

algo (REINFORCE) [Williams 1992]

1° sample  $\{u_n\}_{n=1}^N \sim p_\theta$

2° update  $\theta^+ = \theta - \gamma \cdot \frac{1}{N} \sum_n c(u_n) \cdot D_\theta \log p_\theta(u_n)$

$\hookrightarrow E[c(u) D_\theta \log p_\theta(u)] = \lim_{N \rightarrow \infty} \underbrace{\frac{1}{N} \sum_n c(u_n) \cdot D_\theta \log p_\theta(u_n)}_{\text{by Central Limit Thm}}$