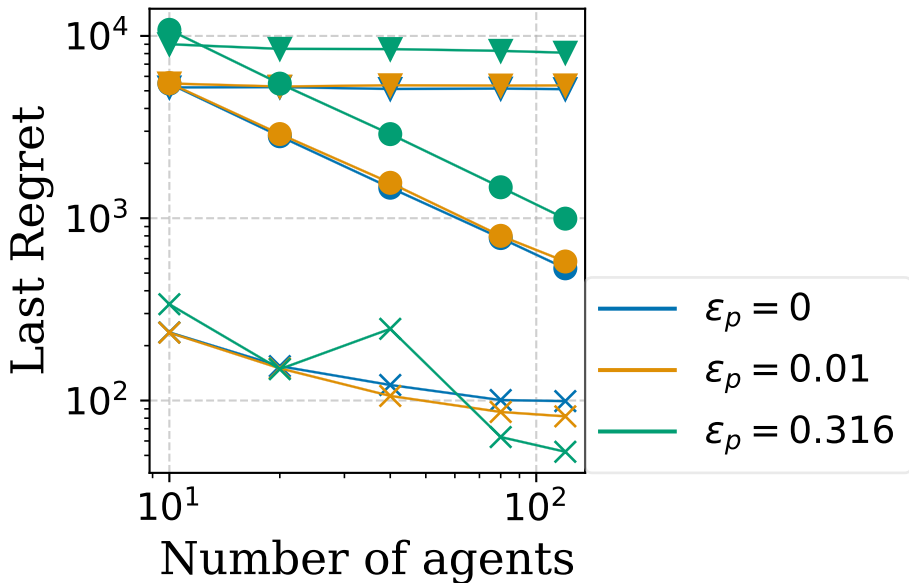


GridWorld environment



Common regret (lower is better), at $T = 10^4$ as a function of M for different heterogeneity levels in a log-log scale:
crosses represent Fed-UCBVI, circles represent FedQ-Bernstein and triangles represent FedQ-Advantage.