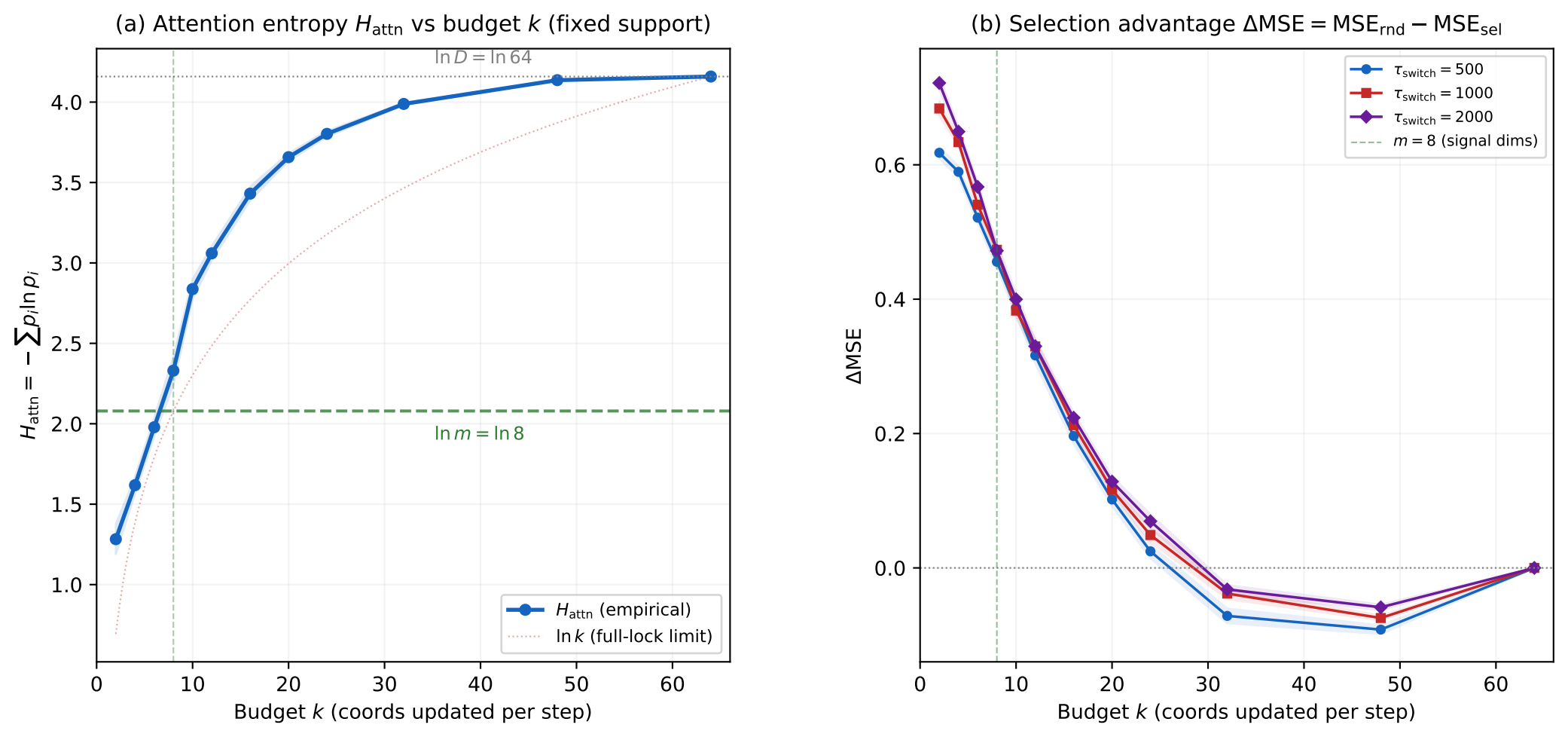
(a) Attention entropy H_{attn} vs budget k (fixed support)
$\ln D = \ln 64$
$\ln m = \ln 8$
H_{attn} (empirical)
$\ln k$ (full-lock limit)
$H_{\text{attn}} = -\sum p_i \ln p_i$
Budget k (coords updated per step)
(b) Selection advantage $\Delta\text{MSE} = \text{MSE}_{\text{rnd}} - \text{MSE}_{\text{sel}}$
$\tau_{\text{switch}} = 500$
$\tau_{\text{switch}} = 1000$
$\tau_{\text{switch}} = 2000$
$m = 8$ (signal dims)
ΔMSE
Budget k (coords updated per step)