Reconstructing stellar DEM and metallicity using high-resolution X-ray Spectra

Victoria Liublinska
Joint work with CHASC: X.-L. Meng, V. Kashyap, D. van Dyk

Harvard University

June 15, 2010
The model

Let $Y_{i}^{obs} \sim Pois(\xi_{i})$, where $\xi_{i}$ is photon intensity in wavelength channel $i$, $i = 1 \ldots I$ and

$$\xi = \xi^{source} + \xi^{bkg} = MD \left( \lambda^{C} + \sum_{k} \lambda_{k}^{L} \right) + \xi^{bkg}$$

$$= MD \left( G^{C} + \sum_{k} \gamma_{k} G_{k}^{L} \right) \mu + \xi^{bkg},$$

Parameters are $\gamma$ and $\mu$. 
Previously estimated Capella DEM
log10(DEM) used for simulations

log10(DEM) used for simulations (16 nodes)

log10(DEM) used for simulations (16 nodes)
Other properties of current simulations

• Includes all censoring, adds informative prior to abundances

• 16 nodes and some smoothing ($\alpha_\rho=20$): Depending on the starting point EM converges in 200-500 iterations

• 16 nodes and some smoothing ($\alpha_\rho=2$): EM converges in 800-1100 iterations
Estimated DEM

Stronger smoothing

No smoothing
Estimated $\log_{10}(\text{DEM})$

**Stronger smoothing**

**No smoothing**
Estimated abundance

Stronger smoothing

No smoothing
Estimated rho

**Stronger smoothing**

![Plot showing estimated rho for different multiscale levels with stronger smoothing.](image)

**No smoothing**

![Plot showing estimated rho for different multiscale levels with no smoothing.](image)
EM results from the simulation
EM results on a real data
Gibbs results from the simulation
Alternative sampling of the DEM

Estimated rho

Real data

Real data

Elements

Abundance

C  N  O  Ne  Mg  Al  Si  S  Ar  Ca  Fe  Ni

log10(T)

DEM

0e+00  2e+18  4e+18

0.0  0.4  0.8

0.0  0.4  0.8
Gibbs in a nutshell

- Initiating parameters $\gamma, \mu$
- 5 steps of sampling different missing data
- Updating abundance $\gamma^{(t+1)}|\mu^{(t)}, \gamma^{(t)}, Y_{mis}$
- 3 more steps of sampling missing data
- Updating (scaled) DEM $\tilde{\mu}^{(t+1)}|Y_{mis}$, where $\tilde{\mu} = f(\gamma)\mu$
- Rescale DEM $\mu^{(t+1)} = \tilde{\mu}^{(t+1)}/f(\gamma^{(t+1)})$
Gibbs in a nutshell

- Initiating parameters $\gamma, \mu$
- 5 steps of sampling different missing data
- Updating abundance $\gamma^{(t+1)}|\mu^{(t)}, \gamma^{(t)}, Y_{mis}$
- 3 more steps of sampling missing data
- Updating (scaled) DEM $\tilde{\mu}^{(t+1)}|Y_{mis}$, where $\tilde{\mu} = f(\gamma)\mu$
- Rescale DEM $\mu^{(t+1)} = \tilde{\mu}^{(t+1)}/f(\gamma^{(t+1)})$
- Properties of a current simulation: minimal smoothing ($\alpha_\rho = 2$), 16 nodes for $\log_{10}(T)$, 10000 simulations including 1000 of burn-in
Rho trace

Scale 1, R = 1

Scale 2, R = 1

Scale 2, R = 1.04
EM results from the simulation

EM results on a real data

Gibbs results from the simulation

Alternative sampling of the DEM

Rho trace

Scale 3, $R = 1.07$

Scale 3, $R = 1.02$

Scale 3, $R = 1$

Scale 3, $R = 1.08$
Rho trace

EM results from the simulation

EM results on a real data

Gibbs results from the simulation

Alternative sampling of the DEM
Mu total trace

Mu total, $R = 1$
Abundance trace

Element N, \( R = 1 \)

Element O, \( R = 1.01 \)

Element Ne, \( R = 1.01 \)

Element Mg, \( R = 1 \)

Element Al, \( R = 1 \)

Element S, \( R = 1 \)

Element Ar, \( R = 1 \)

Element Fe, \( R = 1 \)
Posterior distribution of $\mu | \gamma, Y^{obs}$

$$p(\mu | \gamma, Y^{obs}) \propto p(\mu) \prod_{i} p(Y_{i}^{obs} | \mu, \gamma),$$

where

$$Y_{i}^{obs} | \mu, \gamma \sim \text{Pois} \left( \sum_{j=1}^{J} M_{i,j} d_{j} \left( \sum_{t=1}^{T} G_{j,t} C_{0} + \sum_{t=1}^{T} \sum_{k} \gamma_{k} G_{k,j,t}^{L} \mu_{t} \right) + \xi_{i}^{bkg} \right)$$

$$= \text{Pois} \left( \sum_{t=1}^{T} h_{i,t} \mu_{t} + \xi_{i}^{bkg} \right),$$
Posterior distribution of $\mu|\gamma, Y^{obs}$

$$p(\mu|\gamma, Y^{obs}) \propto p(\mu) \prod_i p(Y_{i}^{obs}|\mu, \gamma),$$

where

$$Y_{i}^{obs}|\mu, \gamma \sim \text{Pois} \left( \sum_{j=1}^{J} M_{i,j}d_{j} \left( \sum_{t=1}^{T} G_{j,t}^{C}\mu_{t} + \sum_{t=1}^{T} \sum_{k} \gamma_{k}G_{k,j,t}^{L}\mu_{t} \right) + \xi_{i}^{bkg} \right) \right.$$  

$$= \text{Pois} \left( \sum_{t=1}^{T} h_{i,t}\mu_{t} + \xi_{i}^{bkg} \right),$$

with $\mu_{1} = \mu_{0} \prod_{r=1}^{R} \rho_{r,0}$, $\mu_{T} = \mu_{0} \prod_{r=1}^{R} (1 - \rho_{r,2^{r-1}})$ and each $\mu_{t}$ is a product of $\mu_{0}$ and $R$ splitting factors, either $\rho_{r,n}$ or $(1 - \rho_{r,n})$. 
Posterior distribution of $\mu|\gamma, Y^{obs}$

$$p(\mu|\gamma, Y^{obs}) \propto p(\mu) \prod_i p(Y_i^{obs}|\mu, \gamma),$$

where

$$Y_i^{obs}|\mu, \gamma \sim \text{Pois} \left( \sum_{j=1}^{J} M_{i,j} d_j \left( \sum_{t=1}^{T} G_{j,t} C \mu_t + \sum_{t=1}^{T} \sum_{k} \gamma_k G_{k,j,t}^L \mu_t \right) + \xi_i^{\text{bkg}} \right)$$

$$= \text{Pois} \left( \sum_{t=1}^{T} h_{i,t} \mu_t + \xi_i^{\text{bkg}} \right),$$

with $\mu_1 = \mu_0 \prod_{r=1}^{R} \rho_{r,0}$, $\mu_T = \mu_0 \prod_{r=1}^{R} (1 - \rho_{r,2r-1})$ and each $\mu_t$ is a product of $\mu_0$ and $R$ splitting factors, either $\rho_{r,n}$ or $(1 - \rho_{r,n})$. The following example clarifies the representation ($R=4$)

$$\mu_{12} = \mu_0 q_{12,R} = \mu_0 (1 - \rho_{0,0}) q_{12,R-1} = \mu_0 (1 - \rho_{0,0}) \rho_{1,1} q_{12,R-2} = \cdots = \mu_0 (1 - \rho_{0,0}) \rho_{1,1} (1 - \rho_{2,3}) (1 - \rho_{2,6})$$
Parametrization of the multiscale smoothing
Posterior distribution of $(\mu_0, \rho) | \gamma, Y^{obs}$

- $\mu_0 \sim Gamma(\alpha_\mu) / \beta_\mu$, where flat prior would correspond to $\alpha_\mu = 1$ and $\beta_\mu = 0$
- $\rho_{r,n} \sim Beta(\alpha_\rho, \alpha_\rho)$, $r = 0, \ldots, R - 1$ and $n = 0, \ldots, 2^r - 1$
- Instead of working with $Y^{obs}$ we can also use $Y$ (counts free from the background contamination)
Posterior distribution of $(\mu_0, \rho) | \gamma, Y^{obs}$

- $\mu_0 \sim Gamma(\alpha_\mu)/\beta_\mu$, where flat prior would correspond to $\alpha_\mu = 1$ and $\beta_\mu = 0$
- $\rho_{r,n} \sim Beta(\alpha_\rho, \alpha_\rho)$, $r = 0, \ldots, R - 1$ and $n = 0, \ldots, 2^r - 1$
- Instead of working with $Y^{obs}$ we can also use $Y$ (counts free from the background contamination)

$$p(\mu_0, \rho | \gamma, Y) \propto p(\mu_0) \prod_{r,n} p(\rho_{r,n}) \prod_{i} p(Y_i | \rho, \mu_0, \gamma).$$
Posterior distribution of \((\mu_0, \rho) | \gamma, Y^{obs}\)

- \(\mu_0 \sim Gamma(\alpha_\mu)/\beta_\mu\), where flat prior would correspond to \(\alpha_\mu = 1\) and \(\beta_\mu = 0\)
- \(\rho_{r,n} \sim Beta(\alpha_\rho, \alpha_\rho)\), \(r = 0, \ldots, R - 1\) and \(n = 0, \ldots, 2^r - 1\)
- Instead of working with \(Y^{obs}\) we can also use \(Y\) (counts free from the background contamination)

\[
p(\mu_0, \rho | \gamma, Y) \propto p(\mu_0) \prod_{r,n} p(\rho_{r,n}) \prod_i p(Y_i | \rho, \mu_0, \gamma).
\]

As it was noted previously

\[
Y_i | \mu_0, \rho, \gamma \sim Pois \left( \sum_{t=1}^{T} h_{i,t} \mu_t \right) = Pois \left( \mu_0 \sum_{t=1}^{T} h_{i,t} q_{t,R} \right) = Pois \left( s_i \mu_0 \right)
\]
Posterior distribution of \((\mu_0, \rho) | \gamma, Y^{obs}\)

- \(\mu_0 \sim Gamma(\alpha_\mu) / \beta_\mu\), where flat prior would correspond to \(\alpha_\mu = 1\) and \(\beta_\mu = 0\)
- \(\rho_{r,n} \sim Beta(\alpha_\rho, \alpha_\rho), r = 0, \ldots, R - 1\) and \(n = 0, \ldots, 2^r - 1\)
- Instead of working with \(Y^{obs}\) we can also use \(Y\) (counts free from the background contamination)

\[
p(\mu_0, \rho | \gamma, Y) \propto p(\mu_0) \prod_{r,n} p(\rho_{r,n}) \prod_i p(Y_i | \rho, \mu_0, \gamma).
\]

As it was noted previously

\[
Y_i | \mu_0, \rho, \gamma \sim Pois\left(\sum_{t=1}^T h_{i,t} \mu_t\right) = Pois \left(\mu_0 \sum_{t=1}^T h_{i,t} q_{t,R}\right) = Pois \left(s_i \mu_0\right)
\]

Therefore, \(\mu_0\) can be updated in closed form without any missing data imputation

\[
\mu_0 | \rho, Y, \gamma \sim Gamma \left(\alpha_\mu + \sum_i Y_i \right) / \left(\beta_\mu + \sum_i s_i\right)
\]
Updating splitting factors

Since $\rho_{r,n} \sim Beta(\alpha_{\rho}, \alpha_{\rho})$ and

$$p(\rho_{0,0}|\rho_-, \mu_0, \gamma, Y) \propto p(\rho_{0,0}) \prod_i p(Y_i|\rho, \mu_0, \gamma).$$

The distribution of background-free counts can be represented as

$$Y_i|\mu_0, \rho, \gamma = \text{Pois} \left( \sum_{t=1}^{T/2} h_{i,t} \mu_0 \rho_{0,0} q_{t,R-1} + \sum_{t=T/2+1}^{T} h_{i,t} \mu_0 (1 - \rho_{0,0}) q_{t,R-1} \right)$$

$$= \text{Pois} \left( s_{i,1} \rho_{0,0} + s_{i,2} (1 - \rho_{0,0}) \right),$$
Updating splitting factors

Since $\rho_{r,n} \sim \text{Beta}(\alpha_\rho, \alpha_\rho)$ and

$$p(\rho_{0,0}|\rho_-, \mu_0, \gamma, Y) \propto p(\rho_{0,0}) \prod_i p(Y_i|\rho, \mu_0, \gamma).$$

The distribution of background-free counts can be represented as

$$Y_i|\mu_0, \rho, \gamma = \text{Pois} \left( \sum_{t=1}^{T/2} h_{i,t} \mu_0 \rho_{0,0} q_{t,R-1} + \sum_{t=T/2+1}^{T} h_{i,t} \mu_0 (1 - \rho_{0,0}) q_{t,R-1} \right)$$

$$= \text{Pois} \left( s_{i,1} \rho_{0,0} + s_{i,2} (1 - \rho_{0,0}) \right),$$

then the conditional distribution of $\rho_{0,0}|\rho_-, \mu_0, Y, \gamma$ has the following form

$$\rho_{0,0}^{\alpha_\rho - 1} (1 - \rho_{0,0})^{\alpha_\rho - 1} \prod_i (s_{i,1} \rho_{0,0} + s_{i,2} (1 - \rho_{0,0})) Y_i e^{-(s_{i,1} \rho_{0,0} + s_{i,2} (1 - \rho_{0,0}))}$$
Updating splitting factors

Possible idea: update splitting factors by scale level. The number of simultaneously updated factors would be 1, 2, 4, 8, 16, etc.
Updating splitting factors

Possible idea: update splitting factors by scale level. The number of simultaneously updated factors would be 1, 2, 4, 8, 16, etc.

For example let

$$\eta_i = s_{i,2,1} \rho_{1,0} + s_{i,2,2}(1 - \rho_{1,0}) + s_{i,2,3} \rho_{1,1} + s_{i,2,4}(1 - \rho_{1,1}),$$

then

$$p(\rho_{1,0}, \rho_{1,1}|\rho_-, \mu_0, \gamma, Y) \propto \rho_{1,0}^{\alpha_{\rho} - 1}(1 - \rho_{1,0})^{\alpha_{\rho} - 1}\rho_{1,1}^{\alpha_{\rho} - 1}(1 - \rho_{1,1})^{\alpha_{\rho} - 1} \prod_i \eta_i^Y e^{-\eta_i}$$
Updating splitting factors

Possible idea: update splitting factors by scale level. The number of simultaneously updated factors would be 1, 2, 4, 8, 16, etc.

For example let
\[ \eta_i = s_{i,2,1} \rho_{1,0} + s_{i,2,2}(1 - \rho_{1,0}) + s_{i,2,3} \rho_{1,1} + s_{i,2,4}(1 - \rho_{1,1}), \]
then
\[ p(\rho_{1,0}, \rho_{1,1} | \rho, \mu_0, \gamma, Y) \propto \rho_{1,0}^{\alpha - 1} (1 - \rho_{1,0})^{\alpha - 1} \rho_{1,1}^{\alpha - 1} (1 - \rho_{1,1})^{\alpha - 1} \prod_i \eta_i Y_i e^{-\eta_i} \]

Implementation questions:

- Which proposal is the best? Truncated MVN (or MVt), others? OR simply evaluate the posterior and sample from the grid (since each \( \rho_{r,n} < 1 \)?)
- Other updating scheme? One-by-one, in pairs etc.
- Update \( \mu \) directly? (the posterior is not that simple since each prior on \( \rho_{r,n} \) contains all \( \mu_1, \ldots, \mu_T \))
- Alternate this scheme with full augmentation?