Comparative Study of Sampling Methods in Bayesian Inference for Gamma-ray Bursts

Kaori Obayashi (Aoyama Gakuin University: AGU) with Ryo Yamazaki (AGU), Yo Kusafuka (ICRR)

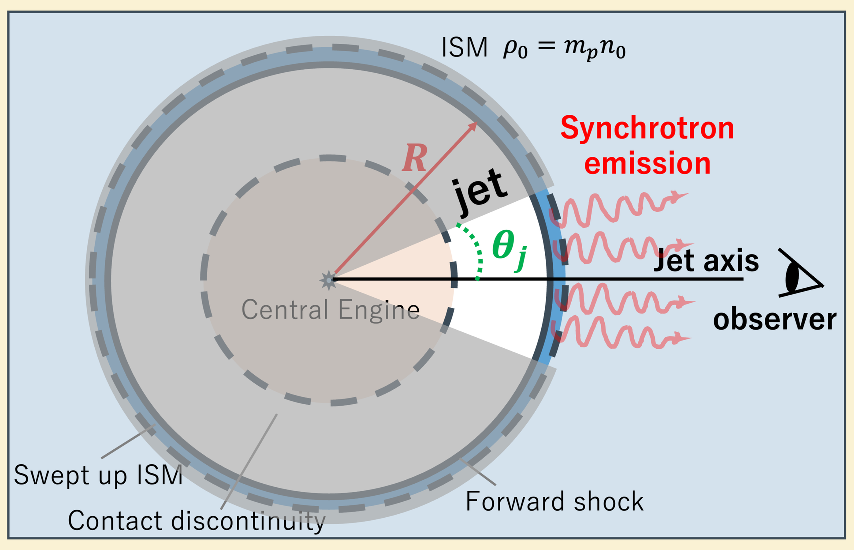
ABSTRACT

In recent years, Bayesian inference has been actively applied to estimate model parameters for gamma-ray bursts, with Markov Chain Monte Carlo (MCMC) methods being widely used for sampling. Alternatively, an approach known as Clustered Nested Sampling (CNS) has also been proposed for this purpose. Since these are independent methods, it is essential to compare their characteristics, advantages, and limitations to determine the optimal approach for parameter estimation. This study aims to evaluate both methods to provide insights that aid in selecting the most suitable technique.

1. Motivation

Gamma-ray burst (GRB) is the most luminous explosion in electromagnetic bands with gamma-ray prompt emission and multi-wavelength afterglow emission. These emissions are thought to originate from radiation from relativistic jet ejected by collapsing massive stars or binary-neutron star mergers, but the radiation mechanism and jet structure are not yet understood.

GRB afterglow theory has a lot of model parameters, therefore, Bayesian Inference to estimate unknown parameters based on observed data.



Standard model parameters	
θ_v	Viewing angle
θ_j	Half-opening angle of jet
Γ_0	Initial Lorentz factor jet central axis
E_0	Initial isotropic equivalent energy on-axis
n_0	Number density of ISM
p	Electron distribution power-law index
ε_B	Thermal energy fraction in electrons
ε_e	Thermal energy fraction in magnetic field
ξ_N	Fraction of electrons that get accelerated
More complicated Model parameters	
θ_c	Half-width of the jet core
k	Spectral index of number density ISM
…	

2. Bayes' theorem

Bayesian Inference is a statistical method for estimating unknown parameters based on observed data. It uses Bayes' theorem to combine prior knowledge (the prior distribution) with new data (likelihood) to update beliefs about the parameter values, resulting in a posterior distribution:

$$p(\vec{\Theta}|\vec{X}) = \frac{L(\vec{X}|\vec{\Theta})\pi(\vec{\Theta})}{p(\vec{X})} = \frac{L(\vec{X}|\vec{\Theta})\pi(\vec{\Theta})}{\int_{\Theta} L(\vec{X}|\vec{\Theta})\,\pi(\vec{\Theta})d\Theta}$$

$\vec{X}$: observed data, $\vec{\Theta} = (\Theta_1, \Theta_2, \dots, \Theta_D)$: D-dimensional model parameters.

$p(\vec{\Theta}|\vec{X})$: Posterior probability distribution.

$L(\vec{X}|\vec{\Theta})$: Likelihood function.

$\pi(\vec{\Theta})$: Prior function.

Useful for inferring the most likely "cause" from observation.

To obtain the posterior distribution using Monte Carlo methods, it is necessary to adopt **efficient sampling techniques**, as exhaustive search across all values in a high-dimensional space is impractical. In this study, we discuss two methods; Markov Chain Monte Carlo and Clustered Nested Sampling method.

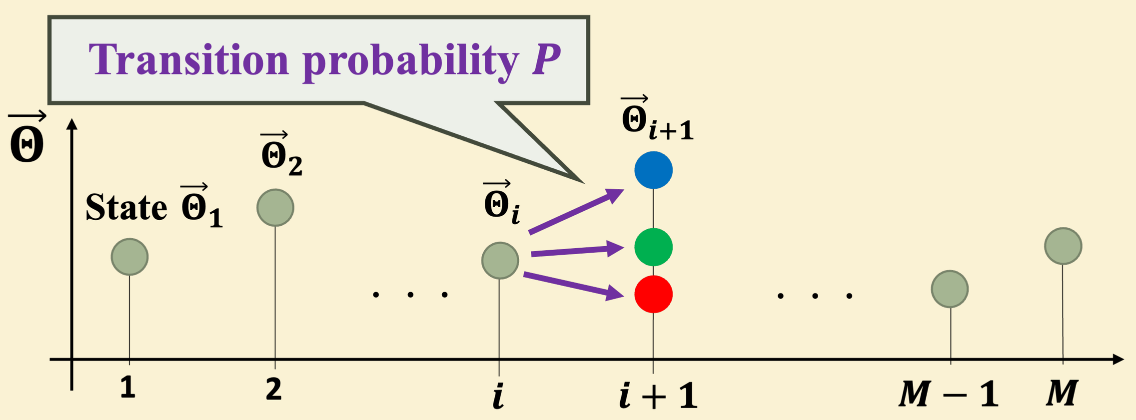
3. Markov Chain Monte Carlo (MCMC)

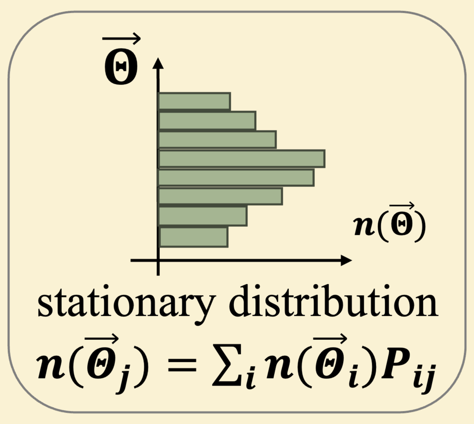
Markov Cain is a chain with non-relation of values except between steps i and $i+1$.

$$P(\overrightarrow{\Theta_{i+1}}|\overrightarrow{\Theta_i}, \overrightarrow{\Theta_{i-1}}, \dots, \overrightarrow{\Theta_1}) = P(\overrightarrow{\Theta_{i+1}}|\overrightarrow{\Theta_i})$$

When the convergence condition is satisfied, to reach a stationary distribution.

- Irreducibility : All states can be reached from any state.
- Aperiodicity : The greatest common divisor of the set of iterations until returning to a certain state is 1.

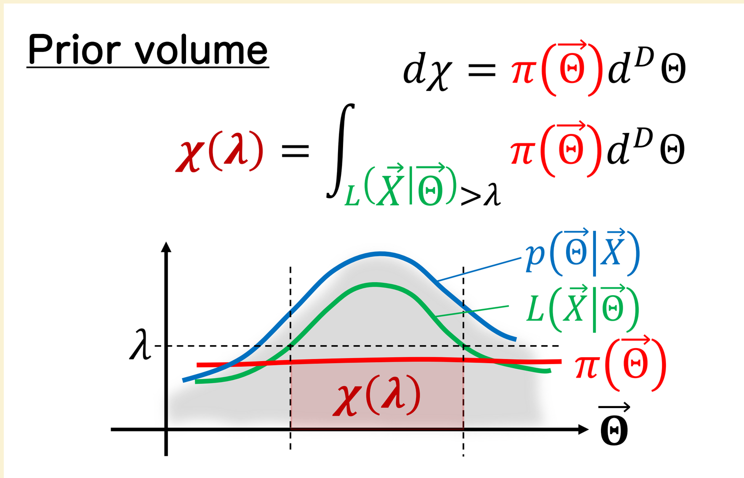




It requires trial and error to tune: Sufficient entire chain length, burn-in phase length, and autocorrelation length.

4. Clustered Nested Sampling (CNS)

Nested Sampling is a Monte Carlo technique aimed at efficient evaluation of the $\boldsymbol{Z}$. $\boldsymbol{Z}$ is estimated by moving **prior volume** χ 1 to 0 and adding up the likelihood.

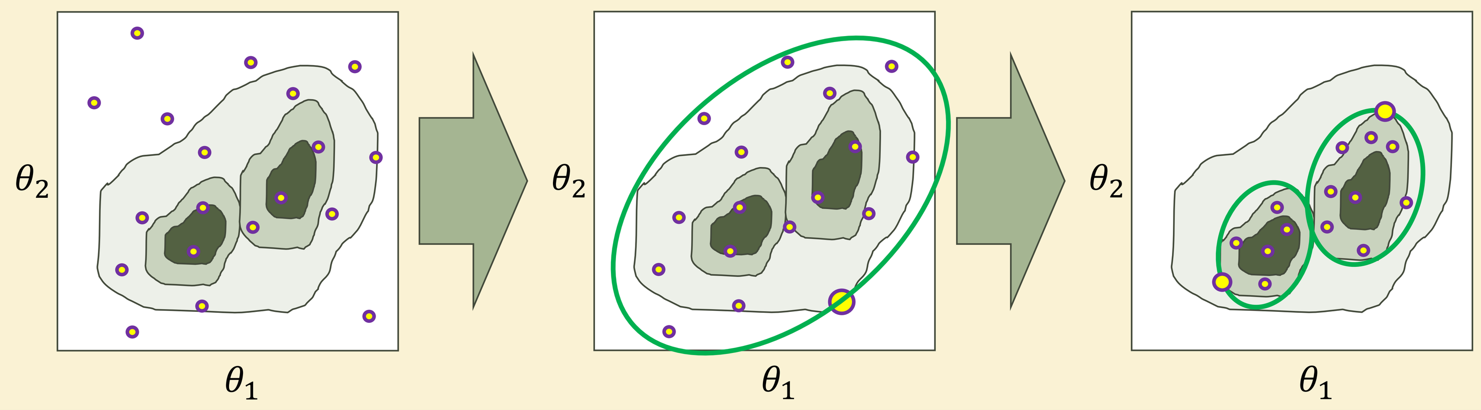


$$Z = \int_{\theta} L(\vec{X}|\vec{\Theta})\pi(\vec{\Theta})d\vec{\Theta}$$

	Prior bound
$\vec{\Theta}$	
λ	$0 \to L_{max}$
χ	$1 \to 0$

$$Z = \int_0^1 L[\chi(\lambda)]\,d\chi = \sum_i^M L(\chi_i)\frac{\chi_{i-1} - \chi_{i+1}}{2}$$

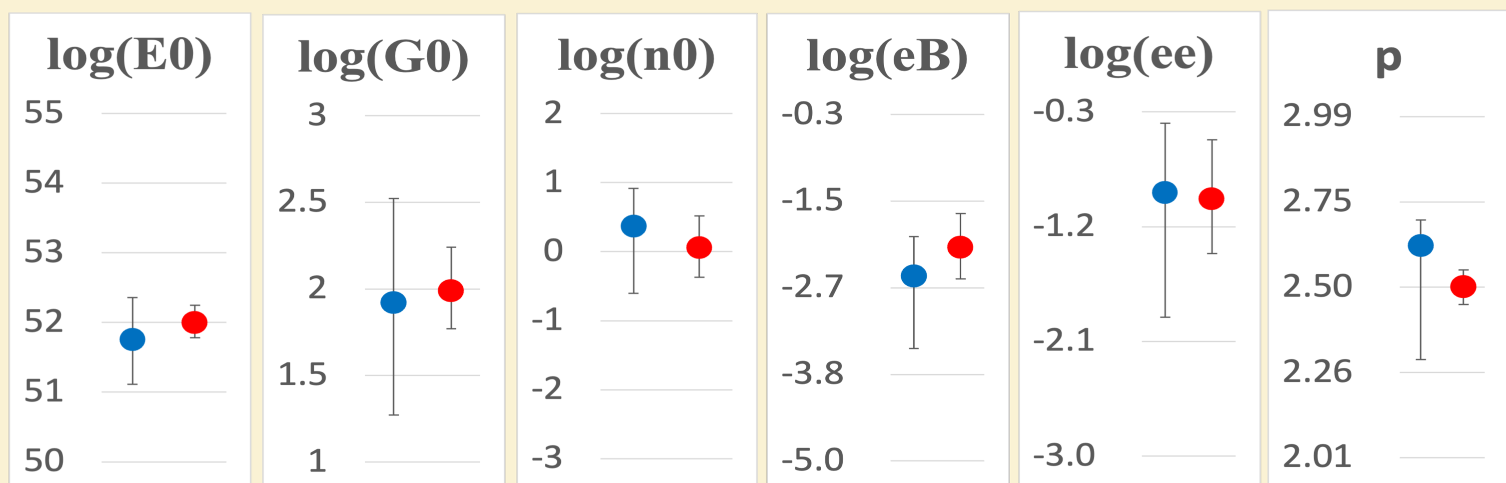
In Clustered Nested Sampling, some ellipsoids in D-dimensional space replace the prior volume χ. This allows sampling from multimodal distributions.



5. Result: Comparison of MCMC and CNS

Using a simulated dataset based on the fireball model for GRB afterglow, generated in collaboration with Yo Kusafuka, we conducted parameter estimation with both MCMC and CNS methods. We present a comparative analysis of these methods, focusing on computational time and interpretability of the results.

MCMC		CNS
Turing.jl[2]	Tool	MULTINEST [1]
16 CPUs	Parallelization	16 CPUs
6	Number of dimensional	6
4.7 hour (1e4 steps)	Convergence Time	～ 0.5 hour



6. Conclusion & Future Work

In this study, we explored the use of MCMC and CNS methods for posterior distribution sampling in high-dimensional parameter spaces, particularly in the context of model parameter estimation for phenomena such as Gamma-Ray Bursts (GRBs). The MCMC method, while relatively simple and versatile across various models, requires careful tuning to ensure convergence, such as adjusting the proposal distribution and handling the complexities of long chains. On the other hand, the CNS method excels in efficiently sampling from multimodal distributions and can provide rapid convergence when the approximate shape of the posterior distribution is known. However, it may be less efficient for cases involving simple posterior distributions. We should understand the characteristics of these methods and use these methods for more systematic theoretical interpretation of GRB.

References

[1] F. Feroz, M. P. Hobson and M. Bridges (2018)

[2] Hong Ge, Kai Xu, and Zoubin Ghahramani, (2018)

[3]物理のためのデータサイエンス入門/植村誠