APPENDIX A
PROOF OF PROPOSITION 1

A. Proposition 1

We show that the expected penalty of the environment, $C_{ji}(\mathbf{P})$, satisfies the required conditions of non-stationary environment Model B [10]. In the following we assume that the initial action probability profiles, $\mathbf{P}(0)$, which are selected by SUs satisfy the stability condition in (1). In the case that this initial probability does not satisfy (1), one or more channels will be highly overloaded and the probabilities of SUs collisions on those channels are increased dramatically. Therefore, the corresponding SUs receive high punishments according to the proposed scheme and gradually tune their channel access probabilities such that the stability condition of (1) is satisfied. After this incurred delay, the learning process will continue until each SU settles in its best strategy and the system becomes stable. That is, without lose of generality, we assume that the initial time is set to zero and the initial probability, $P(0)$, satisfy (1).

The required conditions for non-stationary environment Model B that must be satisfied are [10]:

- Equation (19) is continuous in all of its arguments.
- The value of $\frac{\partial C_{ji}(\mathbf{P})}{\partial p_{ji}}$ is positive as is shown in (27).

$$\frac{\partial C_{ji}(\mathbf{P})}{\partial p_{ji}} = \frac{\lambda_i^{(PU)}}{\lambda_i^{(PU)} + \mu_i} \; \frac{\lambda_j^{(SU)}}{\mu_i}$$

$$\left(\beta_{ji}^1 - \beta_{ji}^3 + (\beta_{ji}^2 - \beta_{ji}^1) \exp\Big(- \frac{p_{ji}\lambda_j^{(SU)} + R_{ji}}{(\mu_i - \lambda_i^{(PU)})(\mu_{ji} - p_{ji}\lambda_j^{(SU)})} \sum_{\substack{k=1\\k\neq j}}^{N} p_{ki}\lambda_k^{(SU)} \Big) \right)$$

$$+ \quad \frac{\lambda_i^{(PU)}}{\lambda_i^{(PU)} + \mu_i} \; \frac{p_{ji}\lambda_j^{(SU)} + \lambda_i^{(PU)} + \sum_{\substack{k=1\\k\neq j}}^{N} p_{ki}\lambda_k^{(SU)}}{\mu_i}$$

$$(\beta_{ji}^1 - \beta_{ji}^2) \frac{\lambda_j^{(SU)}\mu_{ji} + \lambda_j^{(SU)} R_{ji}}{(\mu_i - \lambda_i^{(PU)})(\mu_{ji} - p_{ji}\lambda_j^{(SU)})^2} \; \sum_{\substack{k=1\\k\neq j}}^{N} p_{ki}\lambda_k^{(SU)}$$

$$\exp\Big(- \frac{p_{ji}\lambda_j^{(SU)} + R_{ji}}{(\mu_i - \lambda_i^{(PU)})(\mu_{ji} - p_{ji}\lambda_j^{(SU)})} \sum_{\substack{k=1\\k\neq j}}^{N} p_{ki}\lambda_k^{(SU)} \Big)$$

$$+ \quad \frac{\mu_i}{\lambda_i^{(PU)} + \mu_i} \; \frac{\lambda_j^{(SU)}}{\mu_i}$$

$$\left(\beta_{ji}^4 + (\beta_{ji}^5 - \beta_{ji}^4) \exp\Big(- \frac{p_{ji}\lambda_j^{(SU)} + R_{ji}}{(\mu_i - \lambda_i^{(PU)})(\mu_{ji} - p_{ji}\lambda_j^{(SU)})} \sum_{\substack{k=1\\k\neq j}}^{N} p_{ki}\lambda_k^{(SU)} \Big) \right)$$

$$+ \quad \frac{\mu_i}{\lambda_i^{(PU)} + \mu_i} \; \frac{p_{ji}\lambda_j^{(SU)} + \lambda_i^{(PU)} + \sum_{\substack{k=1\\k\neq j}}^{N} p_{ki}\lambda_k^{(SU)}}{\mu_i}$$

$$(\beta_{ji}^4 - \beta_{ji}^5) \frac{\lambda_j^{(SU)}\mu_{ji} + \lambda_j^{(SU)} R_{ji}}{(\mu_i - \lambda_i^{(PU)})(\mu_{ji} - p_{ji}\lambda_j^{(SU)})^2} \; \sum_{\substack{k=1\\k\neq j}}^{N} p_{ki}\lambda_k^{(SU)}$$

$$\exp\Big(- \frac{p_{ji}\lambda_j^{(SU)} + R_{ji}}{(\mu_i - \lambda_i^{(PU)})(\mu_{ji} - p_{ji}\lambda_j^{(SU)})} \sum_{\substack{k=1\\k\neq j}}^{N} p_{ki}\lambda_k^{(SU)} \Big) \tag{27}$$

Note that using the stability condition (1) for channel F_i, the value of $\mu_{ji} - p_{ji}\lambda_j^{(SU)}$ is positive and hence $\frac{\partial C_{ji}(\mathbf{P})}{\partial p_{ji}} > 0$.

- We have $\frac{\partial C_{ji}(\mathbf{P})}{\partial p_{jm}} = 0, \;\; m = 1, 2, \ldots, M, \;\; m \neq i$ and hence the required condition $\frac{\partial C_{ji}(\mathbf{P})}{\partial p_{jm}} \ll \frac{\partial C_{ji}(\mathbf{P})}{\partial p_{ji}} for \;\; m \neq i$ is satisfied.
- $C_{ji}(\mathbf{P})$ must be continuously differentiable in all its arguments. Equation (27) shows that $C_{ji}(\mathbf{P})$ is continuously differentiable respect to p_{ji}. In the following we compute the

derivative of $C_{ji}(\mathbf{P})$ respect to p_{ki}, $k \neq j$, $k = 1, 2, \ldots, N$.

$$
\begin{aligned}
\frac{\partial C_{ji}(\mathbf{P})}{\partial p_{ki}} &= \frac{\lambda_i^{(PU)}}{\lambda_i^{(PU)} + \mu_i} \; \frac{\lambda_k^{(SU)}}{\mu_i} \\
&\left(\beta_{ji}^1 - \beta_{ji}^3 + (\beta_{ji}^2 - \beta_{ji}^1) \exp(-\frac{2p_{ki}\lambda_k^{(SU)} + R_{ki}}{(\mu_i - \lambda_i^{(PU)})(\mu_{ki} - p_{ki}\lambda_k^{(SU)})} \sum_{\substack{n=1 \\ n \neq k,j}}^{N} p_{ni}\lambda_n^{(SU)}) \right) \\
+ \quad &\frac{\lambda_i^{(PU)}}{\lambda_i^{(PU)} + \mu_i} \; \frac{p_{ki}\lambda_k^{(SU)} + \lambda_i^{(PU)} + \sum_{\substack{n=1 \\ n \neq k}}^{N} p_{ni}\lambda_n^{(SU)}}{\mu_i} \\
&(\beta_{ji}^1 - \beta_{ji}^2) \frac{2\lambda_k^{(SU)}\mu_{ki} + \lambda_j^{(SU)} R_{ji}}{(\mu_i - \lambda_i^{(PU)})(\mu_{ki} - p_{ki}\lambda_k^{(SU)})^2} \; \sum_{\substack{n=1 \\ n \neq k,j}}^{N} p_{ni}\lambda_n^{(SU)} \\
&\exp(-\frac{p_{ki}\lambda_k^{(SU)} + R_{ki}}{(\mu_i - \lambda_i^{(PU)})(\mu_{ki} - p_{ki}\lambda_j^{(SU)})} \sum_{\substack{n=1 \\ n \neq k,j}}^{N} p_{ni}\lambda_n^{(SU)}) \\
+ \quad &\frac{\mu_i}{\lambda_i^{(PU)} + \mu_i} \; \frac{\lambda_j^{(SU)}}{\mu_i} \\
&\left(\beta_{ji}^4 + (\beta_{ji}^5 - \beta_{ji}^4) \exp(-\frac{p_{ki}\lambda_k^{(SU)} + R_{ki}}{(\mu_i - \lambda_i^{(PU)})(\mu_{ki} - p_{ki}\lambda_k^{(SU)})} \sum_{\substack{n=1 \\ n \neq k,j}}^{N} p_{ni}\lambda_n^{(SU)}) \right) \\
+ \quad &\frac{\mu_i}{\lambda_i^{(PU)} + \mu_i} \; \frac{p_{ki}\lambda_k^{(SU)} + \lambda_i^{(PU)} + \sum_{\substack{n=1 \\ n \neq k}}^{N} p_{ni}\lambda_n^{(SU)}}{\mu_i} \\
&(\beta_{ji}^4 - \beta_{ji}^5) \frac{2\lambda_k^{(SU)}\mu_{ki} + \lambda_k^{(SU)} R_{ki}}{(\mu_i - \lambda_i^{(PU)})(\mu_{ki} - p_{ki}\lambda_k^{(SU)})^2} \; \sum_{\substack{n=1 \\ n \neq k,j}}^{N} p_{ni}\lambda_n^{(SU)} \\
&\exp(-\frac{p_{ki}\lambda_k^{(SU)} + R_{ki}}{(\mu_i - \lambda_i^{(PU)})(\mu_{ki} - p_{ki}\lambda_k^{(SU)})} \sum_{\substack{n=1 \\ n \neq k,j}}^{N} p_{ni}\lambda_n^{(SU)})
\end{aligned} \tag{28}
$$

Equation (28) shows that derivative of $C_{ji}(\mathbf{P})$ respect to p_{ki}, $k = 1, 2, \ldots, N$, $k \neq j$ exists and is positive.

- In this item we verify the Lipschitz continuity property of $C_{ji}(\mathbf{P})$ respect to all of its arguments by showing that the derivative of $C_{ji}(\mathbf{P})$ respect to p_{ji} and p_{ki}, $k = 1, 2, \ldots, N$, $k \neq j$ are bounded. As we can see in (27) and (28), $(\mu_{ji} - p_{ji}\lambda_j^{(SU)}) \to 0^+$ and $(\mu_{ki} - p_{ki}\lambda_k^{(SU)}) \to 0^+$ are the critical points, but these points are bounded by the exponential function in the following limits:

$$\lim_{(\mu_{ji}-p_{ji}\lambda_j^{(SU)})\to 0^+} \frac{\lambda_j^{(SU)}\mu_{ji}+\lambda_j^{(SU)}R_{ji}}{(\mu_{ji}-p_{ji}\lambda_j^{(SU)})^2} \exp\big(-\frac{p_{ji}\lambda_j^{(SU)}+R_{ji}}{(\mu_i-\lambda_i^{(PU)})(\mu_{ji}-p_{ji}\lambda_j^{(SU)})}\sum_{\substack{k=1\\k\neq j}}^{N} p_{ki}\lambda_k^{(SU)}\big) = 0 \tag{29}$$

$$\lim_{(\mu_{ki}-p_{ki}\lambda_k^{(SU)})\to 0^+} \frac{2\lambda_k^{(SU)}\mu_{ki}+\lambda_k^{(SU)}R_{ki}}{(\mu_{ki}-p_{ki}\lambda_k^{(SU)})^2} \exp\big(-\frac{p_{ki}\lambda_k^{(SU)}+R_{ki}}{(\mu_i-\lambda_i^{(PU)})(\mu_{ki}-p_{ki}\lambda_k^{(SU)})}\sum_{\substack{n=1\\n\neq k,j}}^{N} p_{ni}\lambda_n^{(SU)}\big) = 0 \tag{30}$$

Therefore, $C_{ji}(\mathbf{P})$ is a Lipschitz function respect to all of its arguments with Lipschitz constant $K = \sup \parallel \nabla C_{ji}(\mathbf{P}) \parallel$.

Therefore, the proposed expected penalty $C_{ji}(\mathbf{P})$ for each SU_j on each channel F_i follows the non-stationary environment Model B properties. As we computed in Proposition 1, derivative of $C_{ji}(\mathbf{P})$ respect to p_{ki}, $k = 1, 2, \ldots, N$, $k \neq j$ is positive. Therefore, it is monotonically increasing function respect to $\mathbf{p}_{-j}$.

Appendix B
Proof of proposition 2

We rewrite (20) for each components of $\triangle\mathbf{P}(t)$ by using (18) as follow:

$$
\begin{aligned}
&\triangle p_{ji}(t) = E\Big[p_{ji}(t+1)|p_{ji}(t)\Big] - p_{ji}(t) \\
= \;& \left(\sum_{l=1}^{6} c_{ji}^{l}\left(-\beta_{ji}^{l}(t)\alpha p_{ji}(t) + \left[1-\beta_{ji}^{l}(t)\right]\alpha(\sum_{m\neq i}^{M} p_{jm}(t))\right)\right)p_{ji} \\
+ \;& \sum_{m\neq i}^{M}\left(\sum_{l=1}^{6} c_{jm}^{l}\left(\beta_{jm}^{l}(t)\left[\frac{\alpha}{M-1} - \alpha p_{ji}(t)\right] - \left[1-\beta_{jm}^{l}(t)\right]\alpha p_{ji}(t)\right)\right)p_{jm} \\
= \;& -\alpha p_{ji}(t)^2 C_{ji}(\mathbf{P}(t)) + \alpha p_{ji}(t)\sum_{m\neq i}^{M} p_{jm}(t)\sum_{l=1}^{6} c_{ji}^{l} - \alpha p_{ji}(t)\sum_{m\neq i}^{M} p_{jm}(t)C_{ji}(\mathbf{P}(t)) \\
+ \;& \alpha\frac{1}{M-1}\sum_{m\neq i}^{M} p_{jm}(t)C_{jm}(\mathbf{P}(t)) - \alpha p_{ji}(t)\sum_{m\neq i}^{M} p_{jm}(t)C_{jm}(\mathbf{P}(t)) \\
- \;& \alpha p_{ji}(t)\sum_{m\neq i}^{M} p_{jm}(t)\sum_{l=1}^{6} c_{jm}^{l} + \alpha p_{ji}(t)\sum_{m\neq i}^{M} p_{jm}(t)C_{jm}(\mathbf{P}(t))
\end{aligned}
\tag{31}
$$

Since $\sum_{l=1}^{6} c_{ji}^{l} = 1, \; i = 1, \ldots, M$ we have:

$$
\begin{aligned}
\triangle p_{ji}(t) \;&=\; -\alpha p_{ji}(t)^2 C_{ji}(\mathbf{P}(t)) - \alpha p_{ji}(t)\sum_{m\neq i}^{M} p_{jm}(t)C_{ji}(\mathbf{P}(t)) \\
&+\; \alpha\frac{1}{M-1}\sum_{m\neq i}^{M} p_{jm}(t)C_{jm}(\mathbf{P}(t))
\end{aligned}
\tag{32}
$$

We can replace $\sum_{m\neq i}^{M} p_{jm}(t) = 1 - p_{ji}(t)$ and therefore:

$$
\begin{aligned}
\triangle p_{ji}(t) \;&=\; -\alpha p_{ji}(t)^2 C_{ji}(\mathbf{P}(t)) - \alpha p_{ji}(t)C_{ji}(\mathbf{P}(t)) + \alpha p_{ji}(t)^2 C_{ji}(\mathbf{P}(t)) \\
&+\; \alpha\frac{1}{M-1}\sum_{m\neq i}^{M} p_{jm}(t)C_{jm}(\mathbf{P}(t)) \\
&=\; -\alpha p_{ji}(t)C_{ji}(\mathbf{P}(t)) + \alpha\frac{1}{M-1}\sum_{m\neq i}^{M} p_{jm}(t)C_{jm}(\mathbf{P}(t))
\end{aligned}
\tag{33}
$$

therefore, equation (21) can be concluded from (33).

APPENDIX C
PROOF OF PROPOSITION 3

Let define $\mathcal{K}$ as the set of action selection probabilities for SUs:

$$\mathcal{K} = \left\{ \mathbf{P} = \left[\mathbf{p}_1^T, \mathbf{p}_2^T, \ldots, \mathbf{p}_N^T \right]^T, \ \forall j, i : \, 0 \le p_{ji} \le 1, \ \forall j : \, \sum_{i=1}^{M} p_{ji} = 1 \right\} \tag{34}$$

It is clear that $\mathcal{K}$ is a compact and convex set. It is also easy to verify that the value of $p_{ji}(t+1)$, $j = 1, \ldots, N$, $i = 1, \ldots, M$ which is computed by components of (23) is in $\mathcal{K}$. Therefore, equation (23) can be considered as a continuous mapping which is denoted by $T(\mathbf{P}) = \alpha \mathbf{f}(\mathbf{P}) + \mathbf{P}$ from $\mathcal{K}$ to $\mathcal{K}$.

By using Brouwer's fix point theorem, $T(\mathbf{P}(t))$ has at least one fix point which is $\mathbf{P}^* = T(\mathbf{P}^*)$ and the sequence (23) converges to this fix point. The fix point of $T(\mathbf{P})$ can be computed as $\mathbf{P}^* = \alpha \mathbf{f}(\mathbf{P}^*) + \mathbf{P}^*$. Since α is a non zero parameter therefore $\mathbf{f}(\mathbf{P}^*) = 0$.

Thus each component $f_{ji}(\mathbf{P}^*) = 0$, $j = 1, \ldots, N$, $i = 1, \ldots, M$. Therefore, for each SU_j there is a system of equations as follow:

$$\begin{cases} \frac{1}{M-1} \sum_{m \ne 1}^{M} p_{jm}^*(t) C_{jm}(\mathbf{P}^*) - p_{j1}^*(t) C_{j1}(\mathbf{P}^*) = 0 \\ \frac{1}{M-1} \sum_{m \ne 2}^{M} p_{jm}^*(t) C_{jm}(\mathbf{P}^*) - p_{j2}^*(t) C_{j2}(\mathbf{P}^*) = 0 \\ \vdots \\ \frac{1}{M-1} \sum_{m \ne M}^{M} p_{jm}^*(t) C_{jm}(\mathbf{P}^*) - p_{jM}^*(t) C_{jM}(\mathbf{P}^*) = 0 \\ \sum_{i=1}^{M} p_{ji} = 1 \end{cases} \tag{35}$$

We can conclude following system of equations from system of equations (35) for each $SU_j, \ j = 1, \ldots, N$:

$$\begin{cases} p_{ji}^* C_{ji}(\mathbf{P}^*) = p_{jm}^* C_{jm}(\mathbf{P}^*) \qquad i, m = 1, \ldots, M \\ \sum_{i=1}^{M} p_{ji} = 1 \end{cases} \tag{36}$$

and therefore, the system of equation (24) can be concluded for whole system. It is clear that the system of equation (24) has at least one solution because this solution is the fix point of the continuous mapping $T(\mathbf{P})$ over $\mathcal{K}$ which based on Brouwer's fix point theorem this point exists for $T(\mathbf{P})$.

Appendix D
Proof of proposition 4 and proposition 5

A. Proposition 4

In order to investigate the Lyapunov stability, first of all, the origin is transferred to equilibrium point $\mathbf{P}^*$, and then a candidate for Lyapunov function is introduced in order to investigate the stability of the discrete-time system (23) which is used in Algorithm 1. By using the following transformation we shift the origin to $\mathbf{P}^*$:

$$\hat{p}_{ji}(t) = p_{ji}(t) - p^*_{ji} \;\; j = 1, \ldots, N, \;\; i = 1, \ldots, M \tag{37}$$

First of all, we show that $\triangle \mathbf{P}\hat{(}t)$ has the components in form of $\triangle \hat{p}_{ji}(t) = -\hat{p}_{ji}(t)C_{ji}(\mathbf{P}(t))$. It is clear that $\triangle \hat{p}_{ji}(t) = \triangle p_{ji}(t)$. Therefore:

$$\begin{aligned}
\triangle \hat{p}_{ji}(t) &= \alpha \left(\frac{1}{M-1} \sum_{m \neq i}^{M} p_{jm}(t) C_{jm}(\mathbf{P}) - p_{ji}(t) C_{ji}(\mathbf{P}) \right) \\
&= \alpha \left(\frac{1}{M-1} \sum_{m \neq i}^{M} (\hat{p}_{jm}(t) + p^*_{jm}) C_{jm}(\mathbf{P}(t)) - (\hat{p}_{ji}(t) + p^*_{ji}) C_{ji}(\mathbf{P}(t)) \right) \\
&= \alpha \left(\frac{1}{M-1} \sum_{m \neq i}^{M} (\hat{p}_{jm}(t) + p^*_{jm}) C_{jm}(\mathbf{P}(t)) - p^*_{ji} C_{ji}(\mathbf{P}(t)) - \hat{p}_{ji}(t) C_{ji}(\mathbf{P}(t)) \right)
\end{aligned} \tag{38}$$

If we consider an action selection probability matrix $\mathbf{Q}$ for SUs in which $q_{ji} = p^*_{ji}$, the value of $\hat{q}_{ji}$ and $\triangle \hat{q}_{ji}$ will be zero and we have:

$$\triangle \hat{q}_{ji}(t) = \left(\frac{1}{M-1} \sum_{m \neq i}^{M} (\hat{q}_{jm}(t) + q^*_{jm}) C_{jm}(\mathbf{Q}(t)) - q^*_{ji} C_{ji}(\mathbf{Q}(t)) \right) = 0 \tag{39}$$

It is clear that equation (39) is not dependent on value $\hat{q}_{ji}$ and therefore this equation is valid for all transformation of action selection probability matrix, i.e. $\hat{\mathbf{P}}$. Using (38) and (39):

$$\triangle \hat{p}_{ji}(t) = -\alpha \hat{p}_{ji}(t) C_{ji}(\mathbf{P}(t)) \tag{40}$$

Now we consider the following Lyapunov function which is used in [37]:

$$V(\hat{\mathbf{P}}) = -\sum_{j} \sum_{i} \hat{p}_{ji} ln(1 - \hat{p}_{ji}) \tag{41}$$

and we investigate the properties of Lyapunov stability of the discrete-time system as follow:

- The Lyapunov function must be zero in equilibrium point.
  $V(\hat{\mathbf{P}})$ is zero only when all of its argument $\hat{p}_{ji} = 0$ which means $V(\hat{\mathbf{P}})$ is zero in $\mathbf{P}^*$.
- $V(\hat{\mathbf{P}}(t+1)) - V(\hat{\mathbf{P}}(t)) < 0 \ \forall \hat{\mathbf{P}}(t)$. We investigate this property in the following equation:

$$\begin{aligned} & -\sum_j \sum_i [\hat{p}_{ji}(t+1)ln(1-\hat{p}_{ji}(t+1)) - \hat{p}_{ji}(t)ln(1-\hat{p}_{ji}(t))] \\ = \ & -\sum_j \sum_i [(\hat{p}_{ji}(t) + \triangle\hat{p}_{ji}(t))ln(1-\hat{p}_{ji}(t) - \triangle\hat{p}_{ji}(t)) - \hat{p}_{ji}(t)ln(1-\hat{p}_{ji}(t))] \\ = \ & -\hat{p}_{ji}(t)\sum_j \sum_i [(1-C_{ji}(\mathbf{P}(t)))ln(1-(1-C_{ji}(\mathbf{P}(t)))\hat{p}_{ji}(t)) - ln(1-\hat{p}_{ji}(t))] \end{aligned} \tag{42}$$

  Since $0 < 1-C_{ji}(\mathbf{P}(t)) < 1$, therefore if $\hat{p}_{ji}(t) > 0$, we can conclude that $(1-C_{ji}(\mathbf{P}(t)))ln(1-(1-C_{ji}(\mathbf{P}(t)))\hat{p}_{ji}(t)) - ln(1-\hat{p}_{ji}(t)) > 0$ and if $\hat{p}_{ji}(t) < 0$, we can conclude that $(1-C_{ji}(\mathbf{P}(t)))ln(1-(1-C_{ji}(\mathbf{P}(t)))\hat{p}_{ji}(t)) - ln(1-\hat{p}_{ji}(t)) < 0$ and therefore $V(\hat{\mathbf{P}}(t+1)) - V(\hat{\mathbf{P}}(t)) < 0 \ \forall \hat{\mathbf{P}}(t)$.
  Therefore, using Lyapunov theorem, it is proved that $\mathbf{P}^*$ is an asymptotically Lyapunov stable equilibrium point of the proposed scheme.

B. Proposition 5

Based on proposition 4, proposed algorithm reaches to $\mathbf{P}^*$ in its convergence which is an asymptotically Lyapunov stable point of the system. Hence, all of the initial points within the domain $\mathcal{K}$ converge to $\mathbf{P}^*$. If another equilibrium point such as $\mathbf{Q}^*$ exists, all of the initial points within the domain $\mathcal{K}$ will also converge to $\mathbf{Q}^*$. This implies that $\mathbf{Q}^* = \mathbf{P}^*$, and therefore, the equilibrium point of the proposed scheme is unique.