Cumulated regrets for different bandit algorithms, averaged 1000 times
9 arms: Bayesian MAB, Bernoulli with means on [0, 1]

Cumulated regret $R_t = t\mu^* - \sum_{s=0}^{t-1}\sum_{k=1}^{9}\mu_k \mathbb{E}_{1000}[T_k(s)]$

Time steps $t = 0 \ldots T-1$, horizon $T = 45678$

Legend:
- KLUCB
- KLUCB^{++} ($T = 45678$)
- DT($T_i = 200 \times 2^i$)[KLUCB^{++}]
- DT($T_i = 200^{2^i}$)[KLUCB^{++}]
- DT($T_i = 200^{1.1^i}$)[KLUCB^{++}]
- DT($T_i = (200/2)2^{2^i}$)[KLUCB^{++}]
- Lai & Robbins lower bound $= 3.74 \log(t)$