VI. Markov Chain Parameter Estimation: Static PU with Markovian correlation across the channel indices with complete information

A. The Estimator

Before diving into the algorithm, let us first define the Forward and Backward probabilities that will be employed in our estimation algorithm. Let,

$$X_i \ = \ x_i \ be\ the\ PU\ Occupancy\ state\ of\ an\ arbitrary\ channel$$

$$X_{i+1} \ = \ X_j \ = \ x_j \ be\ the\ PU\ Occupancy\ state\ of\ the\ channel\ adjacent\ to\ channel\ b_i$$

1) Forward Probabilities: Let, $F(j,\ l)$ represent the probability of being in state $x_j = l$ after observing $y_1,\ y_2,\ y_3,\ \ldots\ldots,\ y_i,\ y_j$.

$$F(j,\ l) \ \triangleq \ \mathbb{P}(y_1,\ y_2,\ y_3,\ \ldots.,\ y_i,\ y_j,\ x_j = l) \tag{51}$$

Using the definition of Marginal Probability, equation (51) can be written as,

$$F(j,\ l) \ = \sum_{r\in\{0,1\}} \mathbb{P}(y_1,\ y_2,\ y_3,\ \ldots.,\ y_i,\ y_j,\ x_j = l,\ x_i = r) \tag{52}$$

Using the definition of conditional probability, equation (52) can be written as,

$$F(j,\ l) \ = \sum_{r\in\{0,1\}} \mathbb{P}(x_j = l,\ y_j \mid y_1,\ y_2,\ y_3,\ \ldots..,\ y_i,\ x_i = r)\mathbb{P}(y_1,\ y_2,\ y_3,\ \ldots..,\ y_i,\ x_i = r) \tag{53}$$

Using the Markov property and definition of Forward Probability outlined in equation (51), we can write equation (53) as follows,

$$F(j,\ l) \ = \sum_{r\in\{0,1\}} \mathbb{P}(x_j = l,\ y_j \mid x_i = r)F(i,\ r) \tag{54}$$

2) Backward Probabilities: Let $B(j,\ r)$ represent the probability of observing $y_j,\ y_{j+1},\ y_{j+2},\ \ldots\ldots,\ y_K$ given state $x_i \ = \ r$.

$$B(j,\ r) \ \triangleq \ \mathbb{P}(y_j,\ y_{j+1},\ y_{j+2},\ \ldots\ldots,\ y_K \mid x_i \ = \ r) \tag{55}$$

Using the definition of Marginal Probabilities,

$$B(j,\ r) \ = \sum_{l\in\{0,1\}} \mathbb{P}(y_j,\ y_{j+1},\ y_{j+2},\ \ldots\ldots,\ y_K,\ x_j = l \mid x_i \ = \ r) \tag{56}$$

Now, re-arranging the terms in equation (56), we get,

$$B(j,\ r) \ = \sum_{l\in\{0,1\}} \mathbb{P}(y_{j+1},\ y_{j+2},\ \ldots\ldots,\ y_K,\ y_j,\ x_j = l \mid x_i \ = \ r) \tag{57}$$

Now, we know that,

$$\mathbb{P}(A,\ B \mid C)\ =\ \mathbb{P}(A \mid B,\ C)\mathbb{P}(B \mid C)$$

Using this, we can write equation (57) as,

$$B(j,\ r)\ =\ \sum_{l \in \{0,1\}} \mathbb{P}(y_{j+1},\ y_{j+2},\ ......,\ y_K \mid y_j,\ x_j = l,\ x_i\ =\ r)\mathbb{P}(y_j,\ x_j = l \mid x_i = r) \tag{58}$$

Now, using the Markov property, equation (58) can be written as,

$$B(j,\ r)\ =\ \sum_{l \in \{0,1\}} \mathbb{P}(y_{j+1},\ y_{j+2},\ ......,\ y_K\ |x_j = l)\mathbb{P}(y_j,\ x_j = l \mid x_i = r) \tag{59}$$

Now, using the definition of Backward Probability outlined in equation (55),

$$B(j,\ r)\ =\ \sum_{l \in \{0,1\}} B(j+1,\ l)\mathbb{P}(y_j,\ x_j = l \mid x_i = r) \tag{60}$$

3) Deriving the analytical expressions for the parameter estimation algorithm: The HMM parameter estimation algorithm for our application can be derived using the Expectation-Maximization route.

The optimization objective is as follows,

$$A\ =\ argmax_A\ \mathbb{P}(\vec{y} \mid A) \tag{61}$$

where,

A is the state transition probability matrix

$$[A]_{lr}\ =\ a_{lr}\ =\ \mathbb{P}(x_j\ =\ r \mid x_i\ =\ l)$$

Converting this into an $argmax$ operation over log,

$$A\ =\ argmax_A\ log\ [\mathbb{P}(\vec{y} \mid A)] \tag{62}$$

Using the definition of Marginal Probabilities,

$$A\ =\ argmax_A\ log\ [\sum_{\vec{x}}\ \mathbb{P}(\vec{x},\ \vec{y} \mid A)] \tag{63}$$

This $argmax$ operation can be done over the joint probability distribution because the conditional is directly proportional to the joint as shown below.

$$A\ =\ argmax_A\ log\ [\sum_{\vec{x}}\ \mathbb{P}(\vec{x},\ \vec{y},\ A)] \tag{64}$$

Explicitly specifying the summation range over the set of all possible state sequences,

$$A \; = \; argmax_A \; log \; [\sum_{k \; = \; 1}^{|X|} \; \mathbb{P}(\vec{x}_k, \; \vec{y}, \; A)] \tag{65}$$

where,

$$|X| \; = \; 2^{|B|} \; = \; 2^K$$

Multiply and divide by α_k where $0 \leq \alpha_k \leq 1$ and $\sum_k \; \alpha_k \; = \; 1$ in order to convert equation (65) into a form of the Jensen's inequality.

$$A \; = \; argmax_A \; log \; [\sum_{k \; = \; 1}^{|X|} \; \alpha_k \frac{\mathbb{P}(\vec{x}_k, \; \vec{y}, \; A)}{\alpha_k}] \tag{66}$$

We know from Jensen's inequality that for any concave function $f(x)$, for any $x_i \; \in \; dom(f) \; (convex)$, and for any $0 \leq \; \theta_i \; \leq \; 1$ such that $\sum_i \; \theta_i \; = \; 1$,

$$f(\sum_i \; \theta_i x_i) \; \geq \; \sum_i \; \theta_i f(x_i)$$

Since the log function is concave, we can apply Jensen's inequality to equation (66) as follows,

$$argmax_A \; log \; [\sum_{k \; = \; 1}^{|X|} \; \alpha_k \frac{\mathbb{P}(\vec{x}_k, \; \vec{y}, \; A)}{\alpha_k}] \; \geq \; argmax_A \; \sum_{k=1}^{|X|} \; \alpha_k \; log \; [\frac{\mathbb{P}(\vec{x}_k, \; \vec{y}, \; A)}{\alpha_k}] \tag{67}$$

Now, α_k can be a Probability Mass Function because,

$$0 \; \leq \; \alpha_k \; \leq \; 1$$

$$\sum_k \; \alpha_k \; = \; 1$$

Here, equation (67) resembles,

$$f(\mathbb{E}[X]) \; \geq \; \mathbb{E}[f(X)] \; for \; a \; concave \; function \; f(x)$$

The inequality holds only when X is a degenerate random variable. So, it is evident from equation (67) that in order to ensure that equality holds,

$$\frac{\mathbb{P}(\vec{x}_k, \; \vec{y}, \; A)}{\alpha_k} \; = \; c \; with \; probability \; 1 \tag{68}$$

where, c is a constant.

Therefore,

$$\alpha_k \; = \; \frac{\mathbb{P}(\vec{x}_k, \; \vec{y}, \; A)}{c} \tag{69}$$

We know that,

$$\sum_{k} \alpha_k \;=\; 1$$

Using this,

$$c \;=\; \sum_{k} \mathbb{P}(\vec{x}_k,\; \vec{y},\; A) \tag{70}$$

Now, using these results,

$$\alpha_k \;=\; \frac{\mathbb{P}(\vec{x}_k,\; \vec{y},\; A)}{\sum_k \mathbb{P}(\vec{x}_k,\; \vec{y},\; A)} \tag{71}$$

Using the definition of Marginal Probabilities,

$$\alpha_k \;=\; \frac{\mathbb{P}(\vec{x}_k,\; \vec{y},\; A)}{\mathbb{P}(\vec{y},\; A)} \;=\; \mathbb{P}(\vec{x}_k \mid \vec{y},\; \hat{A}) \tag{72}$$

So, now the optimization problem becomes,

$$A \;=\; argmax_A \sum_{k\,=\,1}^{|X|} \mathbb{P}(\vec{x}_k \mid \vec{y},\; \hat{A})\; log\; [\frac{\mathbb{P}(\vec{x}_k,\; \vec{y},\; A)}{\mathbb{P}(\vec{x}_k \mid \vec{y},\; \hat{A})}] \tag{73}$$

Here, $\hat{A}$ is the previous estimate of the state transition probability matrix [This will turn out to be an iterative algorithm, i.e. the evaluation and re-estimation repeats until suitable convergence]. We don't care about the denominator in the above optimization problem,

$$A \;=\; argmax_A \sum_{k\,=\,1}^{|X|} \mathbb{P}(\vec{x}_k \mid \vec{y},\; \hat{A})\; log\; [\mathbb{P}(\vec{x}_k,\; \vec{y},\; A)] \tag{74}$$

Using the HMM System Model,

$$A \;=\; argmax_A \sum_{k\,=\,1}^{|X|} \mathbb{P}(\vec{x}_k \mid \vec{y},\; \hat{A})\; log\; [\prod_{i\,=\,1}^{K} \mathbb{P}(y_i \mid x_i)\; \mathbb{P}(x_i \mid x_{i-1},\; \hat{A})] \tag{75}$$

where, if $i \;=\; 1$,

$$\mathbb{P}(x_i \mid x_{i-1},\; \hat{A}) \;=\; \mathbb{P}(X_1 \;=\; x_1)$$

Using the properties of logarithms,

$$A \;=\; argmax_A \sum_{k\,=\,1}^{|X|} \mathbb{P}(\vec{x}_k \mid \vec{y},\; \hat{A}) \sum_{i\,=\,1}^{K} log\; [\mathbb{P}(y_i \mid x_i)] \;+\; log\; [\mathbb{P}(x_i \mid x_{i-1},\; \hat{A})] \tag{76}$$

Using indicator random variables to expand the state and observation associations,

$$A \;=\; argmax_A \sum_{k\,=\,1}^{|X|} \mathbb{P}(\vec{x}_k \mid \vec{y},\; \hat{A}) \sum_{l\,\in\,\{0,\;1\}} \sum_{r\,\in\,\{0,\;1\}} \sum_{i\,=\,1}^{K} \sum_{j\,=\,1}^{K} I\{x_i \;=\; r \;and\; y_j \;=\; y_i\}\; log\; [m_r(y_i)]$$
$$+\; I\{x_{i-1} \;=\; l \;and\; x_i \;=\; r\}\; log\; [a_{lr}] \tag{77}$$

Using Lagrange multipliers to find the solution, the Lagrangian is given as follows,

$$\mathcal{L} = \sum_{k=1}^{|X|} \mathbb{P}(\vec{x}_k \mid \vec{y},\ \hat{A}) \sum_{l \in \{0,\ 1\}} \sum_{r \in \{0,\ 1\}} \sum_{i=1}^{K} \sum_{j=1}^{K} I\{x_i = r \ and\ y_j = y_i\}\ log\ [m_r(y_i)] + I\{x_{i-1} = l \ and\ x_i = r\}\ log\ [a_{lr}] + \sum_{l \in \{0,\ 1\}} \lambda_l (1 - \sum_{r \in \{0,\ 1\}} a_{lr}) \tag{78}$$

Differentiating with respect to a_{lr} and equating it to 0,

$$\frac{\partial}{\partial a_{lr}} \mathcal{L} = \sum_{k=1}^{|X|} \mathbb{P}(\vec{x}_k \mid \vec{y},\ \hat{A}) \frac{1}{a_{lr}} \sum_{i=1}^{K} I\{x_{i-1} = l \ and\ x_i = r\} - \lambda_l = 0 \tag{79}$$

Simplifying this, we get,

$$a_{lr} = \frac{1}{\lambda_l} \sum_{k=1}^{|X|} \mathbb{P}(\vec{x}_k \mid \vec{y},\ \hat{A}) \sum_{i=1}^{K} I\{x_{i-1} = l \ and\ x_i = r\} \tag{80}$$

Differentiating with respect to our Lagrange multiplier λ_l and equating it to 0,

$$\frac{\partial}{\partial \lambda_l} \mathcal{L} = (1 - \sum_{r \in \{0,\ 1\}} a_{lr}) = 0 \tag{81}$$

Simplifying this using equation (80), we get,

$$\lambda_i = \sum_{r \in \{0,\ 1\}} \sum_{k=1}^{|X|} \mathbb{P}(\vec{x}_k \mid \vec{y},\ \hat{A}) \sum_{i=1}^{K} I\{x_{i-1} = l \ and\ x_i = r\} \tag{82}$$

Using the definition of Marginal Probabilities, we can write equation (82) as follows,

$$\lambda_i = \sum_{k=1}^{|X|} \mathbb{P}(\vec{x}_k \mid \vec{y},\ \hat{A}) \sum_{i=1}^{K} I\{x_{i-1} = l\} \tag{83}$$

Using equation (83) in equation (80), we get,

$$a_{lr} = \frac{\sum_{k=1}^{|X|} \mathbb{P}(\vec{x}_k \mid \vec{y},\ \hat{A}) \sum_{i=1}^{K} I\{x_{i-1} = l \ and\ x_i = r\}}{\sum_{k=1}^{|X|} \mathbb{P}(\vec{x}_k \mid \vec{y},\ \hat{A})\ \sum_{i=1}^{K} I\{x_{i-1} = l\}} \tag{84}$$

Let's simplify equation (84) further to remove the summation over all possible state sequences,

$$a_{lr} = \frac{\sum_{k=1}^{|X|} \sum_{i=1}^{K} \mathbb{P}(\vec{x}_k \mid \vec{y},\ \hat{A})\ I\{x_{i-1} = l \ and\ x_i = r\}}{\sum_{k=1}^{|X|} \sum_{i=1}^{K} \mathbb{P}(\vec{x}_k \mid \vec{y},\ \hat{A})\ I\{x_{i-1} = l\}} \tag{85}$$

Using the definition of Conditional Probability,

$$a_{lr} = \frac{\sum_{k=1}^{|X|} \sum_{i=1}^{K} \mathbb{P}(\vec{x}_k,\ \vec{y},\ \hat{A})\ I\{x_{i-1} = l \ and\ x_i = r\}}{\sum_{k=1}^{|X|} \sum_{i=1}^{K} \mathbb{P}(\vec{x}_k,\ \vec{y},\ \hat{A})\ I\{x_{i-1} = l\}} \tag{86}$$

Combining the Indicator with the Joint,

$$a_{lr} = \frac{\sum_{k=1}^{|X|} \sum_{i=1}^{K} \mathbb{P}(\vec{x}_k,\ \vec{y},\ \hat{A},\ x_{i-1} = l,\ x_i = r)}{\sum_{k=1}^{|X|} \sum_{i=1}^{K} \mathbb{P}(\vec{x},\ \vec{y},\ \hat{A},\ x_{i-1} = l)} \tag{87}$$

Using the definition of Marginal Probabilities,

$$a_{lr} \ = \ \frac{\sum_{i\ =\ 1}^{K} \mathbb{P}(\vec{y},\ \hat{A},\ x_{i-1} \ = \ l,\ x_i \ = \ r)}{\sum_{i\ =\ 1}^{K} \mathbb{P}(\vec{y},\ \hat{A},\ x_{i-1} \ = \ l)} \tag{88}$$

Expanding the observation vector and using the definition of Marginal Probabilities to modify the denominator,

$$a_{lr} \ = \ \frac{\sum_{i\ =\ 1}^{K} \mathbb{P}(y_1,\ y_2,\ ....,\ y_{i-1},\ x_{i-1} \ = \ l,\ y_i,\ x_i \ = \ r,\ y_{i+1},\ y_{i+2},\ ....,\ y_K,\ \hat{A})}{\sum_{r\ \in\ \{0,\ 1\}} \sum_{i\ =\ 1}^{K} \mathbb{P}(y_1,\ y_2,\ ....,\ y_{i-1},\ x_{i-1} \ = \ l, x_i \ = \ r,\ y_i,\ y_{i+1},\ y_{i+2},\ ....,\ y_K,\ \hat{A})} \tag{89}$$

Extracting the independent terms and creating conditionals,

$$a_{lr} \ = \ \frac{\sum_{i=1}^{K} \mathbb{P}(y_1, .., y_{i-1}, x_{i-1} = l)\mathbb{P}(y_i, x_i = r|y_1, .., y_{i-1}, x_{i-1} = l, \hat{A})\mathbb{P}(y_{i+1}, .., y_K|x_i = r)}{\sum_{r\in\{0,1\}} \sum_{i=1}^{K} \mathbb{P}(y_1, .., y_{i-1}, x_{i-1} = l)\mathbb{P}(x_i = r, y_i|y_1, .., y_{i-1}, x_{i-1} = l, \hat{A})\mathbb{P}(y_{i+1}, .., y_K|x_i = r)} \tag{90}$$

Using the properties of the assumed Markov Chain model,

$$a_{lr} \ = \ \frac{\sum_{i=1}^{K} \mathbb{P}(y_1, .., y_{i-1}, x_{i-1} = l)\mathbb{P}(y_i, x_i = r|x_{i-1} = l, \hat{A})\mathbb{P}(y_{i+1}, .., y_K|x_i = r)}{\sum_{r\in\{0,1\}} \sum_{i=1}^{K} \mathbb{P}(y_1, .., y_{i-1}, x_{i-1} = l)\mathbb{P}(x_i = r, y_i|x_{i-1} = l, \hat{A})\mathbb{P}(y_{i+1}, .., y_K|x_i = r)} \tag{91}$$

Now, we know that,

$$P(y_i,\ x_i \ = \ r \mid x_{i-1} \ = \ l,\ \hat{A}) \ = \ m_r(y_i)a_{lr}$$

Using this to re-write equation (92),

$$a_{lr} \ = \ \frac{\sum_{i=1}^{K} \mathbb{P}(y_1, .., y_{i-1}, x_{i-1} = l)m_r(y_i)a_{lr}\mathbb{P}(y_{i+1}, .., y_K|x_i = r)}{\sum_{r\in\{0,1\}} \sum_{i=1}^{K} \mathbb{P}(y_1, .., y_{i-1}, x_{i-1} = l)m_r(y_i)a_{lr}\mathbb{P}(y_{i+1}, .., y_K|x_i = r)} \tag{92}$$

We know from our definitions of Forward and Backward Probabilities defined in equations (51) and (55) respectively that,

$$\mathbb{P}(y_1, .., y_{i-1}, x_{i-1} = l) \ = \ F(i-1,\ l)$$

$$\mathbb{P}(y_{i+1}, .., y_K|x_i = r) \ = \ B(i+1,\ r)$$

Using these results in equation (92),

$$a_{lr} \ = \ \frac{\sum_{i=1}^{K} F(i-1, l)m_r(y_i)a_{lr}B(i+1, r)}{\sum_{r\in\{0,1\}} \sum_{i=1}^{K} F(i-1, l)m_r(y_i)a_{lr}B(i+1, r)} \tag{93}$$

B. The Algorithm

Known Parameters:

- Variance of the channel impulse response, i.e. σ_H^2
- Variance of the noise, i.e. σ_V^2
- Emission Probabilities = $m_l(y_i) \ \sim \mathcal{CN}(0,\ \sigma_H^2 l + \sigma_V^2)$

Initialization: Initialize the state transition probability matrix (A) to some random values.

$$\mathbb{P}(X_1 \ = \ 1) \ = \ \Pi$$

$$a_{lr} \ = \ \mathbb{P}(x_j \ = \ r \mid x_i \ = \ l)$$

Iteration: Evaluate the Forward and Backward probabilities using current estimates of A, the state transition probability matrix.

$$F(i-1,\ l)^t \ = \sum_{k \in \{0,1\}} m_l(y_{i-1}) a_{kl}^t F(i-2,\ k)$$

$$B(i+1,\ r)^t \ = \sum_{s \in \{0,1\}} m_s(y_{i+2}) a_{rs}^t B(i+2,\ s)$$

Re-estimate the elements of the state transition probability matrix.

$$a_{lr}^{t+1} \ = \ \frac{\sum_{i=1}^{K} F(i-1,l)^t m_r(y_i) a_{lr}^t B(i+1,r)^t}{\sum_{r \in \{0,1\}} \sum_{i=1}^{K} F(i-1,l)^t m_r(y_i) a_{lr}^t B(i+1,r)^t}$$

Termination:

$$|a_{lr}^{t+1} \ - \ a_{lr}^t| \leq \epsilon$$

where, t is the iteration counter, $\forall\ l,\ r \ \in \ \{0,\ 1\}$, and for any $\epsilon > 0$

C. Simulation Results

The algorithm outlined in the previous subsection has been implemented in Python and its results are detailed in this subsection.

- Number of frequency bands / channels = 18
- Number of observation vectors = 300
- SNR (signal ON) = 19.03 dB
- Convergence Threshold (ϵ) = 10^{-5}
- Initial Assignment of $\mathbb{P}(Occupied \mid Idle) \ = \ p$ = 10^{-5}
- True value of $\mathbb{P}(Occupied \mid Idle) \ = \ p$ = 0.30