

Algorithm *DYNA-RRTPI(N)*

(* Episodic offline version: builds RRT and implements policy iteration using DYNA framework *)

1. $G \leftarrow$ Build a vanilla RRT using Euclidean distance metric
2. $J_o \leftarrow$ NN-TD(G) i.e use the vanilla RRT over the space to initialize the J values
3. $n \leftarrow 0$
4. **while** $n \leq N$
5. $G_n \leftarrow$ Episode($S, ||\cdot||_{J_{n-1}}$)
6. $J_n \leftarrow$ NN-TD(G_n, J_{n-1})
7. $n \leftarrow n+1$
8. End of Algorithm