# CDS&E: Collaborative Research: Machine Learning for Automated Discovery and Control in Turbulent Plasma

**Barnabas Poczos, Jeff Schneider,**
**Axel Brandenburg, Tina Kahniashvili, Alexander Tevzadze, and Hy Trac**

## 1 Introduction

Executions of large-scale simulations that generate enormous amounts of data are now ubiquitous across the sciences. Low-level differential equations are used to drive the simulations and scientists hope to see and understand the larger scale phenomena that arise from them. Traditionally, simulations would be visualized by scientists or summarized by a few simple statistics. As these data sets increase in size, however, scientists are increasingly unable to easily answer even basic questions: Did anything interesting or unusual happen in the simulation? When and where did it happen? What are the most common phenomena comprising the simulation? What are the spatial and temporal distributions of those phenomena? Do they interact with each other? How can we find interesting relationships among different parts of the data? What quantities are conserved in these the experiments?

Recent advances in machine learning (ML) open a new avenue to help scientists make new discoveries in complex datasets. In this work, *we propose new ML methods to help domain scientists understand complicated properties of turbulent plasma.* In particular we will propose data processing methods for i) turbulent simulated plasma (including atmospheric turbulence), ii) solar wind data, and iii) fusion data collected from a tokamak reactor. The unique role of our proposed work is to rigorously and systematically study and develop effective methodologies for data processing in turbulent plasma, and deploy these innovative methods and tools to help the daily practice of computational physicists and enable them to make scientific discoveries faster, easier, and cheaper.

Understanding the nature of turbulence, its development and evolution (including transition, damping, decay and amplification mechanisms) is one of the key questions of fundamental physics. Turbulence, independently of medium, is always associated with chaotic changes of variables (although not all chaotic processes are turbulent). Indeed, the solution of Navier-Stoke equation - equation that governs the simplest form of hydrodynamic turbulent motions remains elusive, despite being declared as one of the Clay Millennium Problems*. The extreme sensitivity of the solution on the initial and boundary conditions makes turbulent motions chaotic, both in time and space. Hence, the turbulent processes have to be treated statistically, rather than deterministically. Turbulence occurs in different media - ranging from cosmic plasma (at extremely large scales order of Mpc) till blood flow in human heart. There are several common features that characterize turbulent processes including: i) random fluctuations of physical variables such as velocity, pressure, density, etc; ii) the rotationality of motions - unavoidable presence of kinematic vorticity; iii) diffusivity - a tendency to mixing media (e.g. *homogenization*) due to eddy motions; iv) dissipation (e.g. damping) - transfer of turbulent energy through viscous process to internal thermal energy. Turbulence produces variety of eddies (in size), while most of the energy is contained in the large-scale eddies: the size of these structures sets the integral scale of turbulence. In standard forward cascade of turbulence the energy of chaotic fluctuations is transferred from large to small scales producing smaller and smaller structures (eddies) down to the viscous scales where dissipative effects become dominant (scale called the Kolmogorov scale) and energy of turbulent fluctuations dissipates into heat. The statistical theory of turbulence proposed by Kolmogorov [1,2] is based on the concept of the energy cascade (originally discussed by Richardson [3], and self-similarity hypothesis. The former assumption has been revisited (see Ref. [4] for a review) leading to reconsideration of turbulence statistical description [5], and the nature of turbulence still remains one of unsolved problems of physics. Some open questions include: i) turbulence non scale-invariance and related issues with rescaling; ii) turbulence non-locality and understanding of the local versus non-local contributions to the magnetohydrodynamic (MHD) turbulence; iii) inverse transfer of energy and difference between

*http://www.claymath.org/millennium-problems non-helical inverse transfer and helical inverse cascade; iv) the nature of large-scale structures and identification of helical structures; v) generation/destruction of helical structures, and helicity amplification/damping in turbulence.

The statistical properties are investigated through different data sets that can be divided into three major classes: i) numerically simulated data (numerical experiments); ii) observational data (for example, solar wind); iii) laboratory experiment data (controlled plasma turbulence). We will make use of the data obtained through direct numerical simulations (DNS) of hydrodynamical (HD) and magnetohydrodynamical (MHD) turbulence, as well as atmospheric turbulence. In particular, we will use ML to analyze decaying turbulence properties, in order to determine the major classes of turbulence. Being independent of ill-defined forcing mechanisms, decaying turbulence has a better chance of displaying generic properties of turbulence. Important applications include grid turbulence [6], turbulent wakes [7], atmospheric turbulence [8], as well as interstellar turbulence [9], galaxy clusters [10], and the early Universe [11].

Analysis of DNS data will be carried out in order to clarify the nature of inverse transfer in non-helical turbulence, properties of partially helical turbulence and specific features of the fully helical turbulence decay. We will identify intermittent and persistent features of turbulence including coherent structures and other anomalies observed in numerical simulations. This will include analysis of non-Kolmogorov component in the turbulence spectrum, and unusual spectral behavior often observed at intermediate scales in MHD turbulence. We also aim to determine the universal laws for turbulence, and make classifications of turbulence classes based on empirically determined conserved quantities. DNS data will be used to find out which invariants of turbulent flows are conserved, what are the specific reasons associated with the violation of these invariants or identify new conservative quantities in a decaying turbulence.

In addition to simulated turbulent data, we will also test our novel data processing methods in important real world problems including solar wind turbulence, and tokamak fusion reactor data.

The solar wind is a flow emanating from the solar corona. It is highly turbulent, covers huge time and length scales, and can be extremely bursty owing to solar flares and coronal mass ejections that send hazardous radiation toward the Earth. Today, numerous satellites monitor routinely the flow speeds and the full magnetic field vector at high time resolution, making these recordings an invaluable tool for studies of MHD turbulence.

Solar wind turbulence is studied extensively with regards to spectral properties both in the inertial range [12] as well as in the dissipation range beyond the ion cyclotron frequency [13]. Even in the middle of the turbulence spectrum, i.e., within the inertial range, the magnetic and kinetic energies spectra have been found to be non-parallel to each other, contrary to what was expected. At large wavenumbers, however, the magnetic and kinetic energy spectra are expected to converge [14]. The strong dominance of magnetic over kinetic energy at larger scales has been identified as being the main course for different spectral indices [15].

The question of magnetic helicity has received particular attention [16] and its understanding is critical for solar dynamo theory [17]. The available data are sufficiently rich to apply ML algorithms to learn about the properties of coherent structures at different locations in the solar wind.

As another real world application, we will demonstrate how our new ML methods can be applied to control plasma in tokamak fusion reactors. The development of fusion power is the most promising long-term path to eliminating our dependence on carbon-emitting energy sources. Its combination of limitless fuel, minimal land use, and lack of radiation safety issues make it superior to the alternatives.

The most successful reactors (tokamaks) use magnetic confinement of plasmas in the shape of a torus. Control of these devices is extremely challenging due to the nonlinear physics involved. At the heart of the problem are the equations of magnetohydrodynamics (MHD) governing the state of the plasma. These result from interlinking Navier-Stokes equations of fluid dynamics and Maxwell's equations of electromagnetism. The nonlinearities yield many instabilities that must be controlled to achieve the temperatures and pressures necessary for fusion.

Considerable progress has been achieved through a significant investment in understanding the physics of fusion and magnetic confinement from first principles. That understanding has been used to design systems that are "easy" to control in the sense that "proportional-integral-derivative" (PID) controller, or other simple, linear controllers are used. When those controllers do not perform well, extensive experimental effort is spent to improve their operation. In spite of the progress, significant unsolved problems remain in the way of achieving the sustained temperatures and pressures required.

At the same time physicists were advancing their understanding of controlled fusion, ML researchers were making gains in their ability to learn models of complex systems from data and control them with learned nonlinear controllers. In this project we will develop new ML algorithms to solve currently outstanding problems in controlled fusion.

As a short summary of this introduction, here we list the key aims of our proposal:

**Aim 1: Method Development and Theoretical Contributions.** We will develop new data driven learning methods for processing large scale turbulent data. These methods will enable domain scientists to i) find known phenomena, regular patterns, and structure (e.g vortices) in turbulent data; Find special patterns that are too complex to parameterize with simple parameters and too burdensome to manually search for. ii) detect anomalies and interesting rare events in turbulent data. iii) Help scientist find new unknown patterns or scientific laws that are too complicated to be manually found, iv) develop new data driven control methods for turbulent plasma.

**Aim 2: Answering Important Open Science Questions.** The new learning methods will be deployed to answer important open questions in i) simulated turbulence data (including atmospheric turbulence), ii) real solar wind, and iii) control plasma in tokamak fusion reactors.

The technical details and the novel computational physics and ML contributions of Aims 1 and 2 will be provided in Sections 2 and 3, respectively. Given the ambitious goals of this project, we anticipate that it will take 3 years to complete the objectives we have outlined. The deliverables of this project will be implementations of the algorithms under a BSD open source license, corresponding documentation and APIs, analytical results on benchmark datasets, academic publications in scientific journals, and conference and workshop proceedings.

**Broad Applications.** In this proposal, our applications will focus on the above listed applications. This research, however, also has the potential to revolutionize our understanding of turbulence, including development of stochastic fluid and magnetized plasma motions, establishment of the stationary turbulence, evolution (decay, damping and amplification) of hydro and MHD turbulence. The results of our research might lead to improved numerical simulation techniques. This project will also contribute to an increased interaction between the ML, plasma physics, and nuclear fusion communities. The students participating in this project will have an interdisciplinary expertise in ML, plasma physics, and nuclear fusion.

**The Team.** Our team members are qualified researchers in ML (Drs. Póczos, Schneider) and physics (Drs. Brandenburg, Kahniashvili, Tevzadze, and Trac).

The PI, Dr. Barnabás Póczos, is a leading expert in statistical ML and will head the statistics and ML research effort. He has successfully applied his ML expertise to neuroscience [18], cosmology [19–22], computer vision [23, 24], civil engineering [25], and other domain sciences [26].

Co-PI Dr. Jeff Schneider is a prominent ML researcher. His past contributions in active learning, reinforcement learning, anomaly detection, learning control, and kernel methods, are all relevant to this proposal. He is well known as a researcher who not only publishes algorithms, but also makes ML work on real systems to get new results both scientifically and commercially. Those efforts include his results in astrophysics [27–29], his development and commercialization of a new ML driven *in vivo* CNS drug discovery system while he was the CIO of Psychogenics, Inc. and his work on the first live demonstration of a learning controller for an autonomous helicopter [30].

Co-PI/Institutional PI Dr. Axel Brandenburg is known mostly for his work on astrophysical turbulence with applications to dynamo theory [31], in particular the solar dynamo [32], galactic magnetic fields [33], accretion disk dynamos [34], and the early Universe [35]. Together with D. Dobler, he developed in 2001 the PENCIL CODE [36] as a community project in public domain (https://github.com/pencil-code). He has published over 300 papers in refereed journals and an $h$ index of 60 on Google Scholar.

Co-PI, Dr. Tina Kahniashvili is mostly known for her work on cosmological magnetic field evolution modeling throughout of the universe expansion [37–41] and its effects on the cosmic microwave background [42, 43], including magnetic helicity effects, [44, 45], and large scale structure [46]. She was the first to study circular polarization of relic gravitational waves from parity violating sources at cosmological phase transitions [47–49].

Collaborator, Dr. Alexander Tevzadze is mostly known for his work on astrophysical fluid dynamics, numerical simulations of the accretion disks of compact objects and processes in protoplanetary discs, as well as on MHD processes in the universe. In particular he studied different instabilities in astrophysical discs [50–52]. He has worked on the magnetic field signatures on the CMB polarization [53] and on the cosmological phase transitions produced magnetic field decay through the evolution of the universe [38–41].

Co-PI Dr. Hy Trac works on theoretical and computational astrophysics and cosmology and has developed and applied N-body, hydrodynamic, and radiative transfer codes to simulate structure formation and evolution [54–56]. He also collaborates with ML experts and statisticians to apply modern approaches to improve multi-wavelength data analysis and numerical simulations [21, 22, 29].

Crossing multiple disciplines (plasma physics, fluid dynamics, hydro and MHD turbulence, simulations, ML, statistics), our proposal describes a uniquely integrative research program that constitutes a crucial first step to automated discovery and control in turbulent plasma datasets. The team members have a successful track record working together, developing innovative methods, providing software packages to the public for free, and publishing new results in top journals and conferences in ML, and physics. Proposers have collaborated fruitfully previously on the MHD turbulence development and evolution modeling through 3D direct numerical simulations using the PENCIL CODE, [38–41, 53, 57] and constraining primordial magnetism [37, 46, 58]. They also successfully applied ML methods for finding anomalies [59] and learning complex scientific laws in astrophysical data [21, 22, 29].

# 2 Data Driven ML Methods

One of the challenges of doing science with modern large-scale simulations is identifying interesting phenomena in the results, finding them, and computing basic statistics about when and where they occurred. Below we will propose ML methods that can help domain scientists i) detect known interesting events and structures, (e.g. Lagrangian Coherent Structures, vortices), ii) find interesting rare anomalies, iii) control turbulence data with new nonlinear control methods, iv) find interesting relationships among different parts of the data and learn new scientific laws, e.g. identify quantities that are conserved in the experiments, statistically dependent, or have a complex nonlinear relationship among them.

## 2.1 Detecting Known Structures in Turbulent Data

In many scientific problems, domain scientists have a collected set of labeled examples and their goal is to find similar examples in the rest of the data. For example, when our goal is to find vortices or other know structures in a turbulent dataset, we might already have a labeled collection of vortices and another labeled set of examples not containing vortices. Our goal is to develop a detection algorithm that can learn the similarities and dissimilarities between these classes and using this acquired knowledge detect more vortices (or other structures) in the rest of the dataset.

This problem belongs to the family of classification problems. There are numerous algorithms proposed for classification problems and they are very well-studied in ML. One crucial challenging issue, however, is that most ML algorithms developed for similar classification problems can only operate on simple finite dimensional feature vectors, and they cannot be directly applied for our scientific problems where the inputs are results of complicated simulations, vector fields, or N-body particles, but not simple finite dimensional vectors. Recent advances in functional data analysis, nonparametric statistics, and ML, however, recently has started to study the question of how to develop learning systems that can process these complex scientific data. This is a very rapidly developing field, with lots of open questions both in theory and practice.

**ML on Functions, Vector Fields, and Distributions.** To develop ML methods that can operate on these complex objects, we need to estimate either the distance, or the inner product between these objects. For example, when dealing with particle sets in physical simulations, we might assume that these particles are sampled from some unknown distributions and we need to estimate a divergence (e.g. Kullback-Leibler or Rényi divergence), inner product, or other density functionals between them, to estimate how these objects are related to each other.

An indirect way to obtain the desired divergence or density functional estimates would be to use a naïve *"plug-in"* estimation scheme: first, apply a consistent density estimator for the underlying densities, and then plug them into the desired formulae. The unknown densities, however, are "nuisance" parameters in our case; we are not interested in them and would prefer to avoid estimating them. Furthermore, density estimators usually have tunable parameters, and we may need expensive cross validation to achieve good performance. Density estimation is among the most difficult problems in statistics, and hence in many cases direct estimators, which do not apply density estimation, can achieve better performance than the "plug-in" methods. The most well-known example is the mean functional ($\int xp(x)dx$), which can be simply estimated with the empirical average ($\frac{1}{n}\sum_{i=1}^{n} x_i$), and usually we do not use sophisticated density estimators for this problem. For more complex functionals such as entropy [60], mutual information [61], and certain divergences [62, 63], empirically it was also observed that direct estimators can perform better than the "plug-in" ones. It is an interesting theoretical question what those density functionals are that can be estimated directly, without estimating the densities.

**Euclidean Graph Optimization Based Estimators.** *Euclidean Graph Optimization* (EGO) can be used to estimate certain functionals of densities, e.g. the $\alpha$-entropy, mutual information, and some divergences. Here we do not provide a complete introduction to this topic, instead we only list a few examples to demonstrate the usage of EGO in practice. Having given a sample from a distribution, these algorithms fit a so-called Euclidean Graph to these sample points, and use the edge lengthes of this graph to estimate functionals of the underlying density. A nice property of these estimators is that in many cases they are simple and can estimate these functionals of the density *without* estimating the density itself.

Let $p$, $q$ be $\mathbb{R}^d \to \mathbb{R}$ density functions, and let $\alpha \in \mathbb{R} \setminus \{1\}$. The Shannon entropy ($H^S$), Rényi-$\alpha$ entropy ($H^R_\alpha$), Kullback-Leibler divergence ($D^{KL}$), and Rényi-$\alpha$ divergence ($D^R_\alpha$) are defined respectively as the following functionals of the densities $p$ and $q$: $H^S(p) \doteq -\int p(x)\log p(x)\mathrm{d}x$, $H^R_\alpha(p) \doteq \frac{1}{1-\alpha}\log\int p^\alpha(x)\mathrm{d}x$, $D^{KL}(p\|q) \doteq \int p(x)\log\frac{p(x)}{q(x)}\mathrm{d}x$, and $D^R_\alpha(p\|q) \doteq \frac{1}{\alpha-1}\log\int p^\alpha(x)q^{1-\alpha}(x)\mathrm{d}x$.

Below we briefly review how these expressions can be estimated by an application of EGO. Let $X_{1:n} \doteq (X_1,\dots,X_n)$ be an i.i.d. sample from a distribution with density $p$, and similarly let $Y_{1:m} \doteq (Y_1,\dots,Y_m)$ be an i.i.d. sample from a distribution having density $q$. Let $\rho_k(i)$ denote the Euclidean distance of the $k$th nearest neighbor of $X_i$ in the sample $X_{1:n}$, and similarly let $\nu_k(i)$ denote the distance of the $k$th nearest neighbor of $X_i$ in the sample $Y_{1:m}$. Let $\bar{c} = \pi^{\frac{d}{2}}/\Gamma(\frac{d}{2}+1)$ be the volume of a $d$-dimensional unit ball. Using these notations, Goria et al. [64] derived the following consistent estimator for the Shannon entropy:

$$\widehat{H}^S(X_{1:n}) \doteq \frac{d}{n}\sum_{i=1}^{n}\log\rho_k(i) + \log(n-1) - \psi(k) + \log\bar{c},$$

where $\psi$ is the digamma function. Using a similar technique, the Rényi entropy can also be estimated consistently, as it was shown by Leonenko et al. [60]:

$$\widehat{H}^R_\alpha(X_{1:n}) \doteq \frac{1}{1-\alpha}\log\frac{1}{n}\sum_{i=1}^{n}\frac{\Gamma(k)}{\Gamma(k+1-\alpha)}\bar{c}^{(1-\alpha)}(n-1)^{(1-\alpha)}\rho_k^{d(1-\alpha)}(i)$$

Recently, Wang et al. [62] extended these ideas and derived an estimator for the KL-divergence, and we proposed an estimator for the Rényi-$\alpha$ divergence (Póczos et al., [65]):

$$\widehat{D}^{KL}(X_{1:n}\|Y_{1:m}) \doteq \frac{d}{n}\sum_{i=1}^{n}\log\frac{\nu_k(i)}{\rho_k(i)} + \log\frac{m}{n-1}$$

$$\widehat{D}^{R}_{\alpha}(X_{1:n}\|Y_{1:m}) \doteq \frac{1}{\alpha-1}\log\frac{1}{n}\sum_{i=1}^{n}\left(\frac{(n-1)\rho_k^d(i)}{m\nu_k^d(i)}\right)^{1-\alpha}\frac{\Gamma(k)^2}{\Gamma(k-\alpha+1)\Gamma(k+\alpha-1)}.$$

In Figure 1, we illustrate how to calculate $\rho_k(i)$ and $\nu_k(i)$ quantities that are needed for these estimators. It is worth noting that divergence estimators can also be used for mutual information estimation, since the mutual information by definition is the divergence between the joint density and the product of the marginal densities. The above estimators are simple; they only use certain $k$th-nearest neighbor distance based statistics that can be efficiently calculated using $k$-d trees [66]. Recently, Székely has proposed an estimator that uses all distances between the sample points (i.e. the complete Euclidean graph of the sample points) to define a new quantity for measuring the dependence between random variables [67].

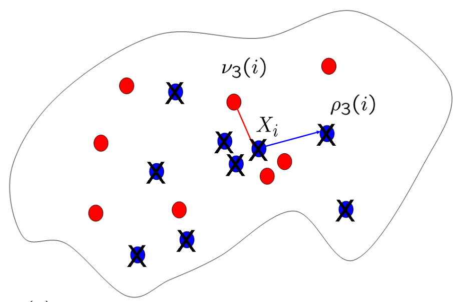



Figure 1: Calculating $\rho_k(i)$ and $\nu_k(i)$. Blue dots with X signs and red dots show samples $X_{1:n}$ and $Y_{1:m}$, respectively. We fixed $k=3$ in this example, i.e. we need to calculate the 3rd nearest neighbors of each $X_i$.

More general Euclidean graphs (e.g. minimum spanning trees, $k$-nearest neighborhood graphs, minimum matching, traveling salesman problem, Steiner graphs) can also be used to estimate the Rényi entropy [68,69] and information [70]. Nonetheless, currently it is an open question what other functionals can be estimated with these graphs.

Let $\mathcal{G}$ be a system of graphs on $n$ nodes numbered from $1$ to $n$ (specific examples will be given below). For a graph $G\in\mathcal{G}$, let $E(G)\subset\{1,\dots,n\}^2$ be its edge-set, and let $G_n$ be the complete graph on $n$ nodes. Thus, each graph $G\in\mathcal{G}$ is a subgraph of $G_n$. Define the estimate of the Rényi-$\alpha$ entropy as

$$\widehat{H}^{R}_{\alpha}(X_{1:n}) = \frac{1}{1-\alpha}\ \log\left(\frac{1}{\gamma_{d,\alpha}\, n^{\alpha}}\min_{G\in\mathcal{G}}\sum_{(i,j)\in E(G)}\|X_i-X_j\|^{d(1-\alpha)}\right), \tag{1}$$

where $\gamma_{d,\alpha}>0$ is a universal constant. Possible systems of Euclidean graph sets $\mathcal{G}$ include $\mathcal{G}_{\mathrm{ST}}$, all spanning trees of $G_n$; $\mathcal{G}_{\mathrm{H}}$, the set of all Hamiltonian cycles of $G_n$, $\mathcal{G}_{\mathrm{R}(k)}$ ($k>0$), the set of all subgraphs of $G_n$ where the out-degree of each node is $k$; and many more. The functional $L_n(X_{1:n})$ is known as *Euclidean functional*, and the minimization in Eq. (1) is called Euclidean Graph Optimization. Different choices of the graph system $\mathcal{G}$ lead to different optimization problems. When $\mathcal{G}=\mathcal{G}_{\mathrm{ST}}$, computing $L_n$ amounts to finding the ($d(1-\alpha)$-weighted) minimal spanning tree (MST). When $\mathcal{G}=\mathcal{G}_{\mathrm{H}}$, we need to solve a traveling salesman problem (TSP), while when $\mathcal{G}=\mathcal{G}_{\mathrm{R}(k)}$, $L_n$ can be computed by finding the $k$-nearest neighbors ($k$-NN) for each node and summing up the $d(1-\alpha)$-power edge-lengths. Using these Euclidean graphs, one can prove that Eq. (1) is an almost surely consistent estimator of the Rényi entropy [68, 71] under certain conditions. By exploiting the fact that the mutual information is the negative entropy of the copula of the underlying distribution, we have shown that Euclidean Graph Optimization can be used for mutual information estimation as well [61, 70].

We have proved that for any fixed $k$ these estimators are consistent, and we have also demonstrated the applicability of these estimators in various important problems including cosmology [22] and computer vision [23]. Currently, however, we do not know the convergence rate of these estimators yet.

**Reproducing Kernel Hilbert Space Based Estimators.** There is a significantly different method that has also been used to measure dependence between random variables and divergence between distributions. This method embeds random variables into a Reproducing Kernel Hilbert Space (RKHS) and estimates dependence or divergence via solving convex optimization tasks in the RKHS function space. Kernel mutual information [72], kernel canonical correlation analysis, and kernel generalized variance [73] have been defined this way to measure dependence. Consistent RKHS based $f$-divergence and likelihood ratio estimators have been proposed as well [63]. Although this approach and the previous Euclidean Graph Optimization approach can be used for similar estimation problems, currently it is an open question how these approaches are related to each other, and what those conditions are when one method can outperform the other [74].

**Preliminary Demonstrations on Turbulence Data.** Using the above results, recently our group generalized the popular support vector machine method to be able to classify distributions [75]. The main idea was to estimate a divergence (e.g. Kullback-Leibler or Rényi) between distributions using the above described tools, then using this divergence create a positive semi-definite kernel function and plug that kernel function into a support vector machine (SVM) that can be used for classification. We call this method support distribution machine (SDM), and have demonstrated in many practical problems (cosmology [22], computer vision [23], and civil engineering [25]) that these algorithms can often outperform the state-of-the art.

To show a preliminary result on the applicability of this method to detect vortices in turbulence datasets, we performed the following experiment. We trained an SDM classifier using turbulence data from the JHU Turbulence Data Cluster[†] [76], (TDC). TDC simulates fluid flow through time on a 3-dimensional grid, calculating 3-dimensional velocities and pressures of the fluid at each step. We trained an SDM on a manually labeled training set of a few positive (containing vortex) and negative examples (not containing vortex). Training examples are shown in Figure 3(a,b,c). This classifier was then used to evaluate groups along $z$-slices of the data. One slice of the resulting probability estimates is shown in Figure 2(a); the arrows represent the mean velocity at each classification point. The high-probability region on the left is a canonical vortex, while the slightly-lower probability region in the upper-right deviates a little from the canonical form. Note that some other areas show somewhat complex velocity patterns but mostly have low probabilities. As one can see the method can accurately detect vortices.

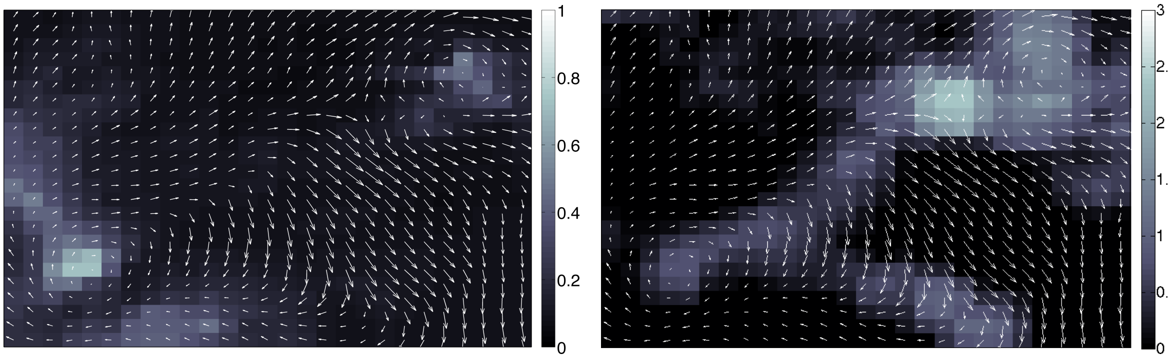
In this proposal we will extend these ideas and show that similar methods can also be used to detect other Lagrangian Coherent Structures in turbulent datasets (Section 3.2). Using recent advances in ML on big data sets, we will also study how to scale these novel classification algorithms up to large datasets. In related function-to-function regression and distribution regression problems we were able to scale up regression methods to big data problems [77, 78].

## 2.2 Finding Anomalous Events

In addition to finding instances of known patterns (Section 2.1), part of the process of exploring the results of a large-scale simulation is looking for *unexpected rare phenomena* that domain scientist can analyze further.

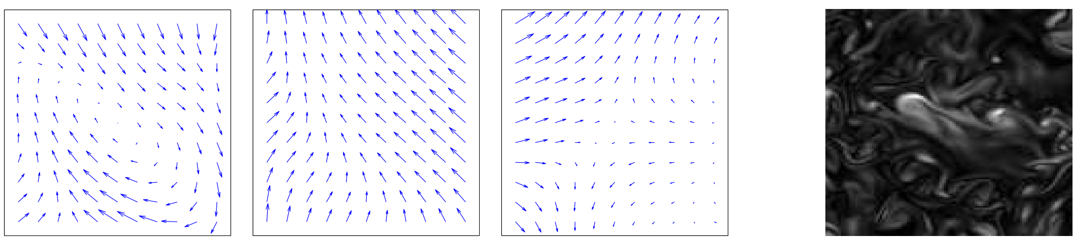
In the above described TDC simulation we consider the vertices in a local cubic region as a group, and the goal is to find groups of vertices whose velocity distributions (i.e. moving patterns) are unusual and potentially interesting. To demonstrate anomaly detection with SDMs, we trained a "one-class SDM" on 100 typical distributions from the turbulence data, with centers chosen uniformly at random. This one-class SDM method then was asked to process the data and find regions that are most untypical and look the most different from the examples in the training set. As we can see in Figure 2(b), the two vortices in the data set are picked out, but the area with the highest score is a diamond-like velocity pattern; these may well be less common and therefore more anomalous in the dataset than vortices. In Figure 3(d) we show another

[†] `http://turbulence.pha.jhu.edu`




(a) Classification probabilities (b) Anomaly scores

Figure 2: Classification and anomaly scores along with velocities for one $54 \times 48$ slice of the turbulence data.



(a) Positive (b) Negative (c) Negative (d) Vorticity

Figure 3: a,b,c: Training examples for the vortex classifier. d: Detected anomaly in the JHU turbulence data.

detected anomaly in the same TDC dataset. As we can see, this is indeed a rare and interesting interaction among three vortices positioned in a straight line.

In this proposal we will extend these ideas to be able to process large datasets and find different types of anomalies in turbulent datasets. We believe that SDM classifiers and anomaly detectors can serve as simulation exploration tools that will allow scientists to iteratively look for anomalous phenomena and label some of them. Classifiers could then find more instances of those phenomena and compute statistics about their occurrence, while anomaly detection would be iteratively refined to highlight only what is truly new. We will use these methods to find anomalies that can lead to the non-Kolmogorov contribution in Atmospheric Turbulence.

### 2.3 Controlled Function Spaces with Applications to Turbulent Plasma

There are many efficient and widely used control methods for finite-dimensional vector state spaces. Without providing a full literature review, we mention PID controllers, (Extended) Kalman filter (EKF) controllers, controlled Autoregressive Moving Average (CARMA, ARMAX), Generalized Linear Models (GLM) models, and Markov Decision Processes (MDP). In these popular methods the controlled state space is a simple finite dimensional vector space. In our scientific applications however, the state space is more complicated: vector fields, N-body particles, or space of continuous density functions. We want the controller to operate on these complex spaces. For example, some specific tasks can be to control a continuously changing distribution to keep certain properties, e.g. its support should always be a 3D torus of a tokamak reactor, or control a 3D vector field such that it has no "big tearing mode islands" in it, etc. In this proposal we will study and develop new control methods that can operate on these complex state spaces.

Recently our team has developed new regression methods that can learn an operator map when the inputs or outputs are both distributions or functions. We call these methods distribution-to-distribution regression (DDR) and function-to-function regression (FFR) methods. DDR and FFR learn a regression operator $F$ that maps a continuous function or distribution $p$ to another function (distribution) $q$, that is $F[p] = q$ in the noiseless case. In this proposal we will further generalize these methods to become applicable in control problems as well.

First, we will generalize controlled autoregressive (AR) models from finite dimensional vectors to function spaces: $X_t = \sum_{i=1}^{k} F_i[X_{t-i}] + + \sum_{j=1}^{l} G_j[u_{t-j}] + N_t$, where $X_i$ is a function, $F_i$, and $G_j$ are operators that map a function to another function, $u_t$ is a control function. Controlled linear and nonlinear dynamical systems (Kalman filter, Extended Kalman filter) can also be generalized to function spaces, for example as $X_t = A[X_{t-1}] + B[u_t] + N_t$, $Y_t = C[X_t] + D[U_t] + M_t$, where $A, B, C, D$ are operators that map functions to functions, $u_t$ is a control function, and $N_t$ and $M_t$ are noise functions in each time step $t$. $X_t$ is an unknown state space function at time step $t$. It is a nice theoretical question what those conditions are when this model is identifiable. Our goal is to estimate the operators $A, B, C, D$, and use these estimated operators for prediction and control. The observation functions $\{Y_1, \ldots, Y_t\}$ are not completely available to us, we can observe them only through some input-output pairs. As a simplest case, one might assume that the unknown operators are parameterized with finite dimensional parameters (e.g. $A = A_\theta$, $\theta \in \mathbb{R}^d$), and then use EM [79], spectral [80] or subspace methods [81] to estimate these parameters. We will also investigate these possibilities when we do not make these parametric assumptions, similarly as we did in our FFR and DDR methods.

### 2.4 Discovering New Scientific Laws with ML Tools

Recently our team has proposed a new ML method that can estimate the dynamical mass of galaxy clusters from the velocities of the individual galaxies in the cluster [22]. This prediction problem was very challenging for astrophysicists, because the scientific law that describes this regress problem is very complicated. Our ML based algorithm, however, was able to learn a prediction rule from the data and achieved better accuracy than the state-of-the-art. In this project, we would like to follow a similar approach: We will develop nonlinear regression methods that operate on turbulence data (instead of galaxy clusters) and can extract features from the data that can reveal important nonlinear relationship between subsets of the turbulence data.

Domain scientists always try to create some hand-made features from the data, (e.g. energy spectra or some other statistics), that can be useful in some applications. Here we will let ML methods construct the features that are the most important in certain prediction problems, revealing structures, or higher-order statistical correlations in the data.

## 3 Open Questions in Turbulent Plasma

In this section we review the science questions of turbulence that we plan to solve using ML methods. As we stated in Section 1, we will consider three different classes of data: i) turbulence data from DNS (including simulations of atmospheric turbulence); ii) solar wind turbulence data; iii) controlled turbulence data (fusion plasma).

### 3.1 Simulated Turbulence Data

The results of our recent high-resolution DNS runs indicate that based on the decay laws the following the classification might be made: i) HD turbulence (with and without helicity); ii) non-helical MHD turbulence with an initial Batchelor spectrum at large scales; iii) non-helical MHD turbulence with an initial spectrum that is obtained by monochromatic electromagnetic forcing; iv) fully helical MHD turbulence. These different classes of decay occupy distinct regions in the $qp$ diagram, where $q$ and $p$ denote the instantaneous rates of change of the correlation scale $\ell$ and the energy $\mathcal{E}$ with time $t$, i.e., $q(t) = d\ln\ell/d\ln t$ and $p(t) = d\ln\mathcal{E}/d\ln t$; see Fig. 4(a). For the fully helical MHD turbulence the decay process is governed by helicity conservation [82], while the HD turbulence decay lies with the Loitsyansky integral invariance [83].

Cases ii) and iii) show similar behavior, i.e., a phenomenon found for the first time by the proposers – the inverse transfer of energy from small to large scales [38] and one of the questions which we plan to answer is what invariant quantity might be associated with the non-helical inverse transfer (at a rate half as strong as in the case of helical MHD turbulence). Our hypothesis was to associate the non-helical inverse transfer with an interplay between the magnetic and velocity fields, but only the careful statistical investigation through using of ML methods can determines the true reason of this previously un-emphasized and puzzling result. We show DNS data of non-helical MHD turbulence decay on Fig. 4(b) for unforced cases. In Ref. [38] we made an attempt to associate the inverse cascade with the conserved vector potential in two dimensions. This result seems to be surprising given that under our DNS setup the flow cannot be regarded as locally two dimensional. Nevertheless, having determined the appropriate values of $q$ and $p$, the energy spectra collapse onto a single curve $\phi$; see Fig. 4(c), were $\mathcal{E}(k,t) = \ell^{-\beta}\phi(\ell k)$, where $\beta = p/q - 1$.

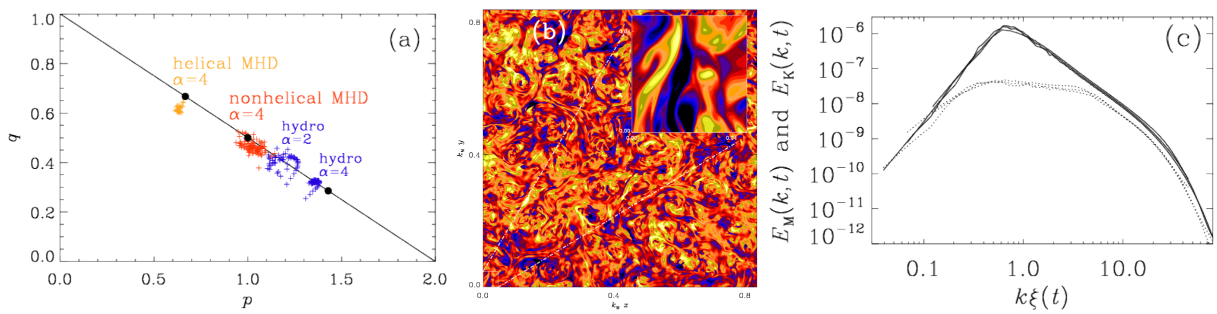



Figure 4: (a) Different classes of MHD and hydrodynamic turbulence in the $qp$ diagram. (b) Contours of vertical velocity from a DNS of non-helical MHD turbulence showing inverse transfer of energy [38]. The inset zooms into the small square in the lower left corner. (c) Collapsed energy spectra from the DNS of non-helical MHD turbulence.

The presence of hydrodynamic and/or magnetic helicity profoundly affects the statistical properties of the turbulence. Existence of the non-zero helicity constraints turbulence properties leading to distinctive spectral properties and the direction of nonlinear cascades in every specific setup. In magnetized turbulence non-zero helicity can be observed in kinetic motions, magnetic field and cross-helicity that describes the correlation of the kinetic and magnetic fields in turbulent flows. One of major questions is the helical turbulence development. Two major scenarios is under discussions [84]: First, so called the *helical Kolmogorov* model, with a forward cascade of both energy and helicity (dominated by energy dissipation on small scales) one has spectral indices (for respective power spectra $-5/3$ and $-8/3$ for energy density and helicity), see p. 243 of Ref. [85]. Second, if helicity transfer and small-scale helicity dissipation dominate (the helicity transfer spectrum) the energy and helicity spectral indices are equal, being $-7/3$ [86]. The helical Kolmogorov spectrum has been observed in the inertial range of weakly helical turbulence [87], while for strongly helical hydrodynamical turbulence the characteristic length scale of helicity dissipation is larger than the Kolmogorov energy dissipation length scale, and the Kraichnan model [88] is realized [89]. The simplified picture (phenomenologically discussed in [47] assumes that the the inertial range consists of two sub-intervals, both with power-law spectra. For smaller wavenumbers, the spectra are determined by helicity transfer (the Kraichnan model), while for larger wavenumbers turbulence becomes non-helical and the more common helical Kolmogorov spectrum is realized. However, numerical experiments show a more complex picture [84] (e.g. see Ref. [90]) and one of our goals is to determine helical turbulence spectral characteristics applying ML methods to simulated data. Brandenburg and collaborators [40,91] have studied the growth of helical structures at large scales. We plan to further extend this analysis by considering the dependence on plasma parameters.

Atmospheric turbulence is one of the best studied examples of turbulent flows. Observations of non-Kolmogorov behavior of turbulent velocity fields is linked to the generation of large-scale coherent structures in the Earth's atmosphere, or the occurrence of intermittent or other persistent anomalous structures in the turbulence. For many years, various scientific communities have studied atmospheric turbulence and important results have been obtained. Nevertheless, the atmosphere's statistical behavior is still not well understood. A good example is the propagation of electromagnetic waves through turbulent media. Turbulence causes fluctuations in the refractive index: Kolmogorov turbulence with a 5/3 power law leads to refractive index fluctuations with a 11/3 power law index in 3 dimensions. However, realistic observations show refractive index fluctuations with power law indices varying between 3 and 5; see e.g. [92, 93]. These observations have prompted the study of electromagnetic wave propagation through turbulence described by non-classical power spectra.

In this proposal we plan to develop ML algorithms to determine the types of deviations from Kolmogorov model using the simulated data from atmospheric turbulence. We plan to analyze the turbulence spectrum and identify what type of anomalies lead to the non-Kolmogorov contribution. Some types of non-Kolmogorov components may lead to the increase, while others to the decrease of the 11/3 power law increment. Knowledge of statistics of atmospheric turbulence is essential for practical design of adaptive optics systems. How effective could adaptive optics be in non-Kolmogorov turbulence remains an answered question, both theoretically and experimentally. Our new approach using ML methods can be a novel step in resolving this problem. Non-Kolmogorov contribution of electromagnetic wave propagation is an important subject for remote sensing, imaging and communication systems that has attracted considerable theoretical and practical interests in the past decades (see e.g. [94, 95]). Results of our study can contribute in the improvement of the atmospheric degradation of high data rate optical transmitters for satellite communication channels, airborne and space communication links. In addition, plasma turbulence can lead to variation of polarization degree or spectral degree of coherence of electromagnetic beams traveling through it. Knowing atmospheric turbulence statistics can increase the precision of measurements from ground based telescopes.

## 3.2 Solar Wind

There are a number of statistical tools that play important roles in turbulence research. Routinely examined are energy spectra, which give information about the distribution of structures across the vast ranges of length scales that are excited in fully developed turbulence. The true spectra can be altered by both physical and numerical effects, especially near the largest and smallest length scales in the system.

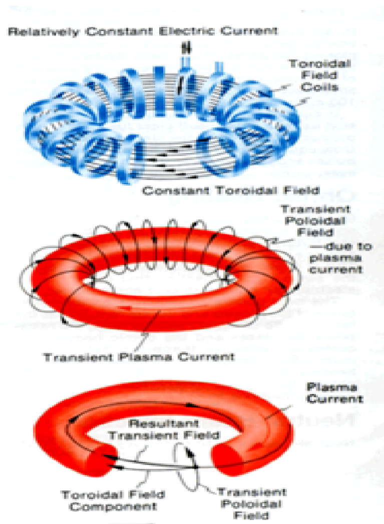
The association of spectral properties with actual physical structures has been appreciated in a number of cases. Indeed, the presence of coherent structures in turbulence is known to be important in the understanding of turbulent energy transport in geophysical [96] and astrophysical convection [97]. The statistical properties of Lagrangian Coherent Structures (LCEs) have been analysed in various circumstances. In the case of MHD turbulence, the LCEs are known to show clear differences in the early linear and late nonlinear phases of magnetic field amplification by a turbulent dynamo [98]. This has also been seen in studies of topological entropies of turbulent flows exhibiting dynamo action [99]. Also the fractal dimensions as well as the multifractal properties of MHD convection are known to show properties (e.g., a fractal dimension of 1.7 for vortex and current structures) that are also found in solar magnetograms [100, 101].

There are a number of concrete questions where detailed analysis of solar wind turbulence has become timely. This is connected with the shedding of magnetic helicity from the Sun. Magnetic helicity is a topological quantity that describes the mutual linkage of magnetic flux structures and is known to be a conserved quantity. Therefore, no net magnetic helicity can be created, but it can be redistributed among different length scales.

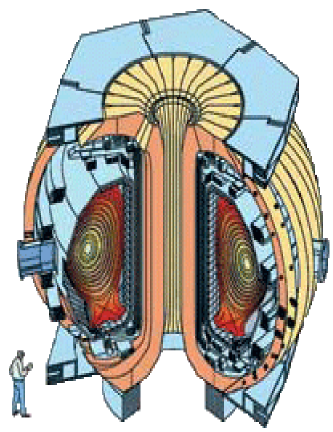
Using machine learning algorithms, we want to study coherent structures within the inertial range of solar wind turbulence. Owing to the proximity of four identical space crafts, we can determine the spatial information to determine magnetic helicity [102]. We will also compare with the method of Matthaeus & Goldstein [103], which makes the assumption of isotropy. The quality of this assumption will be tested. Finally, we will determine the gradual exchange of magnetic helicity between small and large length scales [17]. We will also use measurements of magnetic helicity fluxes densities using a technique developed recently in the context of dynamo simulations.

### 3.3 Controlled Turbulence in Plasma

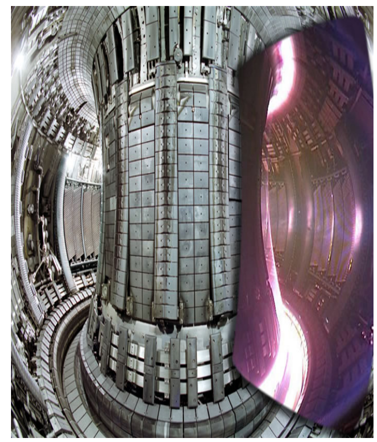
**Difficulties with Fusion Power.** The primary problem with fusion power is that we do not yet know how to produce it in a controlled, commercial scale, economically viable manner. It requires confining a hydrogen plasma at extremely high temperatures and pressures.




(a) Magnetic fields and electric currents in a tokamak.



(b) Schematic of the General Atomics DIII-D Tokamak



(c) Internal view of Joint European Torus (JET) with a visible light image of the plasma during operation superimposed

Figure 5: Tokamak fusion reactors. The plasma is maintained in the shape of a torus by several magnets wrapped around it.

**Tokamak Reactors.** The most successful operating research reactors to date are tokamaks. They use magnetic confinement of plasmas in the shape of a torus (see figure 5). The gas is heated to the point that it becomes an ionized plasma and thus subject to the effects of a magnetic field. A set of controllable magnets is placed around the torus to generate the desired magnetic field. Control of these devices is extremely challenging due to the nonlinear physics involved. At the heart of the problem are the MHD governing the state of the plasma. These result from interlinking Navier-Stokes equations of fluid dynamics and Maxwell's equations of electromagnetism. The nonlinearities yield many instabilities that must be controlled to achieve the temperatures and pressures necessary for fusion.

Considerable progress has been achieved through a significant investment in understanding the physics of fusion and magnetic confinement from first principles [104]. That understanding has been used to design systems that are "easy" to control in the sense that PID, or other simple, linear controllers are used. When those controllers do not perform well, extensive experimental effort is spent to improve their operation. In spite of the progress, significant unsolved problems remain in the way of achieving the sustained temperatures and pressures required.

For much of the history of controlled fusion research, this "physics-driven" approach to control was the only available option. We had neither the methods nor the computational power to design or learn sophisticated non-linear controllers. However, at the same time physicists were advancing their understanding of controlled fusion, ML researchers were making gains in their ability to learn models of complex systems from data and control them with learned nonlinear controllers. We hypothesize that the ML community has now made enough advances that it can significantly aid the pursuit of energy from controlled fusion and impact the timeline to commercial fusion power.

**Tearing Modes.** Tearing modes in a magnetically confined plasma occur at the interface of magnetic fields going in opposite directions. An island forms that becomes isolated from the overall magnetic field and grows until it is disrupted. If the disruption occurs early, it will cause some degradation to the plasma confinement, but the magnetic field lines can return to their original state. If the island gets too large before being disrupted, it will cause a catastrophic loss of plasma confinement and terminate that run of the reactor.

**Open Questions.** The physics of the growth of the magnetic island are understood. What is not understood is what causes the islands to form, how they can be prevented from forming, and how they can be disrupted earlier. A recent attempt to characterize the conditions for their appearance is [105]. A description of tearing modes can be found e.g. in [104].

**The Dataset.** The DIII-D fusion reactor is a tokamak operated by General Atomics that exhibits tearing mode behavior. Here David Humphreys' Lab has collected data for us from experiments done on DIII-D to better understand tearing modes. The observables include the current density profile, plasma pressure profile, and plasma shape. The measurements are taken in a 30x30 or 60x60 grid and have time resolutions ranging from 20 $\mu$s to 2 ms. The time scales of the islands are on the order of 10-100 ms thus giving an ample number of measurements to predict their onset, growth, and disappearance. He has thousands of labeled instances of tearing modes appearing.

**Our Goal.** We propose to do supervised learning on the data to predict the appearance of future islands from the plasma state. Recently developed set kernels will be the basis of the learning algorithms [23, 106]. The kernels compare the collective state of the plasma at different times by considering the data to be samples from an underlying distribution in feature space. They use non-parametric divergence estimators that do not assume those distributions have a pre-defined parametric form. We construct a Gram matrix from these estimates and use it in kernel-based learning methods.

This project looks forward to the possibility of using data driven learned control on a real reactor to find the best ways of managing plasma instabilities. Since such experiments are beyond the scope of this initial effort, we will simulate the learning control process using the labeled data. The "simulation" begins by presenting the learner with the possible controls it might choose (i.e. those used to generate the data) but does not show it the outcomes of the experiments (i.e. whether tearing mode instabilities occur). The learning controller selects which parameters it wants to test sequentially using methods built on those from [107, 108]. As each selection is made, the results of the experiment are revealed to the learner. This will allow us to evaluate its ability to find the highest performing settings with the fewest tearing mode instability problems using the smallest number of experiments.

# 4 Results from prior NSF support

**Barnabas Póczos** is Co-PI on NSF Award 1247658, "III: BIGDATA: Distribution-based machine learning for high dimensional datasets", $1,000,000, duration 1/1/2013-12/31/2015.
*Intellectual Merit:* In this project we use distribution based supervised ML tools to perform data mining in neuroscience data, such as fMRI and DSI brain images. Our main results have been published in [23, 109, 110].
*Broader Impacts:* Beyond neuroscience, the proposed approaches proved to be very efficient in other domains e.g. computer vision and cosmology [23]. Previously Dr. Póczos was a PI on: III-1250350. Results are published in [110–112].

**Jeff Schneider** was a Co-PI on the following recent NSF award: Dubrawski, Clermont, Cooper, Neill, Schneider, NSF Award 0911032, "III: Large: Discovering Complex Anomalous Patterns", $2598153, 2009-2013. This award funds a collaboration of computer scientists and medical doctors to identify complex patterns in medical care data that may indicate unusual events or protocols impacting patient outcomes.
*Intellectual Merit:* Schneider and his students have developed new methods of regularization that allow one to learn complex models with only a small amount of training data [113–115]. Some of the preliminary work on ML for sets that was cited in the previous sections was developed under this grant [23, 106, 116, 117].

*Broader Impacts:* In addition to the publications, all of the software used to produce these results is available to the public at `www.autonlab.org`. Schneider has graduated three PhD students from this award.

**Tina Kahniashvili** is PI and **Axel Brandenburg** is the collaborator on NSF AST-1109180 grant ($459K, 09/2011–08/2015) "Collaborative Research: Cosmic Magnetic Fields: Origin, Evolution and Signatures".
*Intellectual Merit:* The results are published in Refs. [38–41, 44, 46, 57, 118–126] (and Refs. [91, 127–129] in preparation), and include direct numerical MHD simulations of the inflationary [41, 91, 128, 129] and cosmological (electroweak and QCD) phase transitions [38–40, 57] generated magnetic field evolution. We were first to find an inverse transfer for even in nonhelical MHD turbulence [38] with the scaling laws are almost independent of magnetic Prandtl number [39, 40].
*Broader Impacts:* During the reporting period Kahniashvili has advised four undergraduate and four graduate students (three female students). Outreach activities also include popular lectures given by Kahniashvili at the International Science Summer Camp for High School Students, public talks at Allegheny Observatory, Abastumani Astrophysical Observatory, and Ilia State University and Tbilisi State Universities. Kahniashvili and Brandenburg have organized four weeks NORDITA program (including conferences and workshops) on "Origin, Evolution, and Signatures of Cosmological Magnetic Fields".

**Hy Trac** is Co-PI on NSF grant AST-1312991: title "Collaborative Research: Modeling the Reionization of the Intergalactic Medium", PI Anderson, Co-PIs McQuinn and Trac, total amount $428k, CMU amount $51k, period 09/01/2013 - 08/30/2016.
*Intellectual Merit:* The collaborative project uses numerical simulations in comparison with observations to better understand the impact of radiation sources and sinks during the epoch of reionization (EoR). Radiation-hydrodynamic simulations with Trac's RadHydro code have been used to show that the observed large opacity variations in the high-redshift Lyman alpha forest can be explained by temperature fluctuations arising from patchy reionization [130]. N-body simulations with Trac's $P^3M$ code have been used to quantify the abundance and accretion rates of dark matter halos, which are compared against observations of high-redshift galaxies to quantify the galaxy-halo connection in the first billion years [56]. The lightcone effect in the 21cm signal from neutral hydrogen during the EoR has been modeled and quantified using semi-numerical simulations [131].
*Broader Impacts:* Trac supervises graduate student Paul La Plante (CMU) and co-mentors postdoctoral fellow Anson D'Aloisio (Univ. Washington). Trac has given invited talks on "Cosmic Reionization: How the First Galaxies Lit Up the Universe" at Penn State University, University of North Dakota, and University of Utah. Related talks on "Simulating the Universe with Supercomputers" has also been given at the Allegheny Observatory and as a physics undergraduate colloquium at CMU.

# 5 Broader impacts of the proposed work

Educational and outreach efforts will be incorporated into this project at all stages and at all participating universities. The objectives include: (i) involving undergraduates in astronomy and physics research; (ii) graduate student research assistants supervising research done by undergraduates; (iii) graduate student research assistants teaching a few lectures a year in undergraduate physics classes; (iv) working with local partners to develop and nurture interest in sciences in the local community; (v) helping middle and high school teachers develop demonstrations for science classes; (vi) researchers presenting a few astronomy lectures a year to middle and high school students. To achieve these goals:

**CMU, Department of Physics and ML Department.** We will involve undergraduate students in the research, providing crucial experience for them and significant mentoring opportunities for the supported graduate students. Outreach presentations among graduate and undergraduate students are valuable, low stress opportunities to develop and refine their teaching skills. A weekly presentation class for graduate students run by graduate students in the CMU Department of Physics provides a pool of public speakers whose goal it is to effectively communicate their research to the public. Presentations are taped and critiqued, with assistance from expert speakers among the faculty. CMU Physics Department's CONCEPTS program for middle school students from minority groups underrepresented in science is funded by the Grable Foundation. Faculty members have run student-led learning initiatives on the subject of distance scales in the universe, and scale modeling (out of household objects) of the solar system and the local group of galaxies. It will be rewarding for the proposers to explore effective ways to communicate the concepts in this proposal to students.

At Pittsburgh, we are fortunate to have physical and organizational infrastructure for outreach efforts already in place in Pittsburgh, through: (i) the Allegheny Observatory. We plan to widen its already-active public outreach and education program. The astrophysics and astroparticle groups at CMU and Pitt host a monthly public lecture series featuring local faculty members. New astronomy artist and photography STEAM workshops are planned to effect a change of public opinion through art. The art workshops team a visual artist with an astronomer for three-hour classes in painting celestial objects related to gravity research. Artists and researchers will take turns explaining the task at hand and the physics concepts depicted in the paintings. The astrophotography class will have a similar format and also include a lesson on light pollution solutions. (ii) Buhl Planetarium established by the Buhl Foundation in 1939 and being the predecessor organization of Carnegie Science Center. We anticipate extending existing outreach activities through the Buhl Digital Dome, Cafe Scientifique and other programs at the Carnegie Science Center (e.g. Girls Rock Science and SciTech Workshops). CMU and Pitt faculty members are working with the Buhl Planetarium staff to bring an astronomy documentary series to Pittsburgh, featuring newly-released movies such as Starmen, The City Dark and Above and Below to be shown on the Buhl Digital Dome. As past astronomical movie releases have shown, Question and Answer sessions with astronomers after each viewing will inspire lively discussions of astronomical topics. Tina Kahniashvili, Hy Trac, and Diane Turnshek have recently given public lectures at Allegheny Observatory, the Carnegie Science Center and elsewhere (e.g. TEDxPittsburgh). The proposers and funded grad students will continue to be active participants in Pittsburgh outreach activities, which are coordinated through the pghconstellation website and the Astronomy Enthusiasts email list (Diane Turnshek). The Physics Department outreach activities will be coordinated by Diane Turnshek.

**CU-Boulder.** The Department of Astrophysical & Planetary Sciences is participating routinely in a broad range of public outreach activities, most notably through public lectures given at the Fiske Planetarium on campus just a few minutes from the department. Its staff is strongly encouraged to engage in many activities and to speak with the public both at the planetarium and the associated Sommers-Bausch Observatory. Brandenburg has been lecturing to the public and in schools throughout his career and, as new faculty of the Department, is looking forward to continuing this habit. Some of his web pages are regularly being consulted by the public and reporters. He is regularly being interviewed by the press and television. The Department also has links with the Denver Museum of Nature and Science through the participation in film productions that aim at publicizing contemporary science. This poses a particularly exciting avenue for publicizing GW physics and will be pursued.

**Under-represented groups.** Our project will benefit under-represented groups through efforts already underway at CMU. Co-PI Tina Kahniashvili is a female professor and the PIs are regularly advising and have been actively working together with female students. This effort has resulted in many scientific publications with first author female students (See e.g. [19, 21, 22, 25, 122, 126, 132]). The School of Computer Science have made a successful effort to attract women students to CS by holding training camps for high school teachers through CS4HS [133] and research workshops for undergraduate women through OurCS [134].

**Free software.** The CMU Auton Lab and the PIs routinely offer their software for free to the public. As a result of the proposed research, the collection of available software will be substantially enriched. All of the algorithms developed through this grant will be provided to the public on the Auton Lab [‡] and the PIs' websites. [§]

[‡] `http://www.autonlab.org`

[§] `http://www.cs.cmu.edu/~bapoczos/code.html`

**REFERENCES:**


[1] A. Kolmogorov, “The Local Structure of Turbulence in Incompressible Viscous Fluid for Very Large Reynolds’ Numbers,” *Akademiia Nauk SSSR Doklady*, vol. 30, pp. 301–305, 1941.

[2] A. N. Kolmogorov, “Dissipation of Energy in Locally Isotropic Turbulence,” *Akademiia Nauk SSSR Doklady*, vol. 32, p. 16, 1941.

[3] L. F. Richardson, “Weather prediction by numerical process,” *(Cambridge University Press)*.

[4] G. Falkovich and K. R. Sreenivasan, “Lessons from Hydrodynamic Turbulence,” *Physics Today*, vol. 59, no. 4, p. 43, 2006.

[5] S. Y. Chen, B. Dhruva, S. Kurien, K. R. Sreenivasan, and M. A. Taylor, “Anomalous scaling of low-order structure functions of turbulent velocity,” *Journal of Fluid Mechanics*, vol. 533, pp. 183–192, June 2005.

[6] S. R. Stalp, L. Skrbek, and R. J. Donnelly, “Decay of Grid Turbulence in a Finite Channel,” *Physical Review Letters*, vol. 82, pp. 4831–4834, June 1999.

[7] I. P. Castro, “Relaxing wakes behind surface-mounted obstacles in rough wall boundary layers,” *Journal of Fluid Mechanics*, vol. 93, pp. 631–659, August 1979.

[8] D. K. Lilly, “Stratified Turbulence and the Mesoscale Variability of the Atmosphere.,” *Journal of Atmospheric Sciences*, vol. 40, pp. 749–761, March 1983.

[9] M.-M. Mac Low, R. S. Klessen, A. Burkert, and M. D. Smith, “Kinetic Energy Decay Rates of Supersonic and Super-Alfvénic Turbulence in Star-Forming Clouds,” *Physical Review Letters*, vol. 80, pp. 2754–2757, March 1998.

[10] K. Subramanian, A. Shukurov, and N. E. L. Haugen, “Evolving turbulence and magnetic fields in galaxy clusters,” *Mon. Not. R. Astron. Soc.*, vol. 366, pp. 1437–1454, March 2006.

[11] M. Christensson, M. Hindmarsh, and A. Brandenburg, “Inverse cascade in decaying three-dimensional magnetohydrodynamic turbulence,” *Phys. Rev. E*, vol. 64, p. 056405, November 2001.

[12] J. J. Podesta, D. A. Roberts, and M. L. Goldstein, “Spectral Exponents of Kinetic and Magnetic Energy Spectra in Solar Wind Turbulence,” *Astrophys. J.*, vol. 664, pp. 543–548, July 2007.

[13] G. G. Howes, S. C. Cowley, W. Dorland, G. W. Hammett, E. Quataert, and A. A. Schekochihin, “A Model of Turbulence in Magnetized Plasmas: Implications for the Dissipation Range in the Solar Wind,” *J. Geophys. Res.*, vol. A113, p. 5103, 2008.

[14] N. E. L. Haugen and A. Brandenburg, “Hydrodynamic and hydromagnetic energy spectra from large eddy simulations,” *Physics of Fluids*, vol. 18, p. 075106, July 2006.

[15] S. Boldyrev, J. C. Perez, J. E. Borovsky, and J. J. Podesta, “Spectral Scaling Laws in Magnetohydrodynamic Turbulence Simulations and in the Solar Wind,” *Astrophys. J. Lett.*, vol. 741, p. L19, November 2011.

[16] J. J. Podesta and S. P. Gary, “Magnetic Helicity Spectrum of Solar Wind Fluctuations as a Function of the Angle with Respect to the Local Mean Magnetic Field,” *Astrophys. J.*, vol. 734, p. 15, June 2011.

[17] A. Brandenburg, K. Subramanian, A. Balogh, and M. L. Goldstein, “Scale-dependence of magnetic helicity in the solar wind,” *Astrophys. J.*, vol. 734, p. 9, 2011.

[18] J. Oliva, B. Poczos, T. Verstynen, A. Singh, J. Schneider, F. Yeh, and W. Tseng, “Fusso: Functional shrinkage and selection operator,” in *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2014.

[19] X. Xu, S. Ho, H. Trac, J. Schneider, B. Poczos, and M. Ntampaka, “A first look at creating mock catalogs with machine learning techniques,” *Astrophysical Journal (APJ)*, 2013.

[20] K. Kandasamy, J. Schneider, and B. Poczos, “Bayesian active learning for posterior estimation,” in *Astrophysical Journal (International Joint Conference on Artificial Intelligence (IJCAI))*, 2015.

[21] M. Ntampaka, H. Trac, D. J. Sutherland, S. Fromenteau, B. Poczos, and J. Schneider, “Dynamical Mass Measurements of Contaminated Galaxy Clusters Using Machine Learning,” *ArXiv e-prints*, September 2015.

[22] M. Ntampaka, H. Trac, D. J. Sutherland, N. Battaglia, B. Póczos, and J. Schneider, “A Machine Learning Approach for Dynamical Mass Measurements of Galaxy Clusters,” *Astrophys. J.*, vol. 803, p. 50, April 2015.

[23] B. Póczos, L. Xiong, D. Sutherland, and J. Schneider, “Nonparametric kernel estimators for image classification,” in *International Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.

[24] Z. Szabo, B. Poczos, and A. Lorincz, “Online group-structured dictionary learning,” in *International Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011.

[25] I. Jung, M. Berges, J. Garrett, and B. Poczos, “Exploration and evaluation of ar, mpca and kl anomaly detection techniques to embankment dam piezometer data,” *Advanced Engineering Informatics*, 2015.

[26] Z. Szabo, A. Gretton, B. Poczos, and B. Sriperumbudur, “Two-stage sampled learning theory on distributions,” in *International Conference on AI and Statistics (AISTATS)*, 2015.

[27] C. Miller, C. Genovese, R. Nichol, L. Wasserman, A. Connolly, D. Reichart, A. Hopkins, J. Schneider, and A. Moore, “Controlling the false discovery rate in astrophysical data analysis,” *Astronomical Journal*, vol. 122, no. 6, pp. 3492–3505, 2001.

[28] B. Bryan, J. Schneider, C. Miller, R. Nichol, C. Genovese, and L. Wasserman, “Mapping the cosmological confidence ball surface,” *Astrophysical Journal*, 2007.

[29] X. Xu, S. Ho, H. Trac, J. Schneider, B. Poczos, and M. Ntampaka, “A first look at creating mock catalogs with machine learning techniques,” *Astrophysical Journal*, Aug 2013.

[30] J. A. Bagnell and J. Schneider, “Autonomous helicopter control using reinforcement learning policy search methods,” in *Int. Conf. on Robotics and Automation*, 2001.

[31] A. Brandenburg and K. Subramanian, “Astrophysical magnetic fields and nonlinear dynamo theory,” *Phys. Rep.*, vol. 417, pp. 1–209, October 2005.

[32] A. Brandenburg, “The Case for a Distributed Solar Dynamo Shaped by Near-Surface Shear,” *Astrophys. J.*, vol. 625, pp. 539–547, May 2005.

[33] R. Beck, A. Brandenburg, D. Moss, A. Shukurov, and D. Sokoloff, “Galactic Magnetism: Recent Developments and Perspectives,” *Annu. Rev. Astron. Astrophys.*, vol. 34, pp. 155–206, 1996.

[34] A. Brandenburg, A. Nordlund, R. F. Stein, and U. Torkelsson, “Dynamo-generated Turbulence and Large-Scale Magnetic Fields in a Keplerian Shear Flow,” *Astrophys. J.*, vol. 446, p. 741, June 1995.

[35] A. Brandenburg, K. Enqvist, and P. Olesen, "Large-scale magnetic fields from hydromagnetic turbulence in the very early universe," *Phys. Rev. D*, vol. 54, pp. 1291–1300, July 1996.

[36] A. Brandenburg and W. Dobler, "Pencil: Finite-difference Code for Compressible Hydrodynamic Flows." Astrophysics Source Code Library, October 2010.

[37] T. Kahniashvili, A. G. Tevzadze, and B. Ratra, "Phase Transition Generated Cosmological Magnetic Field at Large Scales," *Astrophys. J.*, vol. 726, p. 78, 2011.

[38] A. Brandenburg, T. Kahniashvili, and A. G. Tevzadze, "Nonhelical inverse transfer of a decaying turbulent magnetic field," *Phys. Rev. Lett.*, vol. 114, no. 7, p. 075001, 2015.

[39] T. Kahniashvili, A. G. Tevzadze, A. Brandenburg, and A. Neronov, "Evolution of Primordial Magnetic Fields from Phase Transitions," *Phys. Rev.*, vol. D87, no. 8, p. 083007, 2013.

[40] A. G. Tevzadze, L. Kisslinger, A. Brandenburg, and T. Kahniashvili, "Magnetic Fields from QCD Phase Transitions," *Astrophys. J.*, vol. 759, p. 54, 2012.

[41] T. Kahniashvili, A. Brandenburg, L. Campanelli, B. Ratra, and A. G. Tevzadze, "Evolution of inflation-generated magnetic field through phase transitions," *Phys. Rev.*, vol. D86, p. 103005, 2012.

[42] A. Mack, T. Kahniashvili, and A. Kosowsky, "Microwave background signatures of a primordial stochastic magnetic field," *Phys. Rev. D*, vol. 65, p. 123004, June 2002.

[43] T. Kahniashvili, Y. Maravin, and A. Kosowsky, "Faraday rotation limits on a primordial magnetic field from Wilkinson Microwave Anisotropy Probe five-year data," *Phys. Rev. D*, vol. 80, p. 023009, July 2009.

[44] T. Kahniashvili, Y. Maravin, G. Lavrelashvili, and A. Kosowsky, "Primordial Magnetic Helicity Constraints from WMAP Nine-Year Data," *Phys. Rev.*, vol. D90, no. 8, p. 083004, 2014.

[45] T. Kahniashvili and B. Ratra, "Effects of cosmological magnetic helicity on the cosmic microwave background," *Phys. Rev. D*, vol. 71, p. 103006, May 2005.

[46] T. Kahniashvili, Y. Maravin, A. Natarajan, N. Battaglia, and A. G. Tevzadze, "Constraining primordial magnetic fields through large scale structure," *Astrophys. J.*, vol. 770, p. 47, 2013.

[47] T. Kahniashvili, G. Gogoberidze, and B. Ratra, "Polarized Cosmological Gravitational Waves from Primordial Helical Turbulence," *Physical Review Letters*, vol. 95, p. 151301, October 2005.

[48] T. Kahniashvili, G. Gogoberidze, and B. Ratra, "Gravitational Radiation from Primordial Helical Magnetohydrodynamic Turbulence," *Physical Review Letters*, vol. 100, p. 231301, June 2008.

[49] T. Kahniashvili, L. Campanelli, G. Gogoberidze, Y. Maravin, and B. Ratra, "Gravitational radiation from primordial helical inverse cascade magnetohydrodynamic turbulence," *Phys. Rev. D*, vol. 78, p. 123006, December 2008.

[50] A. G. Tevzadze, G. D. Chagelishvili, G. Bodo, and P. Rossi, "Linear coupling of modes in two-dimensional radially stratified astrophysical discs," *Mon. Not. R. Astron. Soc.*, vol. 401, pp. 901–912, January 2010.

[51] A. G. Tevzadze, G. D. Chagelishvili, and J.-P. Zahn, "Hydrodynamic stability and mode coupling in Keplerian flows: local strato-rotational analysis," *Astron. Astrophys.*, vol. 478, pp. 9–15, January 2008.

[52] G. Bodo, A. Tevzadze, G. Chagelishvili, A. Mignone, P. Rossi, and A. Ferrari, "Stability and non-linear adjustment of vortices in Keplerian flows," *Astron. Astrophys.*, vol. 475, pp. 51–61, November 2007.

[53] T. Kahniashvili, A. Brandenburg, A. G. Tevzadze, and B. Ratra, "Numerical simulations of the decay of primordial magnetic turbulence," *Phys. Rev.*, vol. D81, p. 123002, 2010.

[54] H. Trac and U.-L. Pen, "A moving frame algorithm for high Mach number hydrodynamics," *New Astronomy*, vol. 9, pp. 443–465, July 2004.

[55] H. Trac and R. Cen, "Radiative Transfer Simulations of Cosmic Reionization. I. Methodology and In itial Results," *Astrophys. J.*, vol. 671, pp. 1–13, December 2007.

[56] H. Trac, R. Cen, and P. Mansfield, "SCORCH I: The Galaxy-Halo Connection in the First Billion Years," *Astrophys. J.*, vol. 813, p. 54, November 2015.

[57] T. Kahniashvili, A. Brandenburg, and A. G. Tevzadze, "Evolution of Primordial Magnetic Fields: From Generation Till Today," in *ICTP Summer School on Cosmology 2014 Trieste, Italy, August 4-21, 2014*, 2015.

[58] T. Kahniashvili, A. G. Tevzadze, S. K. Sethi, K. Pandey, and B. Ratra, "Primordial magnetic field limits from cosmological data," *Phys. Rev.*, vol. D82, p. 083005, 2010.

[59] L. Xiong, B. Póczos, J. Schneider, A. Connolly, and J. VanderPlas, "Hierarchical probabilistic models for group anomaly detection," in *International Conference on AI and Statistics (AISTATS)*, 2011.

[60] N. Leonenko, L. Pronzato, and V. Savani, "A class of Rényi information estimators for multidimensional densities," *Annals of Statistics*, vol. 36, no. 5, pp. 2153–2182, 2008.

[61] D. Pál, B. Póczos, and C. Szepesvári, "Estimation of Rényi entropy and mutual information based on generalized nearest-neighbor graphs," in *Proceedings of the Neural Information Processing Systems*, 2010.

[62] Q. Wang, S. Kulkarni, and S. Verdú, "Divergence estimation for multidimensional densities via $k$-nearest-neighbor distances," *IEEE Transactions on Information Theory*, vol. 55, no. 5, 2009.

[63] X. Nguyen, M. Wainwright, and M. Jordan., "Estimating divergence functionals and the likelihood ratio by convex risk minimization," *IEEE Trans. Information Theory*, 2010.

[64] M. Goria, N. Leonenko, V. Mergel, and N. Inverardi, "A new class of random vector entropy estimators and its applications in testing statistical hypotheses," *Journal of Nonparametric Statistics*, vol. 17, pp. 277–297, 2005.

[65] B. Póczos and J. Schneider, "On the estimation of $\alpha$-divergences," in *International Conference on AI and Statistics (AISTATS)*, 2011.

[66] J. Bentley, "Multidimensional binary search trees used for associative searching," *Commun. ACM*, vol. 18, pp. 509–517, 1975.

[67] G. J. Székely, M. L. Rizzo, and N. K. Bakirov, "Measuring and testing dependence by correlation of distances," *Annals of Statistics*, vol. 35, pp. 2769–2794, 2007.

[68] J. M. Steele, *Probability Theory and Combinatorial Optimization*. Society for Industrial and Applied Mathematics, 1997.

[69] J. Steele, "Growth rates of euclidean minimal spanning trees with power weighted edges," *Annals of Probability*, vol. 16, pp. 1767–1787, 1988.

[70] B. Póczos, S. Kirshner, and C. Szepesvári, "REGO: Rank-based estimation of Rényi information using Euclidean graph optimization," in *International Conference on AI and Statistics (AISTATS)*, 2010.

[71] J. Yukich, *Probability Theory of Classical Euclidean Optimization Problems*. Springer, 1998.

[72] A. Gretton, R. Herbrich, and A. Smola, "The kernel mutual information," in *Proc. ICASSP*, 2003.

[73] F. R. Bach and M. I. Jordan, "Kernel independent component analysis," *Journal of Machine Learning Research*, vol. 3, pp. 1–48, 2002.

[74] A. Gretton, K. Fukumizu, and B. Sriperumbudur, "Discussion of: Brownian distance covariance," *The Annals of Applied Statistics*, vol. 3, pp. 1285–1294, 2009.

[75] B. Póczos, L. Xiong, D. J. Sutherland, and J. G. Schneider, "Support distribution machines," *CoRR*, vol. abs/1202.0302, 2012.

[76] E. Perlman, R. Burns, Y. Li, and C. Meneveau, "Data exploration of turbulence simulations using a database cluster," in *Supercomputing SC*, 2007.

[77] B. P. E. X. H. T. S. H. J. S. Junier B. Oliva, Willie Neiswanger, "Fast function to function regression," in *International Conference on AI and Statistics (AISTATS)*, 2015.

[78] J. Oliva, W. Neiswanger, B. Poczos, J. Schneider, and E. Xing, "Fast distribution to real regression," in *International Conference on AI and Statistics (AISTATS)*, 2014.

[79] Z. Ghahramani and G. E. Hinton, "Parameter estimation for linear dynamical systems," tech. rep., 1996.

[80] D. Hsu, S. M. Kakade, and T. Zhang, "A spectral algorithm for learning hidden markov models," *Journal of Computer and System Sciences*, vol. 78, no. 5, pp. 1460 – 1480, 2012.

[81] T. Katayama, *Subspace Methods for System Identification*. Springer, 2005.

[82] D. Biskamp, *Magnetohydrodynamic Turbulence*. February 2008.

[83] P. A. Davidson, "Was loitsyansky correct? a review of the arguments?," *Journal of Turbulence*, vol. 1, pp. 1–14, 2011.

[84] A. S. Monin and A. M. Yaglom, "*Statistical Fluid Mechanics: Mechanics of Turbulence*, vol. 2 (mit press, cambridge, ma," 1975.

[85] M. Lesieur, "*Turbulence in Fluids*, (kluwer academic, dordrecht," 1997.

[86] S. S. Moiseev and O. G. Chkhetiani, "Helical scaling in turbulence," *Soviet Journal of Experimental and Theoretical Physics*, vol. 83, pp. 192–198, July 1996.

[87] V. Borue and S. A. Orszag, "Spectra in helical three-dimensional homogeneous isotropic turbulence," *Phys. Rev. E*, vol. 55, pp. 7005–7009, June 1997.

[88] R. H. Kraichnan, "Helical turbulence and absolute equilibrium," *Journal of Fluid Mechanics*, vol. 59, pp. 745–752, 1973.

[89] P. D. Ditlevsen and P. Giuliani, "Cascades in helical turbulence," *Phys. Rev. E*, vol. 63, p. 036304, March 2001.

[90] A. Rahimi and A. J. Chandy, "Direct numerical simulations of non-helical and helical stratified turbulent flows," *Journal of Turbulence*, vol. 17, pp. 1–29, January 2016.

[91] T. Brandenburg, A. and Kahniashvili and A. G. Tevzadze, "In preparation," 2015.

[92] I. Toselli, L. C. Andrews, R. L. Phillips, and V. Ferrero, "Angle of arrival fluctuations for free space laser beam propagation through non kolmogorov turbulence," in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 6551 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, p. 0, April 2007.

[93] I. Toselli, L. C. Andrews, R. L. Phillips, and V. Ferrero, "Free-space optical system performance for laser beam propagation through non-Kolmogorov turbulence," *Optical Engineering*, vol. 47, no. 2, p. 026003, 2008.

[94] E. Golbraikh, H. Branover, N. S. Kopeika, and A. Zilberman, "Non-Kolmogorov atmospheric turbulence and optical signal propagation," *Nonlinear Processes in Geophysics*, vol. 13, pp. 297–301, July 2006.

[95] N. S. Kopeika, A. Zilberman, and E. Golbraikh, "Generalized atmospheric turbulence: implications regarding imaging and communications," in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 7588 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, p. 8, February 2010.

[96] S. R. De Roode, H. J. J. Jonker, P. G. Duynkerke, and B. Stevens, "Countergradient fluxes of conserved variables in the clear convective and stratocumulus-topped boundary layer: The role of the entrainment flux," *Boundary-Layer Meteorol.*, vol. 112, pp. 179–196, 2004.

[97] A. Brandenburg, "Stellar mixing length theory with entropy rain," *arXiv:1504.03189*, April 2015.

[98] E. L. Rempel, A. C.-L. Chian, A. Brandenburg, P. R. Muñoz, and S. C. Shadden, "Coherent structures and the saturation of a nonlinear dynamo," *Journal of Fluid Mechanics*, vol. 729, pp. 309–329, August 2013.

[99] A. Brandenburg, I. Klapper, and J. Kurths, "Generalized entropies in a turbulent dynamo simulation," *Phys. Rev. E*, vol. 52, p. 4602, November 1995.

[100] J. K. Lawrence, A. A. Ruzmaikin, and A. C. Cadavid, "Multifractal Measure of the Solar Magnetic Field," *Astrophys. J.*, vol. 417, p. 805, November 1993.

[101] A. C. Cadavid, J. K. Lawrence, A. A. Ruzmaikin, and A. Kayleng-Knight, "Multifractal models of small-scale solar magnetic fields," *Astrophys. J.*, vol. 429, pp. 391–399, July 1994.

[102] Y. Narita, F. Sahraoui, M. L. Goldstein, and K.-H. Glassmeier, "Magnetic energy distribution in the four-dimensional frequency and wave vector domain in the solar wind," *Journal of Geophysical Research (Space Physics)*, vol. 115, p. 4101, April 2010.

[103] W. H. Matthaeus and M. L. Goldstein, "Measurement of the rugged invariants of magnetohydrodynamic turbulence in the solar wind," *J. Geophysical Research*, vol. 87, pp. 6011–6028, August 1982.

[104] J. Freidberg, *Plasma Physics and Fusion Energy*. Cambridge University Press, 2007.

[105] F. Turco and T. Luce, "Impact of the current profile evolution on tearing stability of iter demonstration discharges in diii-d," *Nuclear Fusion*, 2010.

[106] B. Poczos, L. Xiong, and J. Schneider, "Nonparametric divergence estimation with applications to machine learning on distributions," in *Uncertainty in Artificial Intelligence (UAI)*, 2011.

[107] M. Tesch, J. Schneider, and H. Choset, "Using response surfaces and expected improvement to optimize snake robot gait parameters," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2011.

[108] M. Tesch, J. Schneider, and H. Choset, "Expensive function optimization with stochastic binary outcomes," in *International Conference on Machine Learning (ICML)*, 2013.

[109] J. Oliva, B. Póczos, and J. Schneider, "Distribution to distribution regression," in *International Conference on Machine Learning (ICML)*, pp. 1049–1057, 2013.

[110] L. Xiong, B. Póczos, and J. Schneider, "Efficient learning on point sets," in *IEEE International Conference on Data Mining (ICDM'13)*, 2013.

[111] B. Póczos, A. Rinaldo, A. Singh, and L. Wasserman, "Distribution-free distribution regression," in *International Conference on AI and Statistics (AISTATS)*, 2013.

[112] S. Reddi and B. Póczos, "Scale invariant conditional dependence measures," in *International Conference on Machine Learning (ICML)*, 2013.

[113] Y. Zhang and J. Schneider, "Projection penalties: Dimension reduction without loss," in *International Conference on Machine Learning (ICML)*, 2010.

[114] Y. Zhang, J. Schneider, and A. Dubrawski, "Learning compressible models," in *In Proceedings of SIAM Data Mining (SDM) Conference*, 2010.

[115] Y. Zhang and J. Schneider, "Multi-label output codes using canonical correlation analysis," in *AISTATS*, 2011.

[116] L. Xiong, B. Poczos, J. Schneider, A. Connolly, and J. VanderPlas, "Hierarchical probabilistic models for group anomaly detection," in *AISTATS*, April 2011.

[117] L. Xiong, B. Poczos, and J. Schneider, "Group anomaly detection using flexible genre models," in *Neural Information Processing Systems (NIPS)*, 2011.

[118] V. A. Acciari, E. Aliu, T. Arlen, T. Aune, M. Beilicke, W. Benbow, D. Boltuch, V. Bugaev, A. Cannon, L. Ciupik, P. Cogan, P. Colin, R. Dickherber, A. Falcone, S. J. Fegan, J. P. Finley, P. Fortin, L. F. Fortson, A. Furniss, D. Gall, G. H. Gillanders, J. Grube, R. Guenette, G. Gyuk, D. Hanna, J. Holder, D. Horan, C. M. Hui, T. B. Humensky, A. Imran, P. Kaaret, N. Karlsson, M. Kertzman, D. Kieda, J. Kildea, A. Konopelko, H. Krawczynski, F. Krennrich, M. J. Lang, S. LeBohec, G. Maier, P. Moriarty, R. Mukherjee, R. A. Ong, A. N. Otte, D. Pandel, J. S. Perkins, A. Pichel, M. Pohl, J. Quinn, K. Ragan, P. T. Reynolds, H. J. Rose, M. Schroedter, G. H. Sembroski, A. W. Smith, D. Steele, S. P. Swordy, M. Theiling, J. A. Toner, A. Varlotta, V. V. Vassiliev, S. Vincent, R. Wagner, S. P. Wakely, J. E. Ward, T. C. Weekes, A. Weinstein, T. Weisgarber, D. A. Williams, S. Wissel, M. Wood, and B. Zitzer, "Multiwavelength Observations of the Very High Energy Blazar 1ES 2344+514," *Astrophys. J.*, vol. 738, p. 169, September 2011.

[119] V. A. Acciari, E. Aliu, T. Arlen, T. Aune, M. Beilicke, W. Benbow, D. Boltuch, S. M. Bradbury, J. H. Buckley, V. Bugaev, K. Byrum, A. Cannon, A. Cesarini, L. Ciupik, W. Cui, R. Dickherber, C. Duke, A. Falcone, J. P. Finley, G. Finnegan, L. Fortson, A. Furniss, N. Galante, D. Gall, G. H. Gillanders, S. Godambe, J. Grube, R. Guenette, G. Gyuk, D. Hanna, J. Holder, C. M. Hui, T. B. Humensky, A. Imran, P. Kaaret, N. Karlsson, M. Kertzman, D. Kieda, A. Konopelko, H. Krawczynski, F. Krennrich, M. J. Lang, G. Maier, S. McArthur, M. McCutcheon, P. Moriarty, R. A. Ong,

A. N. Otte, M. Ouellette, D. Pandel, J. S. Perkins, A. Pichel, M. Pohl, J. Quinn, K. Ragan, L. C. Reyes, P. T. Reynolds, E. Roache, H. J. Rose, A. C. Rovero, M. Schroedter, G. H. Sembroski, G. D. Senturk, D. Steele, S. P. Swordy, M. Theiling, S. Thibadeau, A. Varlotta, V. V. Vassiliev, S. Vincent, R. G. Wagner, S. P. Wakely, J. E. Ward, T. C. Weekes, A. Weinstein, T. Weisgarber, D. A. Williams, S. Wissel, M. Wood, B. Zitzer, A. Garson, III, K. Lee, A. C. Sadun, M. Carini, D. Barnaby, K. Cook, J. Maune, A. Pease, S. Smith, R. Walters, A. Berdyugin, E. Lindfors, K. Nilsson, M. Pasanen, J. Sainio, A. Sillanpaa, L. O. Takalo, C. Villforth, T. Montaruli, M. Baker, A. Lahteenmaki, M. Tornikoski, T. Hovatta, E. Nieppola, H. D. Aller, and M. F. Aller, "TeV and Multi-wavelength Observations of Mrk 421 in 2006-2008," *Astrophys. J.*, vol. 738, p. 25, September 2011.

[120] V. A. Acciari, T. Arlen, T. Aune, M. Beilicke, W. Benbow, M. Böttcher, D. Boltuch, S. M. Bradbury, J. H. Buckley, V. Bugaev, and et al., "Spectral Energy Distribution of Markarian 501: Quiescent State Versus Extreme Outburst," *Astrophys. J.*, vol. 729, p. 2, March 2011.

[121] L. Campanelli, G. L. Fogli, T. Kahniashvili, A. Marrone, and B. Ratra, "Galaxy cluster number count data constraints on cosmological parameters," *Eur. Phys. J.*, vol. C72, p. 2218, 2012.

[122] O. Avsajanishvili, N. A. Arkhipova, L. Samushia, and T. Kahniashvili, "Growth Rate in the Dynamical Dark Energy Models," *Eur. Phys. J.*, vol. C74, no. 11, p. 3127, 2014.

[123] T. Kahniashvili, A. Kar, G. Lavrelashvili, N. Agarwal, L. Heisenberg, and A. Kosowsky, "Cosmic Expansion in Extended Quasidilaton Massive Gravity," *Phys. Rev.*, vol. D91, no. 4, p. 041301, 2015.

[124] L. Kisslinger and T. Kahniashvili, "Polarized Gravitational Waves from Cosmological Phase Transitions," *Phys. Rev.*, vol. D92, no. 4, p. 043006, 2015.

[125] S. Aiola, B. Wang, A. Kosowsky, T. Kahniashvili, and H. Firouzjahi, "Microwave Background Correlations from Dipole Anisotropy Modulation," *Phys. Rev.*, vol. D92, p. 063008, 2015.

[126] O. Avsajanishvili, L. Samushia, N. A. Arkhipova, and T. Kahniashvili, "Testing Dark Energy Models through Large Scale Structure," *ArXiv e-prints*, November 2015.

[127] T. Kahniashvili, A. Brandenburg, A. G. Tevzadze, and T. Vachaspati, "In preparation," 2015.

[128] T. Kahniashvili, A. Brandenburg, R. Durrer, A. G. Tevzadze, and W. Yin, "In preparation," 2015.

[129] T. Brandenburg, A. and Kahniashvili, "In preparation," 2015.

[130] A. D'Aloisio, M. McQuinn, and H. Trac, "Large Opacity Variations in the High-redshift Ly$\alpha$ Forest: The Signature of Relic Temperature Fluctuations from Patchy Reionization," *Astrophys. J. Lett.*, vol. 813, p. L38, November 2015.

[131] P. La Plante, N. Battaglia, A. Natarajan, J. B. Peterson, H. Trac, R. Cen, and A. Loeb, "Reionization on Large Scales. IV. Predictions for the 21 cm Signal Incorporating the Light Cone Effect," *Astrophys. J.*, vol. 789, p. 31, July 2014.

[132] M. Clute, A. Singh, B. Poczos, and T. Verstynen, "The predictive value of functional connectivity," in *Annual Meeting of the Organiz ation for Human Brain Mapping (OHBM)*, 2014.

[133] "Explorations in Computer Science for High School educators," http://www.cs.cmu.edu/cs4hs/.

[134] "Opportunities for Undergraduate Research in Computer Science (OurCS)," http://www.cs.cmu.edu/ourcs/.