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

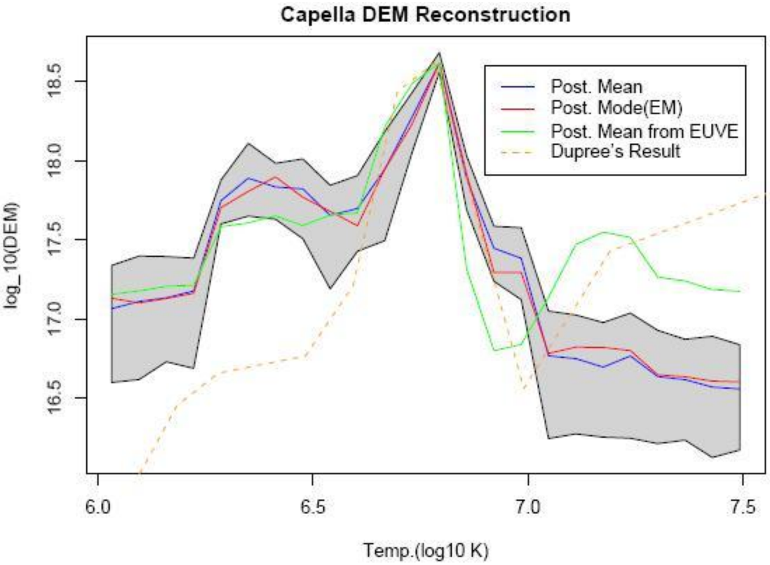
Let $Y_i^{obs} \sim Pois(\xi_i)$, where ξ_i is photon intensity in wavelength channel i, $i = 1 \dots I$ and

$$\boldsymbol{\xi} = \boldsymbol{\xi}^{source} + \boldsymbol{\xi}^{bkg} = MD\left(\boldsymbol{\lambda}^C + \sum_k \boldsymbol{\lambda}_k^L\right) + \boldsymbol{\xi}^{bkg}$$

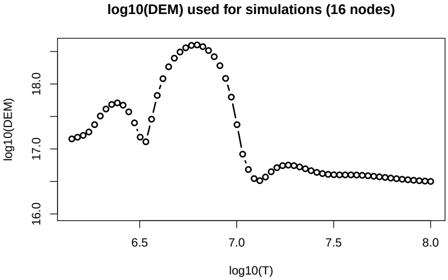
$$= MD\left(G^C + \sum_k \gamma_k G_k^L\right)\boldsymbol{\mu} + \boldsymbol{\xi}^{bkg},$$

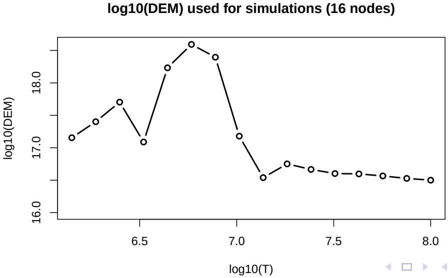
Parameters are $\boldsymbol{\gamma}$ and $\boldsymbol{\mu}$.

Previously estimated Capella DEM



log10(DEM) used for simulations

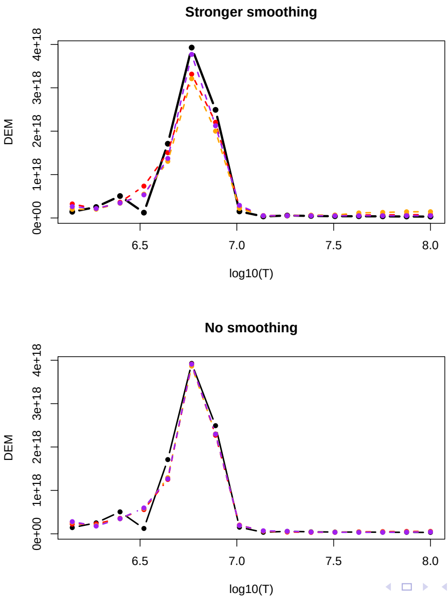




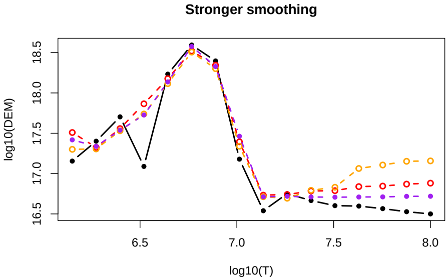
Other properties of current simulations

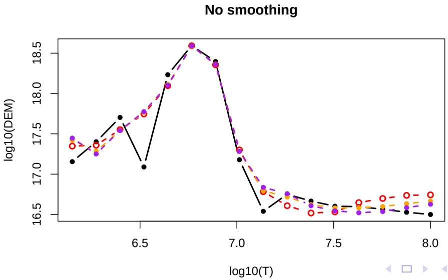
- Includes all censoring, adds informative prior to abundances
- 16 nodes and some smoothing (α_ρ=20): Depending on the starting point EM converges in 200-500 iterations
- 16 nodes and some smoothing (α_ρ=2): EM converges in 800-1100 iterations

Estimated DEM

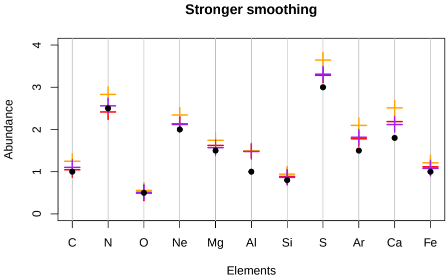


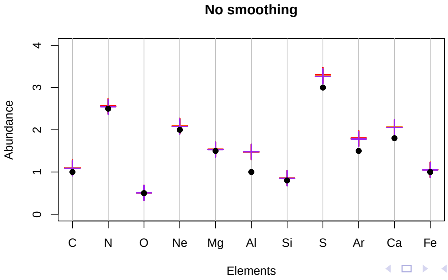
Estimated log10(DEM)



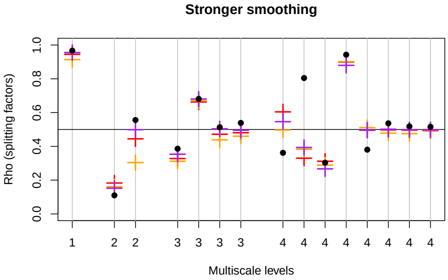


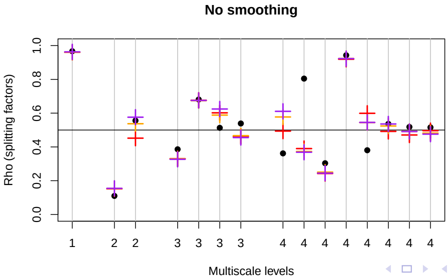
Estimated abundance



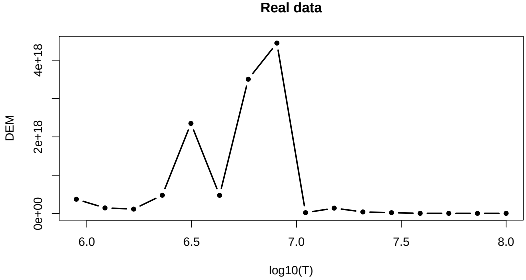


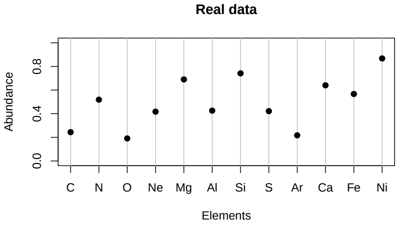
Estimated rho





Estimated rho





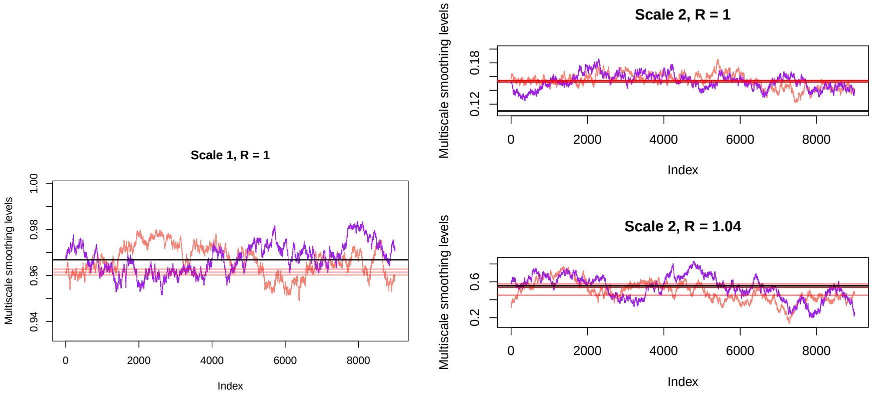
Gibbs in a nutshell

- Initiating parameters γ, μ
- 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 $\mu^{(t+1)} = \tilde{\mu}^{(t+1)}/f\left(\gamma^{(t+1)}\right)$

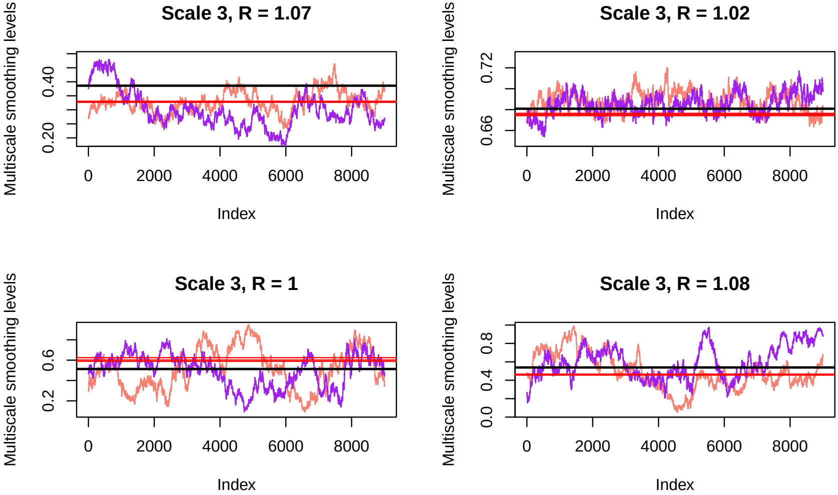
Gibbs in a nutshell

- Initiating parameters γ, μ
- 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 $\mu^{(t+1)} = \tilde{\mu}^{(t+1)}/f\left(\gamma^{(t+1)}\right)$
- Properties of a current simulation: minimal smoothing (α_ρ=2), 16 nodes for $log10(T)$, 10000 simulations including 1000 of burn-in

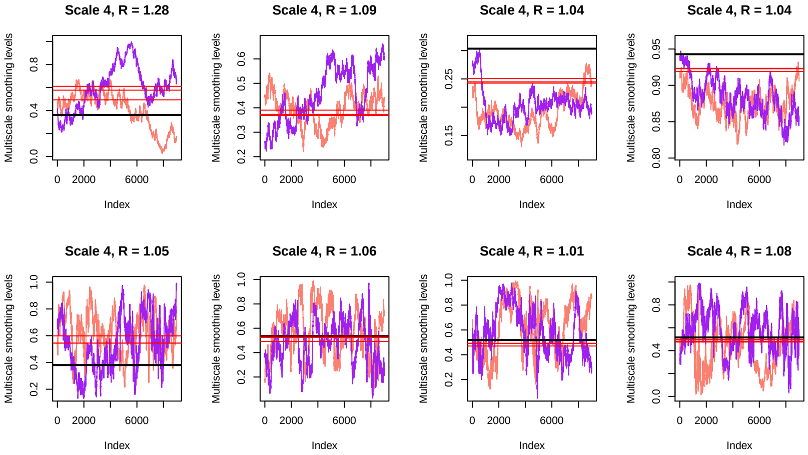
Rho trace


Scale 1, R = 1
Multiscale smoothing levels
0.94
0.96
0.98
1.00
0
2000
4000
6000
8000
Index
Scale 2, R = 1
Multiscale smoothing levels
0.12
0.18
0
2000
4000
6000
8000
Index
Scale 2, R = 1.04
Multiscale smoothing levels
0.2
0.6
0
2000
4000
6000
8000
Index

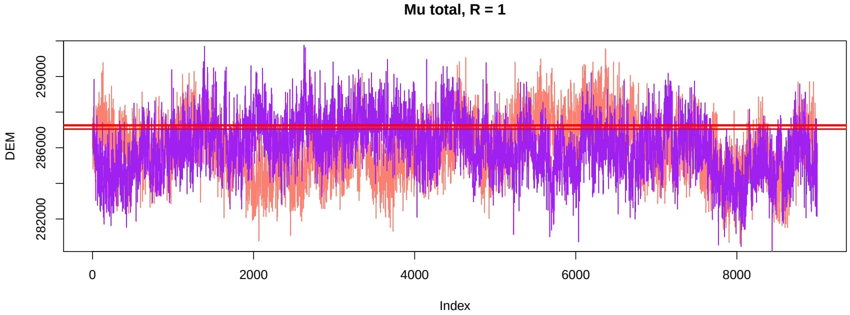
Rho trace


Scale 3, R = 1.07
Scale 3, R = 1.02
Scale 3, R = 1
Scale 3, R = 1.08
Multiscale smoothing levels
Index
0.20
0.40
0.66
0.72
0.2
0.6
0.0
0.4
0.8
0
2000
4000
6000
8000

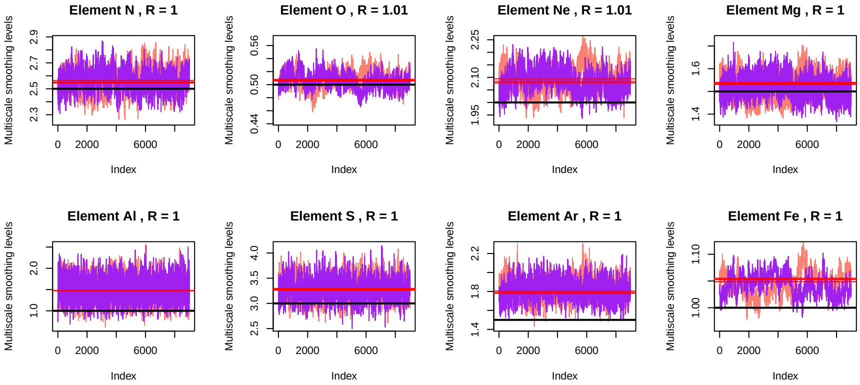
Rho trace


Scale 4, R = 1.28
Scale 4, R = 1.09
Scale 4, R = 1.04
Scale 4, R = 1.04
Scale 4, R = 1.05
Scale 4, R = 1.06
Scale 4, R = 1.01
Scale 4, R = 1.08
Multiscale smoothing levels
Index
0 2000 6000

Mu total trace



Abundance trace



Posterior distribution of $\mu|\gamma, Y^{obs}$

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

where

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

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

Posterior distribution of $\mu|\gamma, Y^{obs}$

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

where

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

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 μ_t is a product of μ_0 and R splitting factors, either $\rho_{r,n}$ or $(1 - \rho_{r,n})$.

Posterior distribution of $\boldsymbol{\mu}|\boldsymbol{\gamma}, Y^{obs}$

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

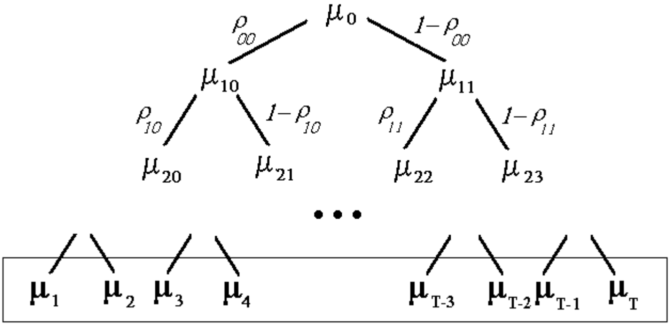
where

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

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 μ_t is a product of μ_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 $(\mu_0, \boldsymbol{\rho})|\boldsymbol{\gamma}, \boldsymbol{Y}^{obs}$

- $\mu_0 \sim Gamma(\alpha_\mu)/\beta_\mu$, 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 $\boldsymbol{Y}^{obs}$ we can also use $\boldsymbol{Y}$ (counts free from the background contamination)

Posterior distribution of $(\mu_0, \boldsymbol{\rho})|\boldsymbol{\gamma}, \boldsymbol{Y}^{obs}$

- $\mu_0 \sim Gamma(\alpha_\mu)/\beta_\mu$, 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 $\boldsymbol{Y}^{obs}$ we can also use $\boldsymbol{Y}$ (counts free from the background contamination)

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

Posterior distribution of $(\mu_0, \boldsymbol{\rho})|\boldsymbol{\gamma}, \boldsymbol{Y}^{obs}$

- $\mu_0 \sim Gamma(\alpha_\mu)/\beta_\mu$, 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 $\boldsymbol{Y}^{obs}$ we can also use $\boldsymbol{Y}$ (counts free from the background contamination)

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

As it was noted previously

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

Posterior distribution of $(\mu_0, \boldsymbol{\rho})|\boldsymbol{\gamma}, \boldsymbol{Y}^{obs}$

- $\mu_0 \sim Gamma(\alpha_\mu)/\beta_\mu$, 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 $\boldsymbol{Y}^{obs}$ we can also use $\boldsymbol{Y}$ (counts free from the background contamination)

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

As it was noted previously

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

Therefore, μ_0 can be updated in closed form without any missing data imputation

$$\mu_0|\boldsymbol{\rho}, \boldsymbol{Y}, \boldsymbol{\gamma} \sim Gamma\left(\alpha_\mu + \sum_i Y_i\right) / \left(\beta_\mu + \sum_i s_i\right)$$

Updating splitting factors

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

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

The distribution of background-free counts can be represented as

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

Updating splitting factors

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

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

The distribution of background-free counts can be represented as

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

then the conditional distribution of $\rho_{0,0}|\boldsymbol{\rho}_{-}, \mu_0, \boldsymbol{Y}, \boldsymbol{\gamma}$ 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 $p(\rho_{1,0}, \rho_{1,1} | \boldsymbol{\rho}_{-}, \mu_0, \boldsymbol{\gamma}, \boldsymbol{Y}) \propto$
$\rho_{1,0}^{\alpha_\rho - 1}(1-\rho_{1,0})^{\alpha_\rho - 1}\rho_{1,1}^{\alpha_\rho - 1}(1-\rho_{1,1})^{\alpha_\rho - 1}\prod_i \eta_i^{Y_i} e^{-\eta_i}$

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 $p(\rho_{1,0}, \rho_{1,1} | \boldsymbol{\rho}_-, \mu_0, \boldsymbol{\gamma}, \boldsymbol{Y}) \propto$
$\rho_{1,0}^{\alpha_\rho - 1}(1-\rho_{1,0})^{\alpha_\rho - 1}\rho_{1,1}^{\alpha_\rho - 1}(1-\rho_{1,1})^{\alpha_\rho - 1}\prod_i \eta_i^{Y_i} e^{-\eta_i}$

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 $\boldsymbol{\mu}$ 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?