# WORKSHOP ON

# STOCHASTIC OPTIMIZATION IN

# NETWORKS AND RELATED TOPICS

**Venue:** Hall 23, Victor Menezes Convention Centre, IIT Bombay, Mumbai

**Organizing committee:** Vivek Borkar, Bruce Hajek, Abhay Karandikar, Nikhil Karamchandani

**Speakers:**

Aditya Gopalan (ECE, IISc)
Avhishek Chattterjee (EE, IIT Madras)
Chandra Murthy (ECE, IISc)
Krishna Jagannathan (EE, IIT Madras)
Piyush Srivastava (STCS, TIFR)
Rajesh Sundaresan (ECE, IISc)
Amitabha Bagchi (CSE, IIT Delhi)
Bruce Hajek (ECE, UIUC)
Ketan Rajawat (EE, IIT Kanpur)
Nikhil Karamchandani (EE, IITB)
Praneeth Netrapalli (Microsoft Research, Bengaluru)

# SCHEDULE

| 23rd Feb. (Friday) | | 24th Feb. (Saturday) | |
|---|---|---|---|
| Time | | Time | |
| 0830-0915 | Registration | | |
| 0915-0930 | Inauguration | | |
| 0930-1020 | Bruce Hajek | 0930-1020 | Krishna Jagannathan |
| 1020-1110 | Rajesh Sundaresan | 1020-1110 | Amitabha Bagchi |
| 1110-1130 | Tea | 1110-1130 | Tea |
| 1130-1220 | Praneeth Netrapalli | 1130-1220 | Ketan Rajawat |
| 1220-1310 | Nikhil Karamchandani | 1220-1310 | Avhishek Chatterjee |
| 1310-1430 | Lunch | 1310-1430 | Lunch |
| 1430-1600 | C. Murthy, S. Khanna (Tutorial) | 1430-1520 | Aditya Gopalan |
| 1600-1620 | Tea | 1520-1610 | Piyush Srivastava |
| 1620-1750 | C. Murthy, S. Khanna (Tutorial) | 1610-1630 | Tea |
| | | 1630-1700 | Short concluding session |

# ABSTRACTS

## Tutorial

*The surprising effectiveness of hierarchical Bayesian methods sparse signal recovery*

Prof. Chandra Murthy, Saurabh Khanna (ECE, IISc)

*Abstract:* Hierarchical Bayesian methods have been empirically shown to significantly outperform their more conventional convex relaxation and greedy optimization based counterparts in several scenarios. In this talk, we will focus on the multiple measurement model, and discuss fundamental results that explain the surprising effectiveness of hierarchical Bayesian methods for sparse signal recovery. Inspired by the results, we discuss several new algorithms based on the principle underlying the hierarchical Bayesian methods, namely, that of covariance matching. These algorithms offer multi-fold improvement in performance and use far lower working memory, without compromising on the performance. Time permitting, we will discuss some applications in wireless communications.

## Talks

1. *Black-box Optimization*

   Prof. Aditya Gopalan (ECE, IISc)

   *Abstract:* How would you maximize an unknown function if you could only query it at chosen inputs and observe the corresponding (noisy) function values? The talk will provide a gentle introduction to the Bayesian optimization paradigm for solving this widely encountered problem. We will also present algorithms for multi-armed bandit problems with continuous action spaces, based on nonparametric Gaussian process models for inference and decision-making under uncertainty (joint work with Sayak Ray Chowdhury).

2. *Decentralized random walk-based data collection in networks*

   Prof. Amitabha Bagchi (CSE, IIT Delhi)

   *Abstract:* We analyze a decentralized random walk-based algorithm for data collection at the sink in a multi-hop sensor network. Our algorithm, Random-Collect, which involves data packets being passed to random neighbors in the network according to a random walk mechanism, requires no configuration and incurs no routing overhead. To analyze this method, we model the data generation process as independent Bernoulli arrivals at the source nodes. We analyze both latency and throughput in this setting, providing a theoretical lower bound for the throughput and a theoretical upper bound for the latency. The main contribution of our paper, however, is the throughput result: we show that the rate at which our algorithm can collect data depends on the spectral gap of the given random walk's transition matrix. In particular, for the simple random walk, we show that the rate also depends on the maximum and minimum degrees of the graph modelling the network. For latency, we show that the time taken to collect data not only depends on the worst-case hitting time of the given random walk, but also depends on the data arrival rate. In fact, our latency bound reflects the data rate-latency tradeoff, i.e., in order to achieve a higher data rate we need to compromise on latency and vice-versa. We also discuss some examples that demonstrate that our lower bound on the data rate is optimal up to constant factors, i.e., there exists a network topology and sink placement for which the maximum stable data rate is just a constant factor above our lower bound.

3. *Energy-Reliability Limits in Nanoscale Boolean Trees and Feedforward Neural Networks*

   Prof. Avhishek Chattterjee (EE, IIT Madras)

   *Abstract:* Nanoscale semiconductor devices are often unreliable, with reliability as a function of energy. We study energy-reliability limits for boolean tree circuits and deep feedforward neural networks (multilayer perceptrons) built using such devices. Pippenger had developed a mutual information propagation technique to characterize the complexity of noisy circuits. Since small circuit complexity need not imply low energy, however, we extend mutual information propagation to obtain energy lower bounds for boolean tree-structured circuits and for deep feedforward neural networks. Many device technologies require all gates to have the same electrical operating point; in circuits of such uniform gates, we show that the minimum energy required to achieve any non-trivial reliability scales superlinearly with the number of inputs. Circuits implemented in emerging device technologies like spin electronics can, however, have gates operate at different electrical points; in circuits of such heterogenous gates, we show energy scaling can be linear in the number of inputs. Building on our extended mutual information propagation technique and using insights from convex optimization theory, we develop an algorithm to compute energy lower bounds for any given boolean tree under heterogeneous gates. This algorithm runs in linear time in number of gates and is therefore useful for designing modern circuits with numerous gates. As part of our development we find a simple procedure for energy allocation across circuit gates with different operating points and neural networks with differently-operating layers.)

4. *On the challenge of gene regulatory network reconstruction from high throughput sequencing data*

   Prof. Bruce Hajek (ECE, UIUC)

*Abstract:* Gene regulatory networks are central to the functioning of biological organisms. High throughput genetic sequencing technology has enabled the production of massive amounts of data pertaining to activation levels of genes, which has the potential to help scientists reverse engineer gene regulatory networks. In this talk I will briefly explain the steps from data collection to network modelling and analysis that go into the process for a particular study examining soybean plants.

In collaboration with Xiaohan Kang, Faqiang Wu, and Yoshie Hanzawa.

5. *Qubits through Queues: The capacity of channels with waiting time dependent errors*

Prof. Krishna Jagannathan (EE, IIT Madras)

*Abstract:* We consider a setting where qubits are processed sequentially, and derive fundamental limits on the rate at which classical information can be transmitted using quantum states that decohere in time. Specifically, we model the sequential processing of qubits using a single-server queue, and derive explicit expressions for the capacity of such a 'queue-channel'. We also demonstrate a sweet-spot phenomenon with respect to the arrival rate to the queue, i.e., we show that there exists a value of the arrival rate of the qubits at which the rate of information transmission (in bits/sec) through the queue-channel is maximized. Next, we consider a setting where the average rate of processing qubits is fixed, and show that the capacity of the queue-channel is maximized when the processing time is deterministic. We also discuss design implications of these results, and conclude by highlighting the importance of explicitly modelling delay induced errors in quantum information processing systems.

Collaborative work with Avhishek Chatterjee (Dept. of Electrical Engineering, IIT Madras) and Prabha Mandayam (Dept. of Physics, IIT Madras).

6. *Beyond Consensus and Synchrony in Decentralized Online Optimization using Saddle Point Method*

   Prof. Ketan Rajawat (EE, IIT Kanpur)

   *Abstract:* We consider online learning problems in multi-agent systems comprised of distinct subsets of agents operating without a common time-scale. Each individual in the network is charged with minimizing the global regret, which is a sum of the instantaneous sub-optimality of each agent's actions with respect to a fixed global clairvoyant actor with access to all costs across the network for all time up to a time-horizon $T$. Since agents are not assumed to be of the same type, the hypothesis that all agents seek a common action is violated, and thus we introduce instead a notion of network discrepancy as a measure of how well agents coordinate their behavior while retaining distinct local behavior. Moreover, agents are not assumed to receive the sequentially arriving costs on a common time index, and thus seek to learn in an asynchronous manner. A variant of the Arrow- Hurwicz saddle point algorithm is proposed to control the growth of global regret and network discrepancy. This algorithm uses Lagrange multipliers to penalize the discrepancies between agents and leads to an implementation that relies on local operations and exchange of variables between neighbors. Decisions made with this method lead to regret whose order is $O(T^{1/2})$ and network discrepancy $O(T^{3/4})$. Empirical evaluation is conducted on an asynchronously operating sensor network estimating a spatially correlated random field. An application to vision based target localization with moving cameras demonstrates the benefits of this approach in practice.

7. *Rumor source identification in networks (or "Who let the dogs out, who, who..?")*

   Prof. Nikhil Karamchandani (EE, IITB)

   *Abstract:* Consider a network where an unidentified source node starts a rumor, which then propagates along the edges of the underlying graph. After sufficient time has elapsed, we get to observe the infected sub-graph and in addition, perhaps some partial information about the infection episodes. Using this information, we would like to design an estimator for the source node to identify the starting point of the rumor. Such an estimator will have applications in several domains, ranging from identifying the source of malware spread in computer networks, to locating the fault which started a cascade of events leading to a power blackout or a financial meltdown, or identifying the starting point of an epidemic.

This problem of 'rumor source identification' has received significant attention recently, several different variants have been studied and various estimators have been proposed. We will begin with a short survey of some of the results and then discuss two recent directions we have pursued in this space:

(a) Persistence of estimators: Does the estimated source change as we wait longer and longer to observe larger and larger infected subgraphs, or does it stabilize after some time?

(b) Pairwise information: We consider the case where we have partial pairwise information about the infection episode, such as A infected B or A was infected before B, and devise efficient estimators for this setting.

8. *Zeros of polynomials and approximation of partition functions*

   Prof. Piyush Srivastava (STCS, TIFR)

   *Abstract:* This talk presents a recent method proposed by Alexander Barvinok which uses the analyticity of the free energy in the complex plane to design algorithms for the approximation of partition functions. We will then look at connections of this work with the classical Lee-Yang theory to give new algorithms for the approximation of the ferromagnetic Ising partition functions in settings where rigorous analyses of algorithms based on traditional ideas such as decay of correlations or Markov chain Monte Carlo are not yet available.

   Based on (partly ongoing) joint work with Jingcheng Liu and Alistair Sinclair

9. *Second moment method and applications*

Dr. Praneeth Netrapalli (Microsoft Research, Bengaluru)

*Abstract:* Consider a hypothesis testing problem where we wish to distinguish whether a random object is being sampled from distribution P or distribution Q. Our ability to answer this question depends on whether P and Q are mutually contiguous with respect to each other or not. Second moment method is a powerful tool that has been applied to show mutual contiguity of two measures for various settings. In this talk, we will introduce the second moment method and illustrate an application to recovery of stochastic block models.

10. *A double-auction mechanism for mobile data-offloading markets with strategic agents*

Prof. Rajesh Sundaresan (ECE, IISc)

*Abstract:* We consider a double-auction mechanism for mobile data-offloading. Network operators (users) derive benefit from offloading their traffic to third party WiFi or femtocell network (link-supplier). A link-supplier experiences costs for the additional capacity that he provides. Users and link-supplier have their utilities and cost function as private knowledge. A system designer decomposes the problem into a network problem (with surrogate utilities and surrogate cost functions) and agent problems (one per agent). The surrogate utilities and cost functions are modulated by the agents' bids. Agents' payoffs and costs are then determined by the allocations and prices set by the system designer.

It is already known that, if the agents do not anticipate the effect of their actions, a competitive equilibrium exists as a solution to the network and agent problems, and this equilibrium optimizes system utility. However, we will show that, if the agents are strategic (and thus price-anticipating), the presence of strategic supplying agents drives the system to an undesirable equilibrium with zero participation. This

is in stark contrast to the setting when link-supplier is not strategic where the efficiency loss is at most 34%.

We will then propose a Stackelberg-type modification to the mechanism to alleviate the efficiency loss problem. The system designer first announces the allocation and payment functions. He then invites the supplying agent to announce his bid. He then invites the users to respond to the suppliers' bids. The resulting efficiency loss is characterized in terms of the supplier's cost function, and is significantly reduced to at most 25% for the quadratic cost function.