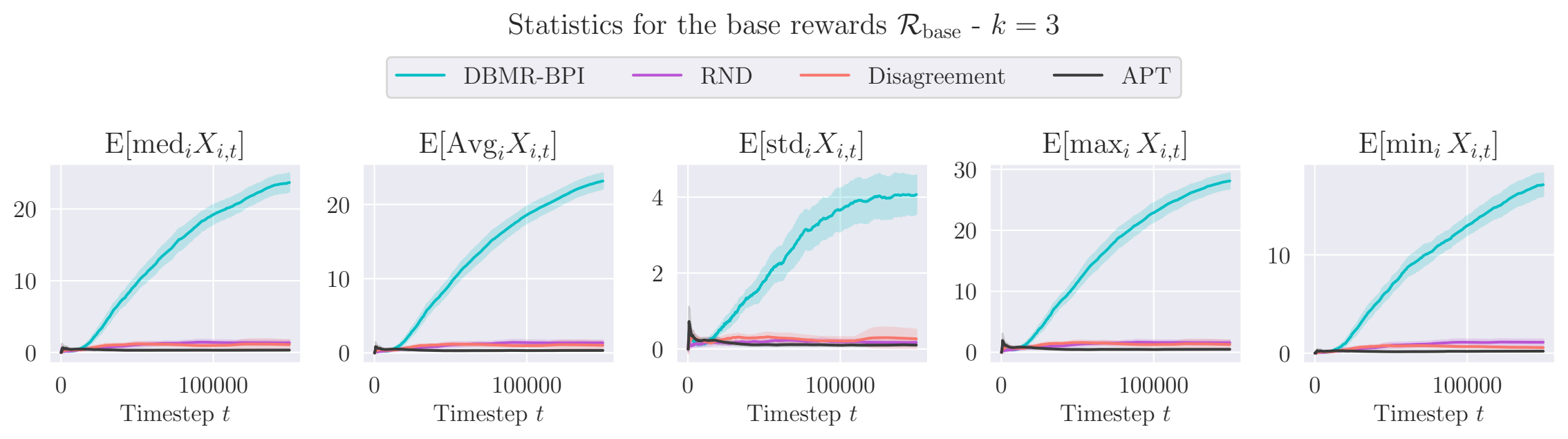
Statistics for the base rewards $\mathcal{R}_{\text{base}}$ - $k = 3$
DBMR-BPI
RND
Disagreement
APT
$\mathrm{E}[\mathrm{med}_i X_{i,t}]$
$\mathrm{E}[\mathrm{Avg}_i X_{i,t}]$
$\mathrm{E}[\mathrm{std}_i X_{i,t}]$
$\mathrm{E}[\max_i X_{i,t}]$
$\mathrm{E}[\min_i X_{i,t}]$
Timestep t
0
100000
0
10
20
2
4
30