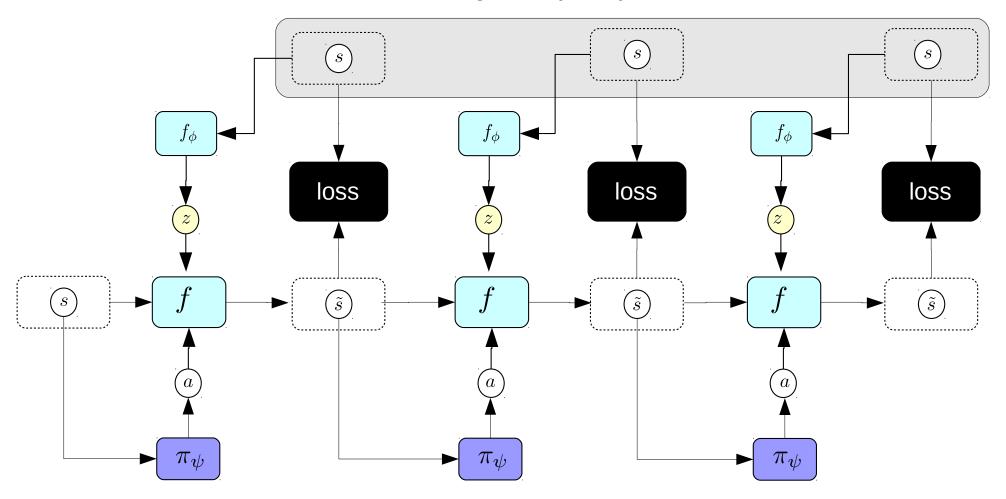Training set trajectory

Stochastic policy
network (optimized)