciences*
5	*Earth Sciences and Civil Engineering (Geo, Water, Climate)*	5	*Smart Materials*
6	*Bio, Stem Cells, Bio Chem, Pharma, Food*	6	*Sustainable Societies*
7	*Semi-Conductors, Optics, Photonics, Networks, Telecomm, Power Eng*	7	*Infrastructure*
8	*HSS, Design, Management*		

The research problem

A common challenge in solving several structured nonlinear discrete optimization problems is the ability of algorithms to efficiently deal with large-scale instances. Many of these problems are well-structured and have approximation algorithms with polynomial running times, but in practice, these algorithms cannot handle large-scale instances, e.g.,

- *Submodular Function Maximization.* Submodular set functions are characterized by the property that incremental values of the function decrease with increasing set size. Submodular functions arise naturally in several applications in Machine Learning, Economics, Game Theory, etc. While Submodular minimization is solvable in polynomial time, submodular function maximization is NP-Hard. Although there is a polynomial-time constant-factor approximation algorithm, several studies have shown that this algorithm does not scale well for large instances.
- *Maximizing strongly log-concave (SLC) set functions.* SLC functions, when normalized, encode discrete distributions with weak log-submodularity properties and certain negative dependence properties. SLC functions can be used in several machine-learning problems that require selecting a diverse set of items. Recently, the first algorithms to approximately maximize these functions (subject to a cardinality constraint) have been developed, based on the greedy algorithm for submodular functions. However, there is significant scope for new algorithmic developments for these optimization problems.
- *Maximum k x k subdeterminant of a PSD matrix.* Given a n x n positive semidefinite matrix, this problem requires the computation of a k x k (where k < n is given) submatrix of the largest determinant. This problem is related to maximum-entropy sampling and can be written as a nonlinear discrete maximization problem with cardinality constraints. Computing optimal solutions for moderate-sized instances is difficult with the current state-of-the-art.

The above problems (and several others) have applications that require solutions of large instances which are not practical at present. Improvements in scalable algorithms for any of these problems would, for instance, allow for use of principled techniques to encourage data-diversity in large scale machine learning systems.

Project aims

The aim of this project is to exploit problem structure and techniques from convex and discrete optimization to develop novel algorithms for these problems that can solve large-scale problem instances. We expect these novel algorithms to have a strong grounding in theory but do not insist on a priori polynomial running time and/or fast convergence rates on all instances as long as the algorithms are practical and return solutions and sub-optimality bounds.

What is expected of the student when at IITB and when at Monash?

Highlight how the project will gain from the students stay at IITB and at Monash

The student is expected to understand the state of the art and improve upon it. At both IIT as well as at Monash, the student will benefit from interaction with supervisors as well as other faculty members and students.

The expectation is that the student will meet weekly with supervisors (online), and will submit (informal) weekly reports on activity (what I've done this week, what I plan to do next week) submitted prior to weekly meetings

Expected outcomes

Highlight the expected outcomes of the project

One or more journal/conference publications in the area(s) of Computational Optimization and/or Machine Learning on novel algorithms (with analysis and convergence proof) that are able to handle large-scale problem instances that are currently intractable.

One or more journal/conference publications demonstrating the computational advantages of the proposed algorithms on practical problems/datasets.

How will the project address the Goals of the above Themes?

Describe how the project will address the goals of one or more of the 6 Themes listed above.

Directly related to the theme of artificial intelligence and optimization models.

Potential RPCs from IITB and Monash

Provide names of the potential research progress committee members (RPCs) and describe why they are most suited for the proposed project

Avinash Bhardwaj (IITB) or Ashutosh Mahajan (IITB). Both of them work on theoretical and computational aspects of optimization problems and are experts in the area.
Mehrtash Harandi (Monash) or Pierre Le Bodic (Monash). Both work on complementary aspects of computational optimization and significant expertise in AI applications.

Capabilities and Degrees Required

List the ideal set of capabilities that a student should have for this project. Feel free to be as specific or as general as you like. These capabilities will be input into the online application form and students who opt for this project will be required to show that they can demonstrate these capabilities.

The student should have a strong background in Linear Algebra, Real Analysis, basic Probability, and some ideas in Optimization. They should be comfortable in dealing with mathematical proofs and should have programming skills.

Necessary Courses

Name three tentative courses relevant to the project that the student should complete during his/her coursework at IITB (the student will require to secure 8 point in these courses)

Optimization Techniques, Probability and Stochastic Processes.

Potential Collaborators

Please visit the IITB website www.iitb.ac.in OR Monash Website www.monash.edu to highlight some potential collaborators that would be best suited for the area of research you are intending to float.

We have already found collaborators: Vishnu Narayanan (IITB) and James Saunderson (Monash)

Select up to **(4)** keywords from the Academy's approved keyword list **(available at http://www.iitbmonash.org/becoming-a-research-supervisor/)** relating to this project to make it easier for the students to apply.

Data Science, Optimization, Algorithms (6), Maths (8)