Advanced Algorithmic Implementations of Bayesian Inference in Cosmology

Jeffrey Jewell
Jet Propulsion Laboratory, California Institute of Technology

Joint Statistical Meeting,
Chicago, Illinois
July 31, 2016

Expansion Timeline of the Universe

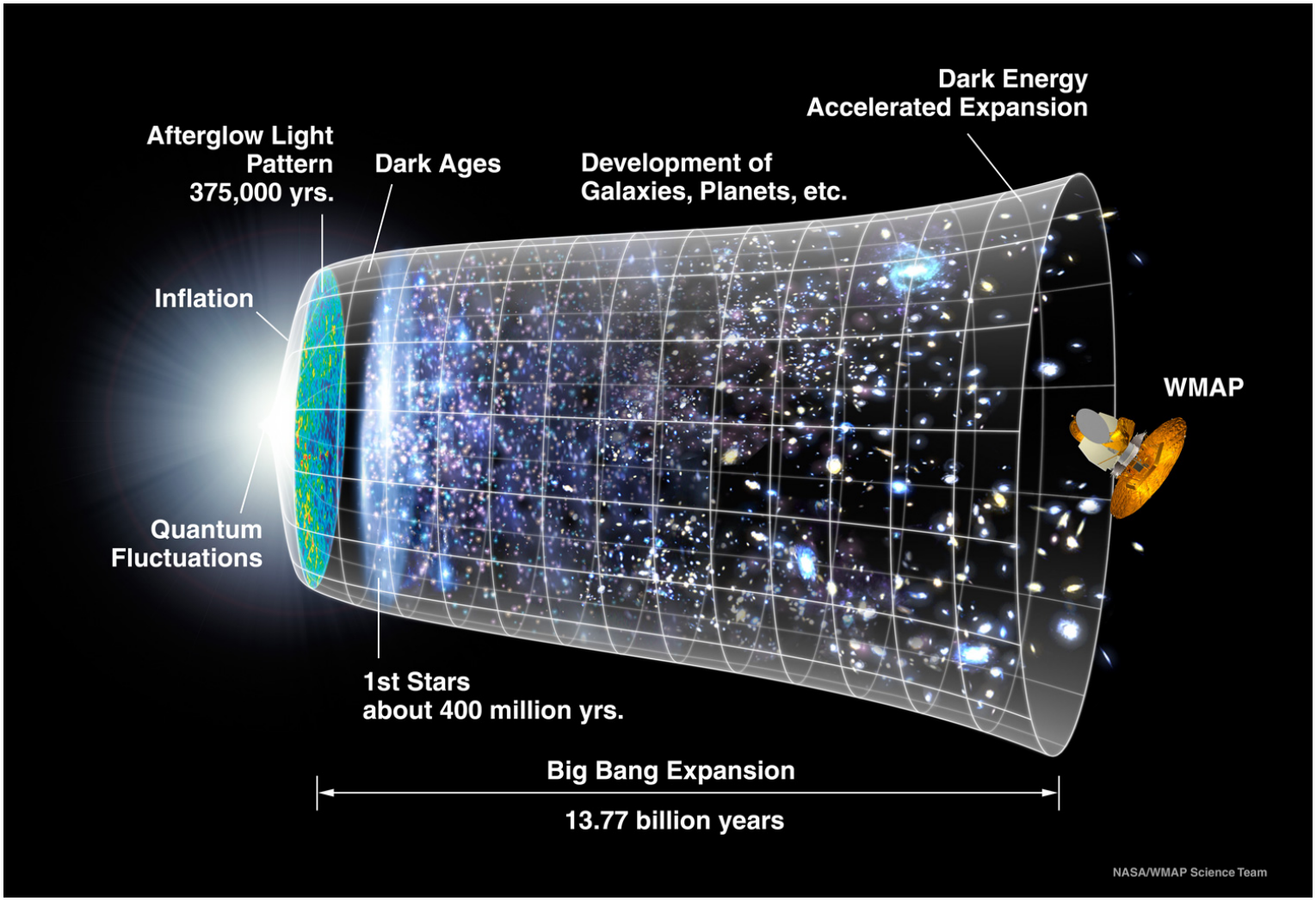


Figure: A representation of the evolution of the universe over 13.77 billion years. Credit: NASA / WMAP Science Team

Introduction

- Current and future ground and space-based missions are designed to not only detect, but map out with increasing precision, details of the universe in its infancy to the present-day.
- Each of these instruments return observations with their own unique insight into the universe we want to understand, as well as their own peculiar response to the signal they are designed to detect.
- As a result, we are faced with the increasingly important challenge of analyzing and interpreting observations from a wide variety of instruments to form a coherent view of the universe.
- How can we make inferences about our universe from observations returned from a diverse collection of instruments spanning a wide range of frequencies and spatial scales?
- Data analysis in cosmology involves challenging problems of Bayesian inference in the context of spatial random fields - including inference for Gaussian spatial random fields, source separation, and data assimilation and parameter estimation for non-linear dynamics.
- For this talk - brief overview of cosmology as it leads to inference problems that have applications in other fields. (examples - data fusion from multiple sources at multiple resolutions, data assimilation in climate science or weather prediction, etc.)

Gravitational Collapse Picture of Large-Scale Structure Formation

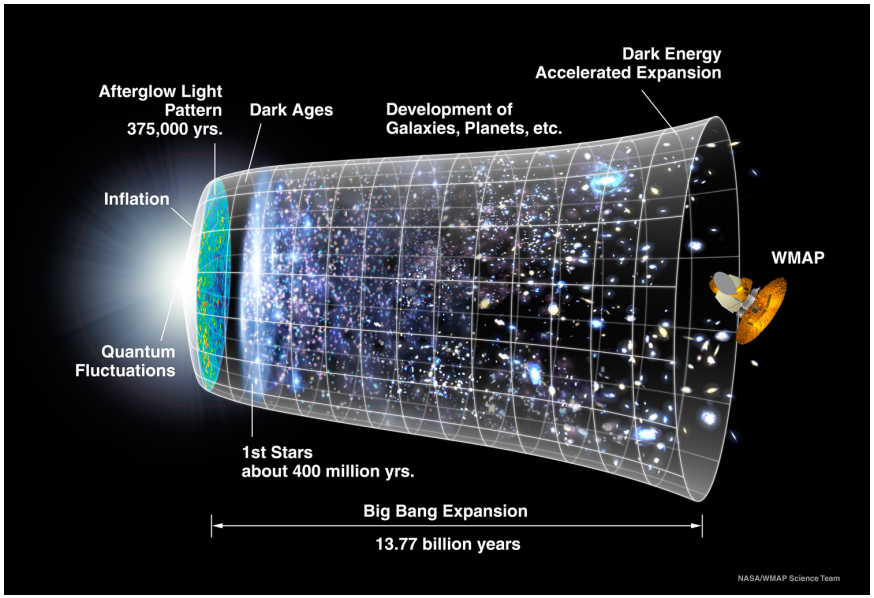


Figure: A representation of the evolution of the universe over 13.77 billion years. Credit: NASA / WMAP Science Team.

- Quantum (Gaussian) fluctuations generate small amplitude initial density perturbations.
- These density perturbations evolve linearly while the universe is in a hot, dense, ionized state (a balance between photon radiation pressure and gravity for small amplitude density variations).
- The universe expands and cools eventually allowing neutral matter to form, and allowing photons to travel long distances without scattering. We observe photons from this epoch in all directions at microwave frequencies - the brightness temperature of the photons show small variations which trace out the initial density perturbations at the time of last scattering.
- The initial (Gaussian) density fluctuations after last-scattering undergo non-linear gravitational collapse (approximately the fluid equations coupled to gravity).
- We observe the resulting 3-dimensional distribution of matter with surveys at visible and other wavelengths.

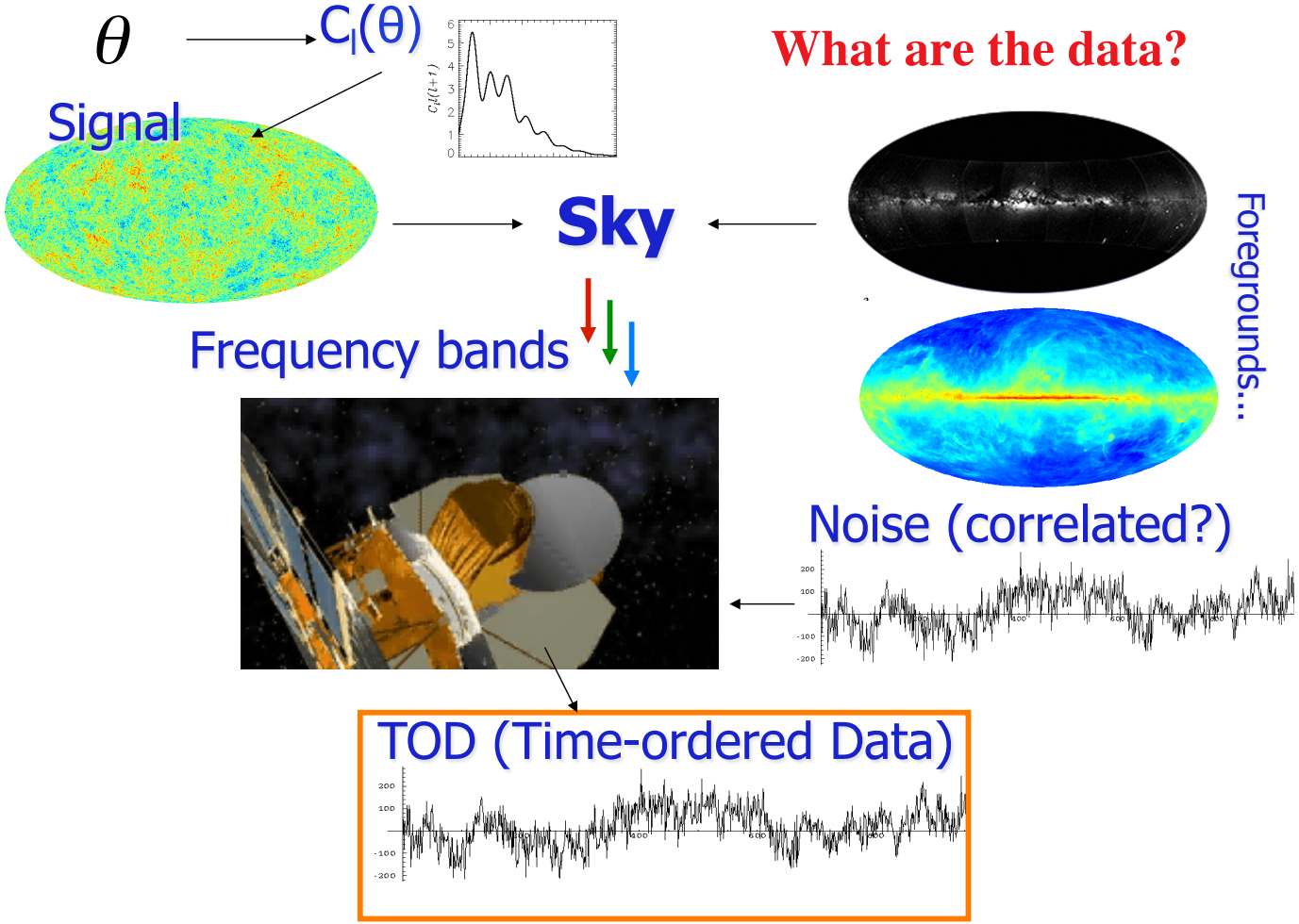
Because of the finite speed of light and the expansion of the universe, we are able to observe the history of the universe by a joint analysis of measurements at different wavelengths.

Bayesian Inference in Cosmology

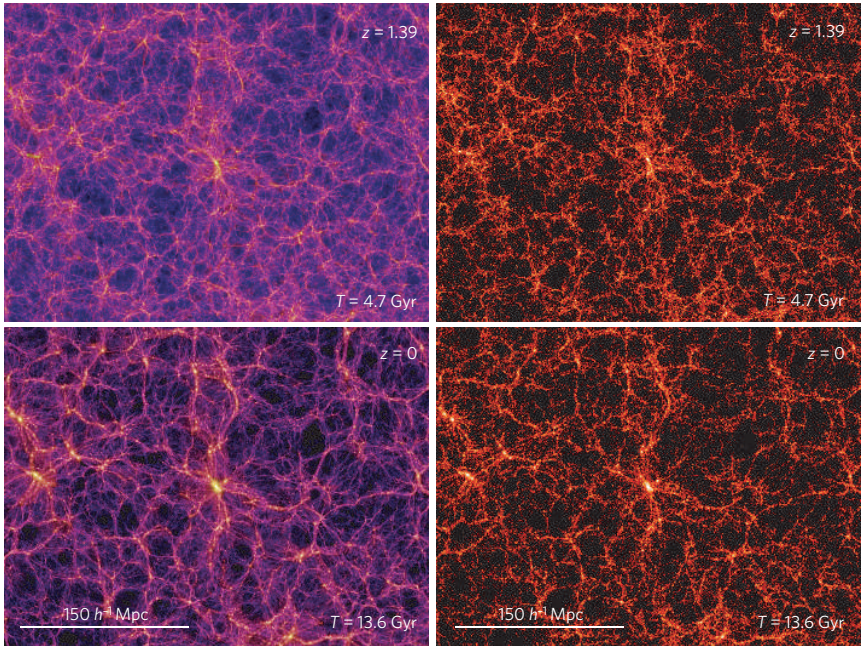
- We have a very detailed, physically motivated, theoretical framework in cosmology - parametric details of which we want to learn from measurements.
- We have *one universe to observe*, so we want to formulate a Bayesian approach which respects the underlying physics and details of the theory as closely as possible!
- A very diverse range of inference problems are encountered when confronting theory and measurement in cosmology - a Bayes' posterior in this situation is *derived in detail from a specification of the simulation of measurement in the context of the theoretical framework*.
- While conceptually straightforward, this quickly leads to computational problems

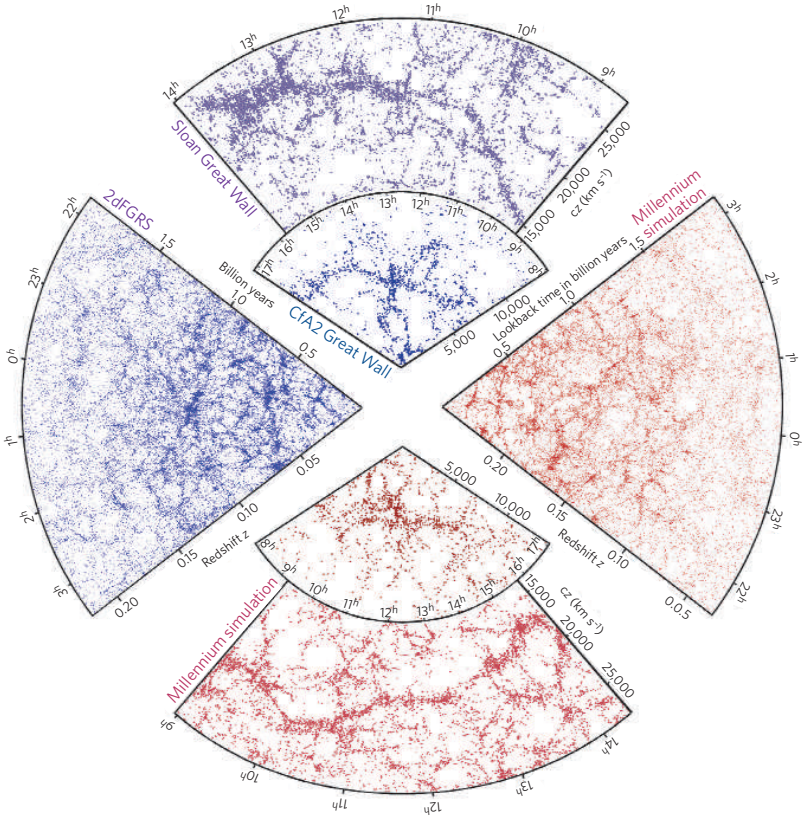
- Two challenges encountered in Bayesian inference for cosmology:
  - Analytically known posterior functional form, yet computationally intractable to evaluate for large datasets
  - Marginal Posterior unknown - and joint sampling for dynamical parameters and state history is extremely challenging due to high-dimensionality and non-Gaussian structure
- We will review these problems and potential solutions along with an indication of breakthroughs needed

Inference Problem 1 - Analysis of the Cosmic Microwave Background



Inference Problem 2 - Analysis of Large-Scale Structure





(Credit: *The Large-Scale Structure of the Universe*, Springel, V.; et al; , Nature, Vol. 440, p. 1137, April 2006)

- *Left Figure:* Slices at different times from an N-body simulation of the dark matter distribution evolving according to gravitational collapse in an expanding universe as determined by a cosmological model.
- *Right Figure:* A semi-analytic model prediction of how luminous galaxies trace this dark matter distribution. The history of the distribution is displayed by the slice outputs at sequentially later times.
- The challenge is to relate the observed redshift survey data to the underlying dark matter distribution as shown on the left, *as well as determine the cosmological parameters governing the evolution.*

Simulation and Inference

- Important first step - write down a "joint density" of all degrees of freedom in the physical model, even if they are not observable!
- We have the joint density given in a "forward" manner determined by specifying how to simulate observations for any parameter in our class of models and the data D)

$$\begin{array}{ll} \text{Model Parameters} & \Omega \\ \text{"True" field, signal, etc.} & X \\ \text{Data} & D \\ \text{Joint Density} & p(D, X, \Omega) = p(D|X)p(X|\Omega)p(\Omega) \end{array} \tag{1}$$

- Simulation and Inference involve different conditional densities inherited from the common joint density on "everything"!
- Simulation conditions on the model -

$$p(D, X|\Omega) = p(D|X)p(X|\Omega) \tag{2}$$

- Inference conditions on the data

$$p(\Omega, X|D) = \frac{p(D|X)p(X|\Omega)p(\Omega)}{\int d(\Omega', X')\ p(D|X')p(X'|\Omega')p(\Omega')} \tag{3}$$

Bayesian Inference for the CMB

- Simplest data model - the observed sky at $k = 1 : K$ frequencies is

$$d_k = B_k s + n_k \tag{4}$$

(with instrument response function ("beam") B_k, underlying true sky s, and instrumental noise n with covariance N_k).

- Factors in the joint density (i.e. $p(d, s, \Omega) = p(d|s)p(s|\Omega)p(\Omega)$) for CMB analysis:

$$\begin{aligned} -2\log p(d|s) &\propto (d - Bs)N^{-1}(d - Bs) \\ -2\log p(s|\Omega) &= sC^{-1}(\Omega)s + \log\det|C(\Omega)| \end{aligned} \tag{5}$$

- Given measurements we have information about details of the cosmological theory and the underlying true sky,

$$p(\Omega, s|d) = \frac{p(d, s, \Omega)}{\int d(\Omega', s')\ p(d, s', \Omega')} = \frac{e^{-(d-Bs)N^{-1}(d-Bs) - sC^{-1}(\Omega)s - \log|C(\Omega)|}}{Z(D)} \tag{6}$$

- Bayes' marginal posterior in cosmological (theory) parameters $\int ds\ p(\Omega, s|d)$ is:

$$-2\log p(\Omega|d) \sim D[N + C(\Omega)]^{-1}D + \log|N + C(\Omega)| \tag{7}$$

- Today's CMB experiments return maps of the sky with $N_{\text{pixel}} > 10^6$ - *It is already intractable to directly evaluate the posterior for any Ω due to the $\mathcal{O}(N^3_{pixel})$ expense of computing the determinant.*

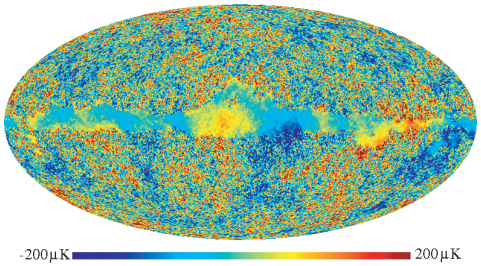
Markov Chain Monte Carlo Sampling

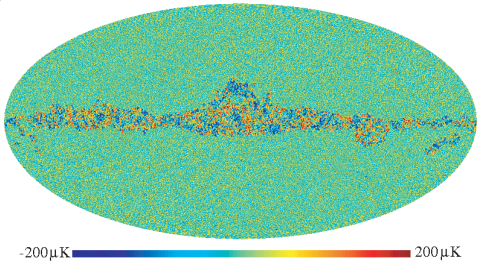
- Instead of direct likelihood evaluation for cosmological parameters, we explore Markov chain algorithms jointly sampling estimates of the CMB map and parameters
- *General MCMC* - We can make proposals for both new parameters and CMB maps, followed by an accept or reject step
- *Gibbs sampling* - (a special case of a Markov chain with accept probability unity) proceeds by interatively sampling from the conditionals

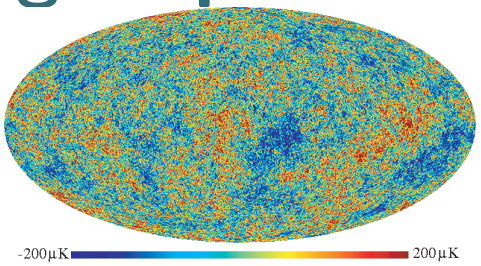
$$\begin{aligned} s_{n+1} &\leftarrow p(s|C(\Omega_n), D) \\ \Omega_{n+1} &\leftarrow p(\Omega|s_{n+1}) \end{aligned} \tag{8}$$

  (where note conditional independence of the data, $p(\Omega|s, d) = p(\Omega|s)$).
- For Gaussian fields, we found that Gibbs sampling works well in the high signal to noise regime (large spatial scales), while it was necessary to develop a special MCMC method for low signal-to-noise (small spatial scales).
- *This will be a common problem when using Gaussian priors for spatial random fields when we have data with different measurement error on different spatial scales!*

Gibbs Sampling - The map making step . . .







- The conditional $p(s|\Omega, D)$ is *Gaussian* -

$$-2\log p(s|\Omega, D) \sim \left(\sum_k (d_k - B_k s) N_k^{-1} (d_k - B_k s)\right) - sC^{-1}(\Omega)s \tag{9}$$

- Exact samples for any signal measurement covariance matrix generated by solving the linear problem(s)

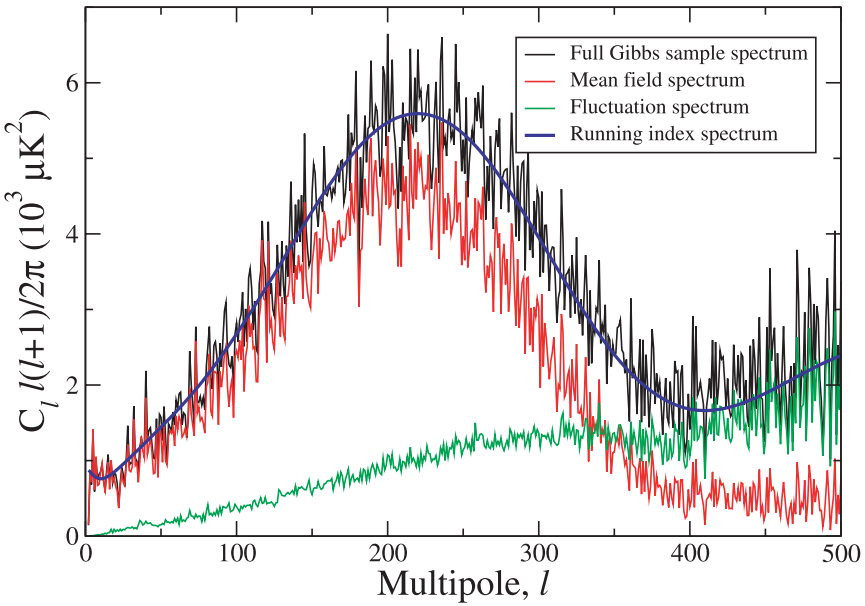
$$\begin{aligned}
\left(C^{-1} + \sum_k B_k^T N_k^{-1} B_k\right) s &= \left(\sum_k B_k^T N_k^{-1} d_k\right) \\
\left(C^{-1} + \sum_k B_k^T N_k^{-1} B_k\right) \xi &= C^{-1/2}\omega_0 + \left(\sum_k B_k^T N_k^{-1/2} \omega_k\right)
\end{aligned} \tag{10}$$

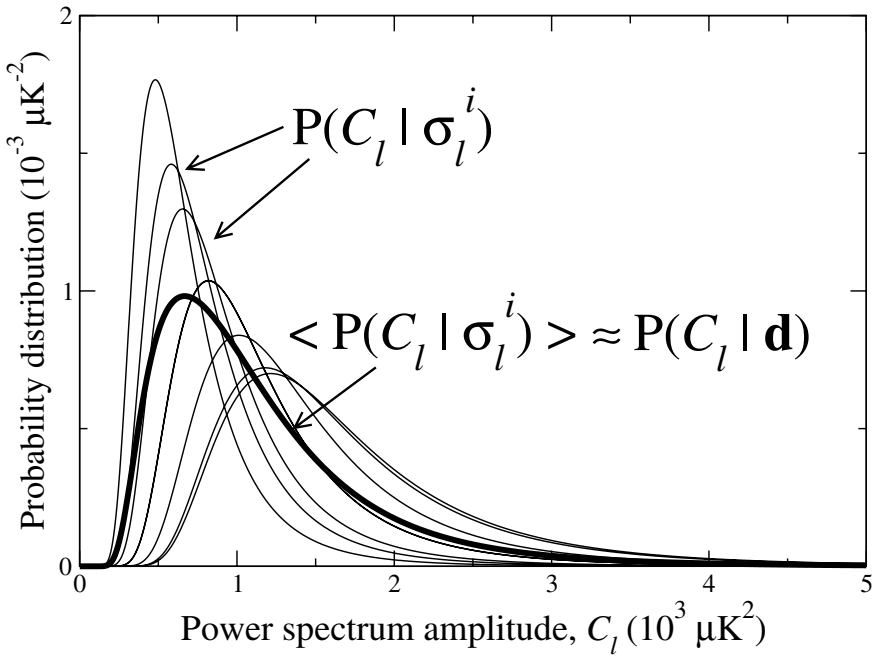
(where $\{\omega_0, \omega_{1:K}\}$ are all independent white noise maps).

- Our exact sample is $s + \xi$, since we can show

$$\text{Cov}(\xi) = \left(C^{-1} + \sum_k B_k^T N_k^{-1} B_k\right)^{-1} \tag{11}$$

Gibbs Sampling - The parameter update step . . .





- The parameter conditional is independent of the data $p(\Omega|s,d) = p(\Omega|s)$ and has the form

$$-logp(\Omega|s) \sim \left(\sum_l (l+1/2) \left(\frac{\sigma_l}{C_l(\Omega)} \right) + \log |C_l(\Omega)| \right) \tag{12}$$

  (with σ_l an empirical power spectrum estimate for inferred signal CMB map s).

- This generalizes to polarization (more generally multi-variate quantities at each pixel) with the *inverse Wishart* distribution.

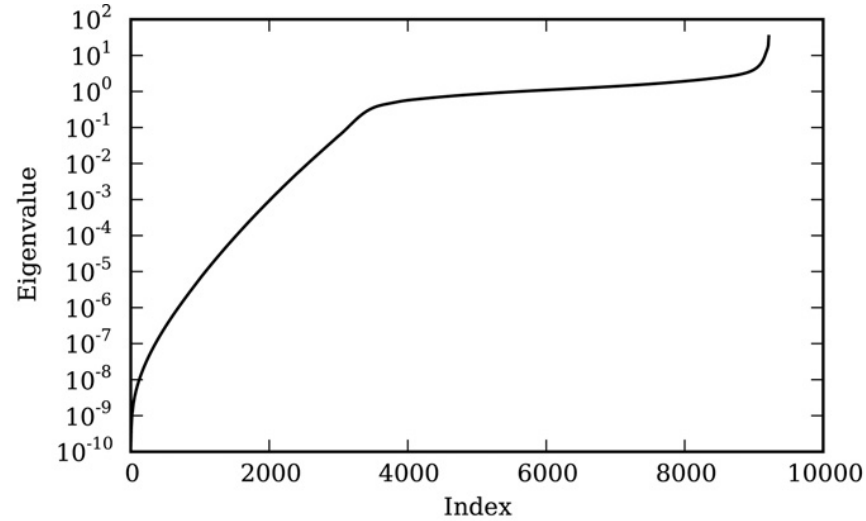
Computational Challenges for Constrained Gaussian Realizations

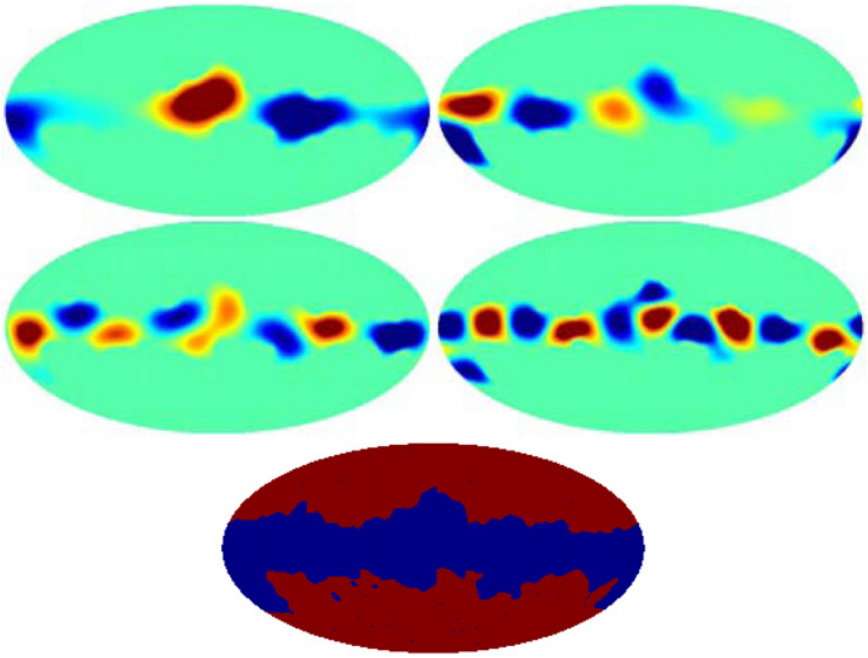
- While we have an analytic solution for exact Gaussian samples, solving the associated linear equations are typically NOT easy for large data sets with missing data (in the CMB analysis context we cannot see through the galaxy). **This will be a very common problem when combining measurements from different experiments with different spatial coverage.**
- The spatial noise N is not uniform due to the scan strategy of the instrument (however, usually close to a diagonal matrix in the pixel basis).
- The signal matrix C here is diagonal in the spherical harmonic basis (a rotationally invariant random field). In other contexts we often want to model a spatial process with eigenmodes that capture spatial correlations that are physically distinct from the spatial properties of measurement error.
- These complications lead to major computational challenges for solving the linear problem (with arbitrary right hand side X)

$$\left(I + \sum_k C^{+1/2} B_k^T N_k^{-1} B_k C^{+1/2} \right) Y \quad = \quad C^{+1/2} X \tag{13}$$

- Solving these types of problems for large spatial datasets, especially with missing data, **requires preconditioning for iterative linear solvers to work efficiently.**

Multi-Grid Approach to Constrained Realizations





- The Constrained Realization system is symmetric and positive definite, which suggests the use of the CG algorithm. For the behavior of CG and other Krylov methods, we are primarily interested in the eigen-spectrum after preconditioning
- Figure on left: Eigen-decomposition of the CR system using a diagonal preconditioner. **Top panel**: the eigenvalues for a (smoothed) 143 GHz Planck channel with a mask covering 40% of the sky. **Bottom panel**: a selection of eigenvectors corresponding to very low eigenvalues. The structure of the mask (bottom) is clearly visible in the eigenvectors.
- Problematic feature is the exponential drop in the eigenvalues (theoretical results indicate that the CG search needs at least one iteration per eigenvalue located in exponentially increasing parts of the eigenspectrum,)
- Exponential spectral feature is due to large-scale modes in the "missing data" regions!
- For this problem, the power spectrum of the signal process is roughly (spatial scale)$^{-2}$. This suggests the *inverse covariance matrix is something like* ∇^2. *This motivated an exploration of a multi-grid approach known to be useful in numerical solutions of elliptic PDE's.*

(For details see: D. S. Seljebotn et al, Astrophysical Journal Supplement Series, 210:24 (13pp), 2014 February)

Problems with Slow Mixing of Gibbs Sampling at Small Spatial Scales

- The problem - infer the signal and covariance matrix of a spatial Gaussian process of physical interest observed with *different experiments with different resolution and spatial coverage - measurement error therefore varies with scale*.
- It is common for global coverage to be at relatively low spatial resolution, while high resolution might only be probed with experiments with smaller regions.
- The end result - it is common for measurement error (and uncertainty) to greatly increase at the smallest spatial scales.
- This presents a problem when jointly inferring physical covariance matrices and the signal with Gibbs sampling! The conditional $p(C|s, d) = p(C|s)$ is typically *very concentrated about the estimator for current sample s, and much more concentrated than the total confidence interval of the posterior for the eigenvalues (associated with "small spatial scale modes") of the covariance matrix*!
- We somehow need to find an MCMC algorithm which can *make large jumps in the covariance matrix in the subspace characterizing small spatial scales - and yet to have a reasonable acceptance probability we need to adjust the signal map as well*.
- *Propose independent of current state* - optimal for mixing length but made difficult by the appearance of intractable determinant ratios in the accept probability, i.e. propose with some approximate $\tilde{p}(\Omega|D)$ followed by exact map update $p(s|\Omega, D)$ gives

$$A = \min\left[\left(\frac{p(\Omega_2, s_2|D)}{p(\Omega_1, s_1|D)}\right)\left(\frac{q(s|\Omega_1, D)\tilde{p}(\Omega_1|D)}{|N + C(\Omega_1)|}\right)\left(\frac{|N + C(\Omega_2)|}{q(s|\Omega_2, D)\tilde{p}(\Omega_2|D)}\right)\right] \tag{14}$$

Specialized MCMC for Faster Mixing at Small Spatial Scales

- Goal - make **large** changes in the covariance matrix to quickly mix in MCMC.
- Consider proposals of the form:

$$\rho(\Omega_2, s_2|\Omega_1, s_1) \quad = \quad \delta\left[s_2 - \mathcal{F}(s_1, \Omega_2, \Omega_1)\right] \; \tilde{p}(\Omega_2|d) \tag{15}$$

  ($\tilde{p}$ any approximate marginal posterior) where we make a *deterministic change in the map variables according to some function* $\mathcal{F}(s_1, \Omega_2, \Omega_1)$.

- Example - consider a move of the form which leaves $sC^{-1}(\Omega_2)s$ invariant,

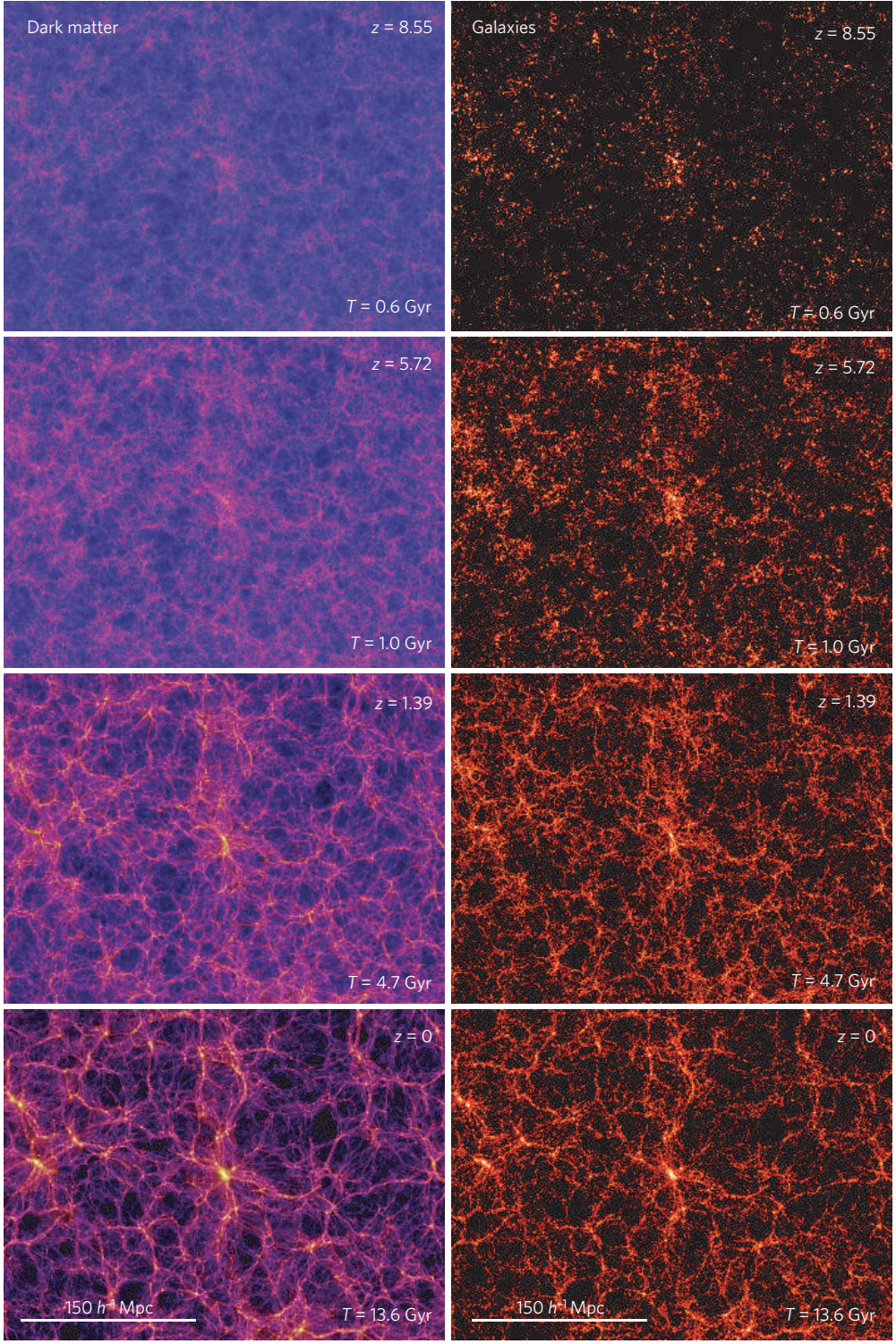
$$s_2 \quad = \quad C^{+1/2}(\Omega_2)C^{-1/2}(\Omega_1)s_1 \tag{16}$$

- The condition of detailed balance including deterministic moves requires the consideration of some technical points (requires including the determinant of the Jacobian of the functional mapping in the derivation of the accept probability).
- For this proposal we have the accept probability

$$A \quad = \quad \min\left[1, \left(\frac{e^{-(d-Bs_2)N^{-1}(d-Bs_2)}}{e^{-(d-Bs_1)N^{-1}(d-Bs_1)}}\right)\left(\frac{\tilde{p}(\Omega_1|D)}{\tilde{p}(\Omega_2|D)}\right)\right] \tag{17}$$

  (For technical details - Jewell et al, The Astrophysical Journal, 697:258268, 2009 May 20)

- **We can therefore make large moves in MCMC with the covariance matrix where the measurement error is large (small spatial scales)**.



Challenges of LSS Inference

- **Figure credit** : *The Large-Scale Structure of the Universe*, Springel, V.; et al; , Nature, Vol. 440, p. 1137, April 2006
- Random initial density perturbations (traced by the CMB) subsequently undergo gravitational collapse in a background uniformly expanding universe to form the observed "cosmic web".
- LSS data composed of the position and redshift of galaxies clustered within dark matter halos.
- Early on, and at large spatial scales, *linearized approximations* work well.
- Both current and future experiments are probing spatial scales of the Universe where the dynamics are nonlinear!
- A wealth of information to extract but mathematically and computationally challenging due to the nonlinearity of the dynamics.
- Unlike CMB (Gaussian) analysis, a computationally feasible implementation of an exact Bayesian approach to the inference of cosmological parameters from these measurements has not yet been fully developed.
- Two approaches - "direct" (as with the CMB, jointly inferring the state history and parameters), and "indirect", in which approximate likelihoods in terms of "test statistics" are used. *Both approaches have to deal with huge computational expense!*

Large Scale Structure - Theory

- Standard framework : we see luminous matter (galaxies) clustered within dark matter halos, which have formed through gravitational collapse in a background uniformly expanding universe.
- Dark matter phase-space distribution evolves according to the Vlasov-Poisson equation (in expanding coordinates)

$$\partial_t f(x,p,t) + \frac{p}{ma(\Omega)} \cdot \partial_x f - a(\Omega) m \nabla\phi \cdot \partial_p f = 0 \tag{18}$$

where the gravitational potential is sourced by the DM density

$$\nabla^2 \phi = \int d^3p \; f(x,p,t) \tag{19}$$

- Observations (spatial distribution of galaxies) probe *moments of the VP equation*, such as

$$\begin{aligned} \rho(x,t) &= \int d^3p \; f(x,p,t) \\ u(x,t)\rho(x,t) &= \int d^3p \; p \; f(x,p,t) \\ \Pi_{ab}(x,t)\rho(x,t) &= \int d^3p \; p_a p_b \; f(x,p,t) \end{aligned} \tag{20}$$

- We can find evolution equations for the moments (but - these equations are not closed!).

$$\begin{aligned} 0 &= \partial_t \rho(x,t) + \nabla \cdot (\rho u) \\ 0 &= \partial_t u(x,t) + u \cdot \nabla u + \nabla\phi + \rho^{-1} \nabla \cdot (T\rho) \end{aligned} \tag{21}$$

(Review in Bernardeau et al. , "Large-Scale Structure of the Universe and Cosmological Perturbation Theory", Physics Reports, 367, (2002), 1-248).

- **Regimes of interest**
  - Linear regime,
  - Intermediate (neglecting tensor term in velocity), treated with perturbation theory,
  - Nonlinear regime - resort to N-body simulations

Comparison of Theory and Observations - Large-Scale Structure

- As discussed, the data model for the observed large-scale structure is that the observed galaxy (smoothed) density field traces the underlying dark matter distribution, which has evolved *deterministically* from Gaussian initial conditions.
- For the purposes of discussion we write

$$d = F(s_0, \Omega) + n \tag{22}$$

  where $F(s_0, \Omega)$ represents the dark matter evolution (according to the Vlasov-Poisson equation) from initial conditions s_0 and cosmological model parameters Ω.
- As before with the CMB example, we have a joint density

$$p(d, s, s_0, \Omega) = p(d|s)\delta[(s - F(s_0, t)]p(s_0|\Omega)p(\Omega) \tag{23}$$

  or a posterior

$$p(\Omega|d) \propto p(\Omega) \int d(s, s_0)\ p(d|s)\delta[(s - F(s_0, t)]p(s_0|\Omega) \tag{24}$$

- The new element here (compared to CMB analysis) is the *nonlinear deterministic evolution*! If the theory were linear, this would not cause much of a problem as the underlying "signal" would stay Gaussian. However, the nonlinear nature of gravitational collapse generates not only a non-Gaussian signal, but one which we cannot solve analytically!

- Therefore two complications for inference given large-scale structure data:
  - Nonlinear dynamics generate non-Gaussian fields - we do not have "sufficient statistics" (i.e., sufficient statistic for Gaussian fields = power spectrum) for comparison of theory and observation
  - We have uncertainty due to *numerical error* which ideally should be taken into account in the final error bars.
- We will review two paths to the likelihood for LSS analysis - the "direct approach", or the "indirect" approach where we construct approximate likelihoods from test statistics.

"Indirect Approach" - Approximate Likelihoods for Large-Scale Structure

- The construction of approximate likelihood functions, quantify the "goodness of fit" of cosmological model parameters to observations through agreement of various chosen "test statistics" (such as power spectra).
- Where does this come from? We go from $p(d, s, s_0, \Omega) = p(d, s, s_0|\Omega)p(\Omega)$ to $p(\Lambda|\Omega)p(\Omega)$ *at potentially a great reduction in information* where

$$p(\Lambda|\Omega) = \int d(d, s, s_0)\ \delta[\Lambda - \Lambda(d)]\ p(d|s)\delta[(s - F(s_0, t)]p(s_0|\Omega) \tag{25}$$

  In the above note that Λ *is a "test statistic", i.e. it is some quantity that we choose to be informative and computationally feasible*.
- The idea is that we take our measurements d, compute the test statistic on those measurements, and have a posterior $p(\Omega|\Lambda(d))$.
- The functional form of the above is analytically unknown, cannot be directly computed, and is (almost always) approximated as a Gaussian with the "inverse Fisher matrix" as the covariance matrix

$$F^{-1}(\Omega) = E_\Omega[(\Lambda - \Lambda(\Omega)) \otimes (\Lambda - \Lambda(\Omega))] \tag{26}$$

  This quantity contains "cosmic variance", "sample variance", and "variance" due to all other systematic effects (quite a powerful matrix!)
- In practice then we need to run *a large number of Monte Carlo simlulations including N-body simulations followed by simulated measurement, to compute the ensemble averages* $\Lambda(\Omega)$ *and the variance* $F(\Omega)$.
- Given data, we can then use MCMC to explore the posterior, however we will also need some means to interpolate between the $\{\Lambda(\Omega_i), F(\Omega_i)\}$ computed on our discrete grid.

Are Gaussian approximations ok?

- By "Gaussian" approximation we essentially mean a second order expansion to $I(\Lambda|\Omega) \equiv -\log p(\Lambda|\Omega)$, i.e.

$$I(\Lambda|\Omega) \sim I(\lambda(\Omega)|\Omega) + (\partial_\Lambda I|_\lambda)(\Lambda - \lambda(\Omega)) + (\Lambda - \lambda(\Omega))\mathbf{F}^{-1}(\lambda)(\Lambda - \lambda(\Omega)) + \ldots \quad (27)$$

- This does not in general have to do with the dynamical linear regime (i.e. large-scales, where linear dynamics imply a Gaussian at all times, although with time varying power spectra). Even if the underlying signal process is Gaussian, we do not have Gaussian measures for the statistics. We seem to rely on the central limit theorem (i.e. many modes in the observational and simulation volumes).
- Note - for any value of Λ, define $\Lambda(u) = (1-u)\Lambda + u\lambda(\Omega)$. Then a Taylor expansion with remainder will give, for some $\hat{u}(\Lambda, \Omega) \in [0, 1]$ (and assuming that $\partial_\Lambda I|_\lambda = 0$)

$$-\log p(\Lambda|\Omega) \sim I(\lambda(\Omega)|\Omega) + (\Lambda - \lambda(\Omega))\mathbf{F}^{-1}(\hat{u})(\Lambda - \lambda(\Omega)) \quad (28)$$

  The inverse covariance matrix is then a function of Λ, i.e. not "stationary".
- Now - we are going to reverse our attention. Fix Λ (i.e. we have collected our data and evaluated $\Lambda(d)$). We have our original identity

$$-\log p(\Lambda(d)|\Omega) \sim I(\lambda(\Omega)|\Omega) + (\Lambda(d) - \lambda(\Omega))\mathbf{F}^{-1}(\hat{u}(d, \Omega))(\Lambda(d) - \lambda(\Omega)) \quad (29)$$

  where we defined implicitly $\lambda(d, \Omega) = (1 - \hat{u})\Lambda(d) + \hat{u}\lambda(\Omega)$ and where $\hat{u} \in [0, 1]$ by the mean value theorem.
- The above suggests that global Gaussian approximations (or more localized mixtures of Gaussians) are perhaps not that bad in "reasonable" regions of parameter space where the data and model parameters make very close predictions of the test statistics.
- Note - the above is meant as a sketch of possible explorations of how well Gaussian approximations perform - it is left as an "open problem" for further exploration to more precisely quantify the convergence in Gaussian approximations, thereby adding precision to inferences of parameters for LSS data.

Back to the Direct Approach - First Steps . . .

- Can we develop sampling algorithms to directly sample $p(s, s_0, \Omega|d)$, including all sources of uncertainty (numerical, cosmic variance, measurement errors and systematics, etc.)?
- **Key New Idea** - Uncertainty due to numerical error is included - the "$\delta[s - F(s_0, d)]$" replaced by $e^{-\beta \| R(s) \|^2}$ where we have defined the residual error

$$R(s) \equiv s - F \tag{30}$$

- Simple example - for a linear ODE $\dot{y} = My$, we would have for data $d = Hy + n$ (i.e. sampled and observed with some linear function H)

$$p(M, y(t), y_0|d) \propto e^{-(d-Hy)N^{-1}(d-Hy)} e^{-\beta \| \dot{y} - My \|^2} p(y_0) p(M) \tag{31}$$

  where $\beta > 0$ is "tuned" so that samples from our signal prior have the same average residual error as a deterministically computed numerical estimate. We could easily solve for the posterior in this case due to the linear dynamics . . .
- Our LSS posterior then becomes

$$p(\Omega|d) \propto p(\Omega) \int d(s, s_0) \; e^{-(d-s)N^{-1}(d-s)} e^{-\beta \| R(s) \|^2} p(s_0|\Omega) \tag{32}$$

- *Work in Progress* How to sample? Or at least, how to estimate the parameters and error bars?
- **What is the computational expense of sampling here as opposed to the massive suite of Monte Carlo simulations required for the computation of the "test statistics" in the "indirect" approximate likelihood approach? What is the optimal trade-off of computational expense vs. information gain??**

Summary

- Current and future ground and space-based missions are designed to not only detect, but map out with increasing precision, details of the universe in its infancy to the present-day.
- Data analysis in cosmology involves challenging problems of Bayesian inference in the context of spatial random fields, source separation, and data assimilation for non-linear dynamics observed with measurements from different instruments with their own resolution and measurement error.
- Computational advances for inference in cosmology can be synergistic with progress in other fields:
  - Data fusion from multiple sources at multiple resolutions
  - Data assimilation and uncertainty quantification in : computational fluid dynamics, climate science, weather prediction - *all involving inference in the context of nonlinear dynamics observed with measurement error.*
- Two paths to the likelihood - the "direct" and "indirect" methods
  - Example of a "direct" approach - Gibbs sampling and MCMC for Bayesian analysis of the Cosmic Microwave Background (more generally, Gaussian spatial processes)
  - Open challenge - quantify accuracy vs. computational expense for "indirect" likelihood approximation for Large-Scale structure analysis, and investigate potential advances allowing a "direct approach".
- **Ultimate goal is continued progress in computational implementations of a Bayesian approach to quantify uncertainty including measurement and computational error given large amounts of data returned from a diverse range of measurements.**

This work was partially performed at the Jet Propulsion Laboratory, California Institute of Technology , under contract with the National Aeronautics and Space Adminstration. Government sponsorship acknowledged.
Copyright, 2015, all rights reserved.