Algorithm 3 Conditional Temporal Difference (CTD) for evaluating policy π

Let θ_1, the nonnegative parameters α and $\{\beta_t\}$ be given.
for $t = 1, \ldots, T$ **do**
Collect α state transition steps without updating $\{\theta_t\}$, denoted as $\{\zeta_t^1, \zeta_t^2, \ldots, \zeta_t^\alpha\}$.
Set

$$\theta_{t+1} = \theta_t - \beta_t \tilde{F}^\pi(\theta_t, \zeta_t^\alpha). \tag{5.11}$$

end for