Technical Report

Liming Shi, Jesper Rindom Jensen, Jesper Kjær Nielsen and Mads Græsbøll Christensen
Audio Analysis Lab, AD:MT, Aalborg University
Emails: {ls, jrj, jkn, mgc}@create.aau.dk

This is the technical report of the paper "Multipitch Estimation Using Block Sparse Bayesian Learning and Intra-block Clustering", submitted to the ICASSP 2018.

I. The derivation of the variational Bayesian inference

A. Hierarchical form of the model

The hierarchical form of the model can be denoted as

$$
\begin{aligned}
\mathbf{y} &= \mathbf{Z}\mathbf{a} + \mathbf{m},\\
\mathbf{m} &\sim C\mathcal{N}(\mathbf{0}, \gamma^{-1}\mathbf{I}_N),\\
\gamma &\sim \Gamma(c, d),\\
\mathbf{a} &= \mathbf{u} \odot \boldsymbol{\theta},\\
\mathbf{u} &\sim C\mathcal{N}(\mathbf{0}, \boldsymbol{\Lambda}^{-1}),\\
\alpha_p &\sim \Gamma(g, h),\ 1 \le p \le P,\\
\theta_{p,l} &\sim \text{Bernoulli}(\pi_{p,l}),\ 1 \le p \le P,\ l \le l \le L_{\text{max}},\\
\pi_{p,l} &= \begin{cases} \pi^0, & \text{if } P0 \\ \pi^1, & \text{if } P1 \\ \pi^2, & \text{if } P2 \\ \pi^3, & \text{if } P3 \end{cases},\ 1 \le p \le P,\ l < l < L_{\text{max}},\\
\pi_{p,1} &= \begin{cases} \pi^1, & \text{if } \theta_{p,2} = 0 \\ \pi^3, & \text{if } \theta_{p,2} = 1 \end{cases},\ 1 \le p \le P,\\
\pi_{p,L_{\text{max}}} &= \begin{cases} \pi^0, & \text{if } \theta_{p,1} = 0 \\ \pi^1, & \text{if } \theta_{p,1} = 1 \text{ and } \theta_{p,L_{\text{max}}-1} = 0 \\ \pi^3, & \text{if } \theta_{p,1} = 1 \text{ and } \theta_{p,L_{\text{max}}-1} = 1 \end{cases},\ 1 \le p \le P,\\
\pi^j &\sim \text{Beta}(e^j, f^j),\ 0 \le j \le 3
\end{aligned}
\tag{1}
$$

where the design parameters are fixed for every experiments as follows:

$$
\begin{aligned}
&g = 1, c = d = h = 1e-6\\
&(e^0, f^0) = (1, 10^6)\\
&(e^1, f^1) = (1/L_{\text{max}}, 1 - 1/L_{\text{max}})\\
&(e^2, f^2) = (1/L_{\text{max}}, 1/L_{\text{max}})\\
&(e^3, f^3) = (1 - 1/L_{\text{max}}, 1/L_{\text{max}})
\end{aligned}
\tag{2}
$$

This work was funded by the Danish Council for Independent Research, grant ID: DFF 4184-00056

B. Variational Bayesian Inference

We first express the joint distribution as

$$\begin{aligned}
&p(\mathbf{y},\mathbf{u},\gamma,\boldsymbol{\theta},\boldsymbol{\alpha},\boldsymbol{\pi})\\
=&p(\mathbf{y}|\mathbf{u},\boldsymbol{\theta},\gamma)p(\mathbf{u}|\boldsymbol{\alpha})p(\gamma)\left[\prod_{p=1}^{P}\left[p(\alpha_p)\prod_{l=1}^{L_{\max}}p(\theta_{p,l}|\pi_{p,l})\right]\right]\left[\prod_{p=1}^{P}p(\pi^1)^{1(l=1,\theta_{p,2}=0)}p(\pi^3)^{1(l=1,\theta_{p,2}=1)}\right]\times\\
&\left[\prod_{p=1}^{P}\prod_{l=2}^{L_{\max}-1}p(\pi^0)^{1(\theta_{p,l}\in P0)}p(\pi^1)^{1(\theta_{p,l}\in P1)}p(\pi^2)^{1(\theta_{p,l}\in P2)}p(\pi^3)^{1(\theta_{p,l}\in P3)}\right]\times\\
&\left[\prod_{p=1}^{P}p(\pi^0)^{1(l=L_{\max},\theta_{p,1}=0)}p(\pi^1)^{1(l=L_{\max},\theta_{p,1}=1,\theta_{p,L_{\max}-1}=0)}p(\pi^1)^{1(l=L_{\max},\theta_{p,1}=1,\theta_{p,L_{\max}-1}=1)}\right]
\end{aligned}\tag{3}$$

The posterior computation is inferred from the variational Bayesian inference [18]. The update equations for the posterior distributions are given as follows:

(1) **the indicator variable** $\theta_{p,l},\ 1\le p\le P,\ 1\le l\le L_{\max}$:

$$\begin{aligned}
q(\theta_{p,l})&\propto\exp(\langle\log p(\mathbf{y}|\mathbf{u},\boldsymbol{\theta},\gamma)p(\theta_{p,l}|\pi_{p,l})\rangle)\\
&=\text{Bernoulli}(\tilde{\pi}_{p,l}),
\end{aligned}\tag{4}$$

where

$$\begin{aligned}
\tilde{\pi}_{p,l}=&[1+\exp\{\langle\log(1-\pi_{p,l})\rangle-\langle\log(\pi_{p,l})\rangle+\langle\gamma\rangle[\langle u_{p,l}^*u_{p,l}\rangle\mathbf{z}_{p,l}^H\mathbf{z}_{p,l}\\
&-2\text{Re}(\langle u_{p,l}\rangle^*\mathbf{z}_{p,l}^H(\mathbf{y}-\sum_{(i,j)\neq(p,l)}\langle\theta_{i,j}\rangle\langle u_{i,j}\rangle\mathbf{z}_{i,j})]\}]^{-1}
\end{aligned}\tag{5}$$

(2) **the complex amplitde** $\mathbf{u}$**:**

$$\begin{aligned}
q(\mathbf{u})&\propto\exp(\langle\log p(\mathbf{y}|\mathbf{u},\boldsymbol{\theta},\gamma)p(\mathbf{u}|\boldsymbol{\alpha})\rangle)\\
&=C\mathcal{N}(\tilde{\boldsymbol{\mu}},\tilde{\boldsymbol{\Sigma}}),
\end{aligned}\tag{6}$$

where

$$\begin{aligned}
\tilde{\boldsymbol{\Sigma}}&=(\langle\boldsymbol{\Lambda}\rangle+\langle\gamma\rangle\langle\text{diag}(\boldsymbol{\theta})\mathbf{Z}^H\mathbf{Z}\text{diag}(\boldsymbol{\theta})\rangle)^{-1},\\
\tilde{\boldsymbol{\mu}}&=\langle\gamma\rangle\tilde{\boldsymbol{\Sigma}}\langle\text{diag}(\boldsymbol{\theta})\rangle\mathbf{Z}^H\mathbf{y},
\end{aligned}\tag{7}$$

where $\langle\text{diag}(\boldsymbol{\theta})\mathbf{Z}^H\mathbf{Z}\text{diag}(\boldsymbol{\theta})\rangle=(\mathbf{Z}^H\mathbf{Z})\odot(\langle\boldsymbol{\theta}\rangle\langle\boldsymbol{\theta}\rangle^T+\text{diag}(\langle\boldsymbol{\theta}\rangle\odot(1-\langle\boldsymbol{\theta}\rangle)))$

(3) **the noise precision** γ**:**

$$\begin{aligned}
q(\gamma)&\propto\exp(\langle\log p(\mathbf{y}|\mathbf{u},\boldsymbol{\theta},\gamma)p(\gamma)\rangle)\\
&=\Gamma(\gamma;\tilde{c},\tilde{d}),
\end{aligned}\tag{8}$$

where

$$\begin{aligned}
\tilde{c}=&c+N,\\
\tilde{d}=&d+\mathbb{E}_{q(\mathbf{u},\boldsymbol{\theta})}[(\mathbf{y}-\mathbf{Z}\mathbf{a})^H(\mathbf{y}-\mathbf{Z}\mathbf{a})].\\
=&d+\text{Tr}\{(\mathbf{y}-\mathbf{Z}\langle\mathbf{a}\rangle)(\mathbf{y}-\mathbf{Z}\langle\mathbf{a}\rangle)^H.\\
&+\mathbb{E}_{q(\mathbf{u},\boldsymbol{\theta})}[(\mathbf{y}-\mathbf{Z}\mathbf{a}-\mathbf{y}+\mathbf{Z}\langle\mathbf{a}\rangle)(\mathbf{y}-\mathbf{Z}\mathbf{a}-\mathbf{y}+\mathbf{Z}\langle\mathbf{a}\rangle)^H]\}\\
=&d+\|\mathbf{y}-\mathbf{Z}\langle\mathbf{a}\rangle\|^2+\text{Tr}\{\mathbf{Z}^H\mathbf{Z}(\langle\mathbf{a}\mathbf{a}^H\rangle-\langle\mathbf{a}\rangle\langle\mathbf{a}\rangle^H)\}\\
=&d+\|\mathbf{y}-\mathbf{Z}(\langle\mathbf{u}\rangle\odot\langle\boldsymbol{\theta}\rangle)\|^2\\
&+\text{Tr}\{\mathbf{Z}^H\mathbf{Z}(\langle\mathbf{u}\mathbf{u}^H\rangle\odot(\langle\boldsymbol{\theta}\boldsymbol{\theta}^T\rangle)\\
&-(\langle\mathbf{u}\rangle\odot\langle\boldsymbol{\theta}\rangle)(\langle\mathbf{u}\rangle\odot\langle\boldsymbol{\theta}\rangle)^H)\},
\end{aligned}\tag{9}$$

where we used the equation

$$\begin{aligned}\mathbb{E}_{q(\mathbf{u},\boldsymbol{\theta})}[\mathbf{a}\mathbf{a}^H] &= \mathbb{E}_{q(\mathbf{u},\boldsymbol{\theta})}[(\mathbf{u}\odot\boldsymbol{\theta})(\mathbf{u}\odot\boldsymbol{\theta})^H] \\ &= \mathbb{E}_{q(\mathbf{u},\boldsymbol{\theta})}[\mathbf{H}], \quad H_{i,j} = u_i u_j^* \theta_i \theta_j \\ &= \mathbb{E}_{q(\mathbf{u},\boldsymbol{\theta})}[\mathbf{H}], \quad H_{i,j} = (\mathbf{u}\mathbf{u}^H)_{i,j}(\boldsymbol{\theta}\boldsymbol{\theta}^T)_{i,j} \\ &= \mathbb{E}_{q(\mathbf{u},\boldsymbol{\theta})}[(\mathbf{u}\mathbf{u}^H)\odot(\boldsymbol{\theta}\boldsymbol{\theta}^T)], \\ &= \langle\mathbf{u}\mathbf{u}^H\rangle \odot \langle\boldsymbol{\theta}\boldsymbol{\theta}^T\rangle, \end{aligned} \tag{10}$$

where $\langle\boldsymbol{\theta}\boldsymbol{\theta}^T\rangle = \langle\boldsymbol{\theta}\rangle\langle\boldsymbol{\theta}\rangle^T + \mathrm{diag}(\langle\boldsymbol{\theta}\rangle\odot(1-\langle\boldsymbol{\theta}\rangle))$.

(4) **the precision** $\alpha_p,\ 1\le p\le P$ **of the sparse complex amplitudes:**

$$\begin{aligned} q(\alpha_p) &\propto \exp(\langle\log p(\mathbf{u}|\boldsymbol{\alpha})p(\alpha_p)\rangle) \\ &= \Gamma(\alpha_p;\tilde{g}_p,\tilde{h}_p), \end{aligned} \tag{11}$$

where

$$\begin{aligned} \tilde{g}_p &= g + L_{\max}, \\ \tilde{h}_p &= h + \langle\mathbf{u}_p^H\mathbf{u}_p\rangle. \end{aligned} \tag{12}$$

(5) **the probability** $\pi_{p,l},\ 1\le p\le P,\ 1<l<L_{\max}$ **of success:**

$$\begin{aligned} q(\pi_{p,l}^j) &\propto \exp(\langle\log p(\theta_{p,l}|\pi_{p,l})p(\pi^0)^{1(\theta_{p,l}\in P0)}p(\pi^1)^{1(\theta_{p,l}\in P1)}p(\pi^2)^{1(\theta_{p,l}\in P2)}p(\pi^3)^{1(\theta_{p,l}\in P3)}\rangle) \\ &= \mathrm{Beta}(\pi_{p,l}^j;\tilde{e}_{p,l}^j,\tilde{f}_{p,l}^j), \end{aligned} \tag{13}$$

where for $j\in\{0,1,2,3\}$,

$$\begin{aligned} \tilde{e}_{p,l}^j &= e^j + p(Pj)\langle\theta_{p,l}\rangle, \\ \tilde{f}_{p,l}^j &= f^j + p(Pj)(1-\langle\theta_{p,l}\rangle), \end{aligned} \tag{14}$$

and

$$\begin{aligned} p(P0) &= 1-\langle\theta_{p,1}\rangle, \\ p(P1) &= \langle\theta_{p,1}\rangle(1-\langle\theta_{p,l-1}\rangle)(1-\langle\theta_{p,l+1}\rangle), \\ p(P2) &= \langle\theta_{p,1}\rangle(\langle\theta_{p,l-1}\rangle(1-\langle\theta_{p,l+1}\rangle)+\langle\theta_{p,l+1}\rangle(1-\langle\theta_{p,l-1}\rangle)), \\ p(P3) &= \langle\theta_{p,1}\rangle\langle\theta_{p,l-1}\rangle\langle\theta_{p,l+1}\rangle. \end{aligned}$$

The Expectation of logarithm function can be calculated as

$$\begin{aligned} \langle\log\pi_{p,l}\rangle &= \sum_{j=0}^{3} p(Pj)\langle\log\pi_{p,l}^j\rangle, \\ \langle\log(1-\pi_{p,l})\rangle &= \sum_{j=0}^{3} p(Pj)\langle\log(1-\pi_{p,l}^j)\rangle. \end{aligned}$$

Similarly, the **probability** $\pi_{p,1},\ 1\le p\le P$ **of success:**

$$\begin{aligned} q(\pi_{p,1}^j) &\propto \exp(\langle\log p(\theta_{p,1}|\pi_{p,1})p(\pi^1)^{1(\theta_{p,2}=0)}p(\pi^3)^{1(\theta_{p,2}=1)}\rangle) \\ &= \mathrm{Beta}(\pi_{p,1}^j;\tilde{e}_{p,1}^j,\tilde{f}_{p,1}^j), \end{aligned} \tag{15}$$

where for $j \in \{1,3\}$,

$$\begin{aligned}\tilde{e}^j_{p,1} =& e^j + p(P^j_1)\langle\theta_{p,1}\rangle,\\ \tilde{f}^j_{p,1} =& f^j + p(P^j_1)(1-\langle\theta_{p,1}\rangle),\end{aligned} \tag{16}$$

and

$$\begin{aligned}p(P^1_1) =& 1-\langle\theta_{p,2}\rangle,\\ p(P^3_1) =& \langle\theta_{p,2}\rangle.\end{aligned}$$

The Expectation of logarithm function can be calculated as

$$\begin{aligned}\langle\log\pi_{p,1}\rangle =& \sum_{j\in\{1,3\}} p(P^j_1)\langle\log\pi^j_{p,1}\rangle,\\ \langle\log(1-\pi_{p,1})\rangle =& \sum_{j\in\{1,3\}} p(P^j_1)\langle\log(1-\pi^j_{p,1})\rangle.\end{aligned}$$

The **probability** $\pi_{p,L_{\max}},\ 1\le p\le P$ **of success:**

$$\begin{aligned}q(\pi^j_{p,L_{\max}}) \propto& \exp(\langle\log p(\theta_{p,L_{\max}}|\pi_{p,L_{\max}})p(\pi^0)^{1(\theta_{p,1}=0)}p(\pi^1)^{1(\theta_{p,1}=1,\theta_{p,L_{\max}-1}=0)}p(\pi^1)^{1(\theta_{p,1}=1,\theta_{p,L_{\max}-1}=1)}\rangle)\\ =& \mathrm{Beta}(\pi^j_{p,L_{\max}};\tilde{e}^j_{p,L_{\max}},\tilde{f}^j_{p,L_{\max}}),\end{aligned} \tag{17}$$

where for $j \in \{0,1,3\}$,

$$\begin{aligned}\tilde{e}^j_{p,1} =& e^j + p(P^j_{L_{\max}})\langle\theta_{p,1}\rangle,\\ \tilde{f}^j_{p,1} =& f^j + p(P^j_{L_{\max}})(1-\langle\theta_{p,1}\rangle),\end{aligned} \tag{18}$$

and

$$\begin{aligned}p(P^0_{L_{\max}}) =& 1-\langle\theta_{p,1}\rangle,\\ p(P^1_{L_{\max}}) =& \langle\theta_{p,1}\rangle(1-\langle\theta_{p,L_{\max}}\rangle),\\ p(P^3_{L_{\max}}) =& \langle\theta_{p,1}\rangle\langle\theta_{p,L_{\max}}\rangle\end{aligned} \tag{19}$$

The Expectation of logarithm function can be calculated as

$$\begin{aligned}\langle\log\pi_{p,L_{\max}}\rangle =& \sum_{j\in\{0,1,3\}} p(P^j_{L_{\max}})\langle\log\pi^j_{p,1}\rangle,\\ \langle\log(1-\pi_{p,L_{\max}})\rangle =& \sum_{j\in\{0,1,3\}} p(P^j_{L_{\max}})\langle\log(1-\pi^j_{p,1})\rangle.\end{aligned}$$

Computational complexity consideration: The main computational load comes from the inversion calculation in (7) ($O(P^3L^3_{\max})$). We can use the Woodbury matrix identity to convert it to $O(N^3)$ if $N \ll PL_{\max}$, i.e.,

$$\begin{aligned}\tilde{\mathbf{\Sigma}} =& (\langle\mathbf{\Lambda}\rangle + \langle\gamma\rangle\langle\mathrm{diag}(\boldsymbol{\theta})\mathbf{Z}^H\mathbf{Z}\mathrm{diag}(\boldsymbol{\theta})\rangle)^{-1},\\ =& (\langle\mathbf{\Lambda}\rangle + \langle\gamma\rangle(\mathbf{Z}^H\mathbf{Z})\odot(\langle\boldsymbol{\theta}\rangle\langle\boldsymbol{\theta}\rangle^T + \mathrm{diag}(\langle\boldsymbol{\theta}\rangle\odot(1-\langle\boldsymbol{\theta}\rangle))))^{-1},\\ =& (\langle\mathbf{\Lambda}\rangle + \langle\gamma\rangle(\mathbf{Z}^H\mathbf{Z})\odot\langle\boldsymbol{\theta}\rangle\langle\boldsymbol{\theta}\rangle^T + \langle\gamma\rangle(\mathbf{Z}^H\mathbf{Z})\odot\mathrm{diag}(\langle\boldsymbol{\theta}\rangle\odot(1-\langle\boldsymbol{\theta}\rangle)))^{-1}\\ =& (\mathbf{\Lambda}' + \langle\gamma\rangle(\mathbf{Z}^H\mathbf{Z})\odot\langle\boldsymbol{\theta}\rangle\langle\boldsymbol{\theta}\rangle^T)^{-1}\\ =& (\mathbf{\Lambda}' + \langle\gamma\rangle\mathrm{diag}(\langle\boldsymbol{\theta}\rangle)\mathbf{Z}^H\mathbf{Z}\mathrm{diag}(\langle\boldsymbol{\theta}\rangle))^{-1}\\ =& \mathbf{\Lambda}'^{-1} - \mathbf{\Lambda}'^{-1}(\mathbf{Z}\mathrm{diag}(\langle\boldsymbol{\theta}\rangle))^H(\langle\gamma\rangle^{-1}\mathbf{I}_{N\times N} + (\mathbf{Z}\mathrm{diag}(\langle\boldsymbol{\theta}\rangle))\mathbf{\Lambda}'^{-1}(\mathbf{Z}\mathrm{diag}(\langle\boldsymbol{\theta}\rangle))^H)^{-1}(\mathbf{Z}\mathrm{diag}(\langle\boldsymbol{\theta}\rangle))\mathbf{\Lambda}'^{-1}\end{aligned} \tag{20}$$

where $\mathbf{\Lambda}' = \langle\mathbf{\Lambda}\rangle + \langle\gamma\rangle(\mathbf{Z}^H\mathbf{Z})\odot\mathrm{diag}(\langle\boldsymbol{\theta}\rangle\odot(1-\langle\boldsymbol{\theta}\rangle))$ which is a diagonal matrix.