Computational Challenges in the Statistical Analysis of Stellar Evolution

Nathan Stein

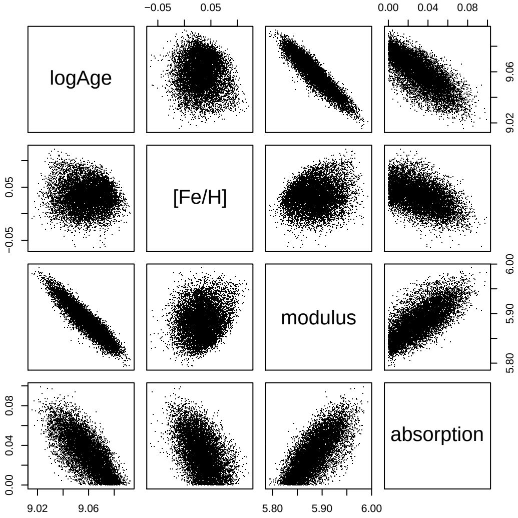
with David van Dyk, Ted von Hippel, Steven DeGennaro,
William H. Jefferys, Elizabeth Jeffery

October 20, 2009

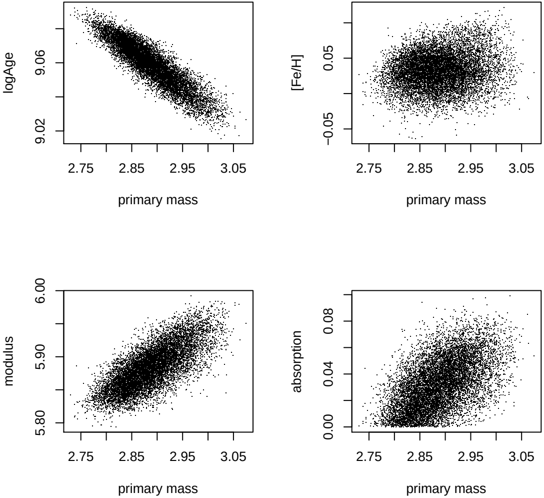
The Model

- $\mathbf{y}_i = (y_{i1}, \ldots, y_{iJ})$ = vector of magnitudes observed through J different filters
- (M_{i1}, M_{i2}) = primary and secondary mass of star i
- $\boldsymbol{\theta}$ = vector of cluster parameters
- $\mathbf{G}(M, \boldsymbol{\theta})$ = deterministic stellar evolution model
- Observational uncertainties $\boldsymbol{\Sigma}_i$ assumed known
- Gaussian errors:
$$\mathbf{y}_i | M_i, \boldsymbol{\theta}, \boldsymbol{\Sigma}_i \sim N(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i),$$
- For single-star systems, $\mu_{ij} = G_j(M_{i1}, \boldsymbol{\theta})$
- For main sequence-main sequence binaries,
$$\mu_{ij} = -2.5 \log_{10} \left(10^{-G_j(M_{i1}, \boldsymbol{\theta})/2.5} + 10^{-G_j(M_{i2}, \boldsymbol{\theta})/2.5} \right)$$
- Mixture model to account for field star contamination
- Informative prior distributions on physical parameters

Posterior Correlations



Posterior Correlations

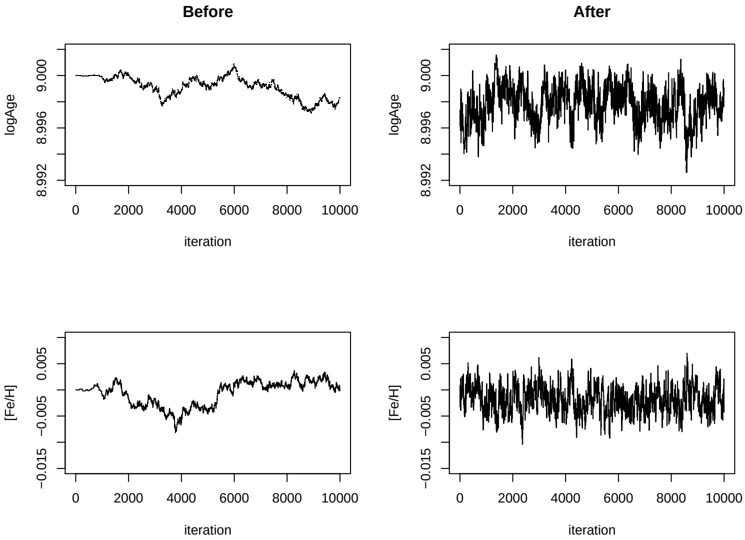


Improving Mixing

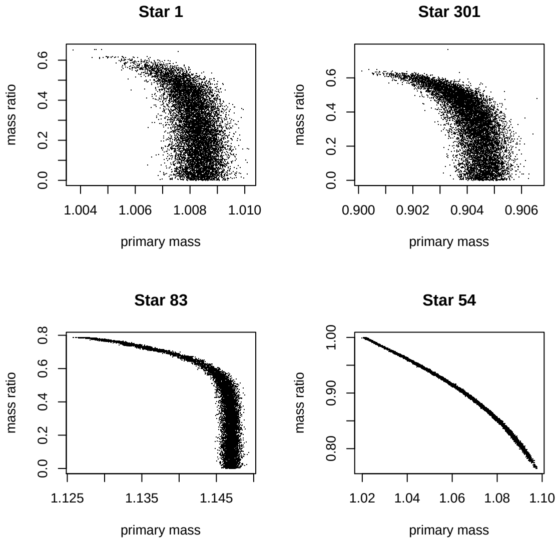
- Widths of proposal distributions (Metropolis jumping rules) are automatically tuned during burn-in
- Parameters are transformed to remove linear correlations

$$
\begin{aligned}
M_{i1} &= U_i + \beta_{R,i}(R_i - \hat{R}_i) + \beta_{\text{age},i}(\theta_{\text{age}} - \hat{\theta}_{\text{age}}) \\
&\quad + \beta_{[\text{Fe/H}],i}(\theta_{[\text{Fe/H}]} - \hat{\theta}_{[\text{Fe/H}]}) + \beta_{m-M_V,i}(\theta_{m-M_V} - \hat{\theta}_{m-M_V}) \\
\theta_{A_V} &= V + \gamma_{[\text{Fe/H}]}(\theta_{[\text{Fe/H}]} - \hat{\theta}_{[\text{Fe/H}]}) + \gamma_{m-M_V}(\theta_{m-M_V} - \hat{\theta}_{m-M_V})
\end{aligned}
$$

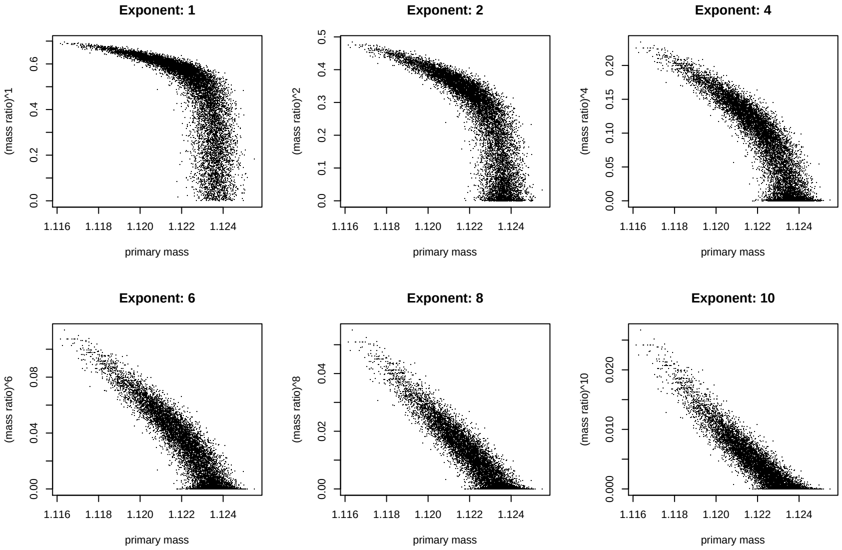
Improved Mixing



Removing Linear Correlations Is Not Enough

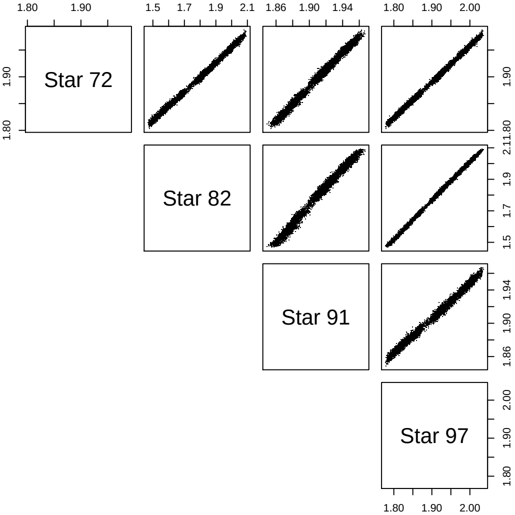


Power Law



Exponent: 1
Exponent: 2
Exponent: 4
Exponent: 6
Exponent: 8
Exponent: 10
(mass ratio)^1
(mass ratio)^2
(mass ratio)^4
(mass ratio)^6
(mass ratio)^8
(mass ratio)^10
primary mass
1.116 1.118 1.120 1.122 1.124

More Correlations: 'Decorrelated' Masses?



Accelerating MCMC

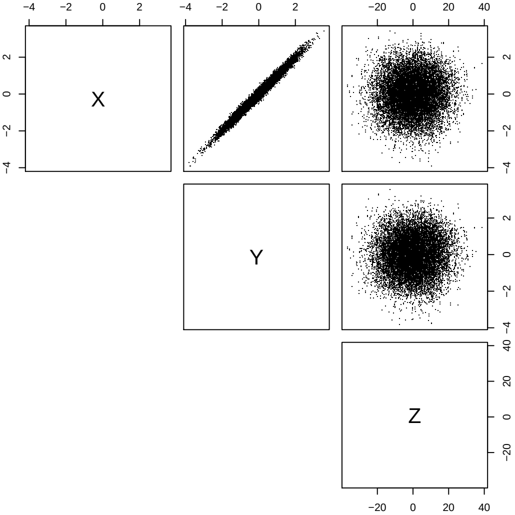
- Want to sample $\pi(\boldsymbol{\theta})$ with $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_D) \in \boldsymbol{\Theta}$
- Can obtain approximate sample (e.g., via trial run of inefficient MCMC sampler)
- Choose threshold $c \in (0, 1)$
- r_{ij} = sample correlation of θ_i and θ_j
- $\mathcal{I} = \{i : |r_{ij}| \geq c \text{ for some } j \neq i\}$
- $M = |\mathcal{I}|$
- $\boldsymbol{\theta} = (\boldsymbol{\theta}_{[\mathcal{I}]}, \boldsymbol{\theta}_{[-\mathcal{I}]})$
- $\{\mathbf{w}_1, \ldots, \mathbf{w}_M\}$ are linearly independent eigenvectors of $\text{cov}(\boldsymbol{\theta}_{[\mathcal{I}]})$
- $\{\mathbf{w}_i\}$ forms orthonormal basis for M-dimensional subspace of $\boldsymbol{\Theta}$
- $\mathbf{W} = M \times M$ matrix with columns $\mathbf{w}_i$
- Alternative parameterization $\phi = \mathbf{W}^T \boldsymbol{\theta}_{[\mathcal{I}]}$

Accelerating MCMC: Algorithm

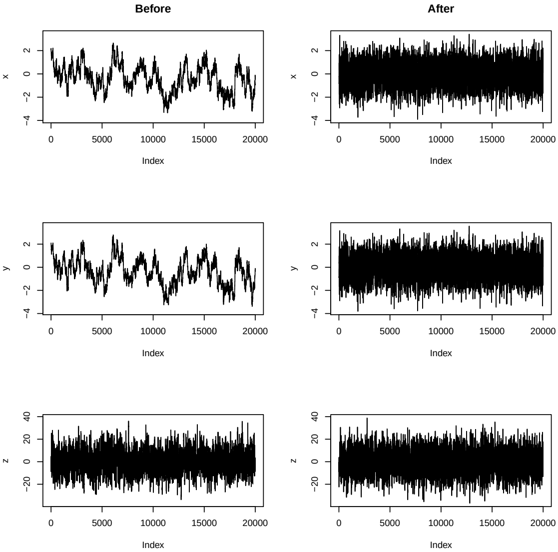
New MCMC scheme:

1. Update $\boldsymbol{\theta}^{(t+0.5)} = \text{MCMC}(\boldsymbol{\theta}^{(t)})$
2. Set $\phi^{(t+0.5)} = \mathbf{W}^T \boldsymbol{\theta}_{[\mathcal{I}]}^{(t+0.5)}$
3. Draw $\phi^{(t+1)} \sim \pi(\phi | \boldsymbol{\theta}_{[-\mathcal{I}]}^{(t+0.5)})$ (e.g., via Metropolis within Gibbs)
4. Set $\boldsymbol{\theta}_{[\mathcal{I}]}^{(t+1)} = \mathbf{W}\phi^{(t+1)}$ and $\boldsymbol{\theta}_{[-\mathcal{I}]}^{(t+1)} = \boldsymbol{\theta}_{[-\mathcal{I}]}^{(t+0.5)}$

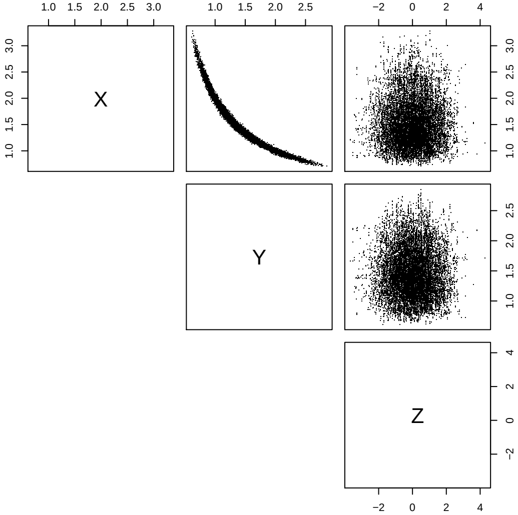
Accelerating MCMC: Illustration



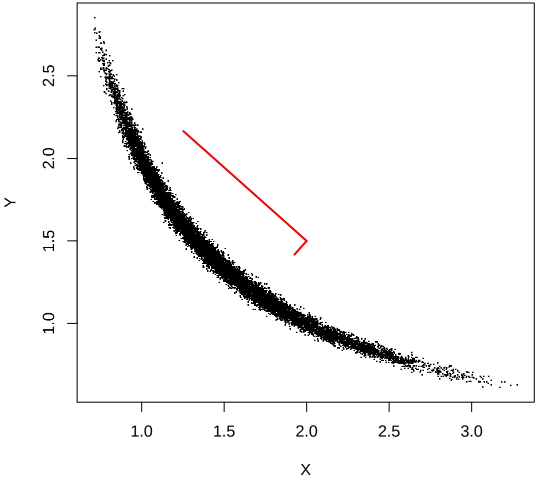
Accelerating MCMC: Illustration



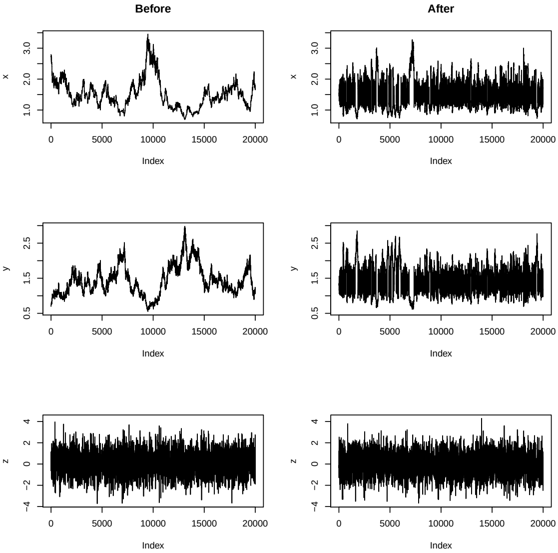
Accelerating MCMC: Illustration



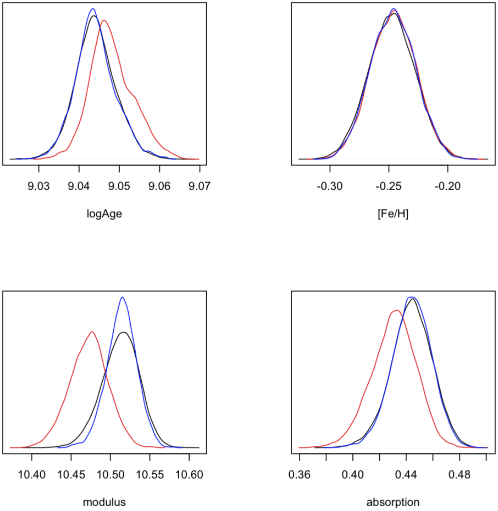
Accelerating MCMC: Illustration



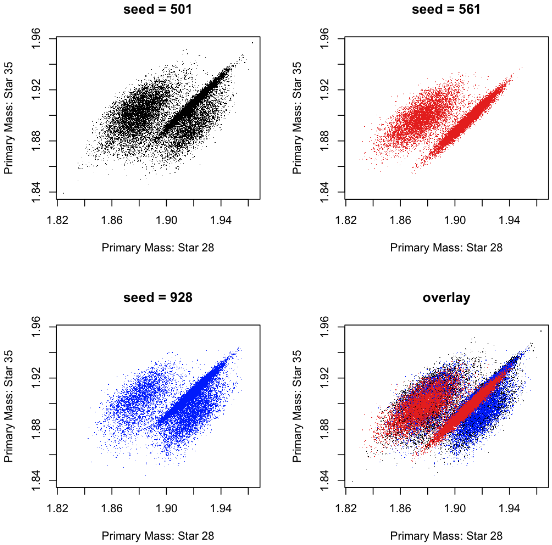
Accelerating MCMC: Illustration



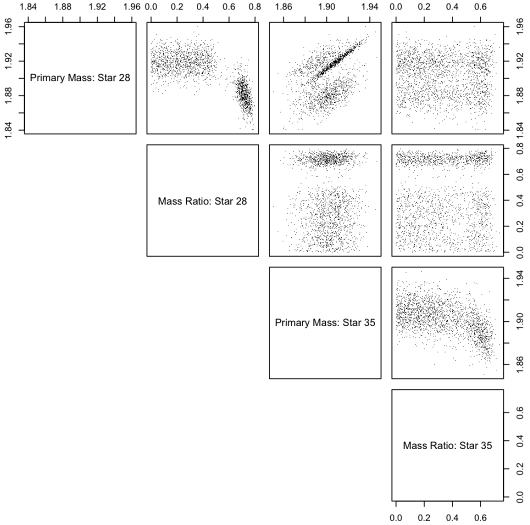
Different Random Seeds



Multiple Modes



Multiple Modes



Multiple Modes