EUGE Mechanism: Late Convergence Stage (Episodes 800-1200)
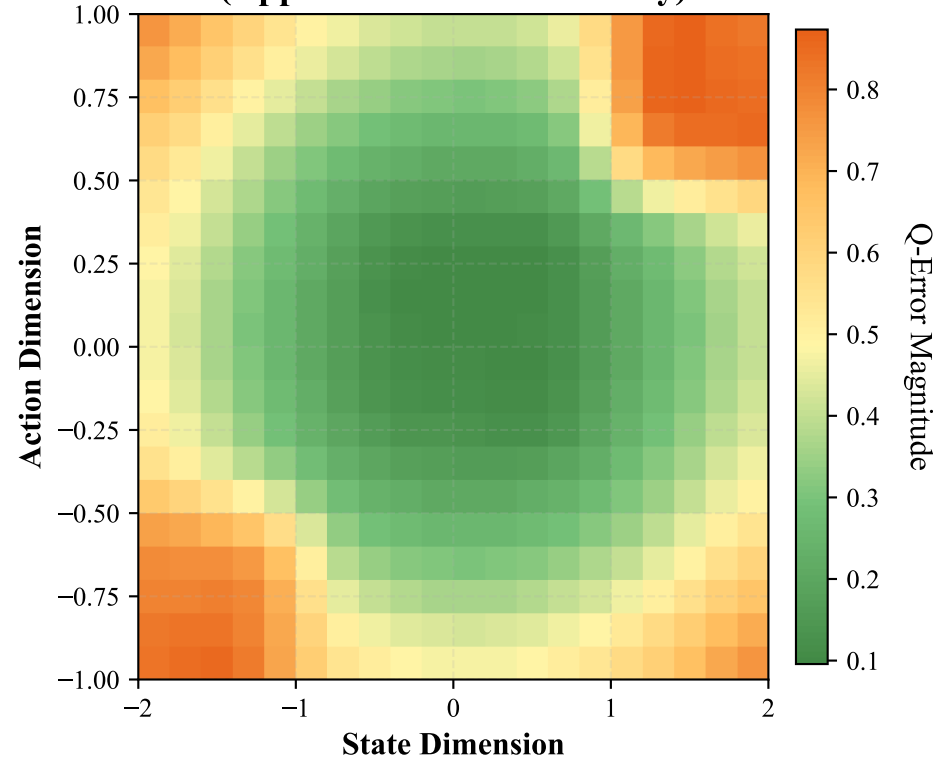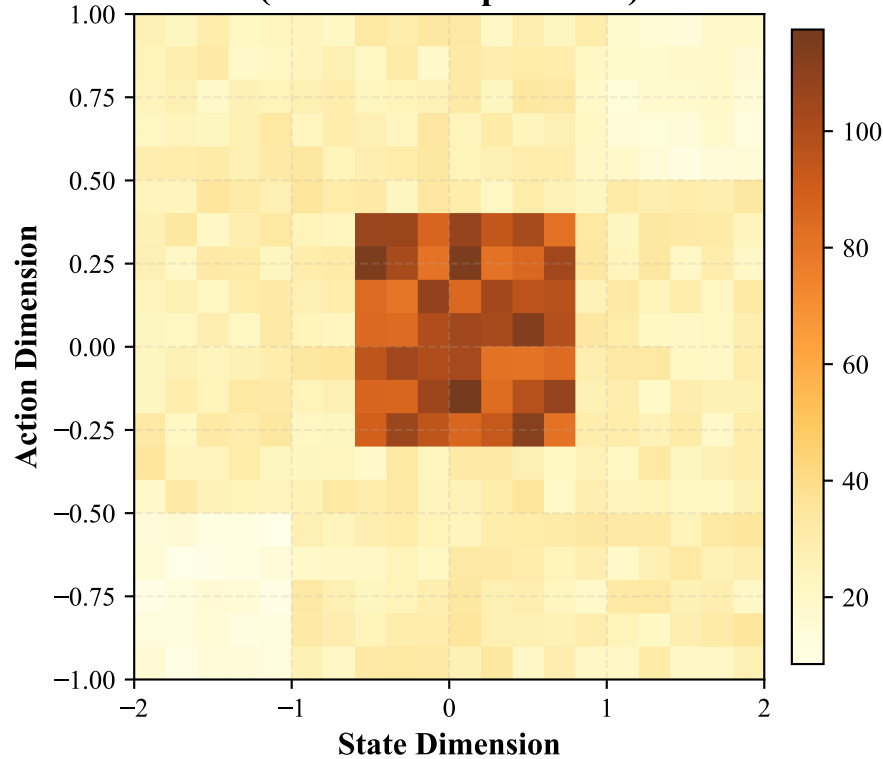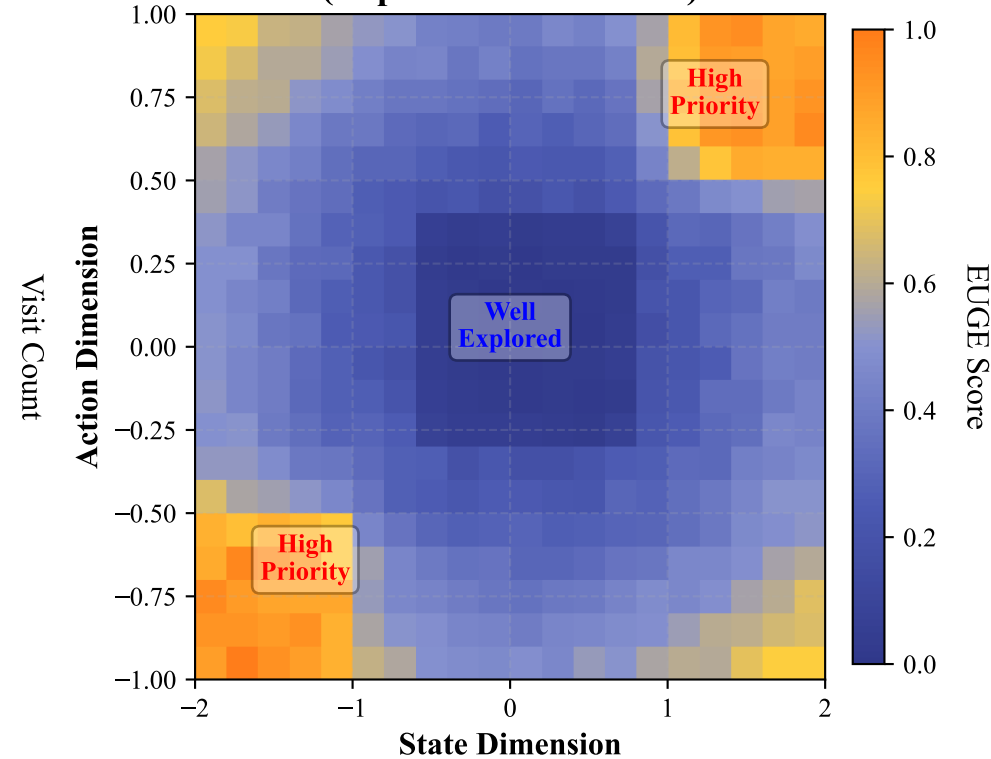

(a) Q-Error Distribution (Approximation Uncertainty) — Q-Error Magnitude

(b) Visit Frequency (Historical Exploration) — Visit Count

(c) EUGE Priority Score (Exploration Guidance) — EUGE Score; High Priority; Well Explored; High Priority

$$EU_t(s, a) = \hat{Q}_t(s, a) + c \cdot \left(Q_{error, t}(s, a) + \sqrt{\frac{\log N_{total, t}}{N_t(s, a) + 1}} \right)$$