

# Learning Neural Observer-Predictor Models for Limb-level Sampling-based Locomotion Planning

Abhijeet M. Kulkarni<sup>1</sup>, Ioannis Poulakakis<sup>2</sup> and Guoquan Huang<sup>1</sup>

**Abstract**—Accurate full-body motion prediction is essential for the safe, autonomous navigation of legged robots, enabling critical capabilities like limb-level collision checking in cluttered environments. Simplified kinematic models often fail to capture the complex, closed-loop dynamics of the robot and its low-level controller, limiting their predictions to simple planar motion. To address this, we present a learning-based observer-predictor framework that accurately predicts this motion. Our method features a neural observer with provable Uniformly Ultimately Bounded (UUB) guarantees that provides a reliable latent state estimate from a history of proprioceptive measurements. This stable estimate initializes a computationally efficient predictor, designed for the rapid, parallel evaluation of thousands of potential trajectories required by modern sampling-based planners. We validated the system by integrating our neural predictor into an Model Predictive Path Integral (MPPI)-based planner on a Vision60 quadruped. Hardware experiments successfully demonstrated effective, limb-aware motion planning in a challenging, narrow passage and over small objects, highlighting our system’s ability to provide a robust foundation for high-performance, collision-aware planning on dynamic robotic platforms.

## I. INTRODUCTION AND RELATED WORK

Quadruped robots have progressed from laboratory prototypes to robust systems capable of real-world deployment. A key driver has been the maturity of low-level, proprioceptive controllers that fuse joint encoders and Inertial Measurement Units (IMUs) to maintain balance while tracking commanded high-level planar base velocities. We term robots with such onboard controllers as Proprioceptively Stabilized Robots (PSRs) and this approach is now prevalent in both the research community [1], [2] and on the commercial platforms [3], [4]. Consequently, the research focus in legged autonomy is shifting from the challenge of achieving dynamic locomotion to enabling safety-guaranteed autonomous navigation in complex environments like industrial inspection sites.

PSR control permits the abstraction of complexity of legged locomotion into a simplified predictive model, such as omnidirectional kinematic models, for task-level planning. These models integrate seamlessly with classical navigation stacks and allow modular collision-aware planning [5]. A

This work supported by the University of Delaware (UD) College of Engineering, and the NSF (IIS-2410019). The work of I. Poulakakis was supported by the project “Applied Research for Autonomous Robotic Systems” (MIS5200632) which is implemented within the framework of the National Recovery and Resilience Plan (NNRP) “Greece 2.0” (Measure: 16618-Basic and Applied Research) and is funded by the European Union-NextGenerationEU.

<sup>1</sup>Department of Mechanical Engineering, University of Delaware, Newark, DE, USA {amkulk, ghuang}@udel.edu

<sup>2</sup>Robotics Institute, Athena Research Center, Marousi, Greece; School of Mechanical Engineering, National Technical University of Athens, Greece; HERON-Center of Excellence in Robotics, Athens, Greece poulakakis@mail.ntua.gr, i.poulakakis@athenarc.gr.

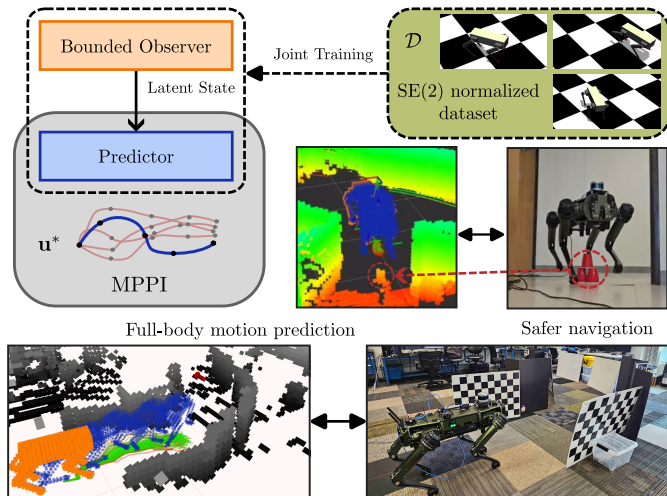


Fig. 1. The pipeline of the proposed system. The dataset is collected from both simulation and real robots and used to jointly train the proposed observer-predictor. The trained predictor is then deployed into the sampling-based MPPI planner for safe navigation. Full-body prediction allows for limb-level collision checking in narrow passages and over small objects.

typical high-level planner then selects velocity commands that accomplish a task in the presence of obstacles by combining perception with simplified dynamics. This paradigm was often adopted for legged PSRs, representing geometry with primitive shapes for collision checks [6]–[11].

However, the effectiveness of these simplifications depends critically on how closely the actual PSR response follows the assumed model under a given velocity command. Mismatch directly undermines the validity of collision checks and degrade task execution performance. To compensate, one can inflate the robot’s shape [11] or employ robust formulations [12]. Data-driven alternatives predict the evolution of specific task related outputs conditioned on velocity commands—for example, forecasting Center of Mass (COM) and footsteps [13] or estimating collision probabilities [14]–[16]. Yet these approaches typically omit full-body motion prediction; moreover, collision probability estimators are often tied to a specific obstacle distribution and requires retraining whenever it changes.

There are works that are able to predict full-body motion based on motion library [17] and even get planned commands [18], but these are suitable for simulation environments where full state of the system is available. In case of real robotic system one does not have access to all states that are required for full-body motion prediction, and thus are to be recovered from measurements. Along with the states of physical body, PSRs also have internal states associated with the low-level controller which are needed to be observed

in order to predict the robot’s motion. For such complex systems, learning based nonlinear stable observers exists [19]–[21]. By constraining the Lipschitz constant of the observer dynamics, one can establish stability guarantees [22], a principle that has also supported stable learned controllers [23]. What remains missing is a unified observer-predictor that recovers the latent state, produces full-body forecasts at planning rates, and slots into sampling-based, collision-aware planners without resorting to overly conservative geometry and simultaneously being agnostic to obstacle distribution.

As such, in this work, we design an efficient learning-based observer–predictor framework for full-body motion prediction of PSRs, see Fig. 1, which is trained only from a history of proprioceptive observations and a sequence of future commands, and thus it is agnostic to obstacles. Once trained, it can readily be deployed on the PSR in *any* environments. In particular, the main contributions of the paper are the following:

- We design a neural observer that reconstructs the latent state of the PSR, including the states of robot and low-level controller, from available measurements. We analytically derive conditions on the Lipschitz constant of the observer dynamics that ensure that it is UUB.
- We build a neural predictor based on Gated Recurrent Unit (GRU), which can perform massive parallel rollouts and enables thousands of trajectories to be evaluated concurrently for sampling-based planning. For instance, starting from a common initial latent state, the model rolls out 1000 command sequences for 200 steps (4 s with sampling rate of  $\Delta t = 20\text{ms}$ ) in 13ms on a NVIDIA RTX T1000.
- We integrate our neural observer/predictor with an MPPI planner [24], which can check limb-level collisions, avoid the conservatism of primitive shapes, and improve goal-pose tracking while retaining real-time performance.
- We validate the proposed system in real settings, demonstrating accurate full-body prediction and safe navigation in cluttered environments, outperforming the baseline Constant Velocity (CV) kinematic model.

## II. SYSTEM OVERVIEW

Consider the discrete-time evolution of a PSR with a sampling period  $\Delta t > 0$ . At each timestep  $k$ , the PSR receives a high-level planar (translational and rotational) velocity command as the input expressed in its body frame:

$$\mathbf{u}_k = [v_k^x \quad v_k^y \quad \omega_k^z]^\top \in \mathbb{R}^{n_u} \quad (1)$$

For collision-aware planning, the planner requires to predict the robot’s full-body configuration, which includes its base position  $\mathbf{p}_k \in \mathbb{R}^3$ , orientation parameterized with roll-pitch-yaw  $\psi_k \in \mathbb{R}^3$ , and joint angles  $\theta_k \in \mathbb{R}^{12}$ . By convention, these quantities are expressed in a fixed, global reference frame. However, the complete full-body configuration is not directly available from onboard sensors. The available proprioceptive measurements consist of the base’s roll and pitch angles (which can be reliably estimated from the

direction of gravity) and the joint angles. We thus define the directly measured output vector  $\mathbf{y}_k$  and the full-body configuration vector  $\mathbf{z}_k$  as follows:

$$\mathbf{y}_k = [\psi_k^{\text{roll}} \quad \psi_k^{\text{pitch}} \quad \theta_k^\top]^\top \quad (2)$$

$$\mathbf{z}_k = [\mathbf{p}_k^\top \quad \psi_k^{\text{yaw}} \quad \mathbf{y}_k^\top] \quad (3)$$

### A. Problem Statement

A state estimator is typically used to infer the unmeasured components (global position and yaw) of  $\mathbf{z}_k$ , which however can be prone to drift over time if no global information is fused. To mitigate the accumulation of global estimation errors, we instead predict future configurations *relative* to the current robot’s reference frame (akin to robocentric visual-inertial odometry [25]). Specifically, by modeling the PSR’s dynamics with a latent state (implicit neural representation)  $\hat{\mathbf{x}}_k \in \mathbb{R}^{n_x}$ , we seek to learn the mapping from historical data to a sequence of future relative configurations  $\hat{\mathbf{z}}'_{k+1:k+T}$ :

$$\left\{ \underbrace{(\mathbf{y}_{k-H:k}, \mathbf{u}_{k-H:k})}_{\text{History}}, \underbrace{\mathbf{u}_{k+1:k+T}}_{\text{Future Commands}} \right\} \xrightarrow{\hat{\mathbf{x}}_k} \underbrace{\hat{\mathbf{z}}'_{k+1:k+T}}_{\text{Relative Predictions}} \quad (4)$$

The core challenge is to recover an estimate  $\hat{\mathbf{x}}_k$  of the latent state from the history and use it to forecast the relative trajectory  $\hat{\mathbf{z}}'_{k+1:k+T}$ . These predictions are essential for a planner to perform limb-level collision checking against a local environment map [26]. The non-smooth nature of this collision-checking problem—combined with the complex dynamics of the PSR—motivates the use of sampling-based planners such as MPPI. This choice, in turn, imposes a critical requirement on the learned model—that is, it has to be highly efficient, capable of generating thousands of trajectory rollouts in parallel to support real-time planning.

To that end, we learn a lightweight data-driven observer-predictor model and efficiently integrate it into a sampling-based planner. Fig. 1 shows the overall pipeline of the proposed system that consists of three key stages: (i) a nonlinear observer that recovers the latent state of the PSR from a history of measurements (outputs) and motion commands (inputs), (ii) a recurrent dynamics predictor that is initialized with this latent state and rolls out future full-body trajectories for candidate command sequences, and (iii) a MPPI planner that uses these predictions to select the optimal command by evaluating many potential outcomes in parallel.

### B. Decoupled Observer-Predictor Architecture

Fig. 2 depicts the proposed decoupled architecture of the state observer and the future dynamics predictor. This decoupling is crucial, as the observer’s dynamics can be constrained to satisfy stability guarantees (see Sec. V), thus ensuring a reliable state estimate. The GRU-based predictor is optimized for the fast parallel rollouts required by the MPPI planner.

In particular, a nonlinear observer is designed to recover the latent state  $\hat{\mathbf{x}}_k$  from the history  $(\mathbf{u}_{k-H:k}, \mathbf{y}_{k-H:k})$ :

$$\hat{\mathbf{x}}_{k+1} = \mathbf{A}\hat{\mathbf{x}}_k + g(\hat{\mathbf{x}}_k, \mathbf{u}_k) + \mathbf{K}(\mathbf{y}_k - \hat{\mathbf{y}}_k) \quad (5)$$

$$\hat{\mathbf{y}}_k = \mathbf{C}_y \hat{\mathbf{x}}_k \quad (6)$$

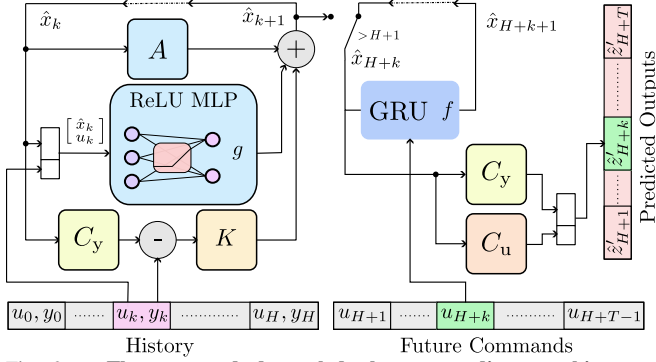


Fig. 2. **The proposed decoupled observer-predictor architecture.** Our framework consists of a correction-based nonlinear observer (left) that produces a stable latent state estimate  $\hat{\mathbf{x}}_k$  from history. This state initializes a recurrent predictor (right) that unrolls future relative full-body configurations  $\hat{\mathbf{z}}^i$  in response to a candidate command sequence.

where  $\mathbf{A} \in \mathbb{R}^{n_x \times n_x}$  is a state transition matrix,  $g(\cdot)$  is an Multi Layer Perceptron (MLP) network that models the underlying nonlinear dynamics,  $\mathbf{K} \in \mathbb{R}^{n_x \times n_y}$  is the learned observer gain, and  $\mathbf{C}_y \in \mathbb{R}^{n_y \times n_x}$  maps the state to measured outputs. It is important to note that this structure allows us to provide formal UUB guarantees on the estimation error (see Thm. V.3).

On the other hand, given the estimated state  $\hat{\mathbf{x}}_k$ , the predictor forecasts a sequence of relative full-body configurations  $\hat{\mathbf{z}}'_{k+1:k+T}$  in response to a future command sequence  $\mathbf{u}_{k+1:k+T}$ . Starting with  $\hat{\mathbf{x}}_k$ , the prediction unrolls sequentially for  $t = 1, \dots, T$ :

$$\hat{\mathbf{x}}_{k+t} = f(\hat{\mathbf{x}}_{k+t-1}, \mathbf{u}_{k+t}) \quad (7)$$

$$\hat{\mathbf{z}}'_{k+t} = \begin{bmatrix} \mathbf{C}_u \\ \mathbf{C}_y \end{bmatrix} \hat{\mathbf{x}}_{k+t} \quad (8)$$

where  $f(\cdot)$  is an efficient GRU network that models the time evolution and  $\mathbf{C}_u \in \mathbb{R}^{(n_z - n_y) \times n_x}$  maps the latent state to the unmeasured components of the relative configuration. The predictor's GRU structure was chosen for its effectiveness in modeling temporal sequences and its computational efficiency for the fast, parallel rollouts required by MPPI.

### C. Joint End-to-End Training

While functionally separated, the model parameters

$$\Theta = \{\mathbf{A}, \mathbf{K}, \mathbf{C}_y, g, f, \mathbf{C}_u\} \quad (9)$$

of the observer and predictor are jointly trained on a quality dataset  $\mathcal{D}$  that comprises diverse trajectories:

$$\mathcal{D} = \left\{ (\mathbf{y}_{0:H}^{(i)}, \mathbf{u}_{0:H+T}^{(i)}, \mathbf{z}'_{H+1:H+T}^{(i)}) \right\}_{i=1}^N \quad (10)$$

To this end, the network is trained to minimize a loss function  $\mathcal{L}(\Theta)$  that penalizes multi-step prediction errors while enforcing the observer's stability condition. This end-to-end training ensures the observer learns a latent representation that is both stable and maximally informative for the predictor's forecasting task. Details are provided in Sec. III.

### D. MPPI Planning

We integrate the learned predictor into an MPPI framework [24] to select safe and task-aware commands. To improve sample efficiency, candidate command sequences

### Algorithm 1 MPPI with Learned Predictor

- 1: **Input:** Current latent state  $\hat{\mathbf{x}}_k$ , nominal control points  $P$
- 2: **for**  $i = 1, \dots, N$  **do** ▷ Sample  $N$  trajectories
- 3:   Sample perturbed control points  $P^{(i)} \sim \mathcal{N}(P, \Sigma)$
- 4:   Generate  $\mathbf{u}_{k+1:k+T}^{(i)}$  from  $P^{(i)}$  using (11)
- 5:   Predict  $\hat{\mathbf{z}}'_{k+1:k+T}^{(i)}$  using (7) from  $\hat{\mathbf{x}}_k$
- 6:   Convert  $\hat{\mathbf{z}}'_{k+1:k+T}^{(i)}$  to global  $\hat{\mathbf{z}}^{(i)}$  using current robot pose
- 7:   Evaluate cost  $\mathcal{J}^{(i)}$  (22) with  $\hat{\mathbf{z}}_{k+1:k+T}^{(i)}$  and  $\mathbf{u}_{k+1:k+T}^{(i)}$
- 8: **end for**
- 9:   Compute weights  $w^{(i)} \propto \exp(-\mathcal{J}^{(i)}/\lambda)$
- 10:   Update control points  $P \leftarrow \sum_i w^{(i)} P^{(i)}$
- 11:   Execute the first command in the new optimal sequence

$\mathbf{u}_{k+1:k+T}$  are parameterized by the control points  $\{P_j\}$  of a degree- $d$  Bézier curve [27]:

$$\mathbf{u}(s) = \sum_{j=0}^d \binom{d}{j} (1-s)^{d-j} s^j P_j, \quad s \in [0, 1]. \quad (11)$$

At each timestep, MPPI samples noisy perturbations of the control points, generates corresponding command sequences, and simulates predicted trajectories using (7). The predicted relative trajectories  $\hat{\mathbf{z}}^{(i)}$  are converted to the global frame  $\hat{\mathbf{z}}^{(i)}$  and evaluated by a cost function  $J$  (detailed in Sec. IV). The control points are then updated via a weighted average. The first command  $\mathbf{u}_{k+1}^*$  of the optimized sequence is executed, and the process repeats (see Alg. 1).

## III. LEARNING OBSERVER AND PREDICTOR

In this section, we present in detail the process of learning our neural observer and predictor. Note that the PSR under consideration is the Ghost Robotics Vision60 quadruped [3] equipped with its proprietary low-level controller that translates high-level velocity commands into joint torques.

### A. Building Datasets

To build quality datasets that capture a rich variety of dynamic behaviors, including the complex internal modes of the low-level controller that may activate during navigation, we have collected 40-minute-long trajectories of the robot executing various commands in simulation and 10-minutes on the real robot in a motion capture environment. In simulations, the inertia parameters are randomized and noise is introduced in the measurements. For each run  $i$ , we logged the command input  $\mathbf{u}_k$ , joint angles  $\boldsymbol{\theta}_k$ , and the ground-truth global base pose  $(\mathbf{p}_k, \boldsymbol{\psi}_k)$  at a frequency of  $1/\Delta t = 50$  Hz resulting in  $N_{\text{dur}}$  timesteps. This results in a raw dataset  $\mathcal{D}_{\text{raw}}^{(i)} = \{(\mathbf{u}_k, \mathbf{p}_k, \boldsymbol{\psi}_k, \boldsymbol{\theta}_k)\}_{k=0}^{N_{\text{dur}}}$ . To ensure the command distribution in our training data mirrors that of the downstream MPPI planner, we generated command sequences using piecewise Bézier segments. At the start of each segment, we sampled a duration  $t_{\text{bez}} \sim \mathcal{U}[3, 7]$  sec and a set of control points  $P_j \sim \mathcal{U}([-0.5, -0.5, -0.5], [0.5, 0.5, 0.5])$ . The continuous command profile was then generated via (11) and discretized at  $\Delta t$ . This strategy produces smooth, non-stationary commands that effectively stimulate the range of the PSR's closed-loop dynamics possible during navigation.

From these collected raw logs, we extract overlapping windows of length  $H + T + 1$  (say  $H = 30$  and  $T = 200$ ). Each window forms a single data sample comprising an observation history  $(\mathbf{y}_{0:H}, \mathbf{u}_{0:H})$ , a future command sequence  $\mathbf{u}_{H:H+T}$ , and a ground-truth future full-body trajectory  $\mathbf{z}_{H+1:H+T}$ . To ensure the model learns motions relative to the robot’s current state rather than the global frame, we transform to robocentric coordinates. This process can improve the sample efficiency and generalization by making learning invariant to the robot’s initial state. For each sample, we define a reference frame at the end of the history window, timestep  $H$ , with pose  $(\mathbf{p}_H, \psi_H)$ . The future full-body configurations  $\mathbf{z}_k$  for  $k > H$  are transformed into this local frame to yield the relative configurations  $\mathbf{z}'_k$ . This transformation is defined as:

$$\mathbf{z}'_k = \mathcal{T}_{(\mathbf{p}_H, \psi_H)}^{-1}(\mathbf{z}_k) = \begin{bmatrix} \mathbf{R}_z(-\psi_H^{\text{yaw}})(\mathbf{p}_k - \mathbf{p}_H) \\ \psi_k^{\text{yaw}} - \psi_H^{\text{yaw}} \\ \psi_k^{\text{roll}} \\ \psi_k^{\text{pitch}} \\ \boldsymbol{\theta}_k \end{bmatrix} \quad (12)$$

where  $\mathbf{R}_z(-\psi_H^{\text{yaw}})$  is the rotation matrix for the yaw angle. Note that roll, pitch, and joint angles are already expressed in the body frame and thus remain unchanged. Applying this transformation process to all extracted windows yields the final processed dataset  $\mathcal{D}$  as defined in (10).

### B. Loss Function

To train the model parameters  $\Theta$  (9) of the observer and predictor, we formulate the following composite loss function that combines a prediction accuracy term with a regularization term to enforce observer stability.

The primary objective is to accurately forecast future trajectories. For each sample in  $\mathcal{D}$  (10), the observer (5) is first unrolled over the history  $(\mathbf{y}_{0:H}, \mathbf{u}_{0:H})$  to compute the latent state estimate  $\hat{\mathbf{x}}_H$ . This state initializes the predictor (7), which is then rolled out for  $T$  steps using the future commands  $\mathbf{u}_{H:H+T-1}$  to predict the relative trajectory  $\hat{\mathbf{z}}'_{H+1:H+T}$ . The prediction loss measures the mean squared error between the predicted and ground-truth relative trajectories:

$$\mathcal{L}_{\text{pred}} = \frac{1}{T} \sum_{t=1}^T \|\mathbf{z}'_{H+t} - \hat{\mathbf{z}}'_{H+t}\|_2^2 \quad (13)$$

To ensure the observer is well-behaved and its state estimate  $\hat{\mathbf{x}}_k$  converges towards the true state  $\mathbf{x}_k$ , we enforce the contraction condition from Thm. V.3. This is achieved by penalizing violations of the condition  $\rho = \|\mathbf{A} - \mathbf{K}\mathbf{C}_y\|_2 + L_g < 1$ . We formulate this as a hinge loss:

$$\mathcal{L}_{\text{stab}} = \max(0, \rho - (1 - \epsilon)) \quad (14)$$

where  $\epsilon > 0$  is a small margin (e.g.,  $10^{-4}$ ) for strict contraction. During training,  $\|\mathbf{A} - \mathbf{K}\mathbf{C}_y\|_2$  is computed as the spectral norm of the matrix. The Lipschitz constant  $L_g$  of the MLP  $g$  is upper-bounded using Lemma V.2 as the product of the spectral norms of its weight matrices.

The models are then trained end-to-end by minimizing the weighted sum of the aforementioned prediction and stability

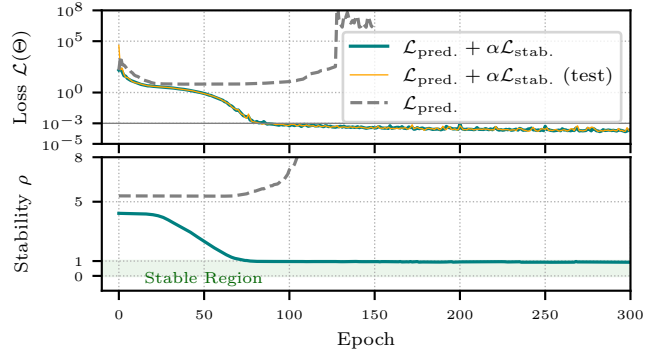


Fig. 3. **Training Behaviors.** Comparison of training with (blue) and without (orange) the stability regularization term (14). The stabilized model maintains the contraction factor  $\rho$  strictly below unity, leading to smoother convergence and lower variance, while the unconstrained model quickly becomes unstable.

losses over the dataset  $\mathcal{D}$  (10):

$$\mathcal{L}(\Theta) = \mathbb{E}_{\mathcal{D}} [\mathcal{L}_{\text{pred}} + \alpha \mathcal{L}_{\text{stab}}] \quad (15)$$

where  $\alpha > 0$  is a hyperparameter balancing prediction accuracy with the observer stability guarantee. We parameterize the predictor’s transition function  $f$  as a GRU network and use the Adam optimizer to find the optimal parameters  $\Theta^*$ . This joint training procedure yields an accurate predictor that is grounded in a provably stable state estimation framework.

### C. Model Analysis

As the proposed neural observer-predictor model is the key enabler for efficient and safe mission-oriented MPPI-based navigation, we perform numerical analysis to quantitatively understand its training and testing behaviors along three key axes: (i) stability during training, (ii) generalization to unseen data, and (iii) prediction accuracy over a finite horizon. To this end, we consider the MLP in the observer has three layers of sizes  $[128, 128, 128]$  and the predictor is a single layer GRU with hidden state of size 128, which is equivalent to latent state of  $n_x = 128$ . The hyperparameter  $\alpha = 0.1$  balances prediction accuracy and observer stability in (15).

The significance of including the stability regularization (14) in the training is clearly demonstrated in Fig. 3. Without it, the observer parameters drift into the unstable regime ( $\rho \geq 1$ ), causing the prediction error to diverge. In contrast, by incorporating the stability penalty, the contraction factor  $\rho$  is constrained to converge strictly below unity. This not only guarantees the observer’s contraction but also leads to smoother loss curves and significantly lower variance across training runs. Importantly, stability enforcement also improves generalization. Across an 80/20 train-test split, the test error of the stabilized model closely tracks the training error with no signs of overfitting. This suggests that the contraction property acts as an effective structural regularizer, preventing the model from learning unstable dynamics and enhancing its robustness to novel trajectories.

Fig. 4 validates the performance of the two core models. The observer (top) consistently recovers the latent state, with the estimation error converging within 30 timesteps (0.6s) despite being initialized from a broad uniform distribution ( $\hat{x}_0 \sim \mathcal{U}([-10, 10]^{n_x})$ ). Once initialized with this stable

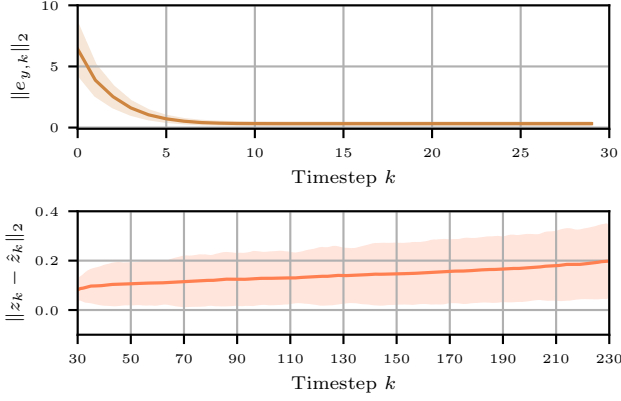


Fig. 4. **Test Performance.** (top) Observer state estimation error converges rapidly from a large initial error distribution. (bottom) The resulting predictor rollout error remains low and bounded over the full 4-second horizon, demonstrating accurate long-term forecasting.

state, the predictor (bottom) achieves low rollout error over the entire  $T = 200$  step horizon. This highlights the synergy between the two components: the observer provides a reliable latent state, enabling the predictor to accurately forecast future trajectories.

Interestingly, it can be seen from Fig. 4 that the empirical convergence rate is faster than the theoretical bound which might suggest for  $\rho \approx 0.97$  (from Fig. 3). This is because our theoretical bound in Thm. V.3 is conservative, relying on worst-case upper estimates of the Lipschitz constants.

#### IV. MPPI PLANING FOR GOAL-POSE NAVIGATION

To demonstrate its utility, we integrate our learned observer-predictor model into the MPPI framework, enabling it to generate collision-aware trajectories for mission-oriented navigation. The specific task is goal-pose tracking, where the objective is to drive the robot’s planar pose,  $\mathbf{q}_k = [p_k^x, p_k^y, \psi_k^{\text{yaw}}]^\top$ , to a user-defined goal,  $\mathbf{q}_{\text{goal}}$ . At each planning cycle  $k$ , the observer provides the latent state estimate  $\hat{\mathbf{x}}_k$  from the recent measurement history. This state serves as the initial condition for the predictor, which generates  $N$  future trajectory rollouts corresponding to perturbed command sequences  $\mathbf{u}_{k+1:k+T}^{(i)}$ . Each predicted relative trajectory,  $\hat{\mathbf{z}}_{k+1:k+T}^{(i)}$ , is projected into the global frame using the robot’s current pose  $(\mathbf{p}_k, \psi_k)$ :

$$\hat{\mathbf{z}}_{k+t}^{(i)} = \mathcal{T}_{(\mathbf{p}_k, \psi_k)}(\hat{\mathbf{z}}_{k+t}^{(i)}), \quad \forall t = 1, \dots, T. \quad (16)$$

These global trajectories are then evaluated using a cost function  $J^{(i)}$ , detailed below. The nominal command is updated via the importance-sampling average, and the first command of the optimized sequence,  $\mathbf{u}_{k+1}^*$ , is applied.

##### A. Cost Function

The cost function  $J$  used in Alg. 1 steers the MPPI planner by aggregating penalties for collision risk, goal tracking, and control effort over the prediction horizon  $T$ . First of all, to ensure safety, we penalize trajectories that risk collision with the environment, which is represented as a voxel map  $\mathcal{M}$ . We use a separately trained<sup>1</sup> MLP,  $h(\theta)$ , that maps joint angles

<sup>1</sup>While its occupied volume can be computed via standard forward kinematics, a learned network offers a simpler, parallelizable implementation.

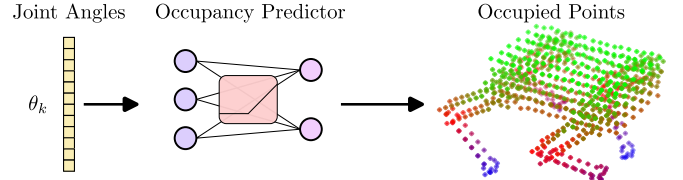


Fig. 5. **Robot Occupancy Predictor.** An ReLU MLP-based occupancy predictor, with layer sizes  $[64, 128]$ , takes joint angles  $\theta_k$  as input and outputs a 780 set of points  ${}^B\mathcal{O}(\theta_k)$  representing occupied space in the robot’s base frame. For collision checking, these points are transformed to the global frame and compared against an environment map  $\mathcal{M}$ .

to a point cloud  ${}^B\mathcal{O}(\theta)$  representing the robot’s occupied volume in its base frame (see Fig. 5). For each predicted state  $\hat{\mathbf{z}}_{k+t}$  along a rollout, these body points are transformed into the global frame. The collision cost is the number of points that fall within occupied voxels:

$$\ell_{\text{coll}}(\hat{\mathbf{z}}_{k+t}) = \sum_{\mathbf{o} \in {}^B\mathcal{O}(\hat{\theta}_{k+t})} \mathbf{1}_{\mathcal{M}}(\mathbf{R}_z(\hat{\psi}_{k+t})\mathbf{o} + \hat{\mathbf{p}}_{k+t}) \quad (17)$$

where  $\mathbf{1}_{\mathcal{M}}(\cdot)$  is an indicator function for the voxel map.

The task cost guides the robot towards a static goal configuration  $\mathbf{q}^{\text{goal}}$ . It is composed of penalties on position and yaw. The position cost at each timestep penalizes the Euclidean distance between the predicted 2D position  $\hat{\mathbf{p}}_{k+t}^{xy}$  and the goal:

$$\ell_{\text{pos}}(\hat{\mathbf{z}}_{k+t}) = \|\hat{\mathbf{p}}_{k+t}^{xy} - \mathbf{p}_{\text{goal}}^{xy}\|_2 \quad (18)$$

The yaw cost dynamically transitions between aligning with the direction of travel when far from the goal and matching the final goal orientation when nearby. A desired heading,  $\psi_{k+t}^{\text{head}}$ , is determined as the direction of the velocity vector. A sigmoid weighting function  $\sigma(\cdot)$ , based on the distance to the goal, smoothly interpolates between these two sub-objectives:

$$\lambda = \sigma\left(\frac{\|\hat{\mathbf{p}}_{k+t}^{xy} - \mathbf{p}_{\text{goal}}^{xy}\|_2 - d_{\text{switch}}}{s_{\text{switch}}}\right) \quad (19)$$

where  $d_{\text{switch}}$  and  $s_{\text{switch}}$  control the transition’s center and sharpness. The per-step yaw cost is a weighted average of the two alignment errors:

$$\ell_{\text{yaw}}(\hat{\mathbf{z}}_{k+t}) = \lambda|\hat{\psi}_{k+t}^{\text{yaw}} - \psi_{k+t}^{\text{head}}| + (1-\lambda)|\hat{\psi}_{k+t}^{\text{yaw}} - \psi_{\text{goal}}^{\text{yaw}}| \quad (20)$$

The total goal-tracking cost at a timestep is  $\ell_{\text{goal},t} = \ell_{\text{pos}} + \ell_{\text{yaw}}$ . To promote smooth motions, the control cost adds a quadratic penalty on the magnitude of the control inputs:

$$\ell_{\text{ctl}}(\mathbf{u}_{k+t}) = \|\mathbf{u}_{k+t}\|_2^2 \quad (21)$$

Finally, the total cost for a sampled trajectory is the weighted sum of these costs over the prediction horizon:

$$J^{(i)} = \sum_{t=1}^T \left[ w_c \ell_{\text{coll}}(\hat{\mathbf{z}}_{k+t}^{(i)}) + w_g \ell_{\text{goal},t}(\hat{\mathbf{z}}_{k+t}^{(i)}) + w_u \ell_{\text{ctl}}(\mathbf{u}_{k+t}^{(i)}) \right] \quad (22)$$

where the positive weights  $w_c$ ,  $w_g$ , and  $w_u$  balance safety, task performance, and control effort. This cost determines the weights used in the MPPI update rule (see Alg. 1).

## V. STABILITY ANALYSIS OF THE NEURAL OBSERVER

At this point, we have demonstrated the practical aspects of the proposed neural observer/predictor in goal-tracking legged navigation. Equally significant from a theoretical standpoint, we establish the conditions under which our neural observer is guaranteed to be UUB. This guarantee is crucial, as it formally ensures that the latent state estimate remains close to the true value, thereby providing a reliable foundation for prediction and planning.

Our analysis relies on a key common assumption about the true system dynamics and leverages the concept of Lipschitz continuity to handle the nonlinearities introduced by the neural network. Specifically, we assume the true closed-loop dynamics of the PSR can be represented by the observer’s learned structure, perturbed by a bounded disturbance term  $\epsilon_k$ . Specifically, the true latent state evolves according to:

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + g(\mathbf{x}_k, \mathbf{u}_k) + \epsilon_k \quad (23)$$

where the disturbance is bounded, i.e.,  $\|\epsilon_k\|_2 \leq \epsilon_{\max}$ . This assumption is justified because the parameters of  $\mathbf{A}$  and the MLP  $g(\cdot)$  are learned to capture the dominant system dynamics from the training data  $\mathcal{D}$ , while  $\epsilon_k$  provides robustness to the inevitable imperfections of the learned model and noise.

### A. Lipschitz Continuity of MLP

Our UUB guarantee requires bounding the effect of  $g(\cdot)$ , which a MLP network achieved with its Lipschitz constant.

**Definition V.1** (Lipschitz Continuity [28]). *A function  $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$  is globally  $L_g$ -Lipschitz continuous if there exists a constant  $L_g > 0$  such that for all  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ , the inequality  $\|g(\mathbf{x}) - g(\mathbf{y})\|_2 \leq L_g \|\mathbf{x} - \mathbf{y}\|_2$  holds.*

For a ReLU MLP as used in our model, the Lipschitz constant can be upper-bounded by the product of the spectral norms of its weight matrices based on the following lemma, which allows us to control  $L_g$  during training:

**Lemma V.2** (Lipschitz constant of ReLU MLP [22]). *Let  $g(\mathbf{x}) = \mathbf{W}^{N+1}\phi(\mathbf{W}^N(\dots\phi(\mathbf{W}^1\mathbf{x} + \mathbf{b}^1)\dots) + \mathbf{b}^N)$ , where each  $\mathbf{W}^n$  is a matrix,  $\mathbf{b}^n$  a bias vector, and  $\phi(\mathbf{x}) = \max\{\mathbf{x}, 0\}$  is applied element-wise. Then an upper bound on the Lipschitz constant of the ReLU MLP is  $L_g \leq \prod_{n=1}^{N+1} \|\mathbf{W}^n\|_2$ .*

*Proof:* See Appendix I.

### B. UUB Guarantee via Contraction Analysis

With these preliminaries, we now have the following key result, analytically establishing the UUB guarantee of the proposed neural observer based on a contraction condition:

**Theorem V.3** (UUB of Observer Error). *Consider the observer from Eq. (5) and assume the true system dynamics are given by (23) with a linear output  $\mathbf{y}_k = \mathbf{C}_y\mathbf{x}_k$ . Let the estimation error be  $\mathbf{e}_k = \mathbf{x}_k - \hat{\mathbf{x}}_k$ . Define the closed-loop error dynamics matrix as  $\mathbf{A}_c = \mathbf{A} - \mathbf{K}\mathbf{C}_y$ . If the spectral norm of  $\mathbf{A}_c$  and the Lipschitz constant  $L_g$  of the MLP  $g$  satisfy the contraction condition*

$$\rho := \|\mathbf{A}_c\|_2 + L_g < 1, \quad (24)$$

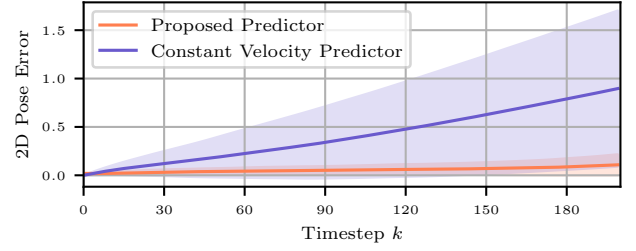


Fig. 6. **Long-Horizon Prediction Error.** Comparison of 2D pose prediction error over a 4-second horizon. Our proposed model (orange, mean with  $\pm 1\sigma$  region) maintains low error. The baseline CV model’s error (blue, mean with  $\pm 1\sigma$  region) accumulates rapidly, highlighting its unsuitability for long-horizon planning.

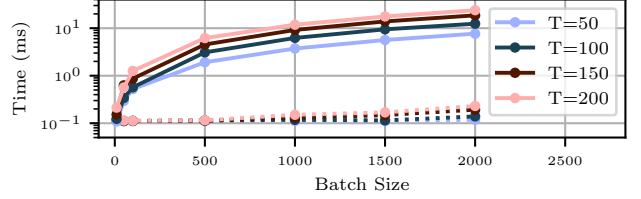


Fig. 7. **Inference Time Comparison.** Computation time versus batch size and prediction horizon for our model (solid lines) and the CV baseline (dashed lines) on an Nvidia RTX T1000 GPU. While more intensive, our model scales efficiently and remains well within real-time limits.

then both the state estimation error  $\mathbf{e}_k$  and the output estimation error  $\mathbf{e}_{y,k} = \mathbf{y}_k - \hat{\mathbf{y}}_k$  are Uniformly Ultimately Bounded. Specifically, their norms are ultimately bounded:

$$\limsup_{k \rightarrow \infty} \|\mathbf{e}_k\|_2 \leq \frac{\epsilon_{\max}}{1 - \rho} \quad \text{and} \quad \limsup_{k \rightarrow \infty} \|\mathbf{e}_{y,k}\|_2 \leq \|\mathbf{C}_y\|_2 \frac{\epsilon_{\max}}{1 - \rho}.$$

*Proof:* See Appendix II.

## VI. EXPERIMENTAL VALIDATION

In this section, we present an experimental validation of our proposed observer-predictor neural models. We first perform a comparative analysis, evaluating our model’s prediction accuracy and computational performance against a standard CV baseline. We then demonstrate the effectiveness of the complete system by integrating our model into the MPPI planner for a challenging goal-pose navigation task on the Vision 60 quadruped robot [3].

### A. Comparative Analysis with Baseline Model

The CV model predicts the future 2D pose given the current pose and commanded velocity:

$$\mathbf{q}_{k+1} = \mathbf{q}_k + \mathbf{R}_z(\psi_k^{\text{yaw}})\mathbf{u}_k\Delta t \quad (25)$$

where  $\mathbf{q} \in \mathbb{R}^3$  is the 2D planar pose. While this model provides a strong baseline, it assumes perfect velocity tracking, an idealization that often fails in practice due to complex robot dynamics and low-level control interactions. We evaluate our proposed model against this baseline for prediction accuracy and computational efficiency.

1) *Prediction Accuracy:* As shown in Fig. 6, our model’s prediction error remains low and bounded over a 4-second horizon. In contrast, the CV model’s error grows without bound. This confirms that our model successfully captures the PSR dynamics, making it far more reliable for use in a predictive planner.

2) *Computational Efficiency*: For real-time planning, computational efficiency is paramount. We benchmarked GPU-accelerated C++ Torch implementations of both models (see Fig. 7). While our GRU-based predictor is naturally more complex than the simple CV model, its inference time scales efficiently with batch size and horizon length. The performance is well within the required limits for real-time sampling-based planners like MPPI, demonstrating significant accuracy gain with a justifiable and manageable computational overhead.

### B. Integrated System Performance on Hardware

To demonstrate practical effectiveness, we integrate the learned predictor into an MPPI planner to perform navigation tasks as described in Sec. IV.

1) *Experimental Setup*: Experiments are conducted on the Vision 60 hardware. The robot’s state is estimated via a contact-aided InEKF [29]. An onboard VLP-16 LiDAR is used to build a voxel map [26] for collision checking. The MPPI planner operates at 25 Hz, using  $N = 1000$  samples over a  $T = 200$  step (4 s) horizon. With our predictor, a full planning cycle (including prediction and collision checks) takes approximately 35 ms on an Nvidia RTX T1000 GPU, well within the 40 ms real-time budget. For all trials, the MPPI constants are held fixed:  $d_{\text{switch}} = 0.5$ ,  $s_{\text{switch}} = 0.5$ ,  $w_c = 500$ ,  $w_g = 10$ , and  $w_u = 0.01$ . We compare the performance of MPPI using our predictor against the CV baseline over 20 trials for each task.

TABLE I  
MPPI PERFORMANCE OVER 20 TRIALS

Predictor Model	Time to Track (s)	Failure Rate (%)
Constant Velocity (Baseline)	$2.57 \pm 0.56$	40%
<b>Proposed Predictor (Ours)</b>	<b><math>2.45 \pm 0.37</math></b>	<b>15%</b>

2) *Goal-Pose Tracking Performance*: We first evaluate tracking a goal pose of [5.0 m, 2.5 m, 1.57 rad] in an open area. As shown in Fig. 8(a) and Table I, our predictor enables faster and more consistent tracking. The CV baseline suffers from poorer performance because it cannot account for complex behaviors, such as the low-level controller’s tendency to enter a “standing mode” at very low velocities, which our model learns from data. This leads to oscillations in the error plot for the baseline as it nears the goal.

3) *Collision Avoidance in a Narrow Passage*: Next, we task the robot with navigating a narrow passage (Fig. 8(b)), where the ability to predict the full-body configuration  $\mathbf{z}_k$ , including limb motion, is critical. Our predictor allows for limb-level collision checking with a robot occupancy predictor (see Fig. 5), whereas the CV baseline is limited to collision checks using a static, zero-configuration body model. For a fair comparison, robot occupancy predictions were not inflated for either model. The results in Table I show a huge improvement in failure rates from 40% to 15%. Failures with the CV model were often catastrophic collisions, while failures with our model were typically minor scrapes, highlighting its ability to generate safer, more aware motion plans.

4) *Navigation Over Obstacles*: Finally, we tested the full system’s ability to plan collision-free paths over small objects. As shown in Fig. 1 and Fig. 9, our predictor’s ability to forecast future limb positions enables the planner to successfully navigate over clutter. We observed that the efficacy of the plan is highly dependent on the quality of the onboard map and the drift of the state estimator [29].

## VII. CONCLUSIONS AND FUTURE WORK

In this work, we introduced a learning-based framework for full-body motion prediction to enable efficient, safe, collision-aware navigation for legged robots. Our approach features a decoupled observer-predictor neural architecture, where a neural observer with provable UUB guarantees provides a reliable latent state estimate to a computationally efficient recurrent predictor. This theoretical guarantee is critical, as it also facilitates stable end-to-end training of the proposed framework. We validated our method through hardware experiments on a Vision 60 quadruped. The results demonstrate that when integrated into an MPPI planner, our model significantly outperforms a kinematic baseline in both goal-pose tracking and collision avoidance. A key limitation of the current work is its focus on flat-ground locomotion; therefore, future work will be directed toward extending this framework to handle complex and uneven terrain.

### APPENDIX I PROOF OF LEM. V.2

For each layer of MLP  $g^n(\mathbf{x}) = \mathbf{W}^n \mathbf{x} + \mathbf{b}^n$ ,  $\|g^n(\mathbf{x}) - g^n(\mathbf{y})\|_2 = \|\mathbf{W}^n(\mathbf{x} - \mathbf{y})\|_2 \leq \|\mathbf{W}^n\|_2 \|\mathbf{x} - \mathbf{y}\|_2$ , we thus have  $L_{g^n} = \|\mathbf{W}^n\|_2$  (note that biases do not affect Lipschitz constants). The ReLU is 1-Lipschitz since, coordinate-wise,  $|\phi(\mathbf{u}_i) - \phi(\mathbf{v}_i)| \leq |\mathbf{u}_i - \mathbf{v}_i|$ , hence  $\|\phi(\mathbf{u}) - \phi(\mathbf{v})\|_2 \leq \|\mathbf{u} - \mathbf{v}\|_2$ . Writing  $g = g^{N+1} \circ \phi \circ g^N \circ \dots \circ \phi \circ g^1$  and noting that Lipschitz constants multiply under composition:  $L_{f \circ g} \leq L_f L_g$  we obtain:

$$L_g \leq \|\mathbf{W}^{N+1}\|_2 \prod_{n=1}^N (1 \cdot \|\mathbf{W}^n\|_2) = \prod_{n=1}^{N+1} \|\mathbf{W}^n\|_2.$$

### APPENDIX II PROOF OF THM. V.3

The dynamics of the estimation error  $\mathbf{e}_k = \mathbf{x}_k - \hat{\mathbf{x}}_k$  are derived by subtracting the observer dynamics from the true system dynamics:

$$\mathbf{e}_{k+1} = (\mathbf{A} - \mathbf{K}\mathbf{C}_y)\mathbf{e}_k + (g(\mathbf{x}_k, \mathbf{u}_k) - g(\hat{\mathbf{x}}_k, \mathbf{u}_k)) + \epsilon_k.$$

Letting  $\mathbf{A}_c = \mathbf{A} - \mathbf{K}\mathbf{C}_y$ , we take the 2-norm and apply the triangle inequality:

$$\begin{aligned} \|\mathbf{e}_{k+1}\|_2 &\leq \|\mathbf{A}_c \mathbf{e}_k\|_2 + \|g(\mathbf{x}_k, \mathbf{u}_k) - g(\hat{\mathbf{x}}_k, \mathbf{u}_k)\|_2 + \|\epsilon_k\|_2 \\ &\leq \|\mathbf{A}_c\|_2 \|\mathbf{e}_k\|_2 + L_g \|\mathbf{e}_k\|_2 + \epsilon_{\max} \\ &= (\|\mathbf{A}_c\|_2 + L_g) \|\mathbf{e}_k\|_2 + \epsilon_{\max}. \end{aligned}$$

With  $\rho = \|\mathbf{A}_c\|_2 + L_g < 1$ , the error norm dynamics are bounded by the scalar sequence  $\|\mathbf{e}_{k+1}\|_2 \leq \rho \|\mathbf{e}_k\|_2 + \epsilon_{\max}$ . By recursive application, this inequality shows the sequence

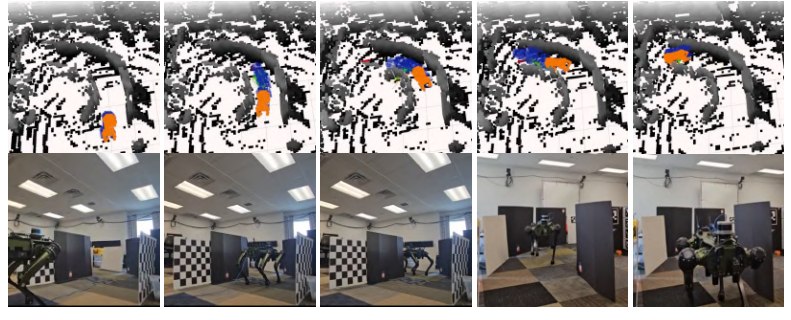
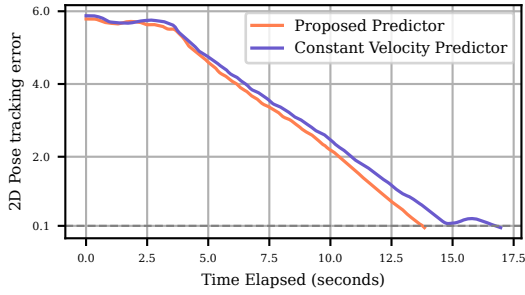


Fig. 8. **Real-Robot Experiments.** (a) Pose tracking error for a representative trial. MPPI with our predictor (orange) converges faster and more smoothly than with the CV baseline (blue), which exhibits oscillations due to unmodeled dynamics. (b) The Vision 60 robot successfully navigating a narrow passage using our full framework, which enables limb-aware, collision-free planning.

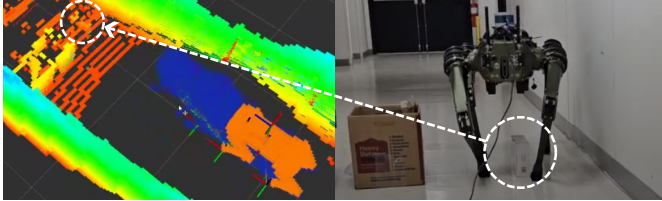


Fig. 9. Limb-aware predictions enables collision-free navigation over small objects.

is bounded, and taking the limit superior yields the ultimate bound for the state error:

$$\limsup_{k \rightarrow \infty} \|\mathbf{e}_k\|_2 \leq \frac{\epsilon_{\max}}{1 - \rho}.$$

We now examine the output error,  $\mathbf{e}_{y,k} = \mathbf{y}_k - \hat{\mathbf{y}}_k$ . By definition of the output map, we have  $\mathbf{e}_{y,k} = \mathbf{C}_y \mathbf{x}_k - \mathbf{C}_y \hat{\mathbf{x}}_k = \mathbf{C}_y \mathbf{e}_k$ . By taking the 2-norm we have:

$$\|\mathbf{e}_{y,k}\|_2 = \|\mathbf{C}_y \mathbf{e}_k\|_2 \leq \|\mathbf{C}_y\|_2 \|\mathbf{e}_k\|_2.$$

With established ultimate bound for  $\|\mathbf{e}_k\|_2$ , we can directly find the ultimate bound for the output error:

$$\limsup_{k \rightarrow \infty} \|\mathbf{e}_{y,k}\|_2 \leq \|\mathbf{C}_y\|_2 \left( \limsup_{k \rightarrow \infty} \|\mathbf{e}_k\|_2 \right) \leq \|\mathbf{C}_y\|_2 \frac{\epsilon_{\max}}{1 - \rho}.$$

#### REFERENCES

- [1] G. B. Margolis, G. Yang, K. Paigwar, T. Chen, and P. Agrawal, "Rapid locomotion via reinforcement learning," *Int. J. Robot. Research*, vol. 43, no. 4, pp. 572–587, 2024.
- [2] J. Norby, Y. Yang, A. Tajbakhsh, J. Ren, J. K. Yim, A. Stutt, Q. Yu, N. Flowers, and A. M. Johnson, "Quad-SDK: Full stack software framework for agile quadrupedal locomotion," in *ICRA Workshop on Legged Robots*, 2022.
- [3] Ghost Robotics Corp., "Vision60," 2025, www.ghostrobotics.io/vision-60.
- [4] Unitree Corp., "Unitree robots," 2025, www.unitree.com/.
- [5] S. Macenski, F. Martín, R. White, and J. G. Clavero, "The marathon 2: A navigation system," in *IEEE Int. Conf. Intell. Robots Syst.*, 2020, pp. 2718–2725.
- [6] J.-K. Huang and J. W. Grizzle, "Efficient anytime clf reactive planning system for a bipedal robot on undulating terrain," *IEEE Trans. Robot.*, vol. 39, no. 3, pp. 2093–2110, 2023.
- [7] K. S. Narkhede, D. A. Thanki, A. M. Kulkarni, and I. Poulakakis, "Overtaking moving obstacles with digit: Path following for bipedal robots via model predictive contouring control," in *IEEE Int. Conf. Intell. Robots Syst.*, 2023, pp. 8590–8596.
- [8] M. Mattamala, N. Chebrolu, and M. Fallon, "An efficient locally reactive controller for safe navigation in visual teach and repeat missions," *IEEE Robot. Autom. Lett.*, vol. 7, no. 2, pp. 2353–2360, 2022.
- [9] Y. Li, X. Tang, K. Chen, C. Zheng, H. Liu, and J. Ma, "Geometry-aware safety-critical local reactive controller for robot navigation in unknown and cluttered environments," *IEEE Robot. Autom. Lett.*, vol. 9, no. 4, pp. 3419–3426, 2024.
- [10] Q. Liao, Z. Li, A. Thirugnanam, J. Zeng, and K. Sreenath, "Walking in narrow spaces: Safety-critical locomotion control for quadrupedal robots with duality-based optimization," in *IEEE Int. Conf. Intell. Robots Syst.*, 2023, pp. 2723–2730.
- [11] M. Gaertner, M. Bjelonic, F. Farshidian, and M. Hutter, "Collision-free mpc for legged robots in static and dynamic scenes," in *IEEE Int. Conf. Robot. Autom.*, 2021, pp. 8266–8272.
- [12] A. Lin, S. Peng, and S. Bansal, "One filter to deploy them all: Robust safety for quadrupedal navigation in unknown environments." arXiv preprint arXiv:2412.09989, 2024.

- [13] I. Taouil, G. Turrisi, D. Schleich, V. Barasuol, C. Semini, and S. Behnke, "Quadrupedal footstep planning using learned motion models of a black-box controller," in *IEEE Int. Conf. Intell. Robots Syst.*, 2023, pp. 800–806.
- [14] J. Guzzi, R. O. Chavez-Garcia, M. Nava, L. M. Gambardella, and A. Giusti, "Path planning with local motion estimations," *IEEE Robot. Autom. Lett.*, vol. 5, no. 2, pp. 2586–2593, 2020.
- [15] Y. Kim, C. Kim, and J. Hwangbo, "Learning Forward Dynamics Model and Informed Trajectory Sampler for Safe Quadruped Navigation," in *Proc. Robot. Sci. Syst.*, New York City, NY, USA, 2022.
- [16] P. Roth, J. Frey, C. Cadena, and M. Hutter, "Learned perceptive forward dynamics model for safe and platform-aware robotic navigation." arXiv preprint arXiv:2504.19322, 2025.
- [17] H. Zhang, S. Starke, T. Komura, and J. Saito, "Mode-adaptive neural networks for quadruped motion control," *ACM Trans. Graph.*, vol. 37, no. 4, 2018.
- [18] J. Kim, T. Li, and S. Ha, "Armp: Autoregressive motion planning for quadruped locomotion and navigation in complex indoor environments," in *IEEE Int. Conf. Intell. Robots Syst.*, 2023, pp. 2731–2737.
- [19] F. Abdollahi, H. Talebi, and R. Patel, "A stable neural network-based observer with application to flexible-joint manipulators," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 17, no. 1, pp. 118–129, 2006.
- [20] A. Chakrabarty and M. Benosman, "Safe learning-based observers for unknown nonlinear systems using bayesian optimization," *Automatica*, vol. 133, p. 109860, 2021.
- [21] A. Jin, F. Zhang, G. Shen, B. Huang, and P. Huang, "Learning-based modeling and predictive control for unknown nonlinear system with stability guarantees," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 36, no. 6, pp. 11 135–11 148, 2025.
- [22] A. Virmaux and K. Scaman, "Lipschitz regularity of deep neural networks: analysis and efficient estimation," in *NIPS*, vol. 31. Curran Associates, Inc., 2018.
- [23] G. Shi, X. Shi, M. O'Connell, R. Yu, K. Azizzadenesheli, A. Anandkumar, Y. Yue, and S.-J. Chung, "Neural lander: Stable drone landing control using learned dynamics," in *IEEE Int. Conf. Robot. Autom.*, 2019, pp. 9784–9790.
- [24] G. Williams, P. Drews, B. Goldfain, J. M. Rehg, and E. A. Theodorou, "Information-theoretic model predictive control: Theory and applications to autonomous driving," *IEEE Trans. Robot.*, vol. 34, no. 6, pp. 1603–1622, 2018.
- [25] Z. Huai and G. Huang, "Robocentric visual-inertial odometry," *Int. J. Robot. Research*, Apr. 2019.
- [26] A. Hornung, K. M. Wurm, M. Bennewitz, C. Stachniss, and W. Burgard, "Octomap: an efficient probabilistic 3d mapping framework based on octrees," *Auton. Robots*, vol. 34, no. 3, p. 189–206, 2013.
- [27] J. Higgins, N. Mohammad, and N. Bezzo, "A model predictive path integral method for fast, proactive, and uncertainty-aware uav planning in cluttered environments," in *IEEE Int. Conf. Intell. Robots Syst.*, 2023, pp. 830–837.
- [28] H. K. Khalil, *Nonlinear systems*. Upper Saddle River, N.J.: Prentice Hall, 2002.
- [29] R. Hartley, M. Ghaffari, R. M. Eustice, and J. W. Grizzle, "Contact-aided invariant extended kalman filtering for robot state estimation," *Int. J. Robot. Research*, vol. 39, no. 4, pp. 402–430, 2020.

#### ACKNOWLEDGMENT

This work was partially funded by the the European Union under Horizon Europe (grant No.101136568-HERON).

