

Observational tests of eternal inflation: Bayesian model selection

Matthew C. Johnson
Perimeter Institute

In collaboration with:
Stephen Feeney
Daniel Mortlock
Hiranya Peiris

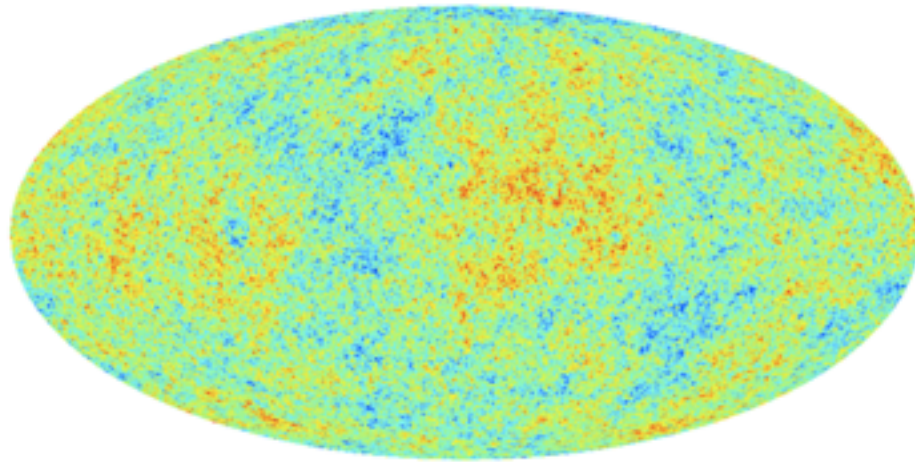
arXiv:1012.1995

arXiv:1012.3667

work in progress

Outline

- How do we decide if there are bubble collisions in CMB data?



- Review of Bayesian statistics.
- What are we testing? (model assumptions)
- An analysis strategy.

Bayesian statistics

- The goal: $P(\text{Model}, \Theta \mid \text{data})$



How should I bet?

- Bayes' Theorem:

$$P(\text{Model}, \Theta \mid \text{data}) = \frac{P(\Theta)P(\text{data} \mid \text{Model}, \Theta)}{P(\text{data} \mid \text{Model})}$$

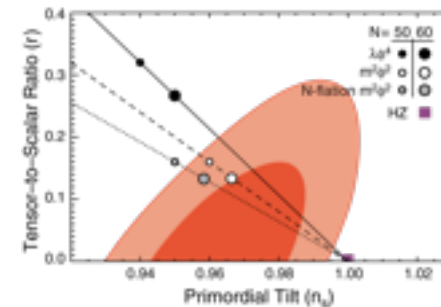
- Theory prior: $P(\Theta) \quad \int P(\Theta)d\Theta = 1$
- Likelihood: $P(\text{data} \mid \text{Model}, \Theta)$
- Evidence (model averaged likelihood): $P(\text{data} \mid \text{Model})$

$$P(\text{data} \mid \text{Model}) = \int d\Theta P(\Theta)P(\text{data} \mid \text{Model}, \Theta)$$

Bayesian statistics

- The likelihood is used to quantify how consistent data is with a set of model parameters.

$P(\text{data} \mid \text{Model}, \Theta)$ \longrightarrow exclusion plots



- This does NOT tell us how we should rank competing theories trying to describe the same data.
- To do so, we can apply Bayes' theorem at the level of Models:

$$P(\text{Model} \mid \text{data}) = \frac{P(\text{Model})P(\text{data} \mid \text{Model})}{P(\text{data})}$$

Bayesian model selection

- Let's say I have a model that fits the data fairly well, should I introduce a more complicated model that might fit it even better?
- We can decide by looking at the evidence ratio:

$$\frac{P(\text{Model 1} \mid \text{data})}{P(\text{Model 0} \mid \text{data})} = \frac{P(\text{Model 1})P(\text{data} \mid \text{Model 1})}{P(\text{Model 0})P(\text{data} \mid \text{Model 0})} = \frac{P(\text{data} \mid \text{Model 1})}{P(\text{data} \mid \text{Model 0})}$$

- The evidence naturally implements Occam's razor: the simpler model should be favored. Tension between volume of parameter space and goodness of fit.

$$P(\text{data} \mid \text{Model}) = \int d\Theta P(\Theta)P(\text{data} \mid \text{Model}, \Theta)$$

Bayesian model selection

- A model is specified both by:
 - The predictions for the data given a particular set of parameters Θ .
 - A prior $P(\Theta)$ that specifies what values these parameters can take.

Model 0: Λ CDM

Ω_m

Ω_Λ

n_s

$\Delta_{\mathcal{R}}^2$

⋮

Model 1: Bubble cosmology with collisions.

Ω_m

Ω_Λ

n_s

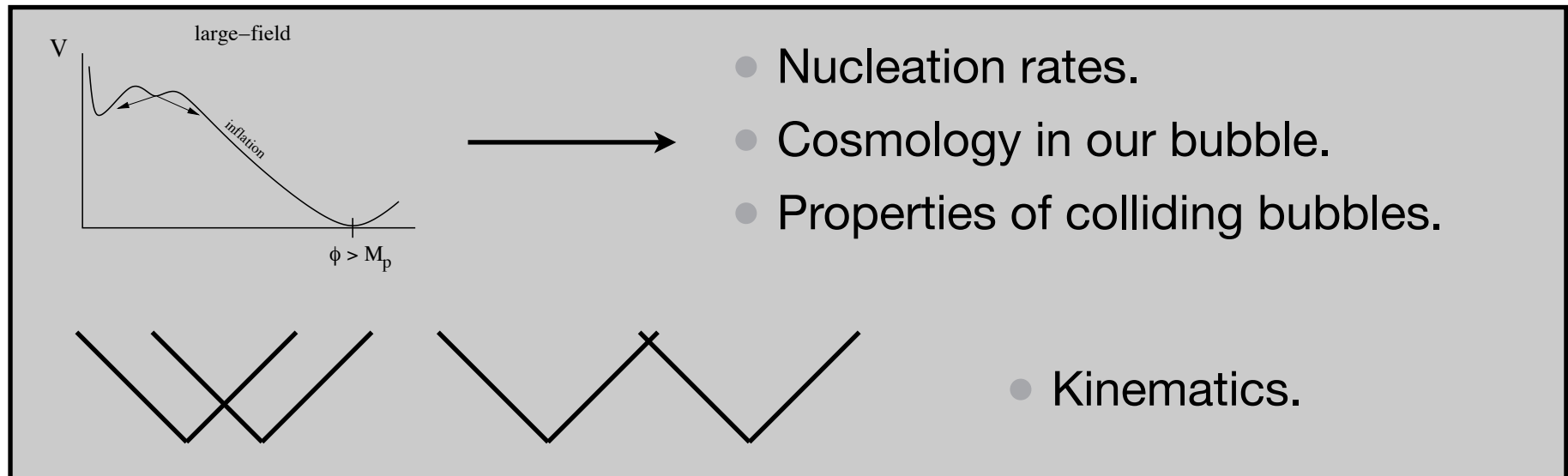
$\Delta_{\mathcal{R}}^2$

⋮

+ $\Theta_{\text{coll}} = ?$

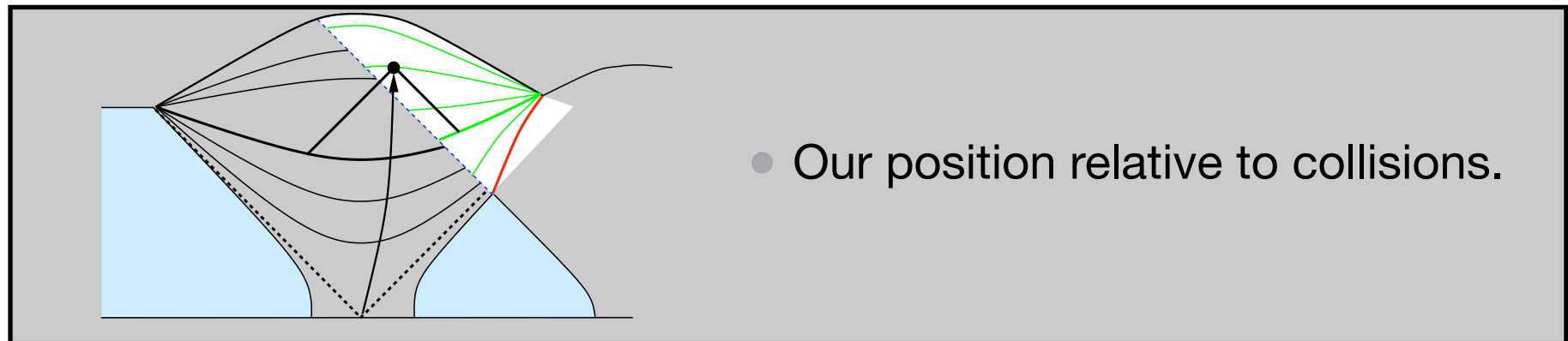
Modeling bubble collisions

- What kinds of parameters specify the collision model?



The diagram shows a potential energy curve V versus a field ϕ . The curve has a local minimum at $\phi > M_p$ and a region labeled "inflation" where the potential is nearly flat. An arrow points from the diagram to a list of parameters. Below the diagram are two diagrams showing the collision of two bubbles, represented by lines forming a V-shape.

- Nucleation rates.
- Cosmology in our bubble.
- Properties of colliding bubbles.
- Kinematics.



The diagram shows a cross-section of two colliding bubbles. The bubbles are represented by blue regions. A black dot marks the point of collision. Green lines represent the trajectories of particles or fields originating from the collision point. A red line indicates a specific trajectory. The diagram illustrates the geometry of the collision and the resulting field configurations.

- Our position relative to collisions.

Modeling bubble collisions

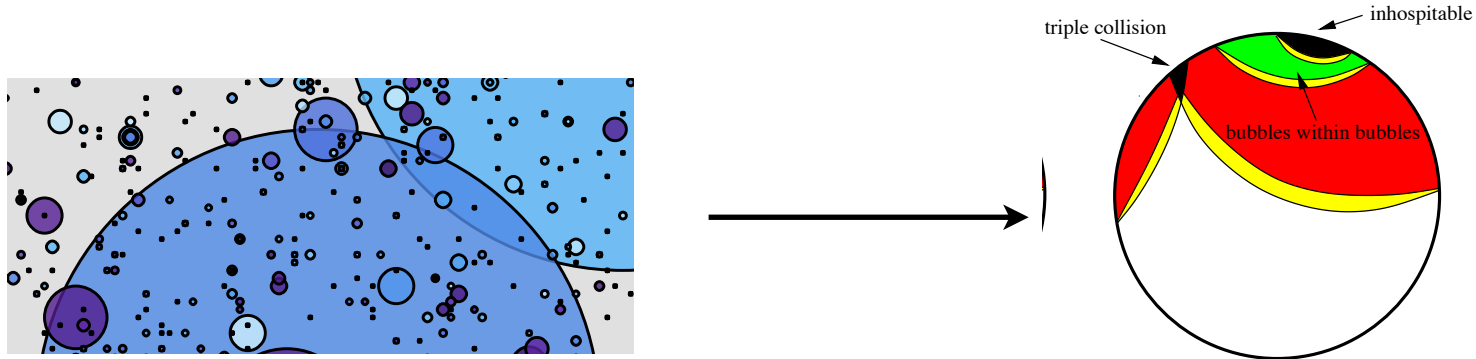
- What kinds of parameters specify the collision model?

- Global properties of the collision spacetime.

- Observed properties of the collision spacetime.

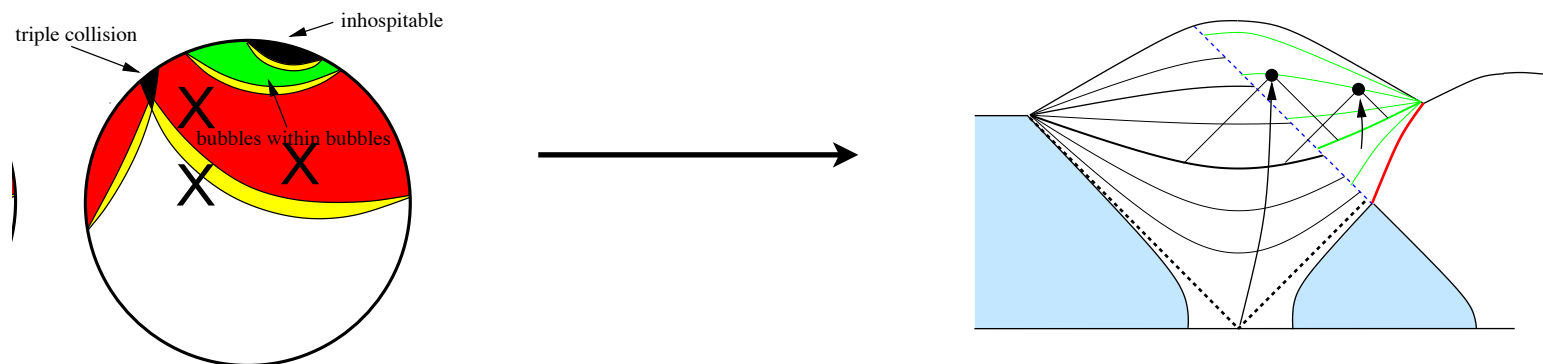
Modeling bubble collisions

- Given unlimited computing power, we could just simulate eternal inflation.



Constant time surface in our bubble.

- Putting observers in different places, we could then ask what they see.
- Counting various observers, we can also (perhaps) determine the prior.

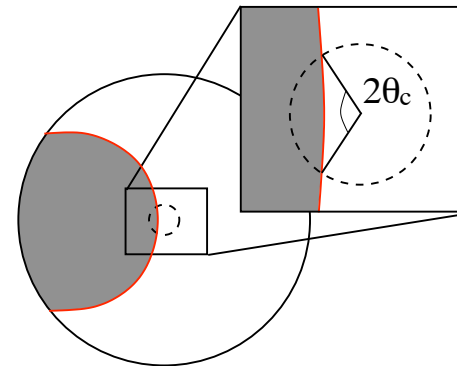


Modeling bubble collisions

- This is impossible.
- It is possibly also wasteful: no one observer can see all this structure.
- To confront data, we also need to determine some useful phenomenological parameters to map the fundamental parameters onto.

A first step:

- Assume the collision is a perturbation on top of inflation: Φ_{coll}
- Thanks to inflation, any observer sees a tiny piece at last scattering:



$$\Phi_{\text{coll}} = \Phi(a) \left(\bar{c}_0 + \bar{c}_1(x - x_{\text{crit}}) + \mathcal{O}((x - x_{\text{crit}})^2) \right) \Theta(x - x_{\text{crit}})$$

global properties are embedded in the \bar{c}_i

observables probe a subset of the \bar{c}_i

Modeling bubble collisions

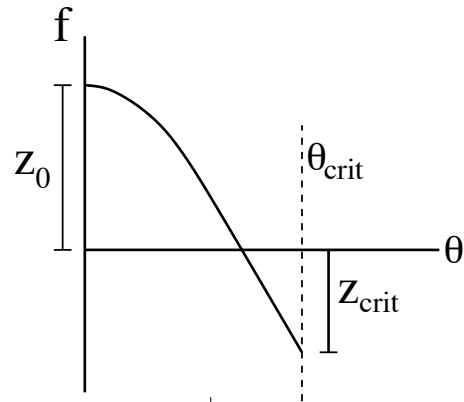
- Projecting onto the past light cone of an observer:

$$\frac{\delta T}{T} \simeq \underbrace{\frac{\Phi_{\text{coll}}(a_{\text{ls}})}{3}}_{\text{SW}} + 2 \underbrace{\int_{a_{\text{ls}}}^1 da \frac{d\Phi_{\text{coll}}}{da}}_{\text{ISW}} + \underbrace{(\vec{v} \cdot \hat{\mathbf{n}} + \mathcal{O}(v^2))}_{\text{doppler}}$$

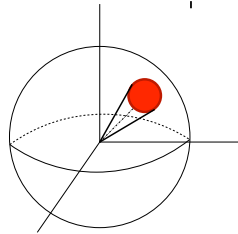
- SW: depends only on the potential at last scattering.
- ISW: depends on how the potential evolves, and how the boundary propagates.
- Doppler: depends on where we formed.

Modeling bubble collisions

- Projecting onto the past light cone of an observer:



$$f(\hat{n}) = \left[\frac{z_{\text{crit}} - z_0 \cos \theta_{\text{crit}}}{1 - \cos \theta_{\text{crit}}} + \frac{z_0 - z_{\text{crit}}}{1 - \cos \theta_{\text{crit}}} \cos \theta \right] \Theta(\theta_{\text{crit}} - \theta)$$



$$\{\Phi(a), \bar{c}_0, \bar{c}_1, x_{\text{crit}}\} \rightarrow \{z_0, z_{\text{crit}}, \theta_{\text{crit}}, \theta_0, \phi_0\}$$

- This form, with $z_{\text{crit}} = 0$, first found by Chang, Kleban, and Levi.
- $z_{\text{crit}} \neq 0$ depends on ISW and doppler contributions: how large?
- How is this template altered by the transfer function?

Model priors

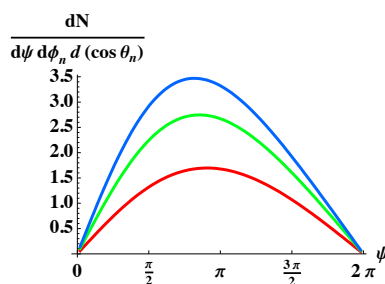
- Previous talks: observable collisions in our PLC are isotropic.

$$\Pr(\theta_0, \phi_0, \theta_{\text{crit}}, z_0, z_{\text{crit}}) = \Pr(\theta_0, \phi_0) \Pr(\theta_{\text{crit}}, z_0, z_{\text{crit}})$$

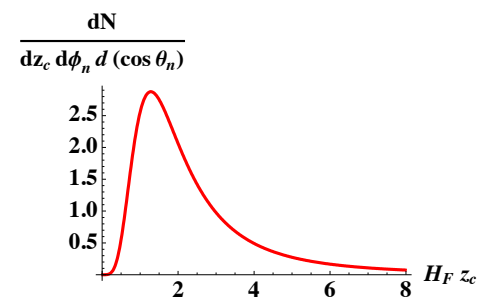
$$\Pr(\theta_0, \phi_0) = \frac{\sin \theta_0}{4\pi}$$

- Kinematics and observer position affects $\{z_0, z_{\text{crit}}\}$ and θ_{crit} .

- Toy model tells us:



For all kinematical configurations.



For all angles.

- Until our understanding of the model improves, we choose flat priors:

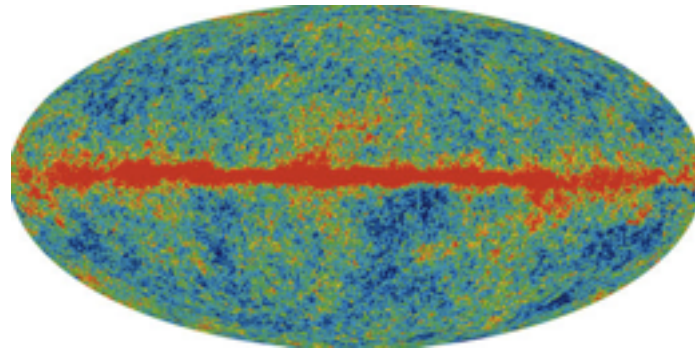
$$\Pr(\theta_{\text{crit}}) = \frac{2}{\pi}, \quad 0 \leq \theta_{\text{crit}} \leq \frac{\pi}{2}$$

$$\Pr(z_{\text{crit}}) = \frac{1}{2 \times 10^{-4}}, \quad -10^{-4} \leq z_{\text{crit}} \leq 10^{-4}$$

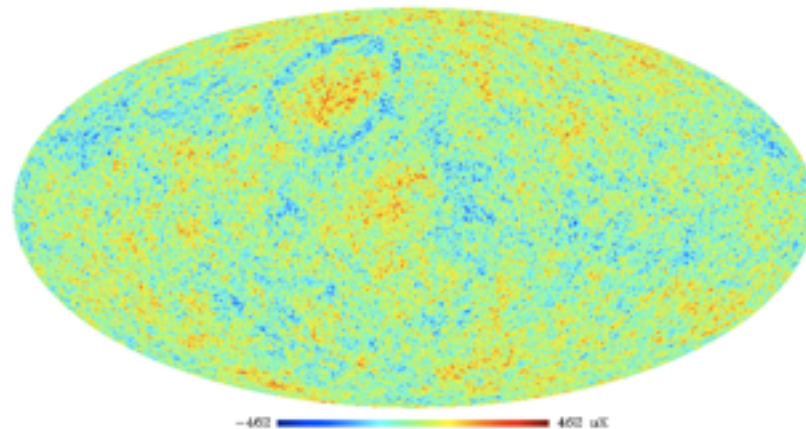
$$\Pr(z_0) = \frac{1}{2 \times 10^{-4}}, \quad -10^{-4} \leq z_0 \leq 10^{-4}$$

The data

- The data under consideration: full-sky CMB maps.



- How do we model select?



Full-sky Bayesian analysis

- Assume collisions can be treated as independent sources on the sky.
- Assume that a theory is specified by the expected number of visible collisions on the full sky:

$$\Pr(\bar{N}_s|\mathbf{d}) = \frac{\Pr(\bar{N}_s) \Pr(\mathbf{d}|\bar{N}_s)}{\Pr(\mathbf{d})}$$

- The evidence ratio we ultimately want to calculate is:

$$\frac{\Pr(\bar{N}_s|\mathbf{d})}{\Pr(0|\mathbf{d})} = \frac{\Pr(\bar{N}_s) \Pr(\mathbf{d}|\bar{N}_s)}{\Pr(0) \Pr(\mathbf{d}|0)}$$

- Assume no theoretical prejudice:

$$\frac{\Pr(\bar{N}_s|\mathbf{d})}{\Pr(0|\mathbf{d})} = \frac{\Pr(\mathbf{d}|\bar{N}_s)}{\Pr(\mathbf{d}|0)}$$

Full-sky Bayesian analysis

- The actual number of collisions is drawn from a Poisson distribution:

$$\Pr(N_s|\bar{N}_s) = \frac{\bar{N}_s^{N_s} e^{-\bar{N}_s}}{N_s!}$$

$$\Rightarrow \Pr(\mathbf{d}|\bar{N}_s) = \sum_{N_s=0}^{\infty} \Pr(N_s|\bar{N}_s) \Pr(\mathbf{d}|N_s)$$

- The evidence is given by:

$$\Pr(\mathbf{d}|N_s) = \int d\mathbf{m}_1 d\mathbf{m}_2 \dots d\mathbf{m}_{N_s} \Pr(\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_{N_s}) \Pr(\mathbf{d}|N_s, \mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_{N_s})$$

$$\mathbf{m}_i = \{ \theta_0^{(i)}, \phi_0^{(i)}, \theta_{\text{crit}}^{(i)}, z_0^{(i)}, z_{\text{crit}}^{(i)} \}$$

- From the independence of each collision:

$$\Pr(\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_{N_s}) = \prod_{i=1}^{N_s} \Pr(\mathbf{m}_i)$$

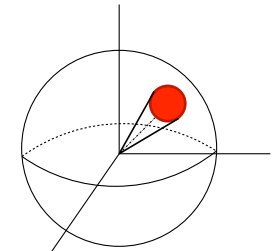
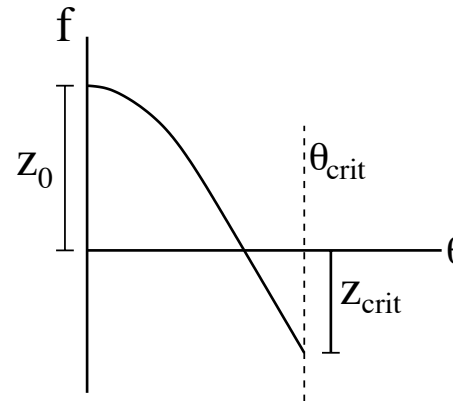
Full-sky Bayesian analysis

- The likelihoods are given by:

$$\Pr(\mathbf{d}|N_s, \mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_{N_s}) = \frac{1}{(2\pi)^{N_{\text{px}}/2} |\mathbf{C}|} \exp \left(-\frac{1}{2} \left[\mathbf{d} - \sum_{i=1}^{N_s} \mathbf{t}(\mathbf{m}_i) \right] \mathbf{C}^{-1} \left[\mathbf{d} - \sum_{i=1}^{N_s} \mathbf{t}(\mathbf{m}_i) \right]^{\text{T}} \right)$$

$$\Pr(\mathbf{d}|0) = \frac{1}{(2\pi)^{N_{\text{px}}/2} |\mathbf{C}|} \exp \left(-\frac{1}{2} \mathbf{d} \mathbf{C}^{-1} \mathbf{d}^{\text{T}} \right)$$

- Templates are defined as before:

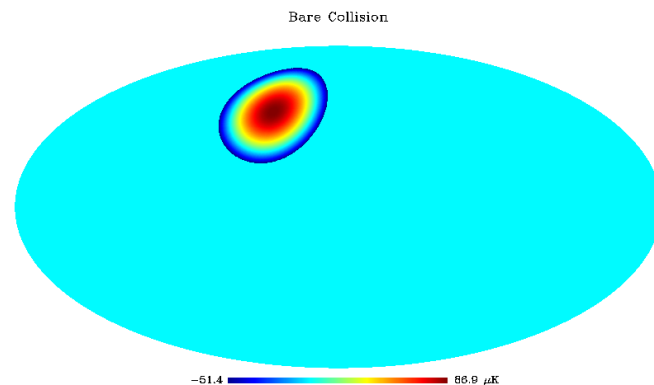


- In the absence of noise and finite instrumental resolution:

$$C_{ij} = \sum_{\ell} \frac{2\ell + 1}{4\pi} C_{\ell} P_{\ell}(\cos \theta_{ij})$$

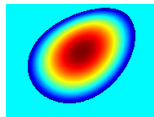
Full-sky Bayesian analysis

- It is impossible to evaluate $\Pr(\mathbf{d}|N_s)$ (let alone $\Pr(\mathbf{d}|\bar{N}_s)$) directly:
 - Inverting C_{ij} at full WMAP resolution is impossible.
 - We must evaluate the likelihood over a $5 N_s$ -dimensional parameter space to find $\Pr(\mathbf{d}|N_s)$; this is impossible for $N_s \gg 1$.
- But, what if we knew something about the possible location and size of candidate collision events?



Full-sky Bayesian analysis

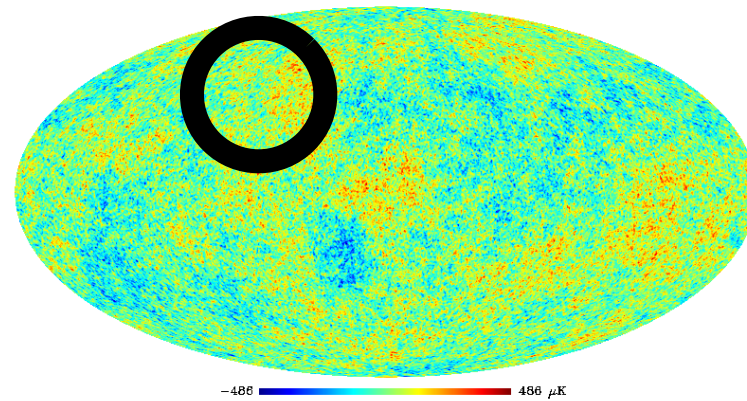
- It is impossible to evaluate $\Pr(\mathbf{d}|N_s)$ (let alone $\Pr(\mathbf{d}|\bar{N}_s)$) directly:
 - Inverting C_{ij} at full WMAP resolution is impossible.
 - We must evaluate the likelihood over a $5 N_s$ -dimensional parameter space to find $\Pr(\mathbf{d}|N_s)$; this is impossible for $N_s \gg 1$.
- But, what if we knew something about the possible location and size of candidate collision events?
 - The compact support for each template means we don't have to do a full-sky analysis, mitigating the above problems!



I'll present the strategy.
Hiranya will present its implementation.

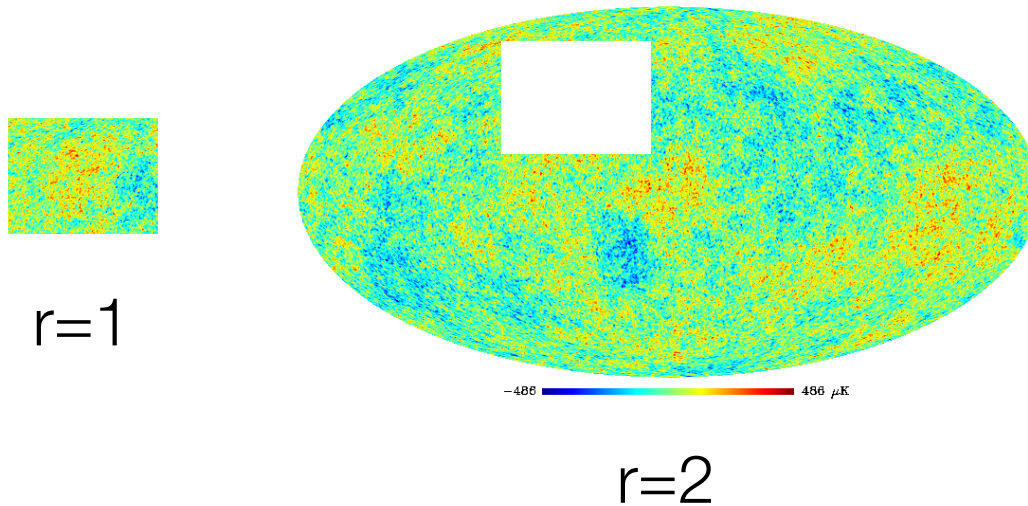
Full-sky Bayesian analysis

- How would this work? Lets say we have one candidate:



Full-sky Bayesian analysis

- How would this work? Lets say we have one candidate:
 - Split the sky into 2 regions: 1 “blob” enclosing the candidate, and the rest of the sky.



- We need to evaluate:

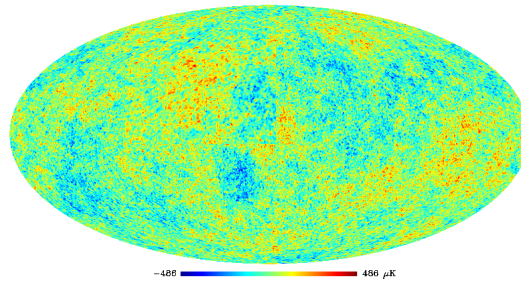
$$\frac{\Pr(\bar{N}_s|\mathbf{d})}{\Pr(0|\mathbf{d})} = \Pr(0|\bar{N}_s) + \Pr(1|\bar{N}_s) \frac{\Pr(\mathbf{d}|1)}{\Pr(\mathbf{d}|0)} + \Pr(2|\bar{N}_s) \frac{\Pr(\mathbf{d}|2)}{\Pr(\mathbf{d}|0)} + \dots$$

- Start with the $N_s = 0$ term: $\Pr(0|\bar{N}_s) = e^{-\bar{N}_s}$

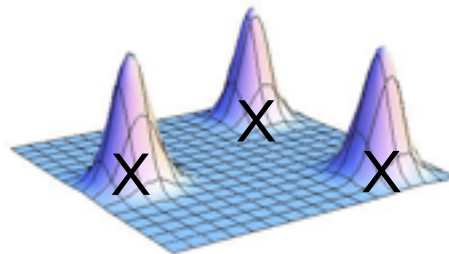
Full-sky Bayesian analysis

- The $N_s = 1$ term.

$$\Pr(\mathbf{d}|1) = \int_{\text{region 1}} d\mathbf{m} \Pr(\mathbf{m}) \Pr(\mathbf{d}|1, \mathbf{m}) + \int_{\text{region 2}} d\mathbf{m} \Pr(\mathbf{m}) \Pr(\mathbf{d}|1, \mathbf{m})$$



- Assume we know the likelihood would be small outside of region 1:

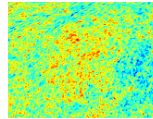


Full-sky Bayesian analysis

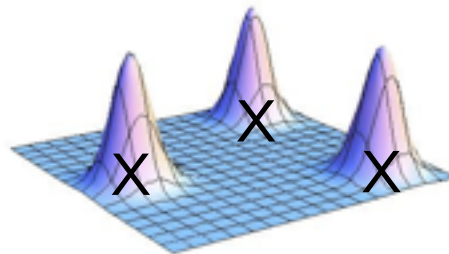
- The $N_s = 1$ term.

$$\Pr(\mathbf{d}|1) = \int_{\text{region 1}} d\mathbf{m} \Pr(\mathbf{m}) \Pr(\mathbf{d}|\mathbf{1}, \mathbf{m}) + \int_{\text{region 2}} d\mathbf{m} \Pr(\mathbf{m}) \Pr(\mathbf{d}|\mathbf{1}, \mathbf{m})$$

likelihood small by assumption

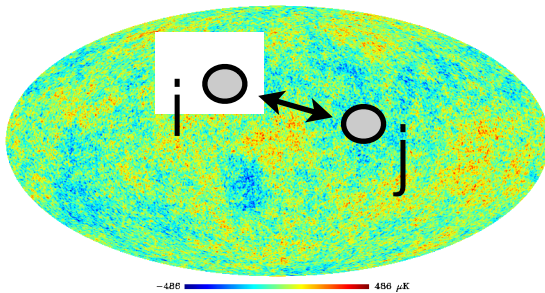


- Assume we know the likelihood would be small outside of region 1:



Full-sky Bayesian analysis

- Although we only integrate over region 1, the covariance still involves the whole sky:



$$C_{ij} = \sum_{\ell} \frac{2\ell + 1}{4\pi} C_{\ell} P_{\ell}(\cos \theta_{ij})$$

- If the blob fully encloses the template, we can approximate:

$$\Pr(\mathbf{d}|1) \propto e^{-[\mathbf{d}_1 - \mathbf{t}_1(\mathbf{m})] \mathbf{C}_1^{-1} [\mathbf{d}_1 - \mathbf{t}_1(\mathbf{m})]^T / 2} \times e^{-\mathbf{d}_2 \mathbf{C}_2^{-1} \mathbf{d}_2^T / 2}$$

$$\Pr(\mathbf{d}|0) \propto e^{-\mathbf{d}_1 \mathbf{C}_1^{-1} \mathbf{d}_1^T / 2} \times e^{-\mathbf{d}_2 \mathbf{C}_2^{-1} \mathbf{d}_2^T / 2}$$

- Under these approximations, we obtain:

$$\frac{\Pr(\mathbf{d}|1)}{\Pr(\mathbf{d}|0)} \simeq \frac{\int_{\text{region 1}} d\mathbf{m} \Pr(\mathbf{m}) e^{-[\mathbf{d}_1 - \mathbf{t}_1(\mathbf{m})] \mathbf{C}_1^{-1} [\mathbf{d}_1 - \mathbf{t}_1(\mathbf{m})]^T / 2}}{e^{-\mathbf{d}_1 \mathbf{C}_1^{-1} \mathbf{d}_1^T / 2}}$$

Full-sky Bayesian analysis

- The $N_s = 2$ term:

$$\Pr(\mathbf{d}|2) = \int \int d\mathbf{m}_1 d\mathbf{m}_2 \Pr(\mathbf{m}_1) \Pr(\mathbf{m}_2) \times \exp\left(-\frac{1}{2} [\mathbf{d} - \mathbf{t}(\mathbf{m}_1) - \mathbf{t}(\mathbf{m}_2)] \mathbf{C}^{-1} [\mathbf{d} - \mathbf{t}(\mathbf{m}_1) - \mathbf{t}(\mathbf{m}_2)]^T\right)$$

- If we know the likelihood will be small in region 2, the highest likelihood occurs when both templates are in region 1.

$$\begin{aligned} \Pr(\mathbf{d}|2) &= \int_{\text{region 1}} \int_{\text{region 1}} (\) + 2 \int_{\text{region 1}} \int_{\text{region 2}} (\) + \int_{\text{region 2}} \int_{\text{region 2}} (\) \\ &\simeq \int_{\text{region 1}} \int_{\text{region 1}} (\) \end{aligned}$$

- But, this is like having one template with twice the parameters!

$$\frac{\Pr(\mathbf{d}|1)}{\Pr(\mathbf{d}|2)} > \text{Vol}(\mathbf{m}) \quad (\text{for flat priors})$$

- We can approximate the full sum by the $N_s = 0$ and $N_s = 1$ terms.

Full-sky Bayesian analysis

- For one blob:

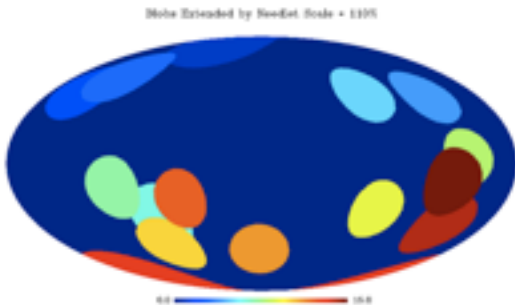
$$\frac{\Pr(\bar{N}_s | \mathbf{d})}{\Pr(0 | \mathbf{d})} \simeq e^{-\bar{N}_s} + \bar{N}_s e^{-\bar{N}_s} \rho_1$$

- For two blobs, assuming they are well separated:

$$\frac{\Pr(\bar{N}_s | \mathbf{d})}{\Pr(0 | \mathbf{d})} \simeq e^{-\bar{N}_s} + \bar{N}_s e^{-\bar{N}_s} (\rho_1 + \rho_2) + \bar{N}_s^2 e^{-\bar{N}_s} \rho_1 \rho_2$$

- For N_b blobs:

$$\frac{\Pr(\bar{N}_s | \mathbf{d})}{\Pr(0 | \mathbf{d})} \simeq \sum_{N_s=0}^{N_b} \frac{\bar{N}_s^{N_s} e^{-\bar{N}_s}}{N_s!} \sum_{b_1, b_2, \dots, b_{N_s}}^{N_b} \prod_{i=1}^{N_s} \rho_{b_i} \prod_{i,j=1}^{N_s} (1 - \delta_{b_i b_j})$$



Full-sky Bayesian analysis

- This method is a calculational trick: we use the full theory priors so there are no a posteriori choices.
- The accuracy of this method relies on how well we can identify candidate collisions. However, it is always a lower bound on the evidence ratio!

$$\frac{\Pr(\bar{N}_s|\mathbf{d})}{\Pr(0|\mathbf{d})} \simeq \sum_{N_s=0}^{N_b} \frac{\bar{N}_s^{N_s} e^{-\bar{N}_s}}{N_s!} \sum_{b_1, b_2, \dots, b_{N_s}}^{N_b} \prod_{i=1}^{N_s} \rho_{b_i} \prod_{i,j=1}^{N_s} (1 - \delta_{b_i b_j})$$

- Need to have both $\bar{N}_s \simeq N_b$ and sizable ρ_{b_i} to favor the collision model.
- Work in progress trying to quantify the accuracy of our approximations.

Full-sky Bayesian analysis

- Easily generalized to include other data sets, i.e. polarization:

$$\begin{aligned} \left[\mathbf{d} - \sum_{i=1}^{N_s} \mathbf{t}(\mathbf{m}_i) \right] \mathbf{C}^{-1} \left[\mathbf{d} - \sum_{i=1}^{N_s} \mathbf{t}(\mathbf{m}_i) \right]^T &= \left[\mathbf{d}^{(T)} - \sum_{i=1}^{N_s} \mathbf{t}^{(T)}(\mathbf{m}_i) \right] \mathbf{C}_{TT}^{-1} \left[\mathbf{d}^{(T)} - \sum_{i=1}^{N_s} \mathbf{t}^{(T)}(\mathbf{m}_i) \right]^T \\ &+ \left[\mathbf{d}^{(T)} - \sum_{i=1}^{N_s} \mathbf{t}^{(T)}(\mathbf{m}_i) \right] \mathbf{C}_{TE}^{-1} \left[\mathbf{d}^{(E)} - \sum_{i=1}^{N_s} \mathbf{t}^{(E)}(\mathbf{m}_i) \right]^T \\ &+ \left[\mathbf{d}^{(E)} - \sum_{i=1}^{N_s} \mathbf{t}^{(E)}(\mathbf{m}_i) \right] \mathbf{C}_{EE}^{-1} \left[\mathbf{d}^{(E)} - \sum_{i=1}^{N_s} \mathbf{t}^{(E)}(\mathbf{m}_i) \right]^T \end{aligned}$$

➔ $\rho_{b_i} = \rho_{b_i}^{(TT)} \rho_{b_i}^{(TE)} \rho_{b_i}^{(EE)}$

- Bayesian methods can be used to rank competing theories of spots: i.e. textures.
- Generalizable to study any features in a full-sky data set.

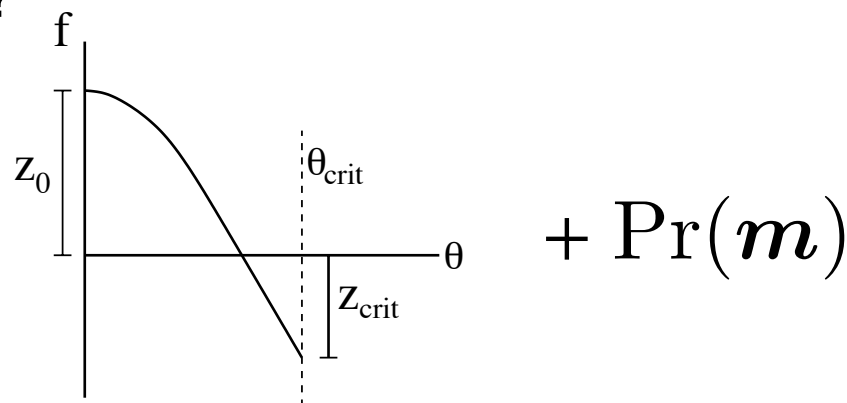
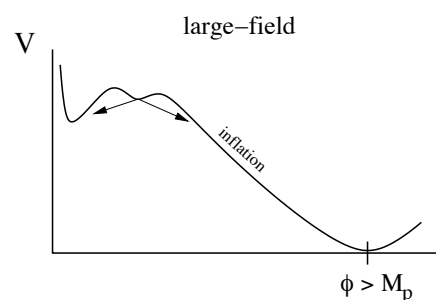
Conclusions

- Bayesian model selection is a consistent framework for determining if we should consider a theory with bubble collisions over one without.
- To do so, it is important to parameterize the theory of bubble collisions, and determine the priors for the parameters.
- A full sky Bayesian analysis can be approximated with a patch-wise analysis if we know something about the likelihood surface.

$$\frac{\Pr(\bar{N}_s|\mathbf{d})}{\Pr(0|\mathbf{d})} \simeq \sum_{N_s=0}^{N_b} \frac{\bar{N}_s^{N_s} e^{-\bar{N}_s}}{N_s!} \sum_{b_1, b_2, \dots, b_{N_s}}^{N_b} \prod_{i=1}^{N_s} \rho_{b_i} \prod_{i,j=1}^{N_s} (1 - \delta_{b_i b_j})$$

Open questions

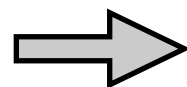
- What is the mapping from a potential to a phenomenological model for the effects of collisions on the CMB?



- Are there any correlations between LCDM parameters and collisions?

bubble collisions \rightarrow detectable r

bubble collisions \rightarrow undetectable Ω_k



$$\Pr(\mathbf{m}, \Lambda\text{CDM}) \neq \Pr(\mathbf{m})\Pr(\Lambda\text{CDM})$$

- What happens in a vast landscape?
 - Bayesian methods inherit the measure problem.
 - Question becomes academic, since the data will probably never be good enough to distinguish models in detail.