Project

Yizhi Fang

$$\nabla_\theta \eta_i(\pi_\theta) = E_\tau[\sum_{t=0}^{T-1} \nabla_\theta \log \pi_i(a_{t,i}|s_t,\theta) A(s_t, a_{t,1}, ..., a_{t,N}) + \beta \sum_{t=0}^{T-1} \nabla_\theta H(s_t|\theta)] \tag{1}$$

$$\hat{A}(s_t, a_{t,1}, ..., a_{t,N}) = \sum_{t=0}^{\infty} (\gamma\lambda)^l \delta_{t+l} \tag{2}$$

$$\begin{aligned}\pi_i(a_{t,i}|s_t) =& \frac{1}{\sqrt{2\pi}\sigma_{1,i}} \exp(-\frac{(a_{t,i}^{(1)} - \mu_{\theta,i}(s_t)^{(1)})^2}{2\sigma_{1,i}^2}) \\ & \cdot \frac{1}{\sqrt{2\pi}\sigma_{2,i}} \exp(-\frac{(a_{t,i}^{(2)} - \mu_{\theta,i}(s_t)^{(2)})^2}{2\sigma_{2,i}^2})\end{aligned}$$