Appendix A
proof of Lemma 1

According to the definition of $Q'_a(t) = \max\{Q_m(t), m \in \mathcal{M}_a\}$, it is can be derived that

$$Pr\left(Q_m(t) \le Q'_a(t)\right) = 1, \forall m \in \mathcal{M}_a, t, \tag{24}$$

which can also be denoted by $Q_m(t) \le Q'_a(t), \forall m \in \mathcal{M}_a, t$. We take expectations of both sides. According to the monotonicity of expectation, there is

$$\mathbb{E}\{Q_m(t)\} \le \mathbb{E}\{Q'_a(t)\}, \forall m \in \mathcal{M}_a, t. \tag{25}$$

Similarly, the inequality can also be extended to

$$\lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T} \mathbb{E}\{Q_m(t)\} \le \lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} \mathbb{E}\{Q'_a(t)\}, \forall m \in \mathcal{M}_a. \tag{26}$$

Thus, if the constraint $\lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} \mathbb{E}\{Q'_a(t)\} < \infty, \forall a \in \mathcal{A}$ is satisfied, there must be

$$\begin{aligned}&\lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} \mathbb{E}\{Q_m(t)\} \le \lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} \mathbb{E}\{Q'_a(t)\} \le \infty,\\ &\forall a \in \mathcal{A}, m \in \mathcal{M}_a.\end{aligned} \tag{27}$$

This proved that if the constraint $\overline{Q'_a(t)} \triangleq \lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} \mathbb{E}\{Q'_a(t)\} < \infty, \forall a \in \mathcal{A}$ is satisfied, the constraint $\overline{Q_m(t)} \triangleq \lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} \mathbb{E}\{Q_m(t)\} < \infty, \forall m \in \mathcal{M}$ satisfies.

This completes the proof.

Appendix B
proof of Lemma 2

Recalling the equation (10), we can denote the $\Delta\left(\boldsymbol{\Theta}(\mathbf{t})\right)$ by

$$\begin{aligned}\Delta\left(\boldsymbol{\Theta}(\mathbf{t})\right) &= \mathbb{E}\left\{\frac{1}{2}\sum_{a=1}^{A} Q'_a(t+1)^2 - \frac{1}{2}\sum_{a=1}^{A} Q'_a(t)^2\right\}\\ &= \mathbb{E}\left\{\frac{1}{2}\sum_{a=1}^{A} \left(y'_a(t) + Q'_a(t)\right)^2 - \frac{1}{2}\sum_{a=1}^{A} Q'_a(t)^2\right\}\\ &= \mathbb{E}\left\{\frac{1}{2}\sum_{a=1}^{A} \left(y'_a(t)\right)^2 + \sum_{a=1}^{A} Q'_a(t) y'_a(t)\right\}.\end{aligned} \tag{28}$$

According to the definition of $y'_a(t)$ in (10), we have

$$\begin{aligned}y'_a(t) &= Q'_a(t+1) - Q'_a(t)\\ &= \max_{m\in\mathcal{M}_a} Q_m(t+1) - \max_{m\in\mathcal{M}_a} Q_m(t)\end{aligned} \tag{29}$$

According to (6), considering that each $m \in m \in \mathcal{M}_a$, there is $y_m(t) \le \max_{m\in\mathcal{M}_a} y_m(t)$, and

$$Q_m(t+1) = Q_m(t) + y_m(t) \le Q_m(t) + \max_{m\in\mathcal{M}_a} y_m(t), \forall m \in \mathcal{M}_a. \tag{30}$$

Based on (30), we have

$$\begin{aligned}\max_{m\in\mathcal{M}_a} Q_m(t+1) &\le \max_{m\in\mathcal{M}_a}\left(Q_m(t) + \max_{m\in\mathcal{M}_a} y_m(t)\right)\\ &\le \max_{m\in\mathcal{M}_a} Q_m(t) + \max_{m\in\mathcal{M}_a} y_m(t).\end{aligned} \tag{31}$$

Recalling (29), there is

$$\begin{aligned}y'_a(t) &= \max_{m\in\mathcal{M}_a} Q_m(t+1) - \max_{m\in\mathcal{M}_a} Q_m(t)\\ &\le \max_{m\in\mathcal{M}_a} y_m(t).\end{aligned} \tag{32}$$

According to (5), considering $x_m(t) \le Q_m(t)$ in (7, there is $Q_m(t) + A_m(t) - x_m(t) \ge 0$ and $Q_m(t+1) = Q_m(t) + A_m(t) - x_m(t)$. Thus, there is

$$y_m(t) = Q_m(t+1) - Q_m(t) = A_m(t) - x_m(t). \tag{33}$$

Associated with (2) and (7), we have

$$\begin{aligned}y_m(t) &= Q_m(t+1) - Q_m(t) = A_m(t) - x_m(t)\\ &\le D_m, \forall m \in \mathcal{M}.\end{aligned} \tag{34}$$

Therefore, the (32) can be rewritten as

$$y'_a(t) \le \max_{m\in\mathcal{M}_a} y_m(t) \le D_m = D^a, \forall m \in \mathcal{M}_a. \tag{35}$$

and also

$$\frac{1}{2}\sum_{a=1}^{A}\left(y'_a(t)\right)^2 \le \frac{1}{2}\sum_{a=1}^{A}\left(D^a\right)^2 = B. \tag{36}$$

Back to (28), we have

$$\Delta\left(\boldsymbol{\Theta}(\mathbf{t})\right) \le \mathbb{E}\left\{B + \sum_{a=1}^{A} Q'_a(t) y'_a(t) | \boldsymbol{\Theta}(\mathbf{t})\right\}. \tag{37}$$

and

$$\Delta_V\left(\boldsymbol{\Theta}(\mathbf{t})\right) \le B + \mathbb{E}\left\{\sum_{a=1}^{A} Q'_a(t) y'_a(t) | \boldsymbol{\Theta}(\mathbf{t})\right\} + V\cdot\mathbb{E}\{P(t)|\boldsymbol{\Theta}(\mathbf{t})\}\}, \tag{38}$$

This completes the proof.

Appendix C
proof of Lemma 3

Due to we have obtained $\Theta(t)$,$\{Q_m(t)\}$,$\{A_m(t)\}$,$\mathbf{b}(\mathbf{t})$ in each slot before we solve the problem, the constraints (4), (7) are both linear constraints. And for the objective $\sum_{a=1}^{A} Q'_a(t) y'_a(t)$, we have:

$$\begin{aligned}&\mathbb{E}\left\{\sum_{a=1}^{A} Q'_a(t) y'_a(t)\right\} = \mathbb{E}\left\{\sum_{a=1}^{A} Q'_a(t)\left(Q'_a(t+1) - Q'_a(t)\right)\right\}\\ &= \mathbb{E}\left\{\sum_{a=1}^{A} Q'_a(t)\left(\max_{m\in\mathcal{M}_a} Q_m(t+1) - \max_{m\in\mathcal{M}_a} Q_m(t)\right)\right\}\\ &= \sum_{a=1}^{A} Q'_a(t)\left(\max_{m\in\mathcal{M}_a} \mathbb{E}\left\{Q_m(t+1)\right\} - \max_{m\in\mathcal{M}_a} Q_m(t)\right)\\ &= \sum_{a=1}^{A} Q'_a(t)\left(\max_{m\in\mathcal{M}_a}\left(Q_m(t) + A_m(t) - x_m(t)\right) - \max_{m\in\mathcal{M}_a} Q_m(t)\right)\\ &= \sum_{a=1}^{A} Q'_a(t) \max_{m\in\mathcal{M}_a}\left(-x_m(t)\right) + C.\end{aligned} \tag{39}$$

Considering the $\max_{m\in\mathcal{M}_a}(-x_m(t))$ is convex, the weighted sum of these convex functions is also convex. Accompanied with the linear constraints, we can prove this problem is convex. This completes the proof.

Appendix D
proof of Theorem 1