

## Algorithm 1 BenchRL-QAS

**Require:** Set of quantum tasks  $\mathcal{T}$  (e.g., VQE, VQSD, VQC, state preparation),  
RL algorithms  $\mathcal{A}$ ,  
Encoding scheme (e.g., tensor-based),  
Illegal action handler,  
Curriculum schedule.

```
1: for each task  $t$  in  $\mathcal{T}$  do
2:   Initialize environment  $\mathcal{E}_t$  with Hamiltonian or objective for  $t$ 
3:   for each RL algorithm  $a$  in  $\mathcal{A}$  do
4:     Initialize agent  $\mathcal{R}_a$  with action and state spaces
5:     Initialize quantum circuit  $C$  (an empty circuit)
6:     while not converged do
7:       Encode current circuit state  $s \leftarrow \text{EncodeState}(C)$ 
8:       Select action  $u \leftarrow \mathcal{R}_a.\text{SelectAction}(s)$ 
9:       if IsIllegalAction( $u$ ) then
10:        Apply penalty or mask action, update agent
11:        Continue to next iteration
12:      end if
13:      Update circuit  $C \leftarrow \text{ApplyAction}(C, u)$ 
14:      Simulate circuit, obtain reward  $r \leftarrow \mathcal{E}_t.\text{Evaluate}(C)$ 
15:      Observe new state  $s' \leftarrow \text{EncodeState}(C)$ 
16:       $\mathcal{R}_a.\text{Update}(s, u, r, s')$ 
17:    end while
18:    Log performance metrics for  $(t, a)$ 
19:  end for
20:  Optionally (if used curriculum learning): Increase task difficulty
21: end for
22: Aggregate and compare results across tasks and algorithms
```

