Research Statement
Parikshit Shah

Overview
Large-scale systems in modern engineering and science present an array of challenging problems involving modeling, identification, analysis and design. Prototypical examples include satellite arrays, vehicle platoons, supply chains, smart power grids, the internet, sensor networks, financial markets, gene regulatory networks, geophysical systems, chemical reaction networks and social networks. The associated computational obstacles range from collective decision-making tasks involving a large number of agents (decentralized control) to formidable model selection and system identification tasks where one may wish to select “simple” models that succinctly describe the relationships between the different components of a large system. Often these questions become more complex due to the presence of dynamics and the time-varying nature of the physical system. My research focuses on developing convex optimization techniques as a means of tackling these computational challenges. A central theme in my research is the use of algebraic structure which provides an expressive modeling framework in a wide array of problem domains, and a rich set of tools that enable a bridge to tractable computation.

Posets in systems theory: Many decision making problems such as leader-follower vehicle platoons and supply chains are hierarchical. I introduce partially ordered sets (or posets) as a new conceptual framework to reason about hierarchical flow of information. This paradigm enables novel ways to exploit the algebraic structure of posets for decision making problems. The algebraic approach provides insight on new control architectures that are intuitive, provably optimal and amenable to efficient computation via convex optimization. Fundamental to the information processing in this architecture are natural notions of integration and differentiation on posets and connections to well-known combinatorial concepts such as Möbius inversion.

Covariance structure with symmetries: Many physical processes possess an algebraic property known as symmetry, which can be studied mathematically using the notion of group invariance. Special cases of this algebraic framework include widely used statistical concepts such as partial exchangeability and stationarity. Inferring covariance properties associated to these processes is an important problem in statistics, engineering and science. In this work I study learning and regularization of covariance matrices that are invariant under the action of a group. My work clarifies how dramatic computational and statistical benefits can be obtained when symmetry is exploited judiciously.

Optimization and statistics in signal/system identification: Model selection in large-scale systems often requires new approaches to deal with the curse of dimensionality. Tools from statistics and optimization provide a powerful framework to deal with such large-scale problems. A representative problem that captures the essence of a much broader class of problems that arise in practice, is the identification of an unknown signal on a general manifold. In [8] I have developed an efficient computational technique with sound statistical properties. In ongoing work, I am investigating optimization based methods for identification of large-scale linear dynamical systems with provable statistical guarantees.

Applications of these methods arise in many problem domains. The insight gained by using posets and the associated algebraic properties to understand hierarchical decision-making can be useful in the context of computing optimal formation controllers in robotics, in inventory management in supply chains and design aspects of complex aircraft systems. My work on exploiting symmetry in large covariance matrices can shed new light on model selection and inferring statistical properties of random fields.
that arise in domains such as oceanography. Convex optimization and algebraic structure will continue to play important conceptual roles in the information and decision sciences; this will be a major theme of my future research. In what follows, I describe how these ideas have already led to exciting new developments, and outline promising future directions.

1 Main Contributions

1.1 Partially Ordered Sets and Decentralized Control

Decision-making in large-scale dynamic systems is a problem of importance which impacts several application domains ranging from congestion management in networks to power grids. For the most general formulations of this problem, obtaining exact solutions is computationally intractable [2]. Seemingly disparate instances have been known to be tractable but a rigorous theory that provides both a unifying viewpoint as well as efficient computational tools and design principles remains a major challenge.

My thesis work focuses on developing a framework to address these challenges using ideas from combinatorics, algebra, and convex optimization. At a high level, it identifies hierarchical decision-making as a tractable class of problems that unifies many disparate instances. Understanding hierarchical decision-making conceptually is important because it arises in numerous applications such as flight-control systems, leader-follower vehicle platoons and the operations of supply chains.

Drawing motivation from trees as a tractable class of objects, I study directed acyclic graphs as a model for hierarchical information flow in systems. The hierarchical structure is modeled by a combinatorial object, called a partially ordered set (or poset), and the resulting system is called a poset-causal system. Posets possess rich algebraic structure such as the incidence algebra [5], which is a collection of matrices that captures the relations of the poset (analogous to the adjacency matrix of a graph). Other algebraic objects that play a distinguished role are the Zeta and the Möbius operators [5]. These objects are generalizations of integration and differentiation on posets that arise in a variety of mathematical contexts, such as the inclusion-exclusion principle and various other problems in combinatorics. Our work on understanding hierarchical control has led to intriguing connections with these algebraic properties of posets.

In joint work with Pablo Parrilo, I have shown deep connections between control over poset-causal systems and these algebraic properties. For example, the incidence algebra can be exploited to exactly reparametrize the (originally nonconvex) problem of optimal decentralized control of a poset-causal system to a convex optimization problem. While these problems are challenging due to their infinite-dimensional nature, I have developed efficient computational techniques to solve them [9]. The optimal controllers also possess intriguing algebraic structure. I have shown that a general architecture that uses the aforementioned Zeta and Möbius operators is in fact optimal [10] for a large class of hierarchical decision-making problems.

Understanding the role of the Möbius operator is significant, because it allows one to quantify the amount of additional information of a subsystem relative to its local neighborhood. While the architecture I have proposed is optimal for hierarchical problems, I anticipate that this approach provides a novel and sound design principle that is applicable well beyond this setting.

1.2 Covariance Structure with Symmetries

An important feature of many modern data analysis problems is the small number of samples available relative to the dimension of the data. Such high-dimensional settings arise in a range of applications in bioinformatics, climate studies, and economics. A fundamental problem that arises in the high-
Figure 1: (a) A poset and the Möbius operation. Posets generalize the notion of causality via a branched notion of time, as indicated by the downward arrows. The function \( f \) is to be thought of as a signal evolving in branched (discrete) time. The nodes of the poset are labeled by the Möbius inverse of \( f \), which is a generalization of differentiation with respect to branched time [10]. (b) An image observed from time-varying camera positions can be abstractly viewed via manifolds. Even though the images themselves are high-dimensional objects, they lie on a low-dimensional submanifold corresponding to the camera locations. (Image courtesy [4]).

dimensional regime is the poor conditioning of sample statistics such as sample covariance matrices [1]. Accordingly, a fruitful and active research agenda over the last few years has been the development of methods for high-dimensional statistical inference and modeling that take into account structure in the underlying model.

In joint work with Venkat Chandrasekaran, I study a notion of algebraic structure known as group symmetry in the context of covariance estimation [7]. Such models naturally arise in settings in which the distribution of a collection of random variables is invariant under certain permutations of the variables. For example, random fields that model physical phenomena such as fluid transport in hydrology, electrostatic fields in transmission lines and ocean circulation in oceanography inherit symmetric statistical properties from their corresponding physical model (often described by a partial differential equation such as the Laplace or Possion equation). Symmetry manifests itself via the notion of exchangeability in Bayesian statistics, and in electrical engineering one often encounters it in the form of cyclostationary processes that describe the behaviour of stochastic phenomena with periodic variations [7].

I systematically investigate the statistical and computational benefits of exploiting symmetry in high-dimensional covariance estimation. Two basic algebraic concepts play important roles in our setup. The first is the notion of a group orbit, which identifies equivalence classes of random variables, and parametrically captures the statistical gains that symmetry allows. The second notion, a fixed point subspace, is the collection of covariance matrices which possess the specific group symmetry, and this notion allows computational techniques to exploit symmetry. A closely related concept is that of Reynolds averaging which ties the two concepts together.

I am able to quantify, in terms of the properties of the underlying group, the statistical gains that one may expect when selecting symmetric covariance or inverse covariance matrices. The proposed regularization allows for statistical gains that may be interpreted as an efficient reuse of samples that exploits the symmetry in the problem.

1.3 Optimization and Statistics in Signal and System Identification

Making inferences about an unknown signal given limited and inaccurate information is a problem frequently encountered in machine learning. In many applications of interest, a significant difficulty arises because the number of measurements available about a signal may be far smaller than the ambient
dimension of the signal. Fortunately, many real-world signals have more structure that restricts their degrees of freedom relative to the ambient dimension. Exploiting such low-dimensional structure in the system of interest is crucial in order to enable reliable and statistically good recovery.

The mathematical notion of a manifold provides a general and useful model for low-dimensional but nonlinear geometric structure; many naturally occurring signals (for example, natural images) are known to empirically fit well within this abstraction. In joint work with Venkat Chandrasekaran [8], I exploit the geometry of manifolds to propose an algorithm that provably and robustly recovers an unknown signal on a given manifold from a small number of measurements. A variant of the algorithm provably tracks time-varying signals on the manifold.

Another important application where low-dimensional structure plays an important role is in the identification of linear dynamical systems. This is a well-studied area because the problem is ubiquitous in engineering, with applications ranging from aircraft design, chemical processes to power systems. However, for large-scale systems (for example, chemical reaction networks) a different theory is required to address the challenges related to computation and robustness to noise when only limited data is available. In ongoing work with Benjamin Recht and Badrinarayan Bhaskar, we study the problem of identifying a low McMillan degree system through the lens of high dimensional statistics and convex optimization [6]. Some partial results that we have obtained provide new and fundamental lower bounds on the statistical accuracy obtainable for system identification. Our approach asks new kinds of questions that lie at the interface of control theory, statistics, and convex optimization, and requires ideas from these fields to be combined in unique ways.

1.4 Dynamic Games, Polynomials, and Computation

Dynamic games are a powerful modeling framework that allows one to model competitive scenarios with dynamics, for example, pursuit evasion games. In joint work with Pablo Parrilo, I have studied the problem of computing equilibria in games [11] with players having infinitely many pure strategies and polynomial payoffs. Drawing on techniques from real algebraic geometry, together with properties of moment spaces of probability measures and their connections to semidefinite programming, I have developed efficient convex optimization based techniques to compute equilibria. This work clarifies the expressive power of semidefinite programming as a computational framework for solving problems related to decision-making, dynamics, and games.

2 Future Research

My future work will revolve around the themes of convex optimization, notions of structure (such as algebraic structure) in signals and systems, and their application to problems in the information sciences including machine learning, statistics and systems theory.

**Posets in the information sciences:** Much of my past work focuses on exploiting the algebraic and combinatorial properties of posets to understand hierarchical control. Similar ideas can have impact for hierarchical decision-making problems arising in operations research, for example in the study of supply chains. More broadly, posets can play an important role in many other problems in information science.

One such data analysis problem that has gained prominence is that of learning user preferences. An active area of research in machine learning focuses on inference of user preferences from extremely limited data. This problem arises in the context of ranking relevancy of a large number of candidate webpages in internet search, or ranking user preferences from a large collection of objects (such as movies, songs, restaurants, etc.) in recommendation systems. Abstractly this can be phrased as constructing a global ranking of a large number of objects from limited comparisons. In many real-world scenarios, some objects may be incomparable (for example, movies corresponding to different genres), and posets are a natural representational concept in this context. I anticipate that posets and the associated machinery
that I have explored in my past work can lead to new insights on this problem.

Another area where posets can play a role is in probabilistic inference in large-scale statistical models. Probabilistic inference in such models, which arise in a wide array of applications ranging from statistical physics to decoding in communication problems, is a hard problem. A major line of research focuses on understanding message passing algorithms to perform approximate inference. One promising direction proposed by McEliece [3] is to organize the computation needed for inference using posets. Concepts such as Möbius inversion [5] in the context of message passing can lead to new techniques for inference.

**Approximation and Computation in Large Scale Dynamic Systems:** A key notion in many areas of engineering and computational mathematics is the notion of approximation for hard problems. Such notions have been studied in many varied settings, for example scaling laws for large-scale communication systems, in high dimensional settings in statistics, and in the study of approximation algorithms in computer science. Can one develop a principled approach to reason about approximation for networks of dynamical systems?

- Given a large scale dynamic system with an underlying network structure, can one identify the structure of the network from a small number of observations? In statistics and machine learning, such problems are often solved using convex relaxations. This would lead to new and exciting connections with the literature on graphical model selection in machine learning and statistics, convex optimization, and system identification in systems theory.

- In graph theory, there is a well-understood notion of approximation of a graph by a simpler graph using the notion of a Laplacian. Is there an analogous notion of approximation of a network of dynamic systems? This agenda would lead to new and interesting connections between model reduction in systems theory and graph theory.

- The study of large random graphs is of fundamental importance in statistical physics. Indeed in certain regimes such graphs are locally tree-like, this observation can be exploited to solve a variety of computational problems. Can one develop a theory for random networks of dynamical systems based on them being locally poset-like? Can this be used for control design?

- In decentralized control, we now understand classes of tractable problems that are amenable to exact solutions and also classes of intractable problems. I plan to commence a study of computing approximately optimal controllers. A related problem is that of understanding fundamental limits (i.e. performance lower bounds) in decentralized control, which would enable us quantify the quality of approximate solutions. In statistics, there are well understood techniques to provide lower bounds on statistical performance via an information-theoretic tool known as Fano’s inequality. A fundamental bridge between control and estimation is the notion of control-estimation duality. One intriguing possibility is to use this bridge to phrase control problems as estimation ones and use statistical techniques to provide performance lower bounds.

- While many of the above problems may be amenable to expressive optimization frameworks such as semidefinite programming, solving the resulting large-scale problems will require new advances in algorithms for optimization, which will be an important component of my future research.

Answering questions related to approximation and model selection for networks of dynamical systems will bring new insights to domains such as chemical reaction networks and power networks. Chemical reaction networks deal with the study of interactions of a large number of chemical compounds, the interactions being modeled by a network. Notions of approximate model selection will allow us to understand the network structure from time-varying and noisy data, and shed insight on the structure
of chemical interactions. Another application domain involves the control of power networks and smart grids. These applications pose difficult control problems due to complex network structure that current control theory fails to address adequately. Approximating these large scale networks will enable design of approximately optimal controllers thus leading to a principled approach to the problem.

References


