Opt算法必定小于每个节点在C个可能边取最大值得期望的累加

因此可以结合固定batch求出竞争比

再证明固定batch也是马尔可夫过程中的一种可能也就是说在强化学习训练的足够好的情况下，其竞争比一定大于固定batch的竞争比

Regret除以opt即可得到与竞争比相关的：竞争比-策略空间中最好的策略的累积/opt。