Draft version September 23, 2025
Typeset using LaTeX **twocolumn** style in AASTeX631

Bayesian Inference and MCMC Methods in Astrophysics

Agastya Gaur[1]

[1] *University of Illinois at Urbana-Champaign*

ABSTRACT

This is the abstract of the paper. It summarizes the work in a concise form.

Contents

1. INTRODUCTION

In the 4th century BC, Hipparchus, attempting to estimate the length of a year, found the middle of the range of a scattered set of Babylonian solstice measurements. Though an acheivement in its own right for the time, Hipparchus's measurement marked the beginning of what would become a long standing marriage between astronomy and statistics. In the centuries to come, a number of breakthroughs in astrostatistics would continue to occur, with Brahe successfully using the mean of a dataset to increase precision of measurements and Laplace rediscovering the work of Thomas Bayes and applying his statistical theories extensively to astronomical problems. Most notably, in the early 1800s, Legendre developed least squares parameter estimation to model the orbit of comets (Feigelson & Babu 2004). By the end of the 19th century, astronomy had firmly established itself as a quantitative science, driven by the refinement of statistical methods to confront the uncertainties inherent in measurement.

The next 100 years brought two developments that reshaped this tradition: the rise of physics as the explanatory foundation of astronomy, and the advent of computing, which enabled unprecedented scales of quantitative analysis. As astronomy grew increasingly intertwined with the theories of physics, the field transformed into what we now call astrophysics. This shift did not replace the statistical tradition but expanded it, integrating new forms of quantitative reasoning with physical modeling. Advances in computing increased the scale of the statistical analysis that was feasible to perform, and since then it has only been rising. While the early history of the field was dominated by statistical reasoning, the growth of physics and computation broadened this into what we now call quantitative analysis (QA): a synthesis of statistical inference, numerical modeling, and data-driven computation.

Today, astrophysics sits in the middle of a universe of complex statistical problems that demand new quantitative approaches and more computing power by the day. In many respects, QA has become the backbone of research in modern astrophysics, and at a critical moment indeed, as the 21st century has ushered in an unprecedented era of astronomical data generation. Sky surveys like Gaia DR3 alone provide astrometry and photometry for nearly two billion stars, plus more than ten million variable sources across dozens of types (Gaia Collaboration et al. 2023). Advances in CCD detectors will see data from sky surveys continue to grow in the next decade from an order of gigabytes to terabytes, and possibly petabytes in the future. The same trend can be seen in data from the Large Synoptic Survey Telescope and NASA's Solar Dynamics Observatory (Borne 2009). This so-called "data deluge" is what makes QA so important to astrophysics today. The ability to extract meaningful insights from these massive datasets in an organized manner is crucial for advancing our understanding of the universe. QA provides a number of powerful tools spanning statistical inference, computational algorithms, and machine learning methodologies to analyze, interpret, and model this data effectively.

Within this broad landscape of quantitative techniques, the one that stands out as particularly impactful is Bayesian inferencing via Monte Carlo Markov Chain (MCMC) methods. For astrophysicists, Bayesian inference with MCMC has become one of the most widely used and versatile approaches for tackling high-dimensional, noisy problems. Von Toussaint (2011) notes the growing applicability of Bayesian inferencing in physics. Computational models are becoming far more complex, with high-dimensional parameter spaces, and the data being analyzed is often noisy and incomplete. Bayesian methods, with their ability to incorporate prior knowledge and handle uncertainty, are well-suited to these challenges. MCMC methods, in particular, provide a practical way to sample from complex posterior distributions that arise in Bayesian analysis. This makes them invaluable for parameter estimation, model comparison, and uncertainty quantification in almost any astrophysical problem.

The rest of the paper will have the following structure: Sec. II will provide a foundational explanation of Bayesian statistics as well as the mathematical and computational methodology behind MCMC methods. Next, Sec. III will introduce three case studies within astrophysics where Bayesian inferencing and MCMC methods are being used to push research forward. These concepts include the direct detection of exoplanets, CMB parameter estimation, and gravity wave fitting. Each case study will include current challenges in the field, how Bayesian inferencing is being used to address it, the pros and cons of the approach, as well as future advancements that could be made. Finally, Sec. IV will include a discussion on how Bayesian inferencing is being used throughout astrophysics overall and how it can address problems in other fields as well.

REFERENCES

Borne, K. D. 2009, Astroinformatics: A 21st Century Approach to Astronomy, arXiv, doi: 10.48550/arXiv.0909.3892

Feigelson, E. D., & Babu, G. J. 2004, Statistical Challenges in Modern Astronomy, arXiv, doi: 10.48550/arXiv.astro-ph/0401404

Gaia Collaboration, Vallenari, A., Brown, A. G. A., et al. 2023, Astronomy & Astrophysics, 674, A1, doi: 10.1051/0004-6361/202243940

Von Toussaint, U. 2011, Reviews of Modern Physics, 83, 943, doi: 10.1103/RevModPhys.83.943