Sample Research Statement, Science

RESEARCH STATEMENT

Monique R. Researcher

PAST AND CURRENT RESEARCH

Theoretical cosmology strives to understand the structure and evolution of the universe. Three pieces of this puzzle with which I have worked are the law of gravity on megaparsec scales, the power spectrum of the initial density fluctuations after the big bang, and the mass function of low-mass galaxy clusters.

The law of gravity has been carefully tested and verified only up to solar-system scales. Persuasive evidence now indicates that the universe is not only expanding, but that the rate of expansion is increasing. Often this acceleration is attributed to an unknown particle or field dubbed "dark energy." Some papers (e.g., [1]) propose that the recent accelerated expansion of the universe could be due to changes in general relativity on large scales, instead of new fields or particles. My research [3] places the first constraints on how gravity behaves at megaparsec scales (scales that span clusters of galaxies).

Naturally, one cannot use orbital motion or laser ranging to test gravity at megaparsec scales. We use the evolution of structure formation: how the overdensities in the universe grow with time. This involves modeling matter as a pressure-less, nearly-homogeneous fluid, and computing how density perturbations change under the influence of gravity. The overdensities grow differently under a different law of gravity. We examine a couple of models in which Newton's law is modified at a huge physical scale, and work to first order and second order in the overdensity. We compare the power spectrum for a modified gravity law with observations, thus placing bounds on how greatly gravity could deviate from an inverse-square law on large scales. We find no deviation, but the constraints are weak. We also compute the change in the bispectrum (the fourier transform of the three-point correlation function), which involves second-order effects in the overdensity.

On the small-scale end of things, cosmology also addresses how the initial fluctuations in the universe were created. The most popular theory currently holds that the inhomogeneities were created by the quantum fluctuations of one or more homogeneous particle fields, and then quickly expanded to classical sizes during a period of rapid inflation of the universe. To test inflation and its competing theories, one needs to know the spectrum of the initial fluctuations. However, we cannot directly observe this primordial power spectrum. Until the universe cooled enough for hydrogen to form, the universe was an opaque, ionized plasma. The WMAP satellite has carefully measured the anisotropies in the cosmic microwave background (CMB), the light freed at the last-scattering surface (when hydrogen formed and the universe became transparent, three hundred thousand years after the big bang) [2]. Since the matter and radiation were strongly coupled before the last scattering, we can therefore infer the matter power spectrum at that time. We also know the physics of how the fluctuations evolved during those first three hundred thousand years: it is a complex set of transfer equations that can be computed with CMBFAST [5].

While a power-law primordial power spectrum fits the data very well, there remains the tantalizing possibility that there could be structure in the primordial power spectrum. Any spectral features would shed light on theories of the universe's birth. In my paper [4], we develop and apply a model-independent method to reconstruct the primordial power spectrum. We adapt and apply the smoothing spline method, a nonparametric technique to reconstruct a continuous function from discrete, noisy data. Using Markov Chain Monte Carlo to explore the likelihood of primordial power spectrum shapes, we find no statistically significant indication of deviations from a power-law spectrum. Our results also examine the degeneracy between the primordial power spectrum shape and the values of the cosmological parameters. We find virtually no difference in the matter density value and error when we give the primordial power spectrum a lot more freedom, but the errors on the optical depth and baryon fraction increase significantly.

For my current project, I am developing a method to help determine the mass function of low-mass galaxy clusters. Since the formation of clusters depends exponentially on a function of the matter density and dark energy, their mass function provides a sensitive probe of cosmological parameters. The Atacama Cosmology Telescope will conduct a large survey of galaxy clusters, observing their Sunyaev-Zel'dovich decrement (SZD). The SZD is caused by electrons in the cluster scattering the CMB photons to higher frequencies. It is roughly proportional to the number of ionized electrons in a cluster, which in turn is proportional to the number of baryons; thus the SZD provides an indication of the cluster's mass. However, we do not know for sure how well the baryons trace the mass for low-mass clusters. For high-mass, low-redshift clusters, weak gravitational lensing can provide a good estimate of the total cluster mass. For a small cluster (10^{14} solar masses), however, the weak-lensing signal is weak, especially at high redshifts. My project aims to combine the weak-lensing signals from several small clusters to increase the signal-to-noise, and thus establish the relationship between SZD and mass for low-mass clusters.

FUTURE RESEARCH

There are several natural extensions of my thesis work I will pursue in the near future. I have already derived an expression for the matter bispectrum under a perturbed law of gravity, which provides an independent constraint on the gravitational force law. I will place stronger limits on possible deviations from the inverse-square law by applying the bispectrum expression to observations, and combining it with my previous results from the power spectrum. The smoothing-spline analysis of the primordial power structure could be enhanced by incorporating large-scale structure data. Additionally, in my previous work, I kept the neutrino mass equal to zero. I plan to perform a nonparametric primordial power spectrum reconstruction while treating the neutrino mass and number of species as free parameters. This will explore degeneracies between the neutrino mass and primordial power spectrum shape, and truly test how much the CMB can say about neutrinos (and vice versa). With my current project, I would like to generalize assumptions about the shapes of galaxy clusters (circular lenses with NFW profiles) to stack their weak-lensing signals. I will also explore ways to improve the signal-to-noise of cosmological parameters from cluster surveys. Rather than translating the observational data into a mass function, and using the mass function to constrain cosmology, perhaps we could find a cosmology-constraining function more directly tied to the observations.

On a longer time scale, I plan to explore Lorentz-covariant modifications to the general relativity (GR) lagrangian at large scales. Many scalar-tensor theories of gravity are mathematically equivalent to modifications of general relativity without a scalar field. If I could find a paramerization of a Lorentz-covariant GR modification consistent with data, I could use the power

M. R. Researcher: Research Statement—3

spectrum and bispectrum from large-scale structure observations to constrain the theory. I'm also fascinated by holography: how the dependence of entropy on surface area, rather than volume, might constrain and imply theories of gravity.

I'm also excited to further develop the statistical tools I encountered while working with the smoothing-spline method. The primordial power spectrum reconstruction poses a challenge to the smoothing-spline, because a complicated nonlinear function lies between the data and the desired function. I chose to use a cross-validation method to select the smoothing parameter for the spline. (The smoothing parameter determines the balance between allowing the reconstructed function to bed to maximize the likelihood, without giving it too much freedom to fit noise in the data.) Since cross-validation is computationally expensive, I would like to derive an analytical approximation in this context. There are also alternative methods worth optimizing and exploring further, such as running another Markov Chain in which the smoothing parameter is a free parameter. Additionally, I plan to learn and apply a more rigorous computation of Bayesian evidence. I expect to find many other problems in astrophysics, and possible other fields, that could greatly benefit from these statistical methods.

UNDERGRADUATE INVOLVEMENT

Touching the extremes of human knowledge on both large and small scales, cosmology is an awe-inspiring field with an influx of data, and it offers many exciting and accessible research projects for undergraduates. I would love to involve bright undergraduates in the simple extensions of my past and current research, described above. Further development and application of the smoothing-spline method can provide many important undergraduate-level projects. One example of such a project would be to create a series of two-dimensional images, along with plots of their correlation functions and power spectra, such that they give an intuition for what the power spectrum means. Such projects would allow freshmen to develop tools they would need for research, without requiring more advanced coursework.

REFERENCES

[1] Sean M. Carroll, Vikram Duvvuri, Mark Trodden, and Michael S. Turner. Is cosmic speed-up due to new gravitational physics? *Phys. Rev.*, D70:043528, 2004.
[2] G. Hinshaw et al. First year Wilkinson microwave anisotropy probe (wmap) observations: Angular power spectrum. *Astrophys. J. Suppl.*, 148:135, 2003.
[3] Monique R. Researcher, Name, and Name. Limits on deviations from the inverse-square law on megaparsec scales. *Phys. Rev.*, D71:083004, 2005.
[4] Monique R. Researcher, Name, and Name. Smoothing spline primordial power spectrum reconstruction. 2005.
[5] U. Seljak and M. Zaldarriaga. A line of sight approach to cosmic microwave background anisotropies, *Astrophys. J.*, 469:437, 1996.

MONIQUE R. RESEARCHER. RESEARCHER, DEPARTMENT OF PHYSICS AND ASTRONOMY, UNIVERSITY OF X, CITY, STATE, ZIP CODE. EMAIL ADDRESS

Sample Research Statement, Mathematics

Research Statement **John Patrick Scholar**

Current and Future Research

I analyze algorithms and data structures using analytic combinatorics and probabilistic methods. I am also greatly interested in information theory (see "Dissertation" below) and game theory (see my Teaching Statement).

My research is currently supported by an NSF grant: Asymptotic Enumeration, Reinforcement, and Effective Limit Theory (2006-2008). Name and I are Co-Principal Investigators. Our goal is the asymptotic estimation of constants for which a multivariate generating function is known. Our research in this direction can be roughly divided into two parts: we use computational symbolic algebra to implement complex analytic methods of analysis; we also continue to apply further techniques of integration of meromorphic forms so that the above analytic methods can be applied to a wider class of generating functions.

Also, I am analyzing the profile of suffix trees. The profile of a tree describes the number of nodes at a certain level of the tree. This parameter is quite powerful because it easily yields a great deal of information about many other parameters of the tree, including depth, fill-up level, height, shortest path, and size. A complete characterization of the profile of suffix trees seems feasible to me for two reasons: Recently, Name (Institution), et al., successfully analyzed the profile of a tree. Also, in my Ph.D. dissertation and in my recent paper with Name (International Institution), I was able to successfully translate several results about trees into new results about suffix trees.

I am collaborating with Name and Name (International Institution) on a problem concerning the unique minima in a collection of independent, identically distributed geometric random variables, where the mean of each depends on the total number of variables. We utilize an exponential diagonal depoissonization scheme.

Name (International Institution) coined the term "scholar-parameter" for the parameter of suffix trees I analyzed in my Ph.D. dissertation. We are working with Name, Name, and Name (International Institution) to analyze the behavior of this new parameter on a variety of other trees, including binary search trees, digital search trees, simply generated trees, and the cumulative scholar-parameter in trees.

Dissertation

My Ph.D. dissertation completed the resolution of an information theory problem: the incorporation of error-correction into an adaptive data compression scheme without the introduction of additional redundancy. Traditionally, adaptive data compression algorithms examine the uncompressed portion of a file, searching for subsequences that occur in a database; compression occurs as subsequences are replaced by pointers into the database. The Lempel-Ziv '77 algorithm is one such scheme with many desirable theoretical and practical properties; its compression ratio is asymptotically optimal, and the algorithm enjoys widespread use. Unfortunately, adaptive data compression algorithms traditionally have no error resilience. Transmission errors in a compressed file are propagated when the file is decompressed (so compressed files are even more susceptible to errors than uncompressed files). The standard remedy is to add redundant bits for error-correction.

My research proves that sufficiently many redundant bits are already present in the Lempel-Ziv '77 algorithm to allow substantial error-correction with no additional overhead. This surprising result is possible because the LZ '77 encoder is unable to completely decorrelate the input sequence. My research was done in collaboration with Name (Institution) and Name (Institution). In my dissertation, I prove that the number of pointers into the database in LZ'77 follows the logarithmic series distribution. My proof includes an extensive analysis and comparison of suffix trees and trees built over independent strings.