Comparisons with related works (Weakness 1 of Reviewer KAmk and Question of Reviewer rZoR)

Table 1: Comparison of the settings in related works.

Algorithm	Noise	Approximation	Timescale	Whittle index
Q-Whittle [3]	*i.i.d.*	✗	***two-timescale***	✓
Q-Whittle-LFA [57]	*i.i.d.*	linear	***two-timescale***	✓
Q-learning-LFA [6, 36, 63]	***Markovian***	linear	*single-timescale*	✗
Q-learning-NFA [13, 15, 22, 58]	***Markovian***	***neural network***	*single-timescale*	✗
TD-learning-LFA [47]	***Markovian***	linear	*single-timescale*	✗
2TSA-IID [19, 21]	*i.i.d.*	✗	***two-timescale***	✗
2TSA-Markovian [20]	***Markovian***	✗	***two-timescale***	✗
Neural-Q-Whittle **(this work)**	***Markovian***	***neural network***	***two-timescale***	✓

1

Experiments with large state space (Weakness 3 of Reviewer rZoR)

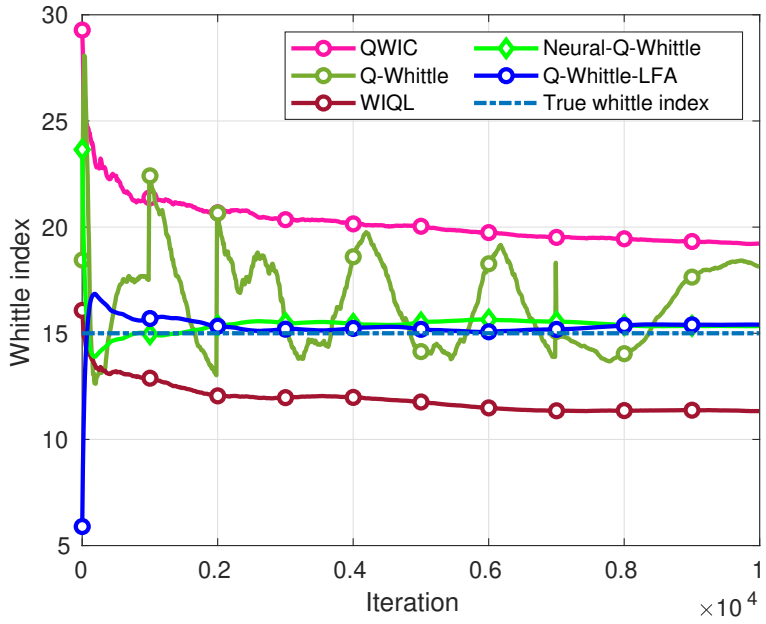

Figure 1: We consider a queuing problem with a state space ranging from 0 to 50. The arrival rate of the packet $\lambda = 30$ and the departure rate is state S and action A dependent as $\mu = vSA$ with $v = 20$. We randomly sample the Whittle index for state $s = 40$ as shown in this figure. Similar observations hold for Whittle indices for other states. This experiment shows similar results as the smaller-state case presented in supplementary materials of our paper.

2