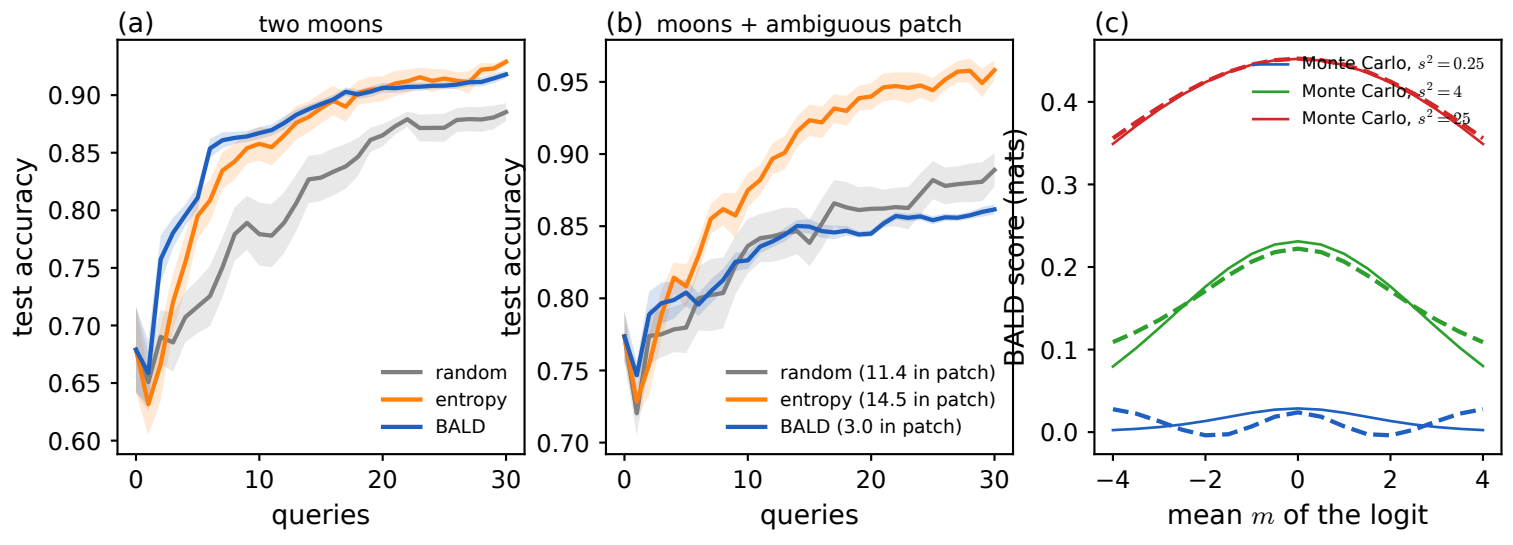

(a)
two moons
test accuracy
queries
random
entropy
BALD
(b)
moons + ambiguous patch
test accuracy
queries
random (11.4 in patch)
entropy (14.5 in patch)
BALD (3.0 in patch)
(c)
BALD score (nats)
mean m of the logit
Monte Carlo, $s^2 = 0.25$
Monte Carlo, $s^2 = 4$
Monte Carlo, $s^2 = 25$