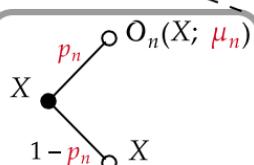


Possible operations:

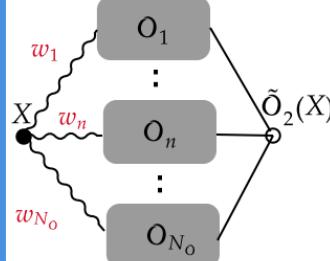


All possible N_O^2 subpolicies



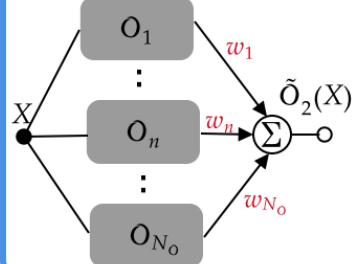
ADDA

ADDA Stage

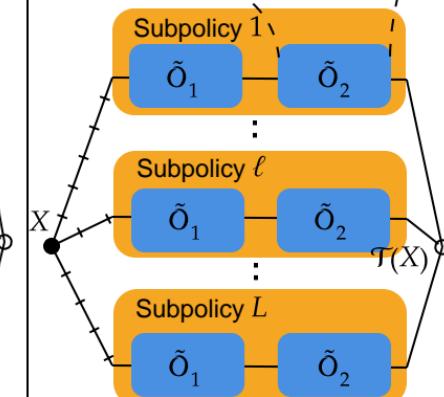
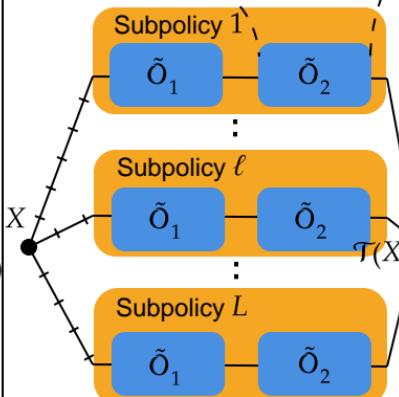
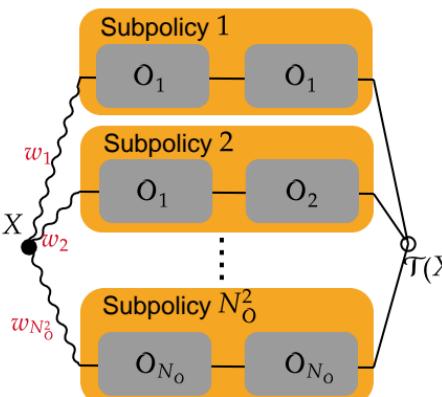


Faster AA

Faster AA Stage



DADA



—●— : uniform sampling

p, μ, w : learnable parameters

—●—
~~~~~  
 $w$  : sampling with log-probability  $w$