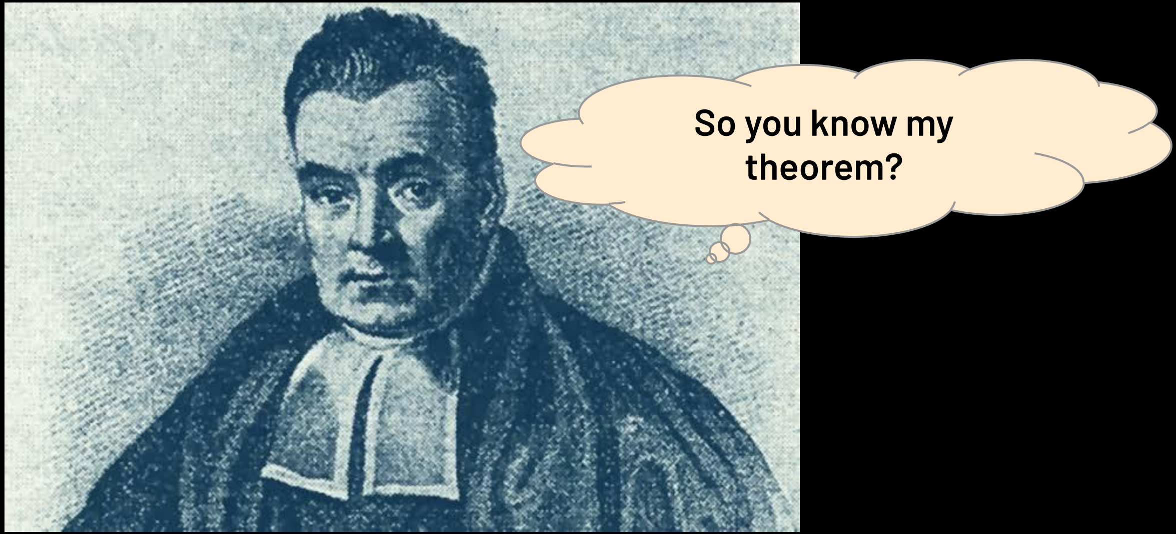
Real Data Meets Bayesian Brilliance

Parth Kothari
M.Sc. Year - II

Principal Investigator: Dr. Suman Majumdar
CSI lab ,DAASE
Course: DSM 517



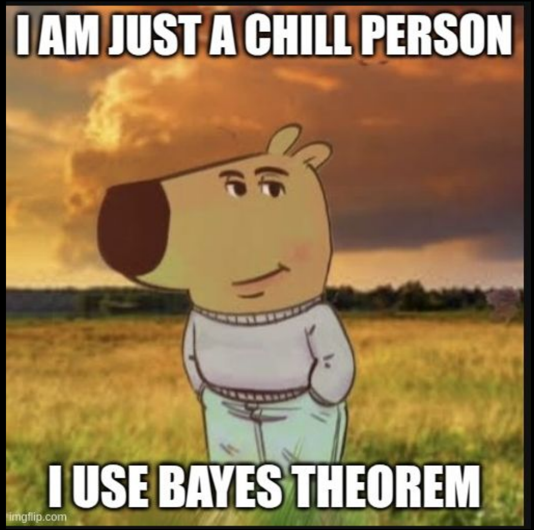
But do you know the story?

01

Bayes Theorem

Bayes Theorem in Life

- Bayes theorem stems from the rationality of human beings
- Mitigating life by always updating your belief
- Conflict resolution through bayes theorem



Bayes theorem in sciences

Posterior

Likelihood

$$p(\theta \mid D) = \frac{p(D|\theta)p(\theta)}{p(D)}$$

Prior

Evidence

02

Markov Chain Monte Carlo

Monte Carlo

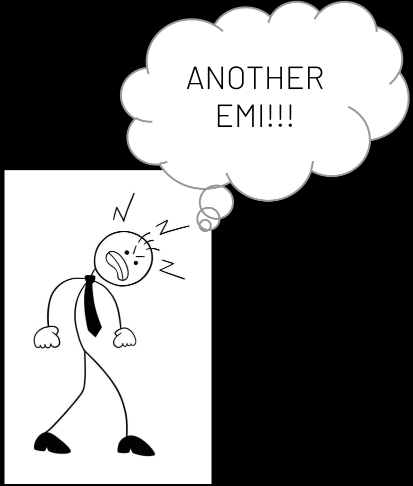
- It is an iterative **Sampling Method**
- Use of **randomly** generated number to predict certain outcomes
- These numbers are generated from a known **probability distribution**

Monte Carlo !

You are at work at 9am

You have a party at 6 pm

Your boss gives you 2 report to complete



Report 1 can take 1-5 hours

Report 2 can take 2-6 hours

How will make an educated guess?

Problem credit: [RiskByNumbers]

Monte Carlo

- Report follows a Uniform distribution

- Sample from these distributions

- Plot the probability distribution

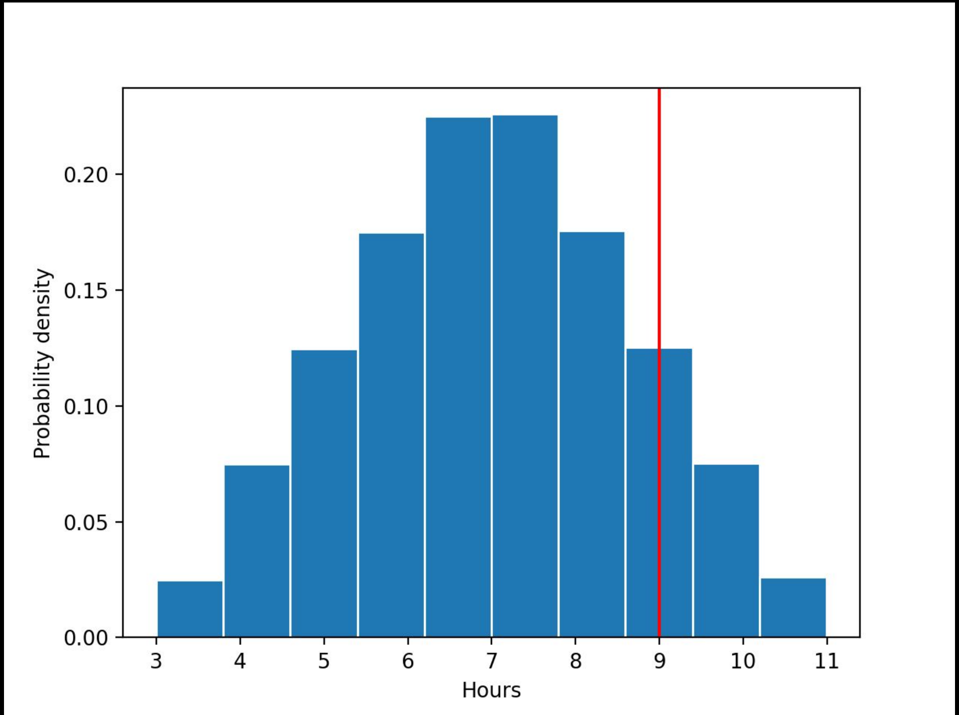
```
sims = 100000

# Report A
A = np.random.uniform(1,5,sims)

# Report B

B = np.random.uniform(2,6,sims)

Duration = A + B #Total time
```



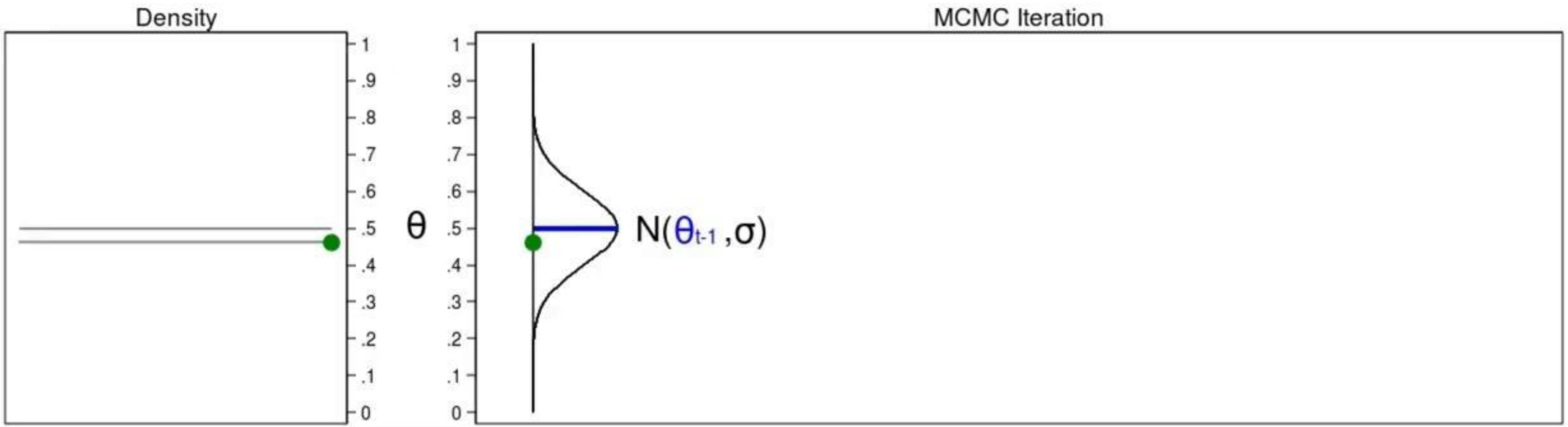
Probability of missing the party : 12 %

Your Choice

Markov Chain Monte Carlo

- Markov chain is an ordered sequence of points $\{x_s\} = \{x_1, x_2, \ldots, x_{Ns}\}$
- x_s (next point) depends on the last point x_{s-1}
- You sample the next point given the previous point via a probability distribution $P(x_s | x_{s-1})$

Visuals are always better!



$$\text{Draw } \theta_t \sim \text{Normal}(\theta_{t-1}, \sigma)$$

$$\text{Normal}(0.500, \sigma) = 0.460$$

Video credit: [StataCorp LLC]()

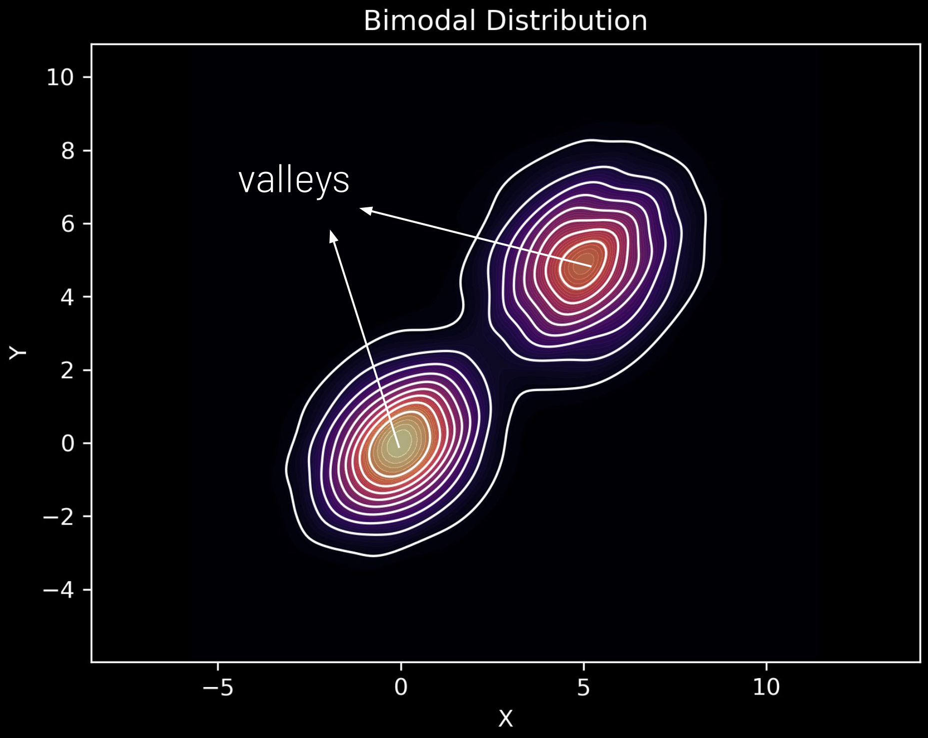
03

Metropolis Hastings -MCMC

MH - MCMC

- Metropolis Hastings is an added condition to MCMC
- You take ratio of the calculated posterior → $R = \frac{P^*(L)\,P^*(\text{prior})}{P(L)\,P(\text{prior})}$
- DO a coin toss for acceptance of the point
- Accept based on the calculation → $U(0,1) < \min(1, R)$
- Considers the less likely samples, exploring the parameter space thoroughly

Drawbacks of MH-MCMC

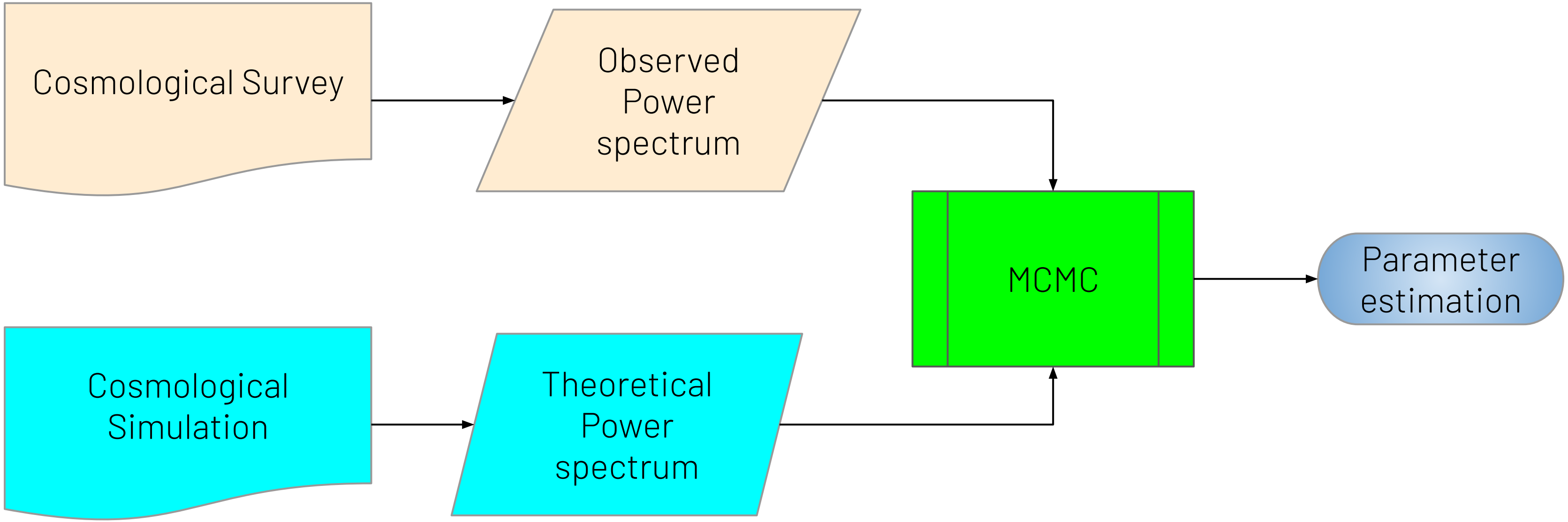


- Smaller step size resulting in walker getting stuck
- Larger step size results in inaccurate sampling of target distribution
- Problem amplifies in high dimensional parameter space

04

My work (Real data)

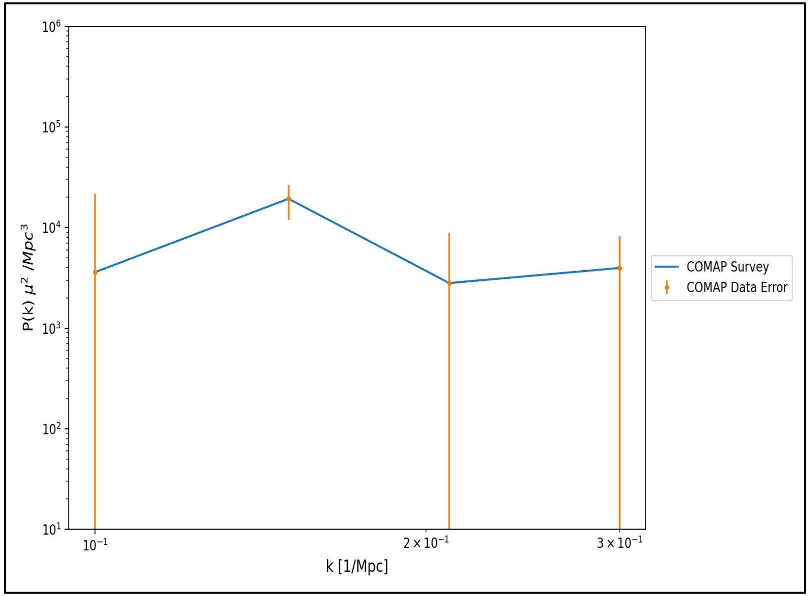
Bayesian Inference pipeline



Observational Data

- Power spectrum at 4 length scales

- The error includes systematics of telescopes, random noise , cosmic variance, etc.



Modeling CO emission

Cosmological simulation

- IllustrisTNG: Halo catalogue
- Dark matter halos in 3D field

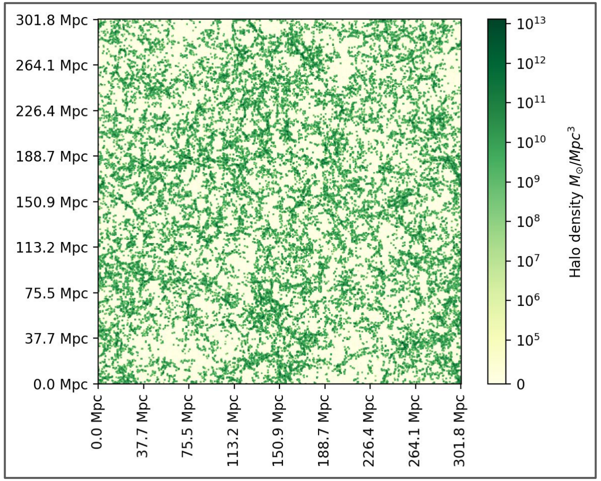
CO Model

- **α** and **β** parameter

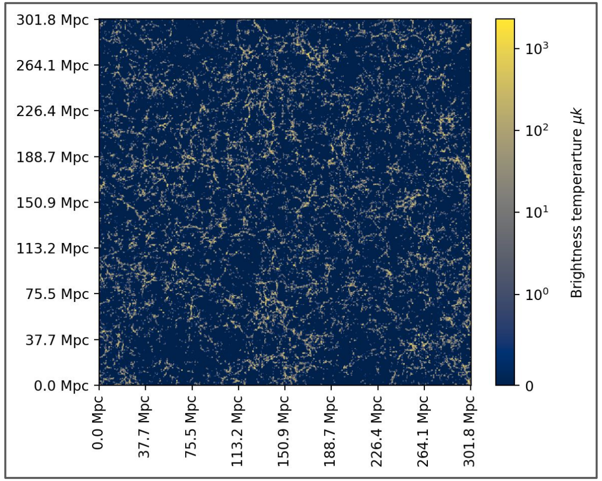
$$\mathrm{logL}_{IR} = \alpha \mathrm{logL}'_{CO} + \beta$$

- Parameters control aggregate CO emission

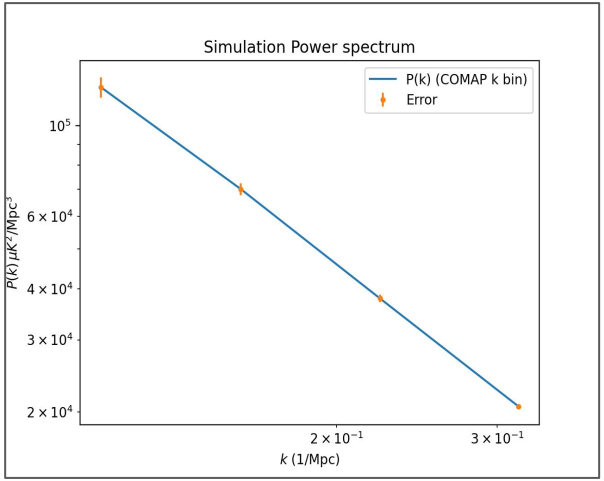
Slice of Halo Field



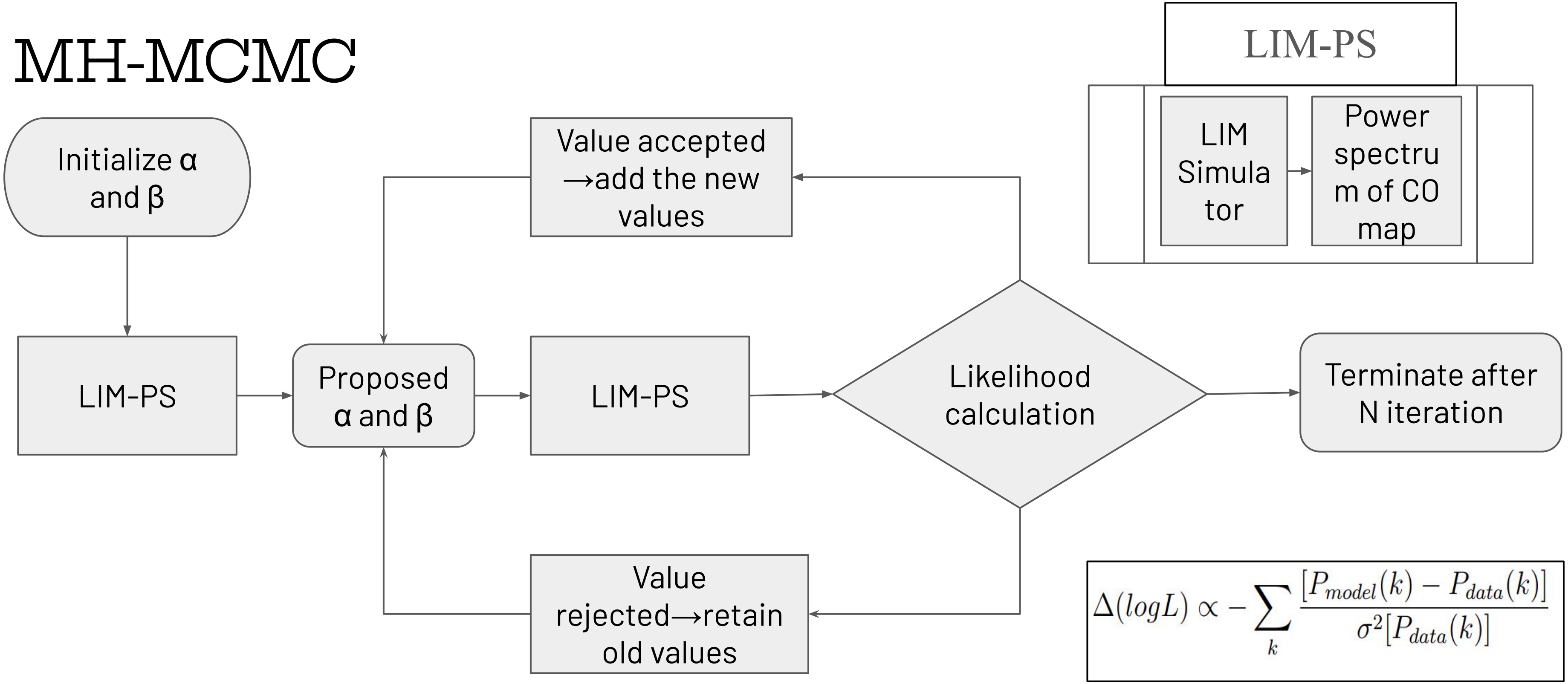
CO simulation



Theoretical Power spectrum

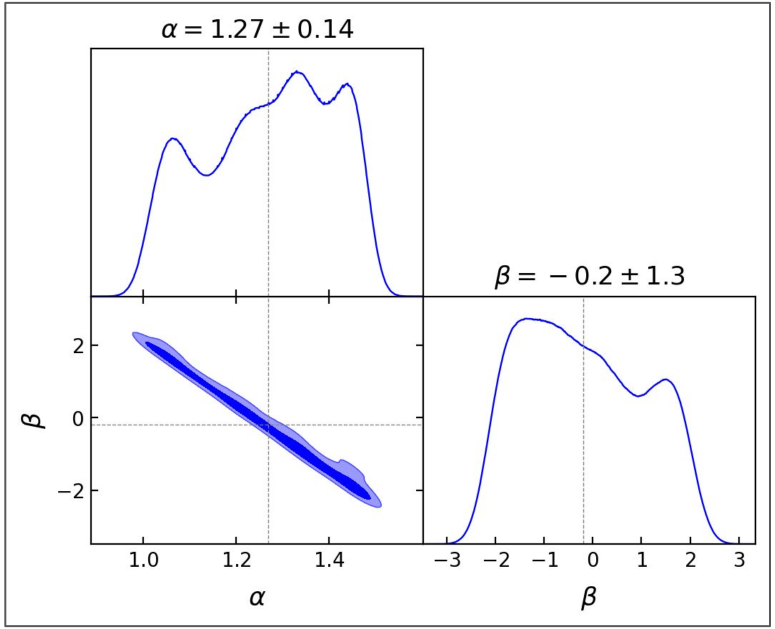


MH-MCMC

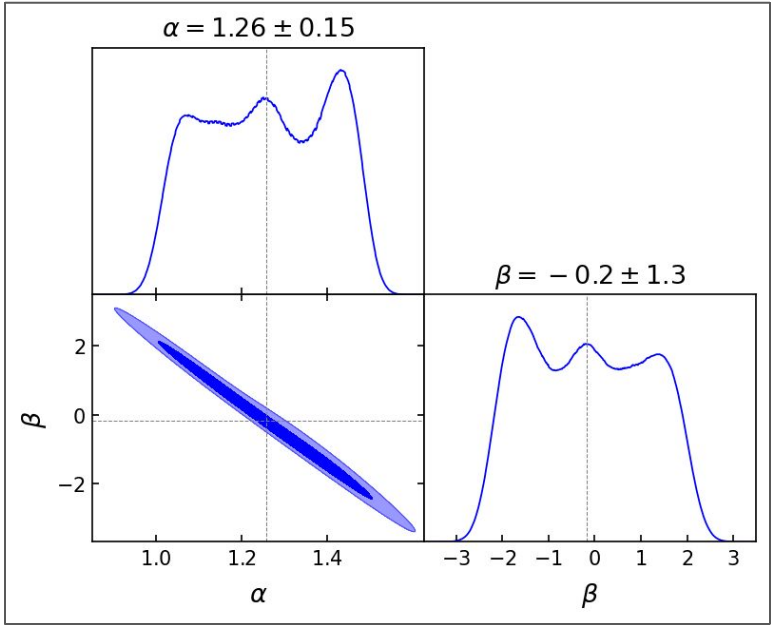


$$\Delta(logL) \propto -\sum_{k} \frac{[P_{model}(k) - P_{data}(k)]}{\sigma^2[P_{data}(k)]}$$

Parameter Estimation



6000 steps with 4 walkers



3000 steps with 4 walkers

Foreman-Mackey et al 2013 PASP 125 306 (emcee)

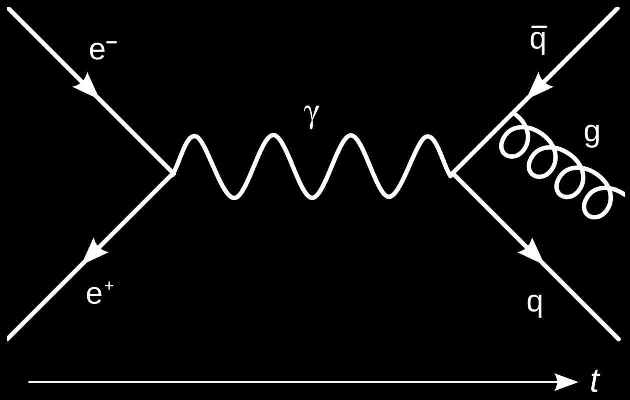
- Real Data is noisy
- Multiple peaks in target distribution
- Inference is biased towards highly local minima

05

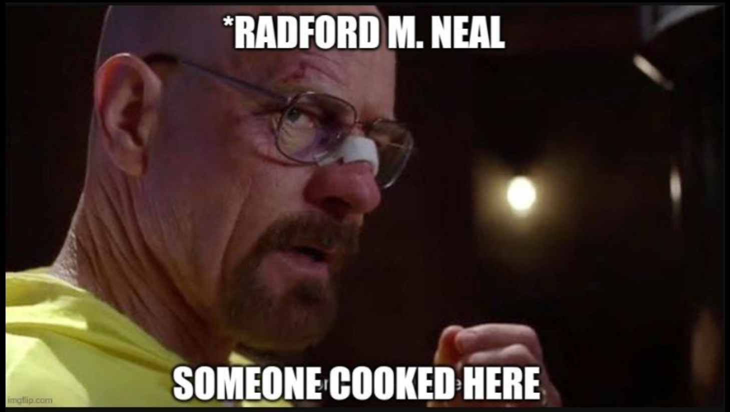
Hamiltonian Monte Carlo

Origin of HMC

- HMC was initially devised for solving the multidimensional integrals in quantum chromodynamics
- Radford borrowed the idea of constructing a physical system into the bayesian inference



Credit: Wikipedia



Hamiltonian dynamics

Hamiltonian, $H(x,u) = U(x) + K(u)$

Potential → $U(x) = -\log[\, p(x)\,]$

Momentum → $K(u) = u^2 / 2$

Momentum sampled from Normal distribution

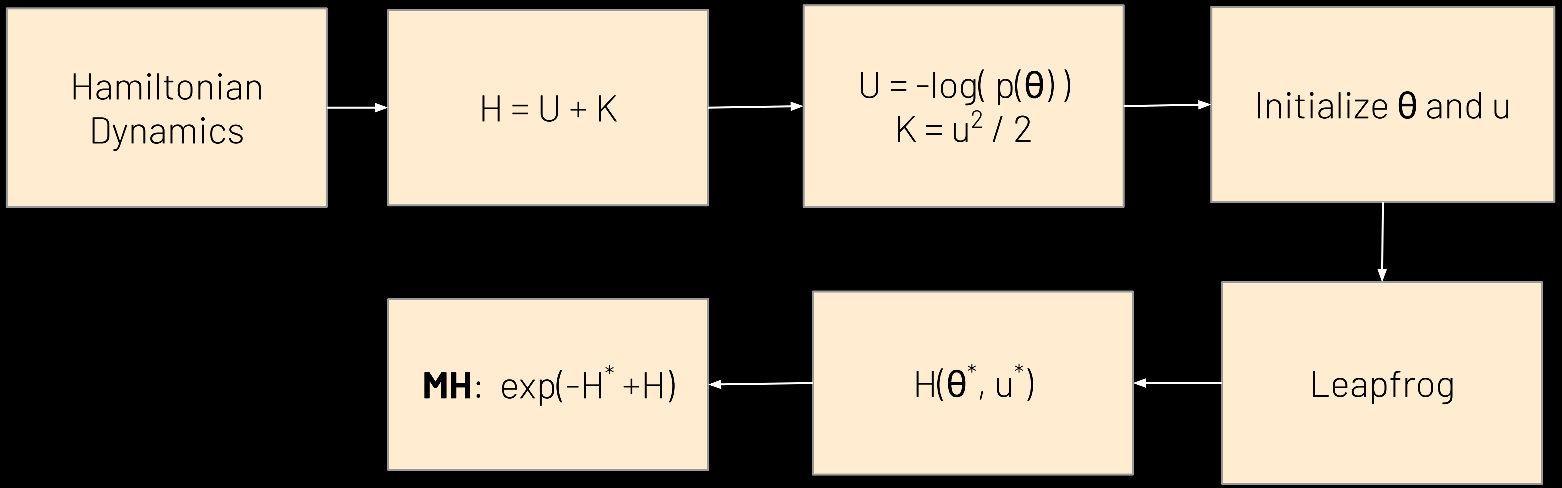
Leapfrog

The KICK → $u_i(t + e/2) = u_i(t) - e/2 * (\partial U / \partial x_i)$

Drift → $x_i(t + e) = x_i(t) + e * u_i(t + e/2)$

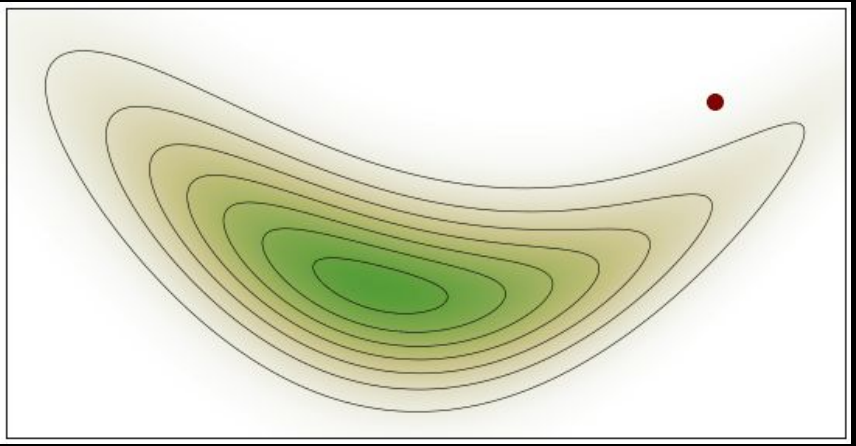
Kick again → $u_i(t + e) = u_i(t + e/2) - e/2 * (\partial U / \partial x_i)$

HMC Flow Chart



Phase space

Position



momentum

- In phase space each point represents position and momentum
- (x1,x2,x3) ; (p1,p2,p3)
- A 6D space on a 2D space

Credit: JustinKunimune

Comparison

MH - MCMC

- Smaller step size resulting in walker getting stuck
- Larger step size results in inaccurate sampling of target distribution
- Problem amplifies in high dimensional parameter space

HMC

- The kick from momentum will overcome the problem of local minima trap
- The equations of motion will help walker cover the entire parameter space as gradient will guide it to different peaks

THANK YOU!

Got questions?
Connect on linkedIn and let's discuss!

TUTORIAL