

Computation Challenges Posed by Highly-Structured Models and Computer Models in Astrophysics

David A. van Dyk

Statistics Section, Imperial College London

BIRS Workshop on Complex MC Computation, March 2012

Outline

- 1 Massive Data and Complex Models in Astronomy
- 2 Application I: X-ray Spectral Analysis and PCG Sampling
- 3 Metropolis within PCG Sampling
- 4 Application II: Stellar Evolution

Massive Data Sets and Data Streams

Dramatic increase in the quality and quantity of data:

- massive new surveys: catalogs containing T/PBs of data,
- high resolution spectrography and imaging across the electromagnetic spectrum,
- incredibly detailed movies of dynamic and explosive processes in the solar atmosphere,
- space-based telescopes tailored to specific scientific goals,
- complex (computer) models link data to underlying physics,
- data volume is growing.... astronomically!!

*Massive challenges in
statistical computation!!*

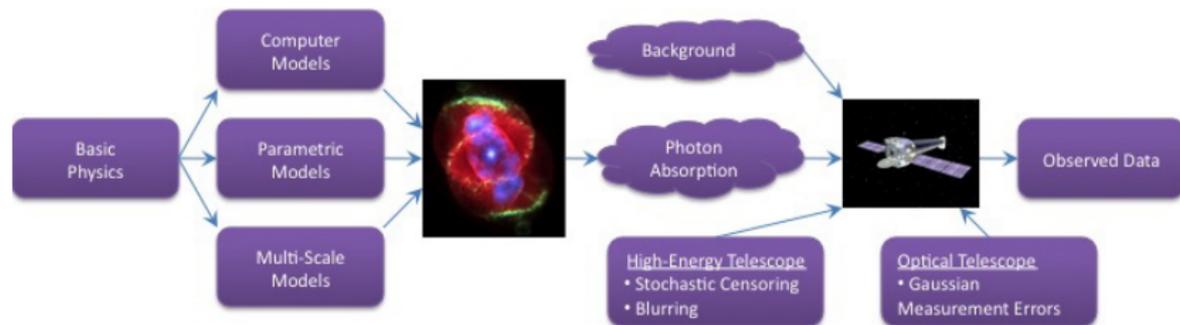


Complex Data and Sophisticated Models



- A great leap forward (~ 2018): *Large Synoptic Survey Telescope* (1.28 petabytes/year).
- But data are *not just massive*: they are rich, deep, & complex.
- LSST: Trigonometric Parallax, Proper Motion, and Photometric data in 5 bands.
- Require specialized models, methods, and computation.
- Idiosyncratic stats challenges.
- MCMC requires careful use of well-known strategies.

Application I: High-Energy Spectral Analysis¹

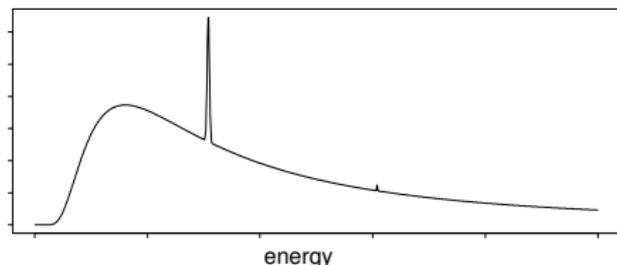


- Embed physics models into multi-level statistical models.
- X-ray and γ -ray detectors count a typically *small number of photons* in each of a *large number of pixels*.
- Must account for complexities of data generation.
- State-of-the-art data and careful computational techniques enable us to fit the resulting complex model.

Basic Spectral Analysis

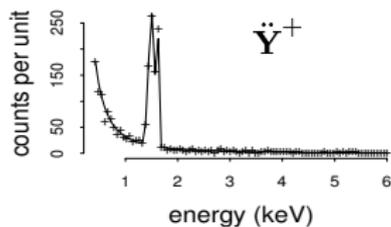
Distrib'n of energy of electromagnetic radiation

- Photon counts modeled with Poisson process.
- The Poisson parameter is a function of energy:
 $\Lambda_i = \text{function}(E_i)$.
- The *continua*, *emission lines*, and *absorption features* give clues as to the physical environment at the source.

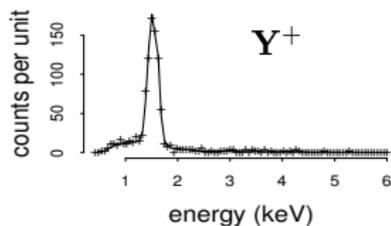


$$\Lambda_i = \left\{ \sum_{k=1}^{K^C} f_k(\theta_k^C, E_i) + \sum_{k=1}^{K^E} h_k(\theta_k^E, E_i) \right\} \prod_{k=1}^{K^A} g_k(\theta_k^A, E_i),$$

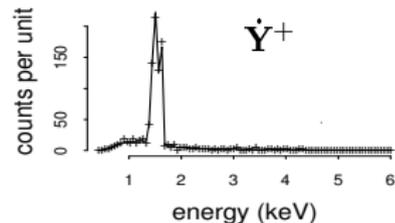
Highly Structured Models



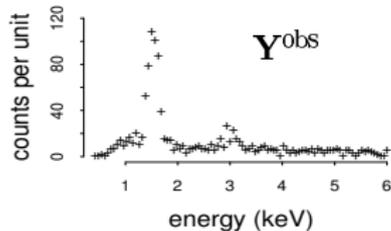
absorption and
submaximal effective
area



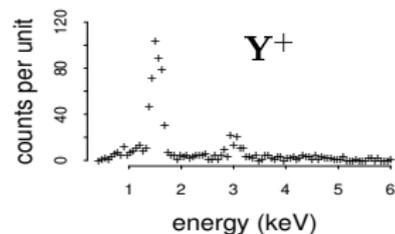
instrument
response



pile-up



background



Searching for Narrow Lines

A simplified *latent* Poisson Process for the scientific model:

$$X_i \sim \text{Poisson} \left(\Lambda_i = \alpha E_i^{-\beta} + \lambda^L \pi_i \right).$$

- With a *delta function* line, we parameterize $\{\pi_i\}$ in terms of $\theta^L =$ the location of the emission line.
- Use *Data Augmentation* to fit this finite mixture model:

Z_{il} = indicator that photon l in bin i is from the line

STEP 1: Given $Z = \{Z_{il}\}$, sample $\theta = \{\alpha, \beta, \lambda^L, \theta^L\}$

STEP 2: Given θ , sample $Z_{il} \sim \text{Ber} \left(\frac{\lambda^L \pi_i}{\alpha E_i^{-\beta} + \lambda^L \pi_i} \right)$

Unfortunately this sampler fails.

Why this Gibbs Sampler Fails

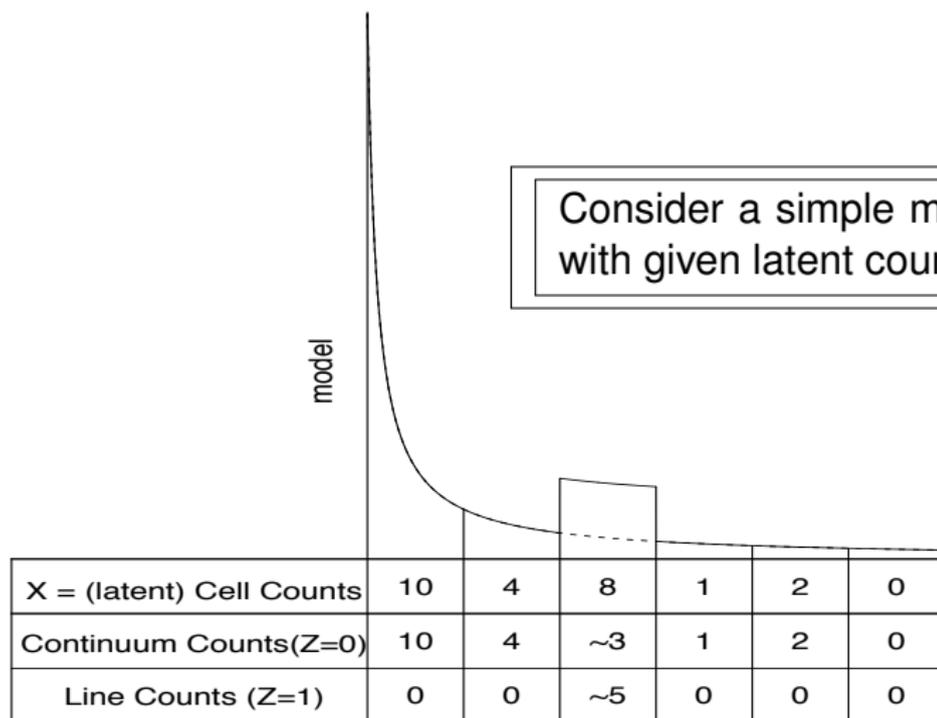
model

Consider a simple model
with given latent counts.

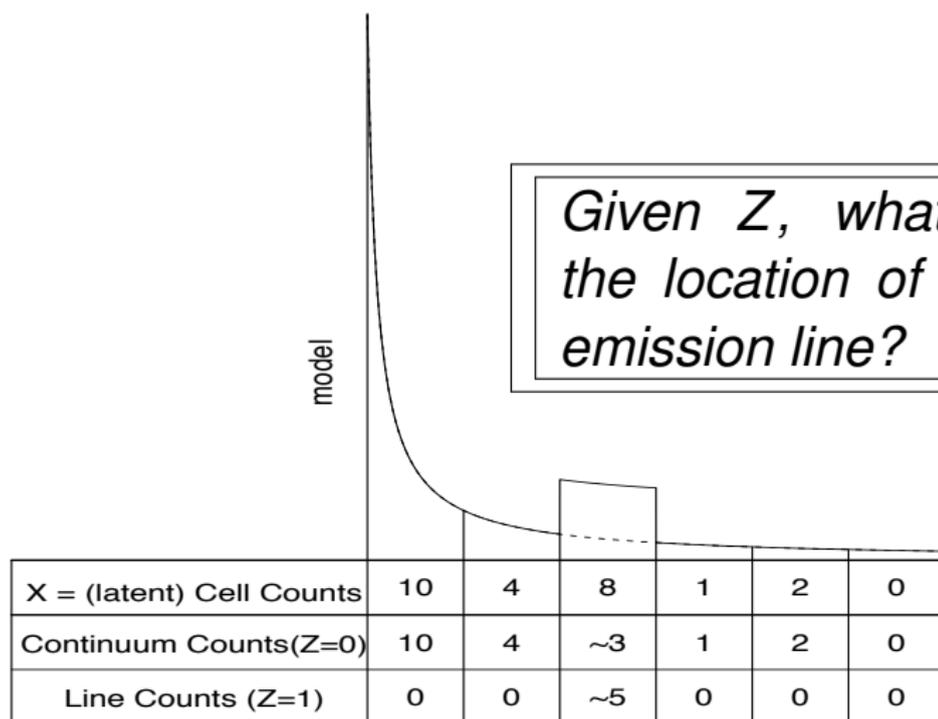
| | | | | | | |
|--------------------------|----|---|---|---|---|---|
| X = (latent) Cell Counts | 10 | 4 | 8 | 1 | 2 | 0 |
| Extend Src Counts (Z=0) | | | | | | |
| Point Src Counts (Z=1) | | | | | | |

Given this Model, what is Z?

Why this Gibbs Sampler Fails



Why this Gibbs Sampler Fails



Standard Gibbs Sampler

We do not observe the latent Poisson Process,

$$X_i \sim \text{Poisson} \left(\Lambda_i = \alpha E_i^{-\beta} + \lambda^E \pi_i \right),$$

rather we observe, $Y_j \sim \text{Poisson} \left(a_j \sum_i P_{ij} \Lambda_i + \xi_j \right)$

| | | | | |
|------------------|---|-----------|---|---------------------------|
| Y_{obs} | = | $\{Y_j\}$ | = | obs cell cnts |
| X | = | $\{X_i\}$ | = | latent cell cnts |
| Z | = | | = | emission line indicators |
| θ^L | = | | = | location of emission line |
| θ^O | = | | = | other model parameters |

Standard Gibbs sampler:

- 1 Sample $p(X, Z | \theta)$
- 2 Sample $p(\theta | X, Z) = p(\theta^O | X, Z) p(\theta^L | X, Z)$

With delta function emission line, this sampler fails.

The Partially Collapsed Gibbs Sampler

Question: Can we improve convergence by reducing the conditioning in *some* of the steps of the Gibbs sampler?

- Update θ^L without conditioning on Z or X ?
- Update θ^L without conditioning on Z ?

Sampler 1: (A Blocked Version of the Original Sampler.)

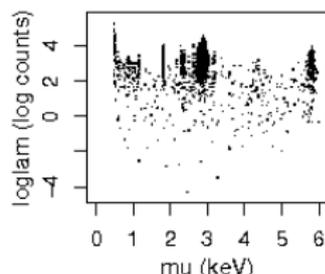
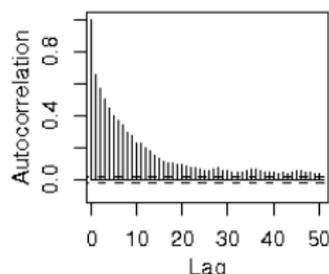
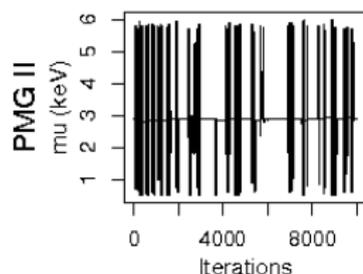
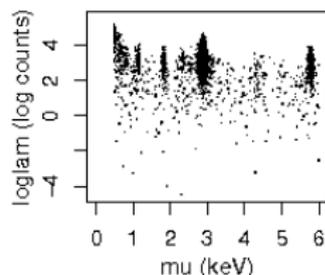
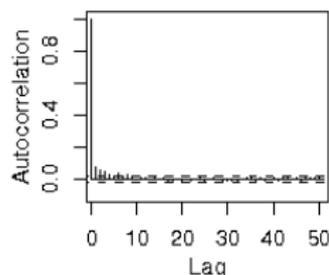
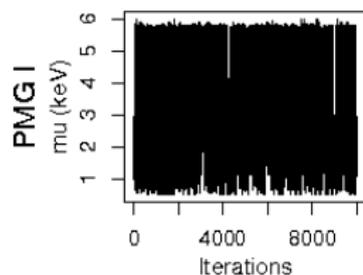
$$\begin{array}{ccccccc}
 p(X, Z|\theta) & & p(X, Z|\theta) & & p(\theta^L|\theta^O) & & \\
 p(\theta^O|\theta^L, X, Z) \longrightarrow & p(\theta^O|\theta^L, X, Z) & \longrightarrow & p(X, Z|\theta) & \longrightarrow & p(\theta^L, X, Z|\theta^O) & \\
 p(\theta^L|\theta^O, X, Z) & p(\theta^L|\theta^O) & & p(\theta^O|\theta^L, X, Z) & & p(\theta^O|\theta^L, X, Z) &
 \end{array}$$

Sampler 2: (Cannot be Blocked: An Incompatible Gibbs Sampler.)

$$\begin{array}{ccccc}
 p(X, Z|\theta) & & p(X, Z|\theta) & & p(\theta^L|\theta^O, X) \\
 p(\theta^O|\theta^L, X, Z) \longrightarrow & p(\theta^O|\theta^L, X, Z) & \longrightarrow & p(X, Z|\theta) & \\
 p(\theta^L|\theta^O, X, Z) & p(\theta^L|\theta^O, X) & & p(\theta^O|\theta^L, X, Z) &
 \end{array}$$

Computational Gains

- Standard sampler doesn't move from its starting value.
- Sampler 1 converges better than Sampler 2.
- However, each iteration of Sampler 1 takes ~ 8 times longer



Why MH within PCG Samplers?

We often require MH updates within the PCG sampler

- van Dyk and Park (2011) suggest using an initial run to generate a MH proposal for θ^L (Path-Adaptive MH).
- f_k , g_k , and h_k may be complex functions \rightarrow MH (PyBL0CXS).
- Absorption features apply to all continua and lines. There is no gain in conditioning on Z .

Recall:

$$\Lambda_i = \left\{ \sum_{k=1}^{K^C} f_k(\theta_k^C, E_i) + \sum_{k=1}^{K^E} h_k(\theta_k^E, E_i) \right\} \prod_{k=1}^{K^A} g_k(\theta_k^A, E_i),$$

More subtle than “Metropolis within Gibbs”.

A Closer Look at the PCG Sampler

A Gibbs sampler:

STEP 1: $\psi_1 \sim p(\psi_1|\psi_2)$

STEP 2: $\psi_2 \sim p(\psi_2|\psi_1)$

A Partially Collapsed Gibbs (PCG) Sampler:

STEP 1: $\psi_1 \sim p(\psi_1|g(\psi_2))$

STEP 2: $\psi_2 \sim p(\psi_2|h(\psi_1))$

- g and/or h are non-invertible functions.
- Generalizes blocking & collapsing, involves incompatibility.
- Step order can effect stationary distribution.
- Improves convergence rate (van Dyk & Park, 2008, JASA).
- Spectral analysis, time series, and multiple imputation (van Dyk & Park; 2011 MCMC Hndbk; 2009 JCGS; 2008 ApJ).

MH within Gibbs & PCG Samplers

An MH within Gibbs sampler:

STEP 1: $\psi_1 \sim \mathcal{K}(\psi_1|\psi)$ via MH with limiting dist. $p(\psi_1|\psi_2)$

STEP 2: $\psi_2 \sim p(\psi_2|\psi_1)$

Using MH within the Partially Collapsed Gibbs Sampler:

STEP 1: $\psi_1 \sim \mathcal{K}(\psi_1|\psi)$ via MH with limiting dist. $p(\psi_1)$

STEP 2: $\psi_2 \sim p(\psi_2|\psi_1)$

- If MH is unnecessary, obtain i.i.d. draws from $p(\psi_1, \psi_2)$.
- With MH we must verify the stationary distribution.
- Improved convergence if ψ_1 and ψ_2 are highly correlated.
- Need only evaluate $p(\psi_1) = p(\psi_1, \psi_2)/p(\psi_2|\psi_1)$.

But... Be Careful!

Another MH within Gibbs Sampler:

STEP 1: $\psi_1 \sim p(\psi_1|\psi_2)$

STEP 2: $\psi_2 \sim \mathcal{M}(\psi_2|\psi_1)$ via MH with limiting dist. $p(\psi_2|\psi_1)$

A *naive* Sampler:

STEP 1: $\psi_1 \sim p(\psi_1)$

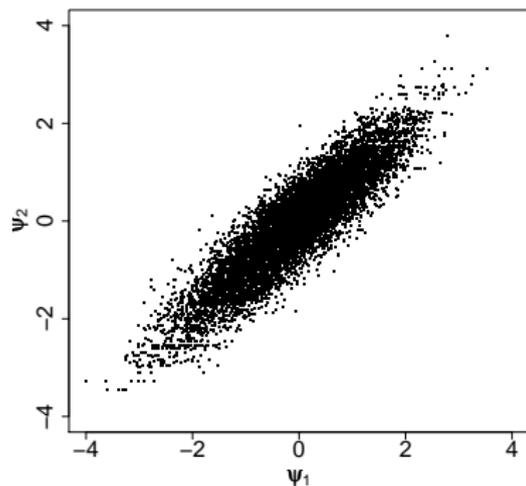
STEP 2: $\psi_2 \sim \mathcal{M}(\psi_2|\psi_1)$ via MH with limiting dist. $p(\psi_2|\psi_1)$

Simulation Study:

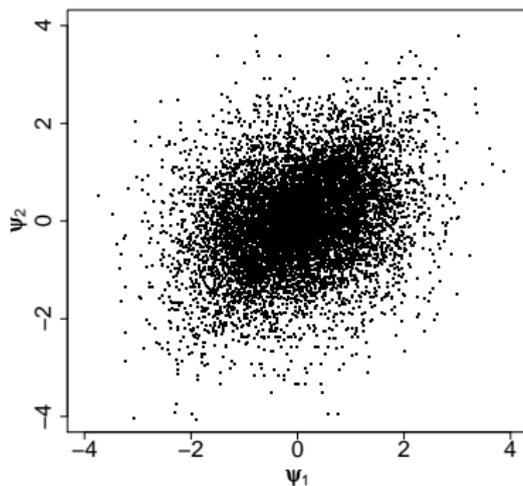
- Suppose $\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \sim N_2 \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0.9 \\ 0.9 & 1 \end{pmatrix} \right]$
- MH: a Gaussian jumping rule centered at previous draw.

Be Careful When Combining MH and PCG Sampling

MH within Gibbs Sampler



The *naive* Sampler



What Goes Wrong

The *naive* Sampler:

STEP 1: $\psi_1^{(t)} \sim p(\psi_1)$

STEP 2: $\psi_2^{(t)} \sim \mathcal{M}(\psi_2 | \psi_1^{(t)}, \psi_2^{(t-1)})$ via Metropolis Hastings

The update of ψ_2 depends on both $\psi_1^{(t)}$ and $\psi_2^{(t-1)}$:

- The limiting distribution of the MH step is $p(\psi_2 | \psi_1^{(t)})$.
- If the proposal is rejected, ψ_2 is set to $\psi_2^{(t-1)}$.

BUT: $\psi_1^{(t)} \sim p(\psi_1)$ —independent of $\psi_2^{(t-1)}$ *at every iteration*.

STEP 2 *will never produce samples from* $p(\psi_2 | \psi_1)$.

Constructing a Legitimate MH within PCG Sampler

1. Marginalizing

$$\begin{array}{l} p(\psi_1|\psi'_2) \\ \mathcal{K}(\psi_2|\psi') \text{ w/} \\ \text{limit } p(\psi_2|\psi_1) \end{array} \longrightarrow \begin{array}{l} p(\psi_1|\psi'_2) \\ \mathcal{K}(\psi_1, \psi_2|\psi') \text{ w/} \\ \text{limit } p(\psi_1, \psi_2) \end{array}$$

Move quantities from the right to the left of the conditioning sign. This does not alter the stationary dist'n, but improves the rate of convergence.

2. Permuting

$$\longrightarrow \begin{array}{l} \mathcal{K}(\psi_1, \psi_2|\psi') \text{ w/} \\ \text{limit } p(\psi_1, \psi_2) \\ p(\psi_1|\psi'_2) \end{array}$$

Permute the order of the steps. This can have minor effects on the rate of convergence, but does not affect the stationary distribution.

3. Trimming

$$\longrightarrow \begin{array}{l} \mathcal{K}(\psi_2|\psi') \text{ w/} \\ \text{limit } p(\psi_2) \\ p(\psi_1|\psi'_2) \end{array}$$

Remove quantities that are not part of the transition kernel. This does not effect the stochastic mapping or the rate of convergence.

Example I: Back to the Spectral Analysis...

We do not observe the latent Poisson Process,

$$X_i \sim \text{Poisson} \left(\Lambda_i = \alpha E_i^{-\beta} + \lambda^E \pi_i \right),$$

rather we observe, $Y_j \sim \text{Poisson} \left(a_j \sum_i P_{ij} \Lambda_i + \xi_j \right)$

| | | | | |
|------------------|---|-----------|---|---------------------------|
| Y_{obs} | = | $\{Y_j\}$ | = | obs cell cnts |
| X | = | $\{X_i\}$ | = | latent cell cnts |
| Z | = | | = | emission line indicators |
| θ^L | = | | = | location of emission line |
| θ^O | = | | = | other model parameters |

Updating θ^L is costly:

- $p(\theta^L | Y_{\text{obs}})$
- $p(\theta^L | X, Y_{\text{obs}})$

must both be evaluated on a grid.

Use an adaptive Metropolis jumping rule instead.

Constructing the MH within PCG Sampler

Updating θ^L without conditioning on X or Z .

- With no Metropolis Steps, we get a blocked version of the original sampler.

Sampler 1 without Metropolis:

$$\begin{array}{ccccccc}
 p(X, Z|\theta) & & p(X, Z|\theta) & & p(\theta^L, X, Z|\theta^O) & & \\
 p(\theta^O|\theta^L, X, Z) & \longrightarrow & p(\theta^O|\theta^L, X, Z) & \longrightarrow & p(X, Z|\theta) & \longrightarrow & p(\theta^L, X, Z|\theta^O) \\
 p(\theta^L|\theta^O, X, Z) & & p(\theta^L, X, Z|\theta^O) & & p(\theta^O|\theta^L, X, Z) & & p(\theta^O|\theta^L, X, Z)
 \end{array}$$

- But *With* a Metropolis step for θ^L , blocking is not possible.

Sampler 1 with Metropolis:

$$\begin{array}{ccccccc}
 p(X, Z|\theta) & & p(X, Z|\theta) & & \mathcal{M}(\theta^L, X, Z|\theta^O) & & \mathcal{M}(\theta^L|\theta^O) \\
 p(\theta^O|\theta^L, X, Z) & \longrightarrow & p(\theta^O|\theta^L, X, Z) & \longrightarrow & p(X, Z|\theta) & \longrightarrow & p(X, Z|\theta) \\
 \mathcal{M}(\theta^L|\theta^O, X, Z) & & \mathcal{M}(\theta^L, X, Z|\theta^O) & & p(\theta^O|\theta^L, X, Z) & & p(\theta^O|\theta^L, X, Z)
 \end{array}$$

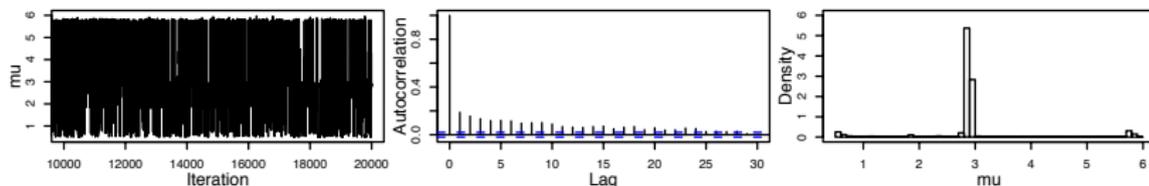
With e.g., $\mathcal{M}(\theta^L, X, Z|\theta^O)$ having jumping rule, $p(X, Z|\theta)J(\theta^L|\theta^{L'}, \theta^O)$.

“**Sampler 1 with Metropolis**” is a generalization, not a special case, of Metropolis within Gibbs Sampler.

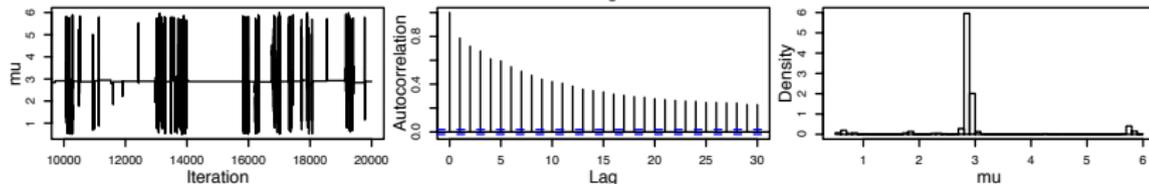
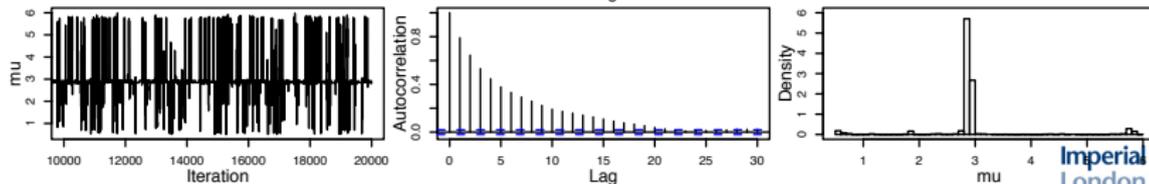
Efficiency of the MH within PCG Sampler

- Sampler 1 converges by far the best, but.....
- each iteration takes ~ 8 times longer than Sampler 2
- and fifty times longer than Sampler 3 (on average).

PCG I



PCG II

PCG 1
w/ MH

Example II: Absorption Features

An absorbed spectrum with two continua

$$\Lambda_i = \left\{ f_1(\theta_1^C, E_i) + f_2(\theta_2^C, E_i) \right\} g(\theta^A, E_i),$$

An MH within PCG sampler

- Again we let Z equal the indicator for the finite mixture.
- Each of θ_1^C , θ_2^C , and θ^A may require MH.
- Z is beneficial for θ_1^C and θ_2^C , but not for θ^A .

Sampler:

$$\begin{array}{ccccccc}
 p(Z|\theta) & & p(Z|\theta) & & \mathcal{M}(\theta^A, Z|\theta_1^C, \theta_2^C) & & \mathcal{M}(\theta^A|\theta_1^C, \theta_2^C) \\
 \mathcal{M}(\theta^A|\theta_1^C, \theta_2^C, Z) \longrightarrow & \mathcal{M}(\theta^A, Z|\theta_1^C, \theta_2^C) \longrightarrow & p(Z|\theta) & \longrightarrow & p(Z|\theta) & \longrightarrow & p(Z|\theta) \\
 \mathcal{M}(\theta_1^C, \theta_2^C|\theta^A, Z) & \mathcal{M}(\theta_1^C, \theta_2^C|\theta^A, Z) & \mathcal{M}(\theta_1^C, \theta_2^C|\theta^A, Z) & & \mathcal{M}(\theta_1^C, \theta_2^C|\theta^A, Z) & & \mathcal{M}(\theta_1^C, \theta_2^C|\theta^A, Z)
 \end{array}$$

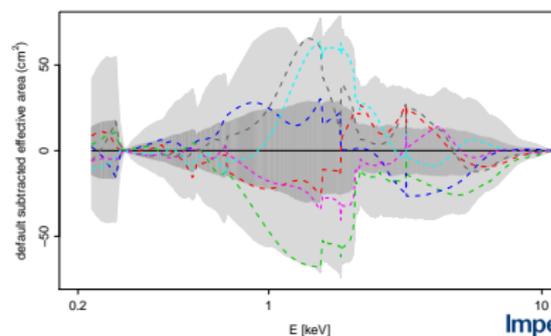
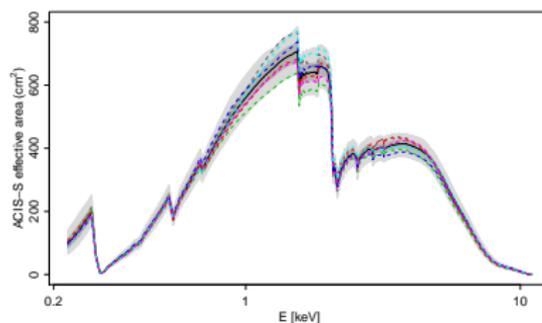
Example III: Instrumental Calibration...

We do not observe the latent Poisson Process,

$$X_i \sim \text{Poisson} \left(\Lambda_i = \alpha E_i^{-\beta} + \lambda^E \pi_i \right),$$

rather we observe, $Y_j \sim \text{Poisson} \left(a_j \sum_i P_{ij} \Lambda_i + \xi_j \right)$ with

Effective Area: $A = (a_1, \dots, a_J)$.



Two Possible Target Distributions

We consider inference under:

A PRAGMATIC BAYESIAN TARGET: $\pi_0(A, \theta) = p(A)p(\theta|A, Y)$.

THE FULLY BAYESIAN POSTERIOR: $\pi(A, \theta) = p(A|Y)p(\theta|A, Y)$.

Concerns:

Statistical Fully Bayesian target is “correct”.

Cultural Astronomers have concerns about letting the current data influence calibration products.

Computational Both targets pose challenges,
but pragmatic Bayesian target is easier to sample.

Practical How different are $p(A)$ and $p(A|Y)$?

With MCMC we sample a different effective area curve at each iteration according to its conditional distribution.

Pragmatic Bayes: PCG Sampler

Simple MCMC for $\pi_0(A, \theta) = p(A)p(\theta|A, Y)$:

- Sample effective area uniformly from calibration sample:
 $A \sim \pi_0(A) = p(A)$.
- Sample model parameters in the usual way, conditioning on the current sample of the effective area:
 $\theta \sim p(\theta|A, Y)$.

Unfortunately, update of θ uses MH (pyBL0CXS in Sherpa).

The naive Sampler Revisited:

STEP 1: $A^{(t)} \sim p(A)$

STEP 2: $\theta^{(t)} \sim \mathcal{M}(\theta|A^{(t)}, \theta^{(t-1)})$ via Metropolis.

Two Simple Solutions

Two possible samplers

- 1 A PCG (Simple Collapsed) Gibbs Sampler:

STEP 1: $A^{(t)} \sim p(A)$

STEP 2: Sample $\theta^{(t-1+\ell/L)} \sim \mathcal{M}(\theta|A^{(t)}, \theta^{(t-1)})$
 L times via MH to obtain $\theta^{(t)} \sim p(\theta|A^{(t)})$.

- 2 A pure MH Sampler:

Jumping Rule: $(A^*, \theta^*) \sim p(A^*)\mathcal{M}(\theta^*|A^*, \theta^{(t-1)})$.

Tradeoff: MH is faster, PCG gives independent draws.

Application II: Stellar Evolution²

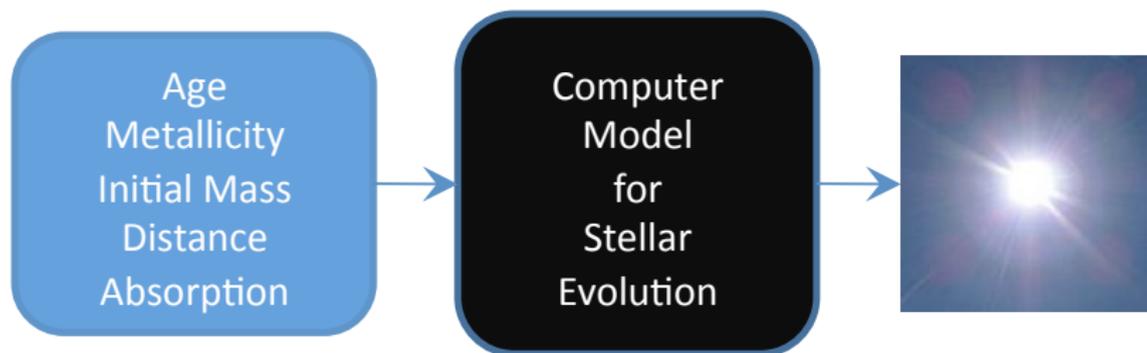
Complex Data and Sophisticated Models

- 1 Complex computer models and simulations are taking the place of the analytic likelihood function.
- 2 Sophisticated data allows us to fit such models, but an entirely new set of methods is required.
- 3 This sort of modeling, computing, and inference is coming to many more areas of Astronomy.
- 4 I will discuss one example in detail: stellar evolution.

Challenge is acute when complex models are combined with massive data streams.

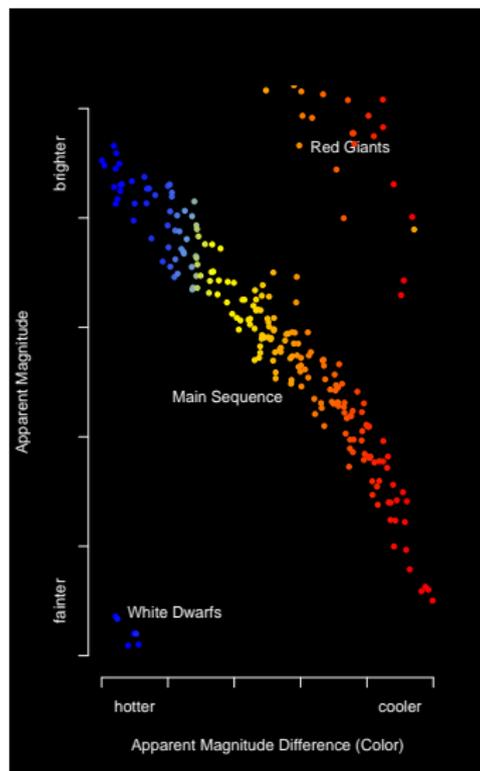
²N Stein, D van Dyk, S DeGennaro, E Jeffery, W Jefferys, and T von Hippel

Computer Model for Sun-Like Stellar Evolution



- Computer model predicts how a the spectrum of a sun-like star evolves as a function of input parameters.
- We aim to embed these models into a sophisticated multi-level model for statistical inference.

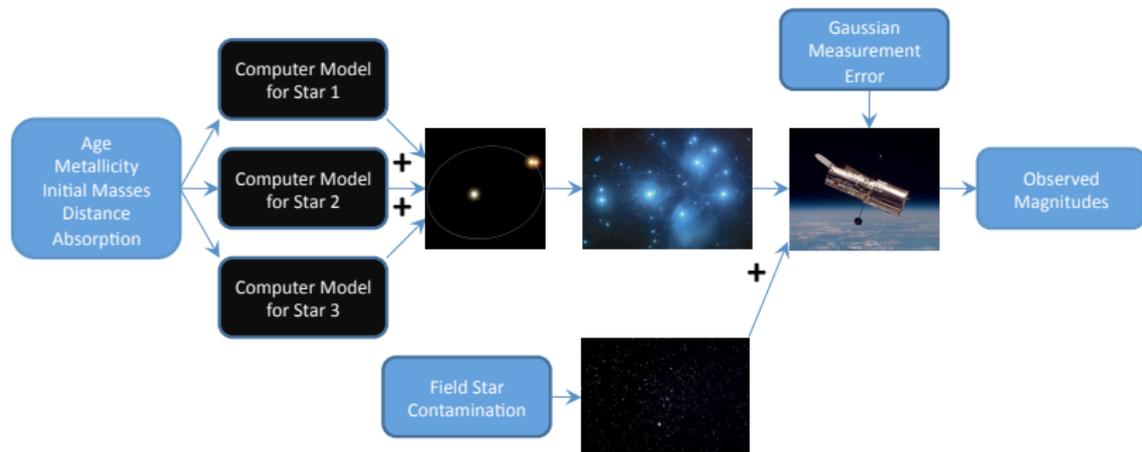
The Data: Color Magnitude Diagrams



Color-Magnitude Diagram

- Plot Magnitude Difference vs. Magnitude.
- Identifies stars at different stages of their lives.
- Evolution of a CMD.
- Facilitates physical intuition as to likely values of parameters.
- “Chi-by-eye” fitting.

Embedding Computer Model into Statistical Model



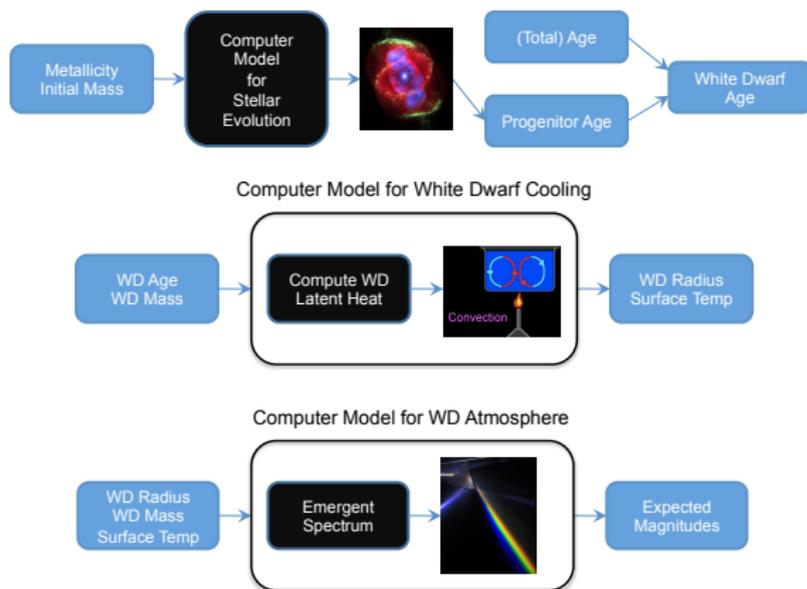
- Between 1/3 and 1/2 of “stars” are unresolved binaries.
- Star clusters: same age, metallicity, distance, & absorption.
- Cluster data is contaminated with field stars.
- Data observed with Gaussian measurement errors.

White Dwarfs Physics



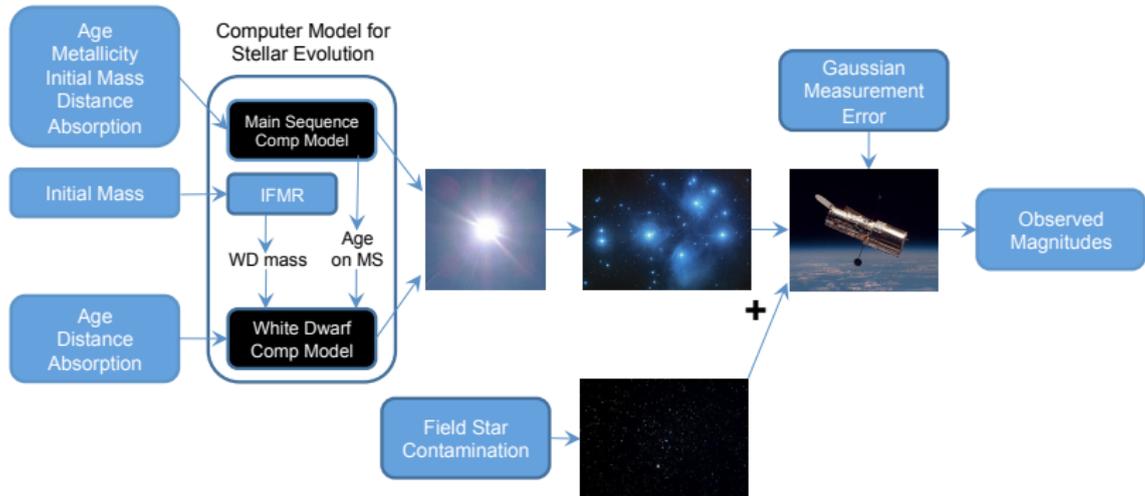
- Sun-like stars are powered by thermal-nuclear reactions.
- White dwarfs are the cooling embers after reactions cease.
- Different physical processes require different models.
- White dwarf colors are highly informative for parameters.

The Missing Link: White Dwarf Mass

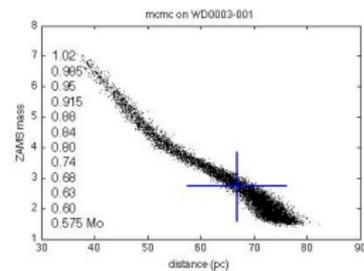
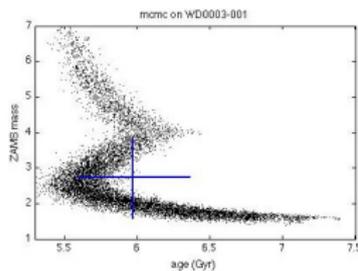
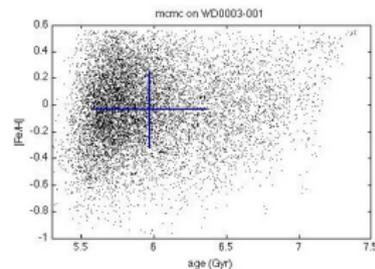
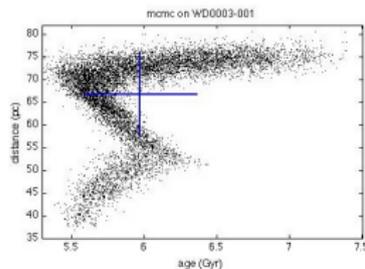


- Must include IFMR: white dwarf mass = $f(\text{initial mass})$.
- Parametric Bridge between Computer Models.

Opening Up the Black Box: The Final Model

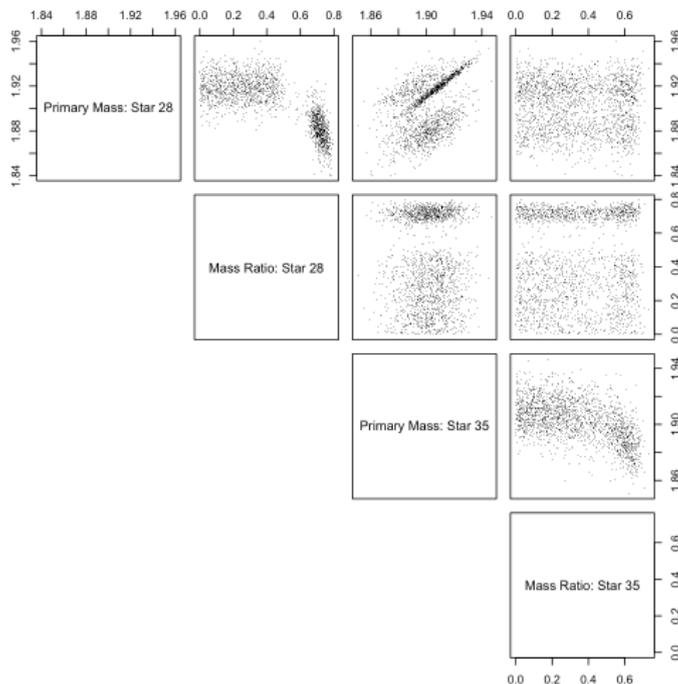


Model Fitting: Complex Posterior Distributions



Highly non-linear relationship among stellar parameters.

Model Fitting: Complex Posterior Distributions



The classification of certain stars as field or cluster stars can cause multiple modes in the distributions of other parameters.

Basic MCMC Strategy

Metropolis within Gibbs Sampling

- Hundreds of parameters
 - **Stellar:** Mass, Mass Ratio, Cluster Membership
 - **Cluster:** Age, Metallicity, Distance, Absorption
- Strong posterior correlations among the parameters.

Evaluation of Computer Stellar Evolution Model is *Very Costly*.

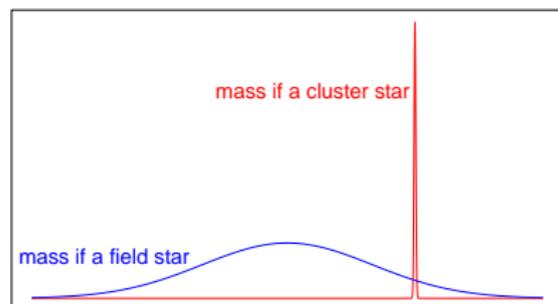
- Instead we use a tabulated form to avoid online evaluation.
- Evaluation points are not evenly spaced, but chosen to capture the complexity of the underlying function.
- Tables provided by developers of computer models.

Correlation Reduction with alternative Prior Dist'n

Field/Cluster Indicator is Highly Correlated with Masses

- Data are uninformative for the masses of field stars.
- Data are highly informative for cluster star masses.
- Cannot easily jump from field to cluster star designation.

Solution: Replace prior for masses given field star membership by approximation of the posterior given cluster star membership.



Does not effect statistical inference & enables efficient mixing.

Correlation Reduction via Dynamic Transformations

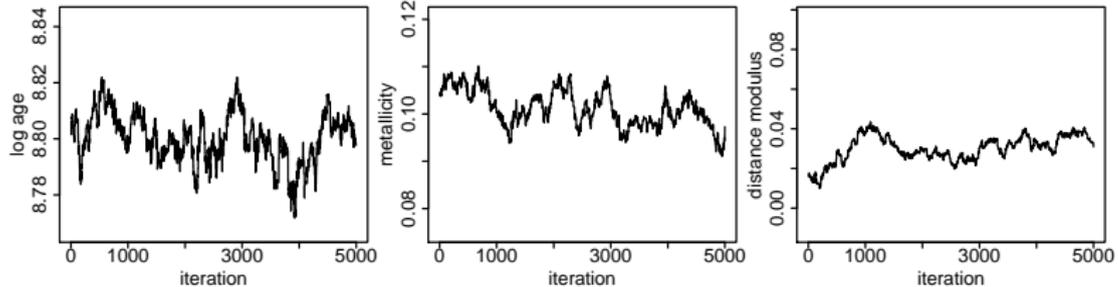
Strong Linear and Non-Linear Correlations Among Parameters

- Static and/or dynamic (power) transformations remove non-linear relationships.
- A series of preliminary runs is used to evaluate and remove linear correlations.
- We tune a linear transformation to the correlations of the posterior distribution on the fly.
- Results in a dramatic improvement in mixing.

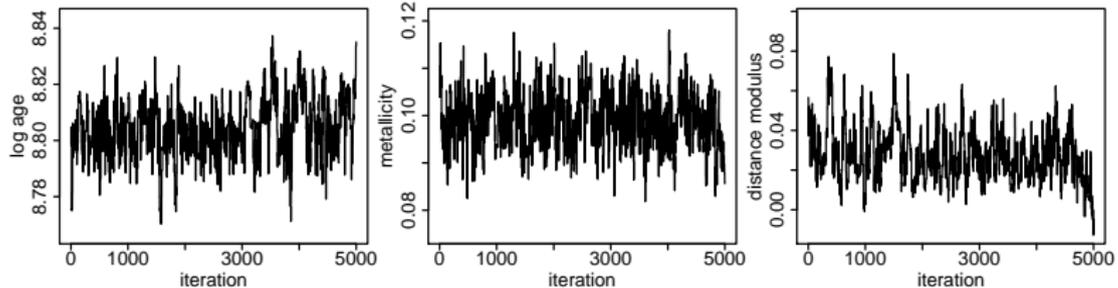
In practice MCMC requires careful implementation of well-known strategies.

Correlation Reduction via Dynamic Transformations

Initial Burn-in Period



After Dynamic Transformation



Collapsing to Handling Multiple Modes

Computation is especially challenging when fitting the IFMR

- Hundreds of parameters

Stellar: Mass, Mass Ratio, Cluster Membership

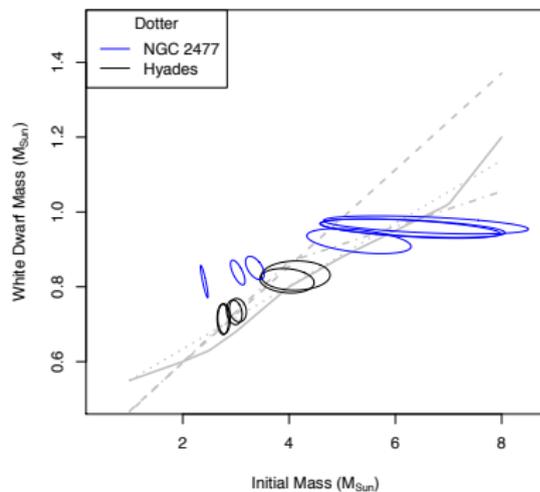
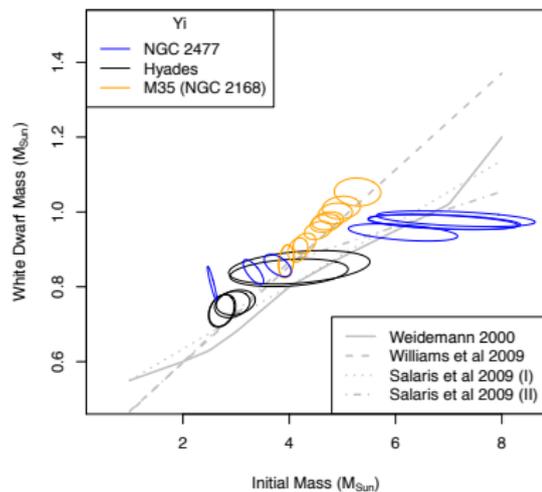
Cluster: Age, Metallicity, Distance, Absorption

General: IFMR slope, IFMR intercept

- Strategy: numerically integrate out stellar parameters and use Metropolis on remaining six parameters.
- Marginal posterior factors into N_{stars} 2D integrals.
- Computer code for MCMC is easy to parallelize.

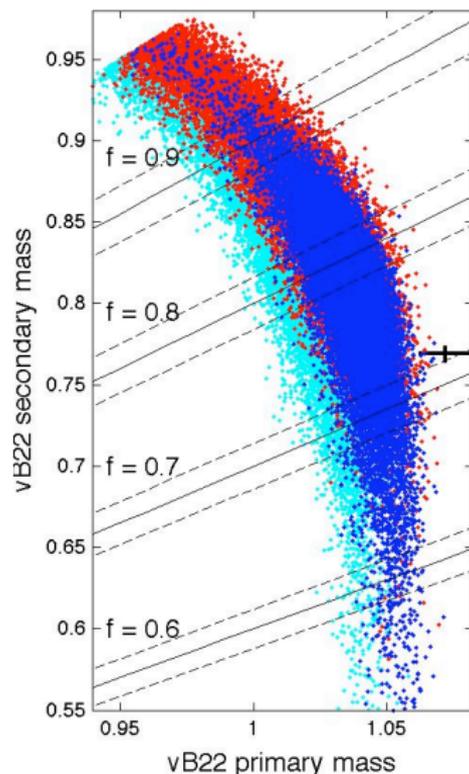
Result: Fast Mixing but computationally expensive code.

Fitting the Initial-Final Mass Relationship



- How best to combine results from three clusters?
- Is there one relationship? Depend on other variables?

Diagnosing Complex Models



- Double-Line Eclipsing Binaries: direct measures of component masses.
- Double line Spectroscopic: direct measure of mass ratio.
- Direct check of a quantity that resides deep in our statistical model and is highly model dependent.
- Use discrepancies to diagnose and tune computer models, and/or build a joint model.

Thanks...

Stellar Evolution:

- Nathan Stein
- Steven DeGennaro
- Elizabeth Jeffery
- William H. Jefferys
- David Stenning
- Ted von Hippel

X-ray Spectral Analysis:

- Taeyoung Park
- Alanna Connors
- Vinay Kashyap
- Aneta Siegminowska

And

iCHASC:

Imperial-California-Harvard
AstroStatistics Collaboration

For Further Reading I



Stein, N, van Dyk, D., von Hippel, T., DeGennaro, S., Jeffery, E., Jeffreys, W. H.

Combining Computer Models in a Principled Bayesian Analysis: From Normal Stars to White Dwarf Cinders.
Under Revision.



Lee, H., Kashyap, V., van Dyk, D., Connors, A., Drake, J., Izem, R., Min, S., Park, T., Ratzlaff, P.,
Siemiginowska, A., and Zezas, A.

Accounting for Calibration Uncertainties in X-ray Analysis: Effective Area in Spectral Fitting.
The Astrophysical Journal, **731**, 126–144, 2011.



van Dyk, D. A. and Park, T.

Partially Collapsed Gibbs Sampling & Path-Adaptive Metropolis-Hastings in High-Energy Astrophysics.
In *Handbook of Markov Chain Monte Carlo* (Editors: Brooks, Gelman, Jones and Meng), Chapman &
Hall/CRC Press, 383-399, 2011.



Park, T. and van Dyk, D. A.

Partially Collapsed Gibbs Samplers: Illustrations and Applications.
Journal of Computational and Graphical Statistics, **18**, 283-305, 2009.



van Dyk, D. A., DeGennaro, S., Stein, N., Jefferys, W. H., von Hippel, T.

Statistical Analysis of Stellar Evolution

The Annals of Applied Statistics **3**, 117-143, 2009.



van Dyk, D. A. and Park, T.

Partially Collapsed Gibbs Samplers: Theory and Methods.

Journal of the American Statistical Association, **103**, 790–796, 2008.