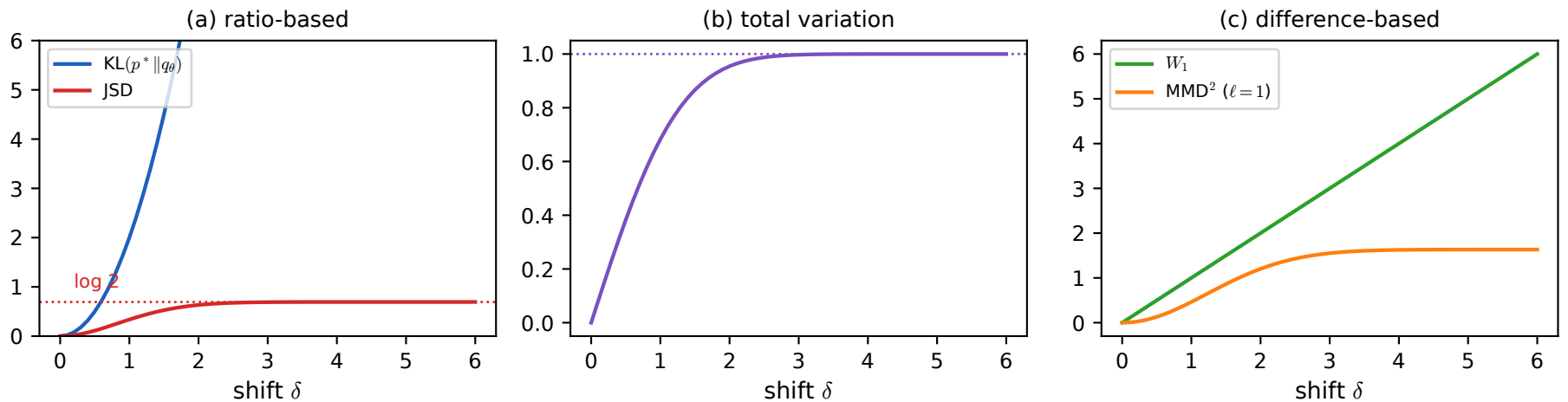

(a) ratio-based
KL($p^* \| q_\theta$)
JSD
log 2
shift δ
(b) total variation
shift δ
(c) difference-based
W_1
MMD^2 ($\ell = 1$)
shift δ