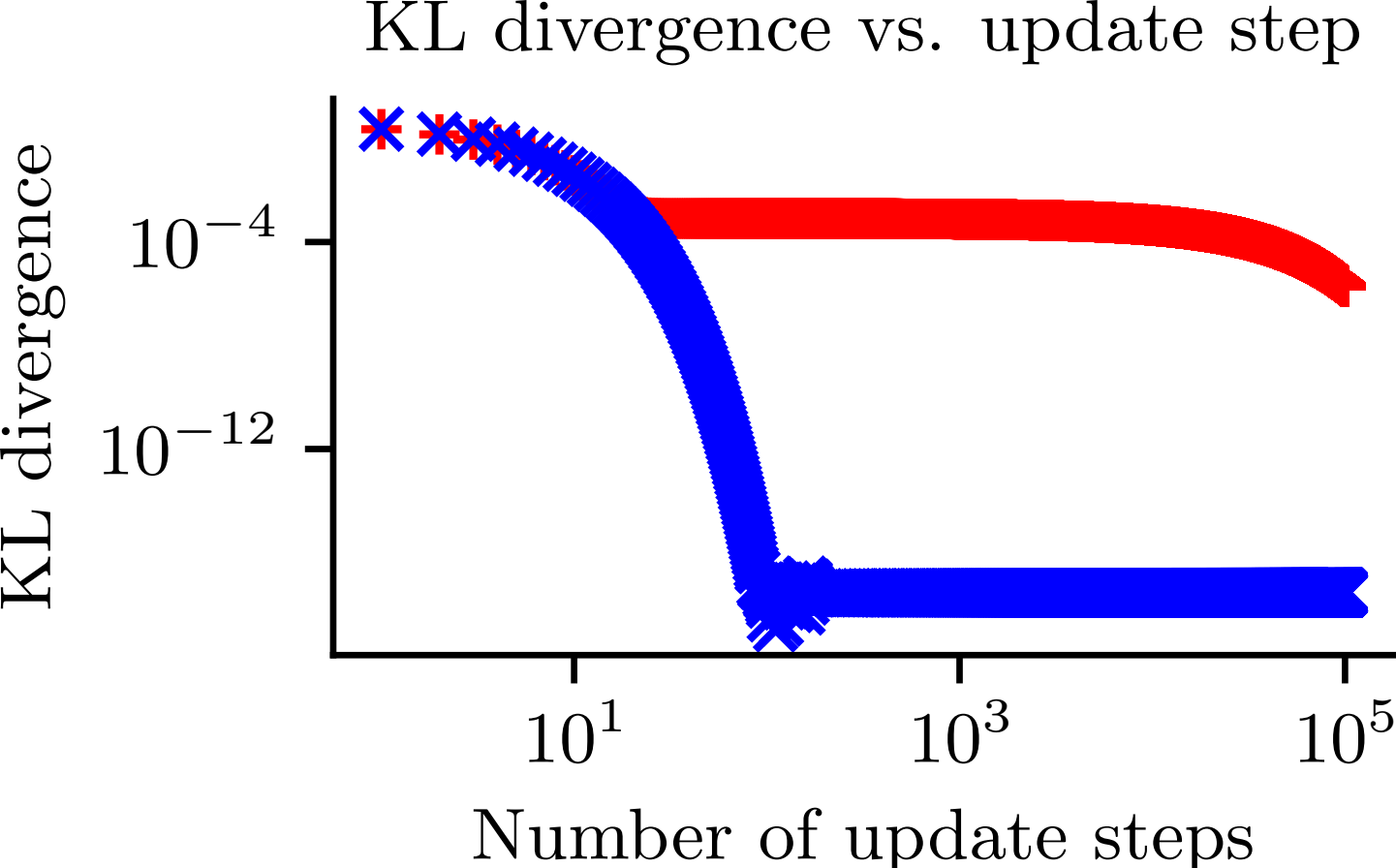
KL divergence vs. update step
KL divergence
10^{-4}
10^{-12}
10^{1}
10^{3}
10^{5}
Number of update steps