Multi-players M=9: Cumulated centralized regret, averaged 200 times

9 arms: $[B(0.1)^*, B(0.2)^*, B(0.3)^*, B(0.4)^*, B(0.5)^*, B(0.6)^*, B(0.7)^*, B(0.8)^*, B(0.9)^*]$ → 9× RandTopM-KLUCB 9× MCTopM-KLUCB $9 \times$ Selfish-KLUCB $9 \times RhoRand-KLUCB$ $\sum_{k=1}^{9} \mu_k \mathbb{E}_{200}[T_k(t)]$ Our lower-bound = $0 \log(t)$ Anandkumar et al.'s lower-bound = $0 \log(t)$ Centralized lower-bound = $0 \log(t)$ Cumulative centralized regret $\sum_{k=1}^9 \mu_k^* t - \sum_{k=1}^9 k_k^* t - \sum_{k=1}^9 k_k^$ 0 2000 4000 6000 10000 8000 Time steps t = 1...T, horizon T = 10000,