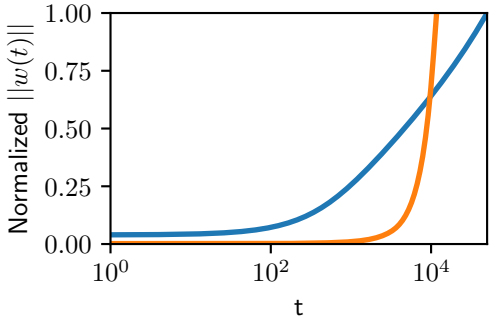
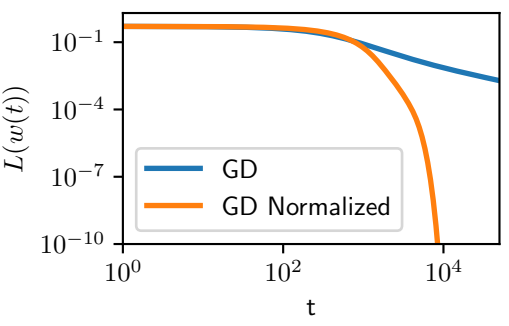
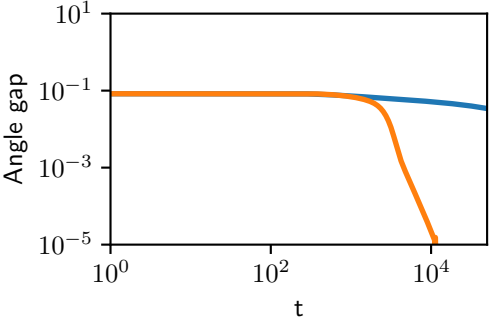
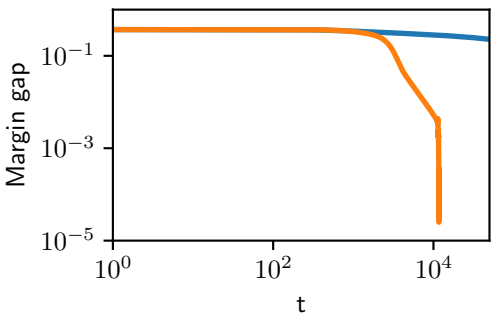

Normalized $||w(t)||$
1.00
0.75
0.50
0.25
0.00
10^0
10^2
10^4
t


$L(w(t))$
10^{-1}
10^{-4}
10^{-7}
10^{-10}
GD
GD Normalized
10^0
10^2
10^4
t


Angle gap
10^1
10^{-1}
10^{-3}
10^{-5}
10^0
10^2
10^4
t


Margin gap
10^{-1}
10^{-3}
10^{-5}
10^0
10^2
10^4
t