Searching for Unspecified Structure in Astrophysical Images:
Maintaining Computational Efficiency and Statistical Validity

Nathan Stein*
University of Pennsylvania, Philadelphia, USA - natstein@wharton.upenn.edu

David A van Dyk
Imperial College London, UK - dvandyk@imperial.ac.uk

Xiao Li Meng
Harvard University, Cambridge, Massachusetts, USA - meng@stat.harvard.edu

Vinay Kashyap
Harvard-Smithsonian Center for Astrophysics, Cambridge, Massachusetts, USA - vkashyap@cfa.harvard.edu

Aneta Siemiginowska
Harvard-Smithsonian Center for Astrophysics, Cambridge, Massachusetts, USA - asiemiginowska@cfa.harvard.edu

Unexpected structure in images of astronomical sources often presents itself upon visual inspection of the image. Researchers would like a formal statistical test of whether such structure corresponds to actual features in the source or can be attributed to noise in the data. To avoid a biased test, the test must be neutral regarding what constitutes structure—tailoring the alternative hypothesis to a particular feature in the data results in a Type I error rate greater than the significance level. We present a neutral test of this sort for pixelated images with Poisson noise. To infer image structure, we conduct a Bayesian analysis of a full model that uses a multi-scale component to allow flexible departures from the posited null model. We consider various test statistics that are summaries of the posterior distribution under this full model. Evaluating the null distribution of such test statistics requires us to sample images under the null hypothesis and fit the full model using MCMC for each sampled image. This is a computationally demanding task and we present two novel techniques to reduce computational time. First, using a tail probability of the posterior distribution under the full model as our test statistic,we estimate an upper bound on a p-value that allows us to increase the significance level with a fixed number of resampled images. Second, for more general test statistics, we consider prematurely terminating the MCMC sampler when computing the test statistic. This strategy can substantially reduce computational costs while providing valid p-values in terms of their Type I error, but may reduce the power of the test. We demonstrate that this trade-off may beworthwhile; a 65% increase in computational power, for example, may be traded for a 15% reduction in statistical power.

Keywords: Image Analysis; Multiscale Models; Monte Carlo P-values; Type I Error; Power; Markov Inequality; Quasars; Jets