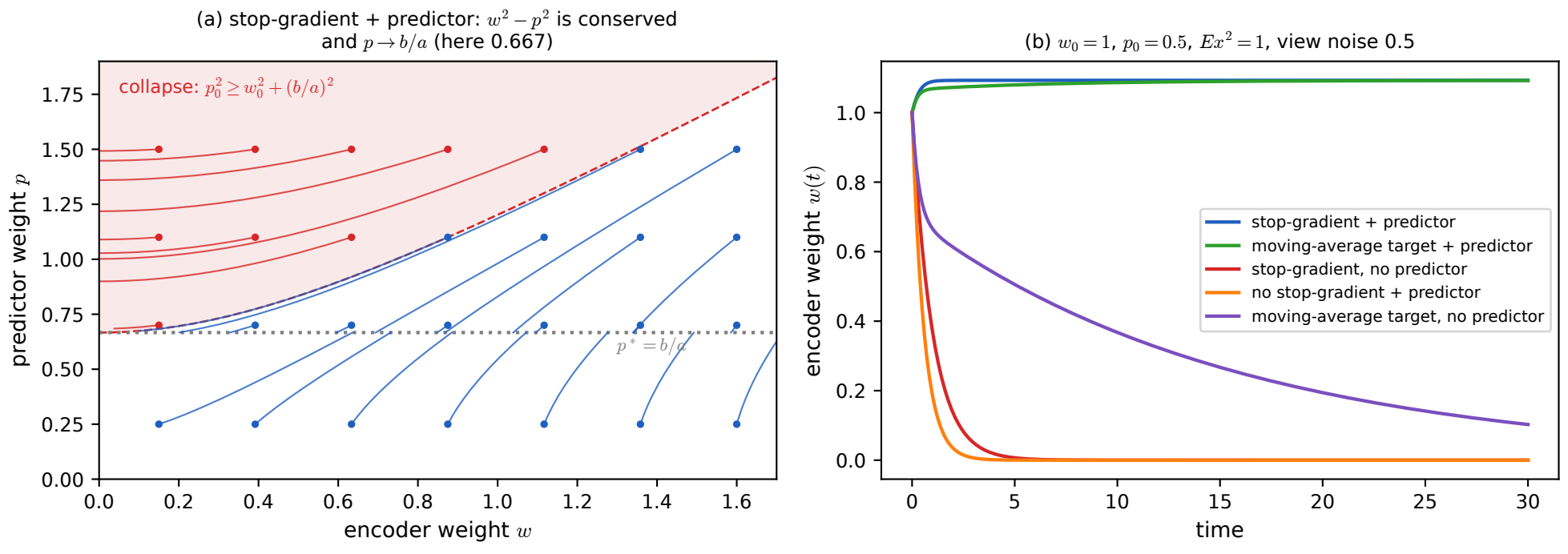

(a) stop-gradient + predictor: $w^2 - p^2$ is conserved
and $p \to b/a$ (here 0.667)
collapse: $p_0^2 \geq w_0^2 + (b/a)^2$
$p^* = b/a$
predictor weight p
encoder weight w
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
(b) $w_0 = 1$, $p_0 = 0.5$, $Ex^2 = 1$, view noise 0.5
encoder weight $w(t)$
time
stop-gradient + predictor
moving-average target + predictor
stop-gradient, no predictor
no stop-gradient + predictor
moving-average target, no predictor
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15
20
25
30