Algorithm *Online-DYNA-RRTPI( )*

(∗ Online version: builds RRT and implements policy iteration using DYNA framework ∗)

1. G← 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. **while** goal is not reached
4. $x_{rand} \leftarrow$ Sample(S);
5. $x_{near} \leftarrow$ Nearest(x_{rand}, V(G) , $||.||_J$);
6. $(x_{new}, a, r) \leftarrow$ Extend(x_{near}, x_{rand}, S);
7. **if** Not-Colliding(x_{new}, x_{near},S)
8. **then** Connect x_{new} to x_{near}
9. V(G) $\leftarrow$ V(G) U x_{new}
10. E(G) $\leftarrow$ E(G) U (x_{new}, a, r)
11. $J(x_{near}) \leftarrow (1-\alpha)J(x_{near} + \alpha(r + \gamma J(x_{new})))$
12. J $\leftarrow$ RRT-Planning(G, J, number of iteration)
13. End of Algorithm