

Fig. 1: **Synthetic data:** We set the parameters as $d = 100, r = 2, T = 100$, and noise variance $= 10^{-6}$. We considered $M = 4$ epochs each with 50 data samples each. We varied the number of tasks as $T = 50, 75, 100$, the dimension of the feature vector as $d = 100, 125, 150$, and the rank of the feature matrix as $r = 2, 5, 10$.

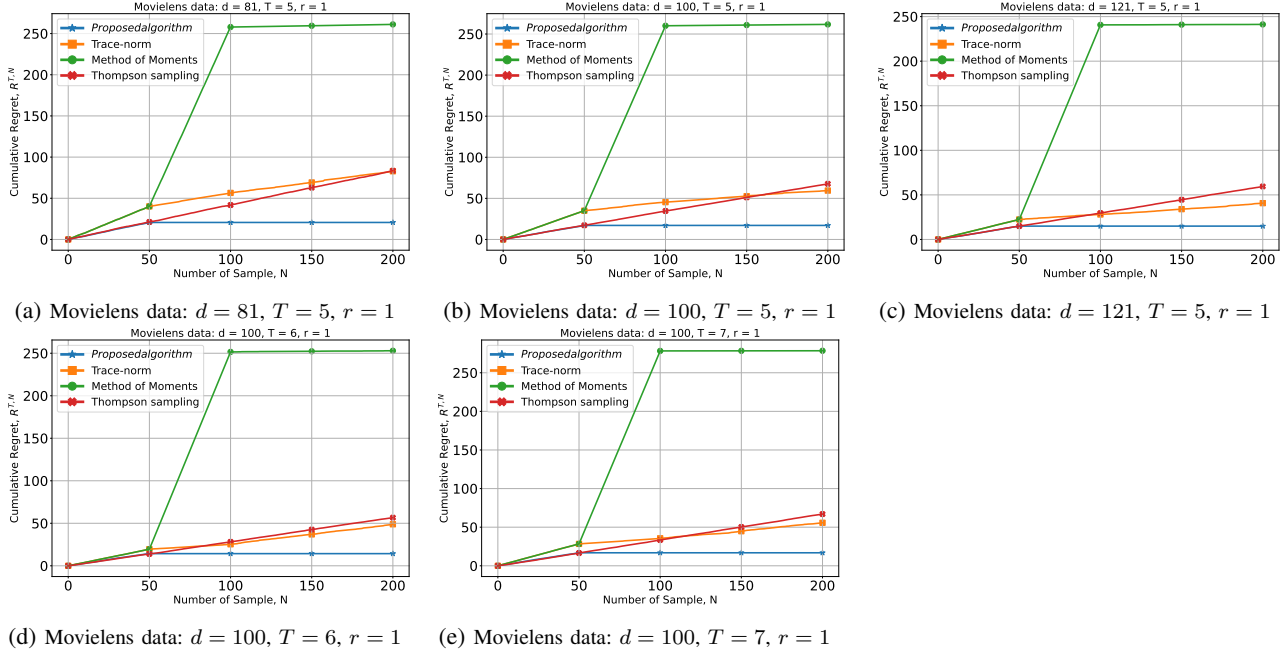


Fig. 2: MovieLens data: We set the parameters as $d = 100, r = 1, T = 5$, and noise variance $= 10^{-6}$. We considered $M = 4$ epochs each with 50 data samples each. We varied the number of tasks as $T = 5, 6, 7$, and the dimension of the feature vector as $d = 81, 100, 121$.