# Bayesian Inference and Monte Carlo Method

**Please submit your homework in pdf form, which can be either a scanned copy of your hand-written answer, or, computer generated document. In either case, please include computer calculated results, necessary scientific figures along with description of the method used and the conclusion/solution to the problem.**

### 1. Is CDM substructure a plausible theory to explain the gravitational "flux-ratio anomaly" observations? (6 pts)

The cold dark matter (CDM) cosmology has been a prevailing theory describing the matter content of the Universe. Together with the "cosmological constant $\Lambda$", which describes the cosmic mass-energy budget in form of dark energy, the theory of $\Lambda$CDM can explain the distribution of large scale structures in the Universe. Galaxies are also expected to form and evolve inside dark matter halos in their large-scale environment. CDM predicts a great number of dark matter halos surviving galaxy merge and manifest themselves as substructures in big galaxy halos. They are often refered to as "subhalos". Their "dark" presence adds density perturbation/fluctuation to the smooth gravitational potential of their host galaxy halos.

Strong gravitational lensing can be used to test the presence of substructures in lensing galaxies, and through the constrained abundance to distinguish the CDM scenario and alternative dark matter cosmologies. In the strong lensing regime, the light from a background galaxy or quasar will be splitted into several paths as it travels through the deep gravitational potential of the foreground lensing galaxy. This simply causes the background source to cast into multiple copies of images, each being magnified or demagnified depending on where it appears (in projection) with respect to the lensing galaxy. Astronomers used to fit the *positions* of the multiple images to constrain the overall mass density distribution of the lensing galaxy. Interestingly, very often the best fitted models (using image positions) failed to reproduce the relative magnifications/fluxes among the multiple images. The phenomena were observed over decades and referred to as the "flux-ratio anomaly" problem. Astronomers came to understand that smaller-scale density perturbations in the lens may produce flux deviations (which are more severe than position fluctuations). In particular, the CDM substructures sitting in the lens galaxy are capable of causing the "flux-ratio anomaly" problem.

Here is the question: cosmological $N$-body simulations predict that the CDM subhalo populations in galaxies may cause a maximum probability of 25% for strong lensing galaxies to exhibit "flux-ratio anomalies". Current observerations have in total observed 5 out of 8 multiply lensed quasar systems with clear evidence of "flux-ratio anomalies" (note, observers know exactly which five out of eight behave so). Using (1) maximum likelihood estimate (assuming Gaussian approximation) and (2) Bayesian analysis assuming a flat prior for your model, can you rule out the proposed theoretical explanation at a confidence level of $\alpha = 5\%$? Please write down the key formula, the relevant statistical distributions, your calculations, reasoning and results.

Note, for both methods your answers will be "yes". Astronomers came to understand that there are other culprits of a baryonic origin may also perturb the lensing potential or affect the radio fluxes as the emission propagates through the interstellar and intergalactic medium.

**2. Sampling a given distribution with a Monte Carlo Markov Chain (6 pts)**

Let us assume we want to generate random numbers from the distribution:

$$p(x) \propto \exp\left(-[2x + 3\cos^2(x)]^2\right). \tag{1}$$

Write your own program to implement this by using a stochastic process constructed with the Metropolis algrorithm:

(a) Start with some random guess $x_0$ for which $p(x_0)$ is not zero.

(b) Make a proposal for $x_i$ in your chain by *adding* a random number drawn uniformly from the interval $[-1, 1]$ to $x_{i-1}$.

(c) Accept the proposal with probability:

$$r = \min\left(1, \frac{p(x_i)}{p(x_{i-1})}\right), \tag{2}$$

i.e., in the case of acceptance make it the entry $x_i$ in your Monte Carlo chain. Otherwise, adopt the unmodified $x_{i-1}$ as your element $x_i$. Then proceed with the next element $i + 1$.

(d) Produce a chain with $N = 10^6$ elements, and make a histogram with bin size $\Delta x = 0.02$ of the entries. In order to verify that they correctly sample the distribution, overplot the given distribution. How many unique points are in your chain?

**3. Generate an NFW distribution for dark matter halo (6 pts)**

N-body simulations of cosmic structure formation under the standard cold dark matter cosmology predict that the radial density distribution of a spherical dark matter halo is given by the Navarro-Frenk-White (NFW) profile (https://arxiv.org/pdf/astro-ph/9707093.pdf):

$$\rho^{\mathrm{DM}}(r) = \rho_{\mathrm{cr}} \frac{\delta_c}{(r/r_s)(1 + r/r_s)^2}, \tag{3}$$

where $r_s$ is a scale radius, $\delta_c$ is characteristic (dimensionless) density, and $\rho_{\mathrm{cr}} = 3H^2/(8\pi G)$ is the critical density of the Universe, determined by the Hubble parameter $H$, both are functions of the cosmic time. We introduce a radius $r_{200}$, within which the mean density is $200\rho_{\mathrm{cr}}$, i.e., the mass enclosed within $r_{200}$ is given by $M_{200} = 200\rho_{\mathrm{cr}}(4\pi/3)r_{200}^3$. With the help of $r_{200}$, we can further define a concentration parameter $C \equiv r_{200}/r_s$ and a normalized radius $x \equiv r/r_{200}$. The cumulative mass distribution $M^{\mathrm{DM}}(\leqslant r)$ can be written as:

$$M^{\mathrm{DM}}(\leqslant r) = 4\pi\rho_{\mathrm{cr}}\delta_c r_s^3 \left[\frac{1}{1 + Cx} - 1 + \ln(1 + Cx)\right]. \tag{4}$$

For any given $C$ and $r_s$, the density normalization is constrained by $M^{\mathrm{DM}}(r_{200}) = M_{200}$. This yields $\delta_c$ given in terms of $C$:

$$\delta_c = \frac{200}{3} \frac{C^3}{\ln(1 + C) - C/(1 + C)}. \tag{5}$$

Now use Monte Carlo method to generate 1,000,000 particles that follow a NFW distribution inside a radius of 30 kpc, for a dark matter halo with $r_s = 20$ kpc and $C = 10$, living in the present day, for which the Hubble constant is $H_0 = 70$ km/s/Mpc. Plot the logarithmic radial density $(\log[\rho^{\mathrm{DM}}(r)/(M_\odot \mathrm{kpc}^{-3})])$ distribution of the particle realization as a function of the logarithmic radius $\log[r/\mathrm{kpc}]$ and compare it to Eq. (3).