Learning Forward Dynamics Model and Informed Trajectory Sampler for Safe Quadruped Navigation

APPENDIX

A. Command sequence sampling to train FDM

To capture broad command sequence distribution, we generated command sequence data via linear time-correlated command sampling (Eqn. [1]), normal time-correlated command sampling (Eqn. [2]), and constant command sampling (Eqn. [3]) in same proportion. For all sampling methods, $c_0 \sim U(c_{\text{min}}, c_{\text{max}})$.

$$\beta \sim U(\beta_{\text{min}}, 1)$$
$$c_{\text{rand}} \sim U(c_{\text{min}}, c_{\text{max}})$$
$$c_{t+1} = \beta \cdot c_t + (1-\beta) \cdot c_{\text{rand}}$$
$$\forall t \in \{0 \ldots T-1\},$$

(1)

$$\sigma \sim U(0, \sigma_{\text{max}})$$
$$c_{t+1} \sim N(c_t, \sigma)$$
$$\forall t \in \{0 \ldots T-1\},$$

(2)

$$c_{\text{rand}} \sim U(c_{\text{min}}, c_{\text{max}})$$
$$c_t = c_{\text{rand}}$$
$$\forall t \in \{0 \ldots T-1\},$$

(3)

B. Hyperparameters

| Parameter                          | Value       |
|-----------------------------------|-------------|
| Number of environments            | 800         |
| Learning rate                     | 3e-4        |
| Batch size                        | 512         |
| Data buffer size                  | 45000 samples |
| Cross entropy loss weight         | 2.0         |
| Mean squared error weight         | 1.7         |
| Single data collection period     | 9 [s]       |
| State encoder hidden units        | [256, 256, 128, 128, 100] |
| Command encoder hidden units      | [32, 64]    |
| Recurrent layer hidden units      | 100         |
| Number of recurrent layers        | 2           |
| Coordinate predictor hidden units | [64, 32, 16, 2] |
| Collision predictor hidden units  | [64, 32, 16, 1] |
| Activation function               | LeakyReLU   |
| Batch Normalization               | True        |
| Dropout                           | 0.2         |

TABLE I: Hyperparameters used for training FDM. Last values in the list of hidden units are output units. Model weights of state encoder and command encoder were fixed with the weights in FDM.

C. Point Goal Navigation in open fields

| Parameter                          | Value       |
|-----------------------------------|-------------|
| Learning rate                     | 3e-4        |
| Batch size                        | 256         |
| State encoder hidden units        | [256, 256, 128, 128, 100] |
| Waypoint encoder hidden units     | [32, 64]    |
| Waypoint recurrent encoder hidden units | 100     |
| Command encoder hidden units      | [32, 64]    |
| Latent vector dimension           | 16          |
| Latent encoder hidden units       | [64, 32, 16] |
| Latent decoder hidden units       | [64, 64, 32] |
| Command recurrent encoder hidden units | 32     |
| Command decoder hidden units      | [32, 3]     |
| Activation function               | LeakyReLU   |
| Batch Normalization               | True        |
| Dropout                           | 0.2         |

TABLE III: Hyperparameters used for sampling-based model-predictive control module

Generated random environments with sampled endpoints did not always show a feasible path. Specifically, the global
planner, BIT* in our work, sometimes failed to find a coarse path connecting the sampled endpoints due to very densely located obstacles. In this case, we sampled a different endpoint or generated a different environment with the random environment generator.