Bayesian foreground cleaning

Dr. Frankenstein

January 19, 2016

1 Setup

Model:

$$\mathbf{d} = \hat{F} \cdot \mathbf{T} + \mathbf{n} \tag{1}$$

- $\mathbf{d}$, $\mathbf{n}$ $\leftarrow (N_p\, N_c\, N_\nu)$: data and noise.
- $\hat{N} \equiv \langle \mathbf{n} \cdot \mathbf{n}^T \rangle \leftarrow (N_p\, N_c\, N_\nu) \times (N_p\, N_c\, N_\nu)$: inverse noise variance.
- $\mathbf{T} \leftarrow (N_p\, N_c\, N_a)$: component amplitudes.
- $\hat{F} \leftarrow (N_p\, N_c\, N_\nu) \times (N_p\, N_c\, N_a)$: component frequency dependence.

Where:

- N_ν: number of frequency channels
- N_p: number of pixels
- $N_c = 3$: number of polarization channels
- $N_a = 3$: number of components (CMB, synchrotron and dust).
- The spectral dependence is:

$$F_{(p'\, c'\, \nu),(p\, c\, a)} \equiv f_a(\nu; \mathbf{b}^c_a(p))\, \delta_{p,p'} \delta_{c,c'} \tag{2}$$

  and $\mathbf{b}$ is a set of spectral parameters. Explicitly:

$$\text{CMB}: \; f_C(\nu) = B_\nu(\Theta_{\rm CMB}), \tag{3}$$

$$\text{synch.}: \; f_s(\nu; \beta_s) = \left(\frac{\nu}{\nu_0^s}\right)^{\beta_s} \tag{4}$$

$$\text{dust.}: \; f_d(\nu; \beta_d, \theta_d) = \left(\frac{\nu}{\nu_0^d}\right)^{\beta_d} \frac{B_\nu(\theta_d)}{B_{\nu_0^d}(\theta_d)} \tag{5}$$

- The noise variance is diagonal in frequency and polarization channels:

$$N_{(p\, c\, \nu),(p'\, c'\, \nu')} = N^{c\nu}_{p,p'} \delta_{cc'} \delta_{\nu\nu'} \tag{6}$$

This method consists on sampling the distribution of the parameters:

- $\mathbf{T} \rightarrow 3 \times 3 \times N_p$ amplitudes.
- $\mathbf{b}^c_a$: $\beta^c_s(p)$, $\beta^c_d(p)$, $\theta^c_d(p)$ spectral indices. If they were completely free they'd be $3 \times 3 \times N_p$ free parameters, and therefore the system could easily be overparametrized. We will use the following simplifying assumptions:
  1. Tier 1: The spectral parameters vary only over larger pixels $N'_p \ll N_p$.

2. Tier 1: The spectral parameters are the same for Q and U.
3. Tier 2: The spectral parameters are the same for T and P.
4. Tier 3: The dust temperature is constant across the sky.

Thus, in total we have $9\times N_p + 6\times N'_p$ parameters for Tier 1, $9\times N_p + 3\times N'_p$ for Tier 2 and $9\times N_p + 2\times N'_p + 1$ for Tier 3.

2 Sampling the likelihood

The joint posterior is:

$$p(\mathbf{T},\mathbf{b}|\mathbf{d}) \propto p_l(\mathbf{d}|\mathbf{T},\mathbf{b})\, p_p(\mathbf{T},\mathbf{b}), \tag{7}$$

where p_p is the prior and p_l is the Gaussian likelihood, given by

$$-2\log p_l(\mathbf{d}|\mathbf{T},\mathbf{b}) = c + \left[\mathbf{d} - \hat{F}(\mathbf{b})\cdot\mathbf{T}\right]^T \cdot \hat{N} \cdot \left[\mathbf{d} - \hat{F}(\mathbf{b})\cdot\mathbf{T}\right] \tag{8}$$

The Gibbs-sampling algorithm is:

1. $\mathbf{T}_{n+1} \leftarrow p(\mathbf{T}|\mathbf{d},\mathbf{b}_n) \propto p(\mathbf{d}|\mathbf{T},\mathbf{b}_n)\, p(\mathbf{T})$
2. $\mathbf{b}_{n+1} \leftarrow p(\mathbf{b}|\mathbf{d},\mathbf{T}_{n+1}) \propto p(\mathbf{d}|\mathbf{T}_{n+1},\mathbf{b})\, p(\mathbf{b})$

Obs: I'm not sure about the second equality in both cases.

2.1 Sampling the amplitudes T

In this case the probability distribution is

$$-2\log p(\mathbf{T}|\mathbf{d},\mathbf{b}) = c + (\mathbf{d} - \hat{F}\mathbf{T})^T \hat{N}^{-1}(\mathbf{d} - \hat{F}\mathbf{T}) + \mathbf{T}^T \hat{N}_P^{-1}\mathbf{T}, \tag{9}$$

where $\hat{N}_P$ is the prior matrix. The distribution is therefore Gaussian in $\mathbf{T}$, and hence the amplitudes can be sampled analytically from

$$\mathbf{T} \leftarrow N\left(\mathbf{m} = \hat{C}\hat{F}^T\hat{N}^{-1}\mathbf{d}\,;\,\hat{C}\right), \quad \hat{C} \equiv (\hat{F}^T\hat{N}^{-1}\hat{F} + \hat{N}_P^{-1})^{-1} \tag{10}$$