

# Variable-Frequency Model Learning and Predictive Control for Jumping Maneuvers on Legged Robots

Chuong Nguyen, Abdullah Altawaitan, Thai Duong, Nikolay Atanasov, and Quan Nguyen

**Abstract**—Achieving both target accuracy and robustness in dynamic maneuvers with long flight phases, such as high or long jumps, has been a significant challenge for legged robots. To address this challenge, we propose a novel learning-based control approach consisting of model learning and model predictive control (MPC) utilizing a variable-frequency scheme. Compared to existing MPC techniques, we learn a model directly from experiments, accounting not only for leg dynamics but also for modeling errors and unknown dynamics mismatch in hardware and during contact. Additionally, learning the model with variable-frequency allows us to cover the entire flight phase and final jumping target, enhancing the prediction accuracy of the jumping trajectory. Using the learned model, we also design variable-frequency to effectively leverage different jumping phases and track the target accurately. In a total of 92 jumps on Unitree A1 robot hardware, we verify that our approach outperforms other MPCs using fixed-frequency or nominal model, reducing the jumping distance error 2 – 8 times. We also achieve jumping distance errors of less than 3% during continuous jumping on uneven terrain with randomly-placed perturbations of random heights (up to 4 cm or 27% the robot's standing height). Our approach obtains distance errors of 1 – 2 cm on 34 single and continuous jumps with different jumping targets and model uncertainties. Code is available at [https://github.com/DRCL-USC/Learning\\_MPC\\_Jumping](https://github.com/DRCL-USC/Learning_MPC_Jumping).

**Index Terms**—Model Learning for Control, Legged Robots, Whole-Body Motion Planning and Control

## I. INTRODUCTION

AGGRESSIVE jumping maneuvers with legged robots have received significant attention recently, and have been successfully demonstrated using trajectory optimization [1]–[4], model-based control [5]–[10], and learning-based control [11]–[16]. Unlike walking or running, aggressive motions are particularly challenging due to (1) the extreme underactuation in the mid-air phase (the robot mainly relies on the force control during contact to regulate its global position and orientation), (2) significant dynamics model error and uncertainty that are inherently hard to model accurately, especially with contact and hardware load during extreme motions,

Manuscript received: July 18, 2024; Revised October 20, 2024; Accepted November 19, 2024.

This paper was recommended for publication by Editor Jaydev P. Desai upon evaluation of the Associate Editor and Reviewers' comments.

Chuong Nguyen and Quan Nguyen are with the Department of Aerospace and Mechanical Engineering, University of Southern California, Los Angeles, CA 90007, USA, e-mails: {vanchuong.nguyen, quann}@usc.edu.

Abdullah Altawaitan, Thai Duong, and Nikolay Atanasov are with the Department of Electrical and Computer Engineering, University of California San Diego, La Jolla, CA 92093, USA, e-mails: {aaltawaitan, tduong, natanasov}@ucsd.edu. A. Altawaitan is also affiliated with Kuwait University as a holder of a scholarship.

Digital Object Identifier (DOI): see top of this page.

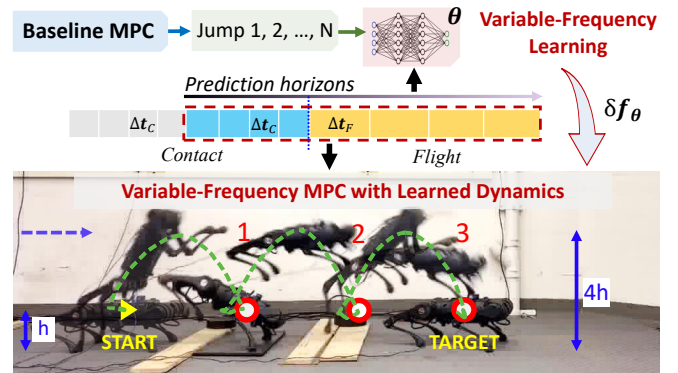


Fig. 1: A Unitree A1 robot performs *continuous jumps* on unknown uneven terrain, achieving both target accuracy and robustness. The target distance for each jump is 0.6 m. The flight phase covers a vertical height of up to 4× the robot's normal height. MPC with a nominal single rigid-body model is used to collect data for training a neural-network residual dynamics model following using a variable-frequency scheme. The learned model is then used in a variable-frequency MPC to execute jumping motions. The variable-frequency scheme varies the time step sizes in the contact phase ( $\Delta t_c$ ) and flight phase ( $\Delta t_f$ ) to dedicate more model capacity and more MPC optimization steps to the contact phase, which improves the jumping robustness and accuracy. The green dashed line is the actual robot trajectory. Supplemental video: [https://youtu.be/yUqJ\\_MBOC6Q](https://youtu.be/yUqJ_MBOC6Q).

and (3) the trade-off between model accuracy and efficient computation for real-time execution. Achieving both target accuracy and robustness for long-flight maneuvers, such as high or long jumps, therefore still presents an open challenge. In this work, we address this challenge by developing a real-time MPC for quadruped jumping using a robot dynamics model learned from experiments.

Many control and optimization techniques have been developed for jumping motions. Trajectory optimization (TO) with full-body nonlinear dynamics is normally utilized to generate long-flight trajectories in an offline fashion (e.g., [1], [2]). Many MPC approaches sacrifice model accuracy to achieve robust maneuvers by using simplified models that treat the trunk and legs as a unified body [5], [7]–[9], [17]. Our recent iterative learning control (ILC) work [18] handles model uncertainty to realize long-flight jumps via multi-stage optimization that optimizes the control policy after each jump offline until reaching a target accurately. It also relies on a simplified model for computational efficiency. However, this work focuses on target jumping rather than robustness, e.g. requiring the same initial condition for all trials. Different from existing works, we learn a robot dynamics model from experiments and develop real-time MPC using the learned dynamics to achieve both target accuracy and robustness in continuous quadruped jumping.

Learning robot models from experiments to capture complex dynamics effects and account for model errors with real hardware has become a popular approach [19]–[22]. Many frameworks learn residual dynamics using neural networks [19], [23] or Gaussian processes [20], [24] and use MPC to control the learned systems [23], [25] to improve tracking errors. Most existing works, however, primarily investigate dynamical systems without contact. Recently, Sun et al. [26] proposed a notable method to learn a linear residual model, addressing unknown dynamics during walking stabilization. Pandala et al. [27] propose to close the gap between reduced- and full-order models by using deep reinforcement learning to learn unmodeled dynamics, which are used in MPC to realize robust walking on various terrains. Meanwhile, we use supervised learning to learn unmodelled dynamics with real data from hardware. Also, we learn dynamics for maneuvers with long-flight periods to tackle (1) the switching between multiple dynamic models due to contact, including flight phase where control action has very little effect on the body dynamics, (2) disturbances and uncertainty in dynamic modeling due to hard impact in jumping, and (3) the effect of intermittent control at contact transitions on state predictions.

In this letter, we propose a residual learning model that uses a variable-frequency scheme, i.e., varying the coarseness of integration time step, to address the aforementioned challenges and enhance long-term trajectory predictions over different jumping phases. MPC has been commonly used for jumping by optimizing the control inputs during the contact phase based on future state prediction during the flight phase (e.g., [5], [7], [8], [10], [28]). A major challenge in MPC for long flight maneuvers is to utilize a limited number of prediction steps yet still effectively cover the entire flight phase and especially the final jumping target. Another challenge is to obtain an accurate model for complex dynamic maneuvers, unknown dynamics, and model mismatch with the real hardware to improve the jumping accuracy, while still ensuring real-time performance. Some recent works have addressed these challenges partially. Many methods use conventional single-rigid-body dynamic (SRBD) models, ignoring the leg dynamics, to achieve real-time execution [5]–[7]. Using the SRBD model in the contact phase can lead to inaccurate predictions of the robot's state at take-off, which is the initial condition of the projectile motion in the flight phase. Thus, the trajectory prediction error can accumulate significantly over long flight periods. Some methods account for leg inertia [10], [29], [30], however, disturbance and uncertainty in dynamic modeling have not been considered. Recently, planning with multi-fidelity models and multi-resolution time steps has emerged as an effective strategy to enhance target accuracy and robustness [31]–[33]. Li et al. [33] adopt a less accurate model in the far horizon and a more accurate but expensive model in the near future. Norby et al. [32] adapt model complexity based on the task complexity along the prediction horizon. Li et al. [33] also design multi-resolution time steps to cover the whole flight phase. While we vary the coarseness of the time step as [33], we design a novel residual model learning approach combined with variable-frequency MPC to address the aforementioned challenges.

**Contributions:** The contributions of this letter are summarized as follows.

- We learn a *residual dynamics model* directly from a small real experiment dataset. The model accounts for nonlinear leg dynamics, modeling errors, and unknown dynamics mismatch in hardware and during contact.
- We propose learning the model in an *variable-frequency scheme* that leverages different time resolutions to capture the entire flight phase, the jumping target, and the contact transitions over a few-step horizon, thereby significantly improving the accuracy of long-term trajectory prediction.
- We develop *variable-frequency MPC using the learned model* to synthesize controls that improve both target accuracy and robustness in dynamic robot maneuvers.
- *Extensive hardware experiments* validate the effectiveness and robustness of our approach with single and consecutive jumps on uneven terrains. Comparisons with other MPC techniques using a nominal model or fixed-frequency are also provided.

## II. PROBLEM STATEMENT

Consider a legged robot modelled by state  $\mathbf{x}_k$ , consisting of the pose and velocity of the body's center of mass (CoM), foot positions  $\mathbf{r}_k$  relative to the body's CoM, and ground force control input  $\mathbf{u}_k$  at the legs, sampled at time  $t_k$ . We augment the nominal SRBD model (e.g., [5], [7], [8]) with a learned residual term to account for leg dynamics, model mismatch with the real robot hardware, as well as capture complex dynamic effects in contact and hardware:

$$\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k, \mathbf{r}_k, \Delta t_k) + \delta \mathbf{f}_\theta(\mathbf{x}_k, \mathbf{u}_k, \mathbf{r}_k) \Delta t_k, \quad (1)$$

where  $\mathbf{f}$  represents the nominal SRBD model,  $\delta \mathbf{f}_\theta$  with parameters  $\theta$  approximates the residual dynamics, and  $\Delta t_k$  is the sampling interval. Our first objective is to learn the residual dynamics  $\mathbf{f}_\theta$  using a dataset of jumping trajectories.

*Problem 1:* Given a set  $\mathcal{D} = \{\mathbf{x}_{0:K}^{(i)}, \mathbf{x}_{0:K}^{(i)}, \mathbf{r}_{0:K}^{(i)}, \mathbf{u}_{0:K}^{(i)}\}_{i=1}^D$  of  $D$  sequences of states, foot positions, and control inputs, find the parameters  $\theta$  of the residual dynamics  $\delta \mathbf{f}_\theta$  in (1) by rolling out the dynamics model and minimizing the discrepancy between the predicted state sequence  $\tilde{\mathbf{x}}_{1:K}^{(i)}$  and the true state sequence  $\mathbf{x}_{1:K}^{(i)}$  in  $\mathcal{D}$ :

$$\begin{aligned} \min_{\theta} \quad & \sum_{i=1}^D \sum_{k=1}^K \mathcal{L}(\mathbf{x}_k^{(i)}, \tilde{\mathbf{x}}_k^{(i)}) + \mathcal{L}_{\text{reg}}(\theta) \\ \text{s.t.} \quad & \tilde{\mathbf{x}}_{k+1}^{(i)} = \mathbf{f}(\tilde{\mathbf{x}}_k^{(i)}, \mathbf{u}_k^{(i)}, \mathbf{r}_k^{(i)}, \Delta t_k) \\ & \quad + \delta \mathbf{f}_\theta(\tilde{\mathbf{x}}_k^{(i)}, \mathbf{u}_k^{(i)}, \mathbf{r}_k^{(i)}) \Delta t_k, \\ & \tilde{\mathbf{x}}_0^{(i)} = \mathbf{x}_0^{(i)}, \quad \forall i = 1, \dots, D, \end{aligned} \quad (2)$$

where  $\mathcal{L}$  is an error function in the state space, and  $\mathcal{L}_{\text{reg}}(\theta)$  is a regularization term to avoid overfitting. Note that it is not necessary for  $\Delta t_k = t_{k+1} - t_k$  to be fixed.

After learning to improve the accuracy of model prediction, our second objective is to use the learned model (1) for MPC to track a desired state trajectory  $\mathbf{x}_{0:K}^*$ .

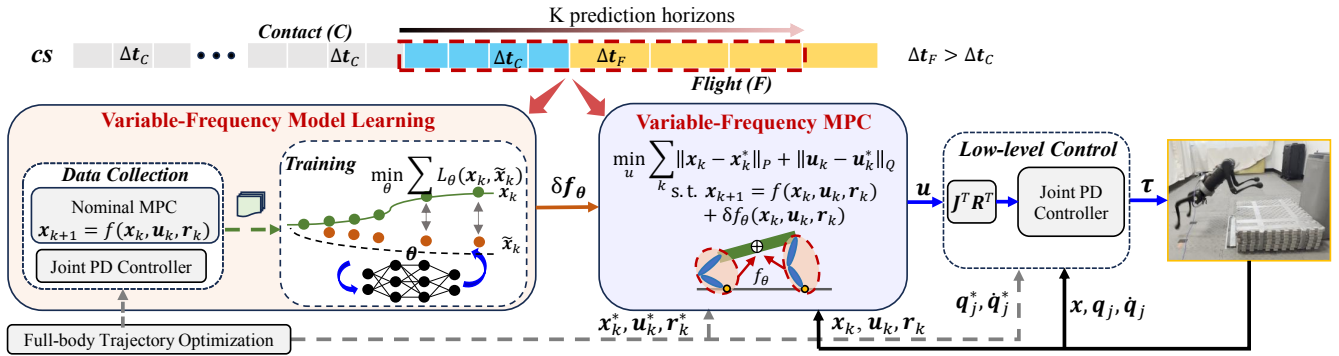


Fig. 2: **System Architecture.** The learning procedure and MPC execution are paired with the same integration timesteps  $\{\Delta t_C, \Delta t_F\}$ , prediction horizon  $K$ , and predefined contact schedule ( $cs$ ). The MPC and low-level joint PD controllers are updated at  $40Hz$  and  $1kHz$ , respectively.

**Problem 2:** Given the dynamics model (1), a current robot state  $\mathbf{x}_0$  and foot positions  $\mathbf{r}_0$ , a desired trajectory of states  $\mathbf{x}_{0:K}^*$ , foot positions  $\mathbf{r}_{0:K}^*$ , and control  $\mathbf{u}_{0:K}^*$ , design a control law  $\mathbf{u}_0 = \pi(\mathbf{x}_0, \mathbf{r}_0, \mathbf{x}_{0:K}^*, \mathbf{r}_{0:K}^*, \mathbf{u}_{0:K}^*; \theta)$  to achieve accurate and robust tracking of the desired state trajectory via shifting-horizon optimization:

$$\begin{aligned} \min_{\mathbf{u}_{0:K-1}} & \sum_{k=1}^K \|\mathbf{x}_k - \mathbf{x}_k^*\|_{\mathbf{Q}_k} + \|\mathbf{u}_{k-1} - \mathbf{u}_{k-1}^*\|_{\mathbf{R}_k} \\ \text{s.t. } & \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k, \mathbf{r}_k, \Delta t_k) + \delta \mathbf{f}_\theta(\mathbf{x}_k, \mathbf{u}_k, \mathbf{r}_k) \Delta t_k, \\ & \mathbf{u}_k \in \mathcal{U}_k, \quad \forall k = 0, \dots, K-1, \end{aligned} \quad (3)$$

where  $\mathbf{Q}_k$  and  $\mathbf{R}_k$  are positive definite weight matrices, and  $\mathcal{U}_{0:K-1}$  represents an input constraint set.

Achieving accurate target jumping with MPC requires not only an accurate model but also the coverage of the final state upon landing, which represents the accuracy of the jumping target. Using a fixed-frequency has shown to be efficient for locomotion tasks with limited aerial phases; however, this can face two challenges in long-flight maneuvers. On the one hand, using a fine time-step discretization enhances the model prediction accuracy but requires a large number of steps to capture the entire flight phase, thereby increasing the optimization size and computational cost. On the other hand, using a coarse time-step discretization allows capturing the entire flight phase efficiently but can sacrifice model prediction accuracy. For jumping tasks, different phases may require different model resolutions, e.g., fine time resolution during the contact phase but coarser time resolution during the flight phase as the model complexity is reduced due to unavailable force control and contact. Therefore, we propose to learn a model for dynamic maneuvers (Problem 1) with an variable-frequency scheme that uses a coarse time discretization in the flight phase and a fine time discretization in the contact phase. Importantly, this variable-frequency scheme is also synchronized with the MPC control of the learned model (Problem 2) by utilizing the same time steps for both fitting the learned model and performing the MPC optimization. This synchronization ensures that the same discretization errors are leveraged when using the model for MPC predictions. Thus, in our formulation  $t_{0:K}$  does not have equal time steps and is a mixture of fine and coarse discretization capturing the contact and flight phases. Note that we use a constant time step  $\Delta t_k$  during the different jumping phases. It is possible

to use  $\Delta t_k$  as input to the residual model neural networks. However, we need to collect a substantial amount of data for a wide range of  $\Delta t_k$  to avoid overfitting, which would also increase the size of the neural network and computational cost.

### III. SYSTEM OVERVIEW

Fig. 2 presents an overview of our system architecture. Our approach consists of two stages: variable-frequency model learning (Sec. IV) and variable-frequency MPC (Sec. V), which solve Problem 1 and Problem 2, respectively.

We synchronize the variable-frequency scheme for both model learning and MPC execution with the same (1) variable prediction timesteps  $\Delta t_C$  and  $\Delta t_F$  for contact and flight phase, respectively, (2) the same horizon length  $K$  for all data collection, training, and MPC, and (3) the same contact schedule. Full-body trajectory optimization (TO) [1] is utilized to generate jumping references for various targets, including body states  $\mathbf{x}^*$ , joint states  $\{\mathbf{q}_J^*, \dot{\mathbf{q}}_J^*\}$ , ground contact force  $\mathbf{u}^*$ , and foot positions  $\mathbf{r}^*$ . For data collection, we combine a baseline MPC using a nominal SRBD model and joint PD controller, generating diverse motions under disturbances. For training, we design a neural network to learn the discretized residual model  $\delta \mathbf{f}_\theta$  with variable sampling time via supervised learning. For control, we design an variable-frequency MPC using the learned dynamics to track a desired reference trajectory obtained from full-body TO. The feedback states from the robot include global body's CoM  $\mathbf{x}$ , and joint states  $\mathbf{q}_J, \dot{\mathbf{q}}_J \in \mathbb{R}^4$ , and foot positions  $\mathbf{r} \in \mathbb{R}^4$ .

### IV. VARIABLE-FREQUENCY MODEL LEARNING

In this section, we describe how to learn the residual dynamics  $\delta \mathbf{f}_\theta$  from data with a variable-frequency scheme that can cover the entire flight phase, the final state upon landing, and the contact transitions between jumping phases.

#### A. Learning Dynamics with Variable-Frequency

We consider a 2D jumping motion on a legged robot with point foot contact, e.g., quadruped robot, with generalized coordinates  $\mathbf{q} = [\mathbf{p}^\top \phi]^\top \in \mathbb{R}^3$ , where  $\mathbf{p} \in \mathbb{R}^2$  is the CoM's position, and  $\phi \in \mathbb{R}$  is the body pitch angle. We define the generalized robot's velocity as  $\zeta = [\mathbf{v}^\top \omega]^\top \in \mathbb{R}^3$ , where  $\mathbf{v} \in \mathbb{R}^2$  and  $\omega \in \mathbb{R}$  are the linear and angular



velocity. Both  $\mathbf{q}$  and  $\zeta$  are expressed in world-frame coordinates. The robot state is  $\mathbf{x} = [\mathbf{q}^\top \ \zeta^\top \ g]^\top \in \mathbb{R}^7$ , where the (constant) gravity acceleration  $g$  is added to obtain a convenient state-space form [34]. We define  $\mathbf{R}(\phi) = [\cos(\phi) \ -\sin(\phi); \sin(\phi) \ \cos(\phi)] \in \mathbb{R}^{2 \times 2}$  as a rotation matrix of the main body, which converts  $\mathbf{r}_{i,b}$  (a relative foot  $i \in \{1, 2\}$  position relative to the body's CoM in the body frame) to the world frame  $\mathbf{r}_i = \mathbf{R}\mathbf{r}_{i,b}$ . We denote  $\mathbf{r} = [\mathbf{r}_1^\top \ \mathbf{r}_2^\top]^\top \in \mathbb{R}^4$ . With the force control input for the front and rear legs as  $\mathbf{u} = [\mathbf{u}_f^\top \ \mathbf{u}_r^\top]^\top \in \mathbb{R}^4$ , the nominal discrete-time SRBD model can be written as:

$$\mathbf{f}(\mathbf{x}_k, \mathbf{u}_k, \mathbf{r}_k, \Delta t_k) = \mathbf{A}_k \mathbf{x}_k + \mathbf{B}_k(\mathbf{r}_k) \mathbf{u}_k, \quad (4)$$

where  $\mathbf{A}_k = \mathbf{I}_7 + \mathbf{A}_{ct} \Delta t_k$ ,  $\mathbf{B}_k(\mathbf{r}_k) = \mathbf{B}_{ct}(\mathbf{r}_k) \Delta t_k$ ,  $\Delta t_k$  is the time step (e.g.,  $\Delta t_C$  or  $\Delta t_{\mathcal{F}}$  for the contact or flight phase, respectively), and  $\mathbf{A}_{ct}$  and  $\mathbf{B}_{ct}$  are obtained from the continuous-time robot dynamics:

$$\mathbf{A}_{ct} = \begin{bmatrix} \mathbf{0} & \mathbf{I}_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{e}_3 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}, \mathbf{B}_{ct}(\mathbf{r}) = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{I}_2/m & \mathbf{I}_2/m \\ [\mathbf{r}_1]_{\times}^\top / J & [\mathbf{r}_2]_{\times}^\top / J \\ \mathbf{0} & \mathbf{0} \end{bmatrix},$$

where  $\mathbf{A}_{ct} \in \mathbb{R}^{7 \times 7}$ ,  $\mathbf{B}_{ct} \in \mathbb{R}^{7 \times 4}$ ,  $\mathbf{e}_3 = [0 \ -1 \ 0]^\top$ ,  $m$  and  $J$  are mass and moment inertial of the body,  $[\mathbf{r}_i]_{\times} = [r_{iz} \ -r_{ix}]^\top \in \mathbb{R}^2$ . The residual term  $\delta \mathbf{f}_\theta(\cdot)$  in (1) is

$$\delta \mathbf{f}_\theta(\cdot) = \mathbf{h}_\theta(\mathbf{x}_k, \mathbf{r}_k) \Delta t_k + \mathbf{G}_\theta(\mathbf{x}_k, \mathbf{r}_k) \Delta t_k \mathbf{u}_k, \quad (5)$$

where  $\mathbf{h}_\theta(\mathbf{x}_k, \mathbf{r}_k)$  and  $\mathbf{G}_\theta(\mathbf{x}_k, \mathbf{r}_k)$  are represented by neural networks with learning parameters  $\theta$ . Since  $\mathbf{u}_k = \mathbf{0}$  during the flight phase,  $\mathbf{G}_\theta \mathbf{u}_k = \mathbf{0}$ . We have two separate models for the contact ( $\mathcal{C}$ ) and flight phase ( $\mathcal{F}$ ):

$$(\mathcal{C}) : \mathbf{x}_{k+1} = \mathbf{A}_C \mathbf{x}_k + \mathbf{B}_C(\mathbf{r}_k) \mathbf{u}_k + \mathbf{h}_{\theta_1}(\mathbf{x}_k, \mathbf{r}_k) \Delta t_C + \mathbf{G}_{\theta_1}(\mathbf{x}_k, \mathbf{r}_k) \Delta t_C \mathbf{u}_k, \quad (6a)$$

$$(\mathcal{F}) : \mathbf{x}_{k+1} = \mathbf{A}_F \mathbf{x}_k + \mathbf{h}_{\theta_2}(\mathbf{x}_k, \mathbf{r}_k) \Delta t_{\mathcal{F}} \quad (6b)$$

where  $\mathbf{A}_C = \mathbf{I}_7 + \mathbf{A}_{ct} \Delta t_C$ ,  $\mathbf{B}_C(\mathbf{r}_k) = \mathbf{B}_{ct}(\mathbf{r}_k) \Delta t_C$ , and  $\mathbf{A}_F = \mathbf{I}_7 + \mathbf{A}_{ct} \Delta t_{\mathcal{F}}$ . We roll out the dynamics based on (6), starting from initial state  $\mathbf{x}_0$  with given control input sequence  $\mathbf{u}_{0:K-1}$  to obtain a predicted state sequence  $\tilde{\mathbf{x}}_{1:K}$ . Using the variable-frequency scheme, the state prediction accounts for contact transitions (feet taking off the ground), a long flight phase, and the final robot state upon landing. We define the loss functions in Problem 1 as follows:

$$\begin{aligned} \mathcal{L}(\mathbf{x}, \tilde{\mathbf{x}}) &= \|\mathbf{q} - \tilde{\mathbf{q}}\|_2^2 + \|\zeta - \tilde{\zeta}\|_2^2 \\ \mathcal{L}_{reg}(\theta) &= \alpha_1 \|\mathbf{h}_\theta\| + \alpha_2 \|\mathbf{G}_\theta\|. \end{aligned} \quad (7)$$

The parameters  $\theta = [\theta_1, \theta_2]$  for each phase are updated by gradient descent to minimize the total loss.

### B. Data Collection

For model learning, we directly collect state-control trajectories from hardware experiments by implementing an MPC controller with the nominal dynamics model (4) and a reference body trajectory  $\mathbf{x}^*$  obtained from full-body TO. The TO assumes jumping from flat and hard ground with point foot contact. We generated various jumps to different

targets under different disturbances (e.g., blocks under the robot feet with random height) to obtain a diverse dataset.

While the MPC aims to track the body reference trajectory  $\mathbf{x}_{0:K}^*$ , a joint PD controller is used to track the joint reference trajectory  $(\mathbf{q}_J^*, \dot{\mathbf{q}}_J^*)$  from the full-body TO via  $\tau_{pd, setpoint} = \mathbf{K}_p(\mathbf{q}_J^* - \mathbf{q}_J) + \mathbf{K}_d(\dot{\mathbf{q}}_J^* - \dot{\mathbf{q}}_J)$ . Thus, the evolution of the robot states is governed by the combination of MPC and the joint PD controller. We collected the trajectory dataset  $\mathcal{D}$  with inputs  $\mathbf{u} = \mathbf{u}_{mpc} + \mathbf{u}_{pd}$ , where  $\mathbf{u}_{pd} = (\mathbf{J}(\mathbf{q}_J)^\top \mathbf{R}^\top)^{-1} \tau_{pd, setpoint}$ ,  $\mathbf{J}(\mathbf{q}_J)$  is the foot Jacobian..

The dataset is collected at the different time steps for the contact and flight phases, i.e.,  $\Delta t_C$  and  $\Delta t_{\mathcal{F}}$ , respectively. The data for each jump is then chunked by shifting a sliding window, size of  $K$  by 1 timesteps. Let  $N$  be the number of collected state-control data points for each jump and  $H$  be the number of jumps. We then obtain  $D = H \times (N - K + 1)$  state-control trajectories in total.

## V. VARIABLE-FREQUENCY MPC WITH LEARNED DYNAMICS

In this section, we design a variable-frequency MPC controller for the learned dynamics (6) to track a desired jumping reference trajectory obtained from full-body TO. For a given robot state  $\mathbf{x}_0$ , we formulate MPC as:

$$\min_{\mathbf{u}_{0:K-1}} \sum_{k=1}^K \|\mathbf{x}_k - \mathbf{x}_k^*\|_{\mathbf{Q}_k} + \|\mathbf{u}_{k-1} - \mathbf{u}_{k-1}^*\|_{\mathbf{R}_k}, \quad (8a)$$

$$\text{s.t. } (\mathcal{C}) : \mathbf{x}_{k+1} = \mathbf{A}_C \mathbf{x}_k + \mathbf{B}_C(\mathbf{r}_k) \mathbf{u}_k + \mathbf{h}_{\theta_1}(\mathbf{x}_k, \mathbf{r}_k) \Delta t_C + \mathbf{G}_{\theta_1}(\mathbf{x}_k, \mathbf{r}_k) \Delta t_C \mathbf{u}_k, \quad (8b)$$

$$(\mathcal{F}) : \mathbf{x}_{k+1} = \mathbf{A}_F \mathbf{x}_k + \mathbf{h}_{\theta_2}(\mathbf{x}_k, \mathbf{r}_k) \Delta t_{\mathcal{F}}, \quad (8c)$$

$$\underline{\mathbf{c}}_k \leq \mathbf{C}_k \mathbf{u}_k \leq \bar{\mathbf{c}}_k, \quad \forall k = 0, \dots, K-1, \quad (8d)$$

$$\mathbf{D}_k \mathbf{u}_k = \mathbf{0}, \quad \forall k = 0, \dots, K-1, \quad (8e)$$

where (8d) represents input constraints related to friction cone and force limits and (8e) aims to nullify the force on the swing legs based on the contact schedule.

With the MPC horizon including  $K_C$  steps in the contact phase and  $K_{\mathcal{F}} = K - K_C$  steps in the flight phase, we define  $\tilde{\mathbf{U}}_C = [\tilde{\mathbf{u}}_0^\top, \tilde{\mathbf{u}}_1^\top, \dots, \tilde{\mathbf{u}}_{K_C-1}^\top]^\top \in \mathbb{R}^{K_C \times 4}$  as a concatenation of the control inputs in the contact phase. With this notation, the predicted trajectory is:

$$(\mathcal{C}) : \tilde{\mathbf{X}}_C = \mathcal{A}_C \mathbf{x}_0 + (\mathcal{B}_C + \mathcal{B}_C \theta) \tilde{\mathbf{U}}_C + \mathbf{H}_C \theta \quad (9a)$$

$$(\mathcal{F}) : \tilde{\mathbf{X}}_{\mathcal{F}} = \mathcal{A}_F \tilde{\mathbf{x}}_{K_C} + \mathbf{H}_{\mathcal{F}} \theta \quad (9b)$$

where  $\tilde{\mathbf{X}}_C = [\tilde{\mathbf{x}}_1^\top, \tilde{\mathbf{x}}_2^\top, \dots, \tilde{\mathbf{x}}_{K_C}^\top]^\top$  and  $\tilde{\mathbf{X}}_{\mathcal{F}} = [\tilde{\mathbf{x}}_{K_C+1}^\top, \tilde{\mathbf{x}}_{K_C+2}^\top, \dots, \tilde{\mathbf{x}}_K^\top]^\top$  denote the concatenation of predicted states belonging to the contact and flight phase, respectively. The matrices  $\mathcal{A}_C$  and  $\mathcal{A}_F$  are computed as

$$\begin{cases} \mathcal{A}_C = [(\mathbf{A}_C)^\top (\mathbf{A}_C^2)^\top \dots (\mathbf{A}_C^{K_C})^\top] \in \mathbb{R}^{7K_C \times 7}, \\ \mathcal{A}_F = [(\mathbf{A}_F)^\top (\mathbf{A}_F^2)^\top \dots (\mathbf{A}_F^{K_{\mathcal{F}}})^\top] \in \mathbb{R}^{7K_{\mathcal{F}} \times 7}. \end{cases}$$

While we use the current robot state for training the neural networks  $\{\mathbf{G}_\theta, \mathbf{h}_\theta\}$  as presented in Sec. IV, we utilize the reference trajectory  $(\mathbf{x}_k^*, \mathbf{r}_k^*)$  in future MPC horizons for guiding the MPC since the future robot states are not available at each MPC update ([5]–[9]). The reference will

be used to compute the output of the neural networks  $\mathbf{G}_\theta$  and  $\mathbf{h}_\theta$  in future horizons. In particular, we still use the current robot state and foot position  $(\mathbf{x}_0, \mathbf{r}_0)$  to evaluate the residual term  $\mathbf{B}_{\theta,0} = \mathbf{G}_\theta(\mathbf{x}_0, \mathbf{r}_0)\Delta t_0$  for each MPC update, then define  $\mathbf{B}_{\theta,k} = \mathbf{G}_\theta(\mathbf{x}_k^*, \mathbf{r}_k^*)\Delta t_k$ ,  $k = 1, \dots, K_C$ . In (9a), we obtain  $[\mathcal{B}_C]_{\{m,n\}} = \mathbf{A}_C^{m-n}\mathbf{B}_{n-1}$  and its residual  $[\mathcal{B}_{C,\theta}]_{\{m,n\}} = \mathbf{A}_C^{m-n}\mathbf{B}_{\theta,n-1}$  if  $m \geq n$ , and  $\mathbf{0}$  otherwise.

For the sake of notation, we define  $\chi_k = (\mathbf{x}_k, \mathbf{r}_k)$ ,  $\chi_k^* = (\mathbf{x}_k^*, \mathbf{r}_k^*)$  as the actual and reference robot state and foot trajectory. The residual matrix  $\mathbf{H}_{C,\theta}$  in (9) is  $\mathbf{H}_{C,\theta} = \Delta t_C \mathbf{S}_C [\mathbf{h}_\theta(\chi_0)^\top \mathbf{h}_\theta(\chi_1^*)^\top \dots \mathbf{h}_\theta(\chi_{K_C-1}^*)^\top]^\top$ , where  $\mathbf{S}_C$  is lower triangular matrix with each entry  $[\mathbf{S}_C]_{\{m,n\}} = \mathbf{A}_C^{m-n} \in \mathbb{R}^{7 \times 7}$ ,  $m \geq n$ . The residual matrix  $\mathbf{H}_{\mathcal{F},\theta} = \Delta t_{\mathcal{F}} \mathbf{S}_{\mathcal{F}} [\mathbf{h}_\theta(\chi_{K_C}^*)^\top \dots \mathbf{h}_\theta(\chi_K^*)^\top]^\top$  where  $\mathbf{S}_{\mathcal{F}}$  is lower triangular matrix with each entry defined  $[\mathbf{S}_{\mathcal{F}}]_{\{m,n\}} = \mathbf{A}_{\mathcal{F}}^{m-n} \in \mathbb{R}^{7 \times 7}$ ,  $m \geq n$ . The state prediction at the end of  $K_C$  steps can be computed as

$$\tilde{\mathbf{x}}_{K_C} = \mathbf{A}_C^{K_C} \mathbf{x}_0 + (\mathbf{S} + \mathbf{S}_\theta) \tilde{\mathbf{U}}_C + \phi_\theta, \quad (10)$$

where  $\mathbf{S} = \begin{bmatrix} \mathbf{A}_C^{K_C-1} \mathbf{B}_0 & \mathbf{A}_C^{K_C-2} \mathbf{B}_1 & \dots & \mathbf{B}_{K_C-1} \end{bmatrix}$ ,  $\mathbf{S}_\theta = \begin{bmatrix} \mathbf{A}_C^{K_C-1} \mathbf{B}_{\theta,0} & \mathbf{A}_C^{K_C-2} \mathbf{B}_{\theta,1} & \dots & \mathbf{B}_{\theta,K_C-1} \end{bmatrix}$ , and  $\phi_\theta = \begin{bmatrix} \mathbf{A}_C^{K_C-1} \mathbf{h}_{\theta,0} & \mathbf{A}_C^{K_C-2} \mathbf{h}_{\theta,1} & \dots & \mathbf{h}_{\theta,K_C-1} \end{bmatrix}$ . Substituting (10) into (9) yields the state prediction as

$$\tilde{\mathbf{X}}_{\mathcal{F}} = \mathbf{A}_{\mathcal{F}} \mathbf{A}_C^{K_C} \mathbf{x}_0 + \mathbf{A}_{\mathcal{F}} (\mathbf{S} + \mathbf{S}_\theta) \tilde{\mathbf{U}}_C + \mathbf{A}_{\mathcal{F}} \phi_\theta + \mathbf{H}_{\mathcal{F},\theta}. \quad (11)$$

$$\tilde{\mathbf{X}} = \begin{bmatrix} \tilde{\mathbf{X}}_C \\ \tilde{\mathbf{X}}_{\mathcal{F}} \end{bmatrix} = \mathbf{A}_{\text{qp}} + (\mathbf{B}_{\text{qp}} + \mathbf{B}_{\theta,\text{qp}}) \tilde{\mathbf{U}} + \mathbf{H}_\theta \quad (12)$$

where,  $\mathbf{A}_{\text{qp}} = \begin{bmatrix} \mathbf{A}_C \\ \mathbf{A}_{\mathcal{F}} \mathbf{A}_C^{K_C} \end{bmatrix}$ ,  $\mathbf{B}_{\text{qp}} = \begin{bmatrix} \mathcal{B}_C & \mathbf{0} \\ \mathbf{A}_{\mathcal{F}} \mathbf{S} & \mathbf{0} \end{bmatrix}$ ,  $\mathbf{B}_{\theta,\text{qp}} = \begin{bmatrix} \mathcal{B}_{C,\theta} & \mathbf{0} \\ \mathbf{A}_{\mathcal{F}} \mathbf{S}_\theta & \mathbf{0} \end{bmatrix}$ ,  $\mathbf{H}_\theta = \begin{bmatrix} \mathbf{H}_{C,\theta} \\ \mathbf{H}_{\mathcal{F},\theta} + \mathbf{A}_{\mathcal{F}} \phi_\theta \end{bmatrix}$ , and  $\tilde{\mathbf{U}} = \begin{bmatrix} \tilde{\mathbf{U}}_C \\ \tilde{\mathbf{U}}_{\mathcal{F}} \end{bmatrix}$ .

The objective of the MPC in (8) can be rewritten as:

$$J(\mathbf{U}) = \|\tilde{\mathbf{X}} - \mathbf{X}^*\|_{\mathbf{P}}^2 + \|\tilde{\mathbf{U}} - \mathbf{U}^*\|_{\mathbf{Q}}^2, \quad (13)$$

leading to a quadratic program (QP):

$$\min_{\mathbf{U}} \mathbf{U}^\top \alpha \mathbf{U} + \beta^\top \mathbf{U} \quad \text{s.t. } \mathbf{U}_k \in \mathcal{U}_k, \quad (14)$$

where

$$\begin{aligned} \alpha &= (\mathbf{B}_{\text{qp}} + \mathbf{B}_{\theta,\text{qp}})^\top \mathbf{Q} (\mathbf{B}_{\text{qp}} + \mathbf{B}_{\theta,\text{qp}}) + \mathbf{R}, \\ \beta &= 2(\mathbf{B}_{\text{qp}} + \mathbf{B}_{\theta,\text{qp}})^\top \mathbf{Q} (\mathbf{A}_{\text{qp}} \mathbf{x}_k + \mathbf{H}_\theta - \mathbf{X}^*) - \mathbf{R} \mathbf{U}^*, \\ \mathcal{U}_k &= \{\mathbf{u} \in \mathbb{R}^4 \mid \underline{\mathbf{c}}_k \leq \mathbf{C}_k \mathbf{u}_k \leq \bar{\mathbf{c}}_k, \mathbf{D}_k \mathbf{u}_k = \mathbf{0}\}. \end{aligned}$$

## VI. EVALUATION

### A. Experiment Setup

We evaluated our approach using a A1 robot. To get data for real-time control, we utilized a motion capture system for estimating the position and orientation of the robot trunk at 1 kHz with position errors of 1 mm. We obtain velocity data from position data using forward finite-difference method.

We used *Pytorch* to implement and train the residual dynamics model. The trained *Pytorch* model was converted to a torch script trace, which in turn, was loaded in our MPC implementation in C++. The MPC is solved by a QP solver

TABLE I: Training Parameters

Parameter	Symbol	Value
Variable time steps	$\Delta t_C, \Delta t_{\mathcal{F}}$	{25, 100}(ms)
Prediction steps	$K$	10
NN architecture	$\mathbf{h}_{\theta_1}(q, r)$	7 - 400 Tanh - 400 Tanh - 3
	$\mathbf{G}_{\theta_1}(q, r)$	7 - 1000 Tanh - 1000 Tanh - 12
	$\mathbf{h}_{\theta_2}(q, r)$	7 - 400 Tanh - 400 Tanh - 3
Learning rate	$\gamma$	$2.10^{-4}$
Regulation weights	$\alpha_1, \alpha_2$	$10^{-3}, 10^{-3}$
Total training steps	$N_{train}$	$2.10^4$

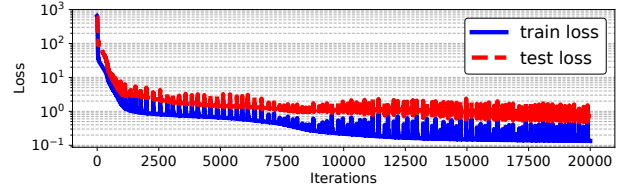


Fig. 3: Training loss and testing loss (Log scale).

(qpOASES). We utilized the optimization toolbox *CasADi* [35] to set up and solve the full-body TO for various jumping targets within [0.3, 0.8] m on a flat ground. We considered jumping motions with contact phase (C) consisting of an all-leg contact period (e.g., 500 ms) and a rear-leg contact period (e.g., 300 ms). With the flight phase (F) scheduled for 400 ms, the front legs and rear legs become swing legs (SW) up to 58% and 33% of the entire jumping period, respectively. Jumping maneuvers feature long flight phases, i.e., the robot can jump up to 4 times its normal standing height during the mid-air period. All jumping experiments are executed with sufficient battery level (> 90%)

### B. Data Collection and Training

We aim to learn a residual dynamics model that is not trajectory-specific and can generalize to a whole family of forward jumps. To collect a sufficient dataset for training, we utilized a baseline MPC controller with a nominal model (Sec. IV-B) to generate diverse trajectories under a variety of unknown disturbances (box under the front feet with random height within [0, 8] cm) and zero-height jumping targets. We collected data from  $H = 20$  jumps with 80% for training and 20% for testing. The data points were sampled with variable sampling time  $\Delta t_C = 25$  ms,  $\Delta t_{\mathcal{F}} = 100$  ms, and horizon length of  $K = 10$ .

The training parameters are listed in Table I. We used fully-connected neural networks with architectures listed in Table I: the first number is the input dimension, the last number is the output dimension, and the numbers in between are the hidden layers' dimensions and activation functions. The output of  $\mathbf{G}_{\theta_1}$  is then converted to a  $3 \times 4$  matrix. The training and testing losses are illustrated in Fig. 3.

### C. Comparative Analysis

1) *Evaluation on testing dataset*: We rolled out the learned dynamics with variable frequency to predict the future state trajectories and compared with (i) those of the variable-frequency nominal model, (ii) those of the fixed-frequency learned model, and (iii) the ground-truth states in Fig. 4. The fixed-frequency learned model is trained with  $K = 10$  and

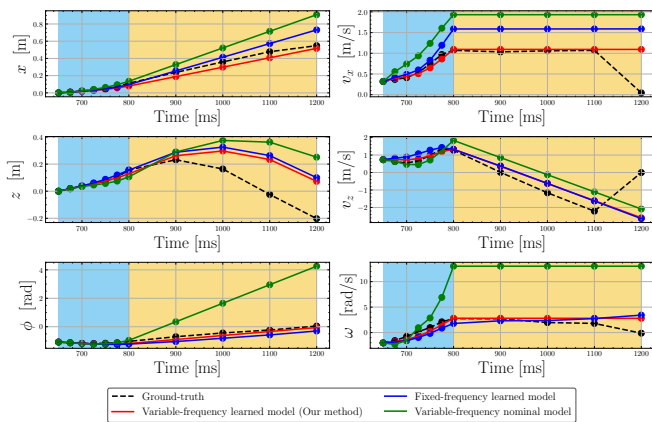


Fig. 4: **State prediction comparison on a random testing dataset.** It starts at 650 ms, consisting of  $K_c = 6$  steps in the contact phase and  $K_f = 4$  steps covering the flight period of 400 ms. The blue and yellow areas represent the contact and flight periods, respectively. We compare variable-frequency learned model (our method), fixed-frequency learned model, variable-frequency nominal model, and the ground-truth trajectories. The ground-truth data is directly obtained from hardware experiments.

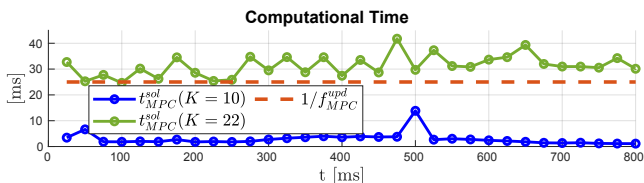


Fig. 5: MPC solving time ( $t_{MPC}^{sol}$ ) with various horizons  $K = 10$  and  $K = 22$ . The red dashed line denotes the MPC update time ( $1/f_{MPC}^{upd}$ ) of 25 ms. Real-time performance is achieved if  $t_{MPC}^{sol} \ll 1/f_{MPC}^{upd}$ .

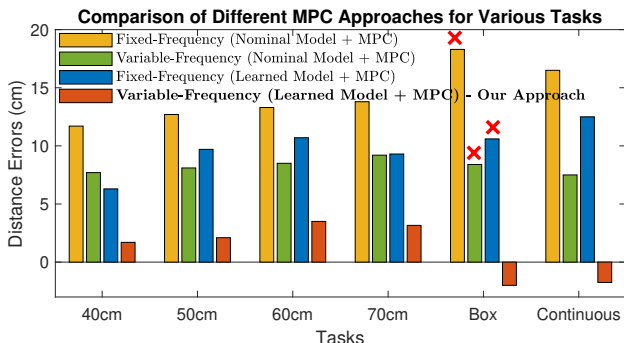


Fig. 6: Comparison of error of final target jumping (in cm) for four real-time MPC setups with different models and timesteps resolutions in a total of 92 jumps with robot hardware. Positive values indicate that the jumps fall short of the target. The cross ( $\times$ ) denotes failure to jump on the box. Supplemental video: <https://youtu.be/2QjZkARs1mU>.

$\Delta t = 25$  ms. The figure shows that our proposed learned model (red lines) outperforms variable-frequency nominal model (green lines) and fixed-frequency learned model (blue lines). We highlight a large deviation in trajectory prediction of variable-frequency nominal model with the others. This inaccuracy can be explained by the use of conventional SRBD for the entire prediction horizon and the coarse Euler integration for the flight phase. Our variable-frequency learned model helps address the inaccuracies introduced by the nominal model and keeps sufficient accuracy of trajectory prediction while allowing real-time MPC by keeping a small number of decision variables.

2) *Effect of prediction horizon:* Fig. 5 compares the solving time for MPC with different horizons. With our variable-

frequency scheme, we can use a few prediction horizons, e.g.,  $K = 10$  (6 for contact and 4 for the flight phase), allowing QP solving time for each MPC update ((14)) to be 2 ms only on average. This efficient computation enables real-time performance for MPC. Using a fixed-frequency scheme with sample time of 25 ms requires MPC to use a large number of steps, e.g.  $K = 22$  (6 for contact and 16 for the whole flight phase). This long horizon yields a large optimization problem, which takes  $\sim 30$  ms for each MPC update and deteriorates the real-time performance.

3) *Execution:* We verified that our proposed approach enables both robust and accurate target jumping to various targets. We compared (a) nominal model & fixed-frequency MPC, (b) nominal model & variable-frequency MPC, (c) fixed-frequency learned model & fixed-frequency MPC, and (d) variable-frequency learned model & variable-frequency MPC (*our approach*). The fixed-frequency scheme uses  $(\Delta t_c, \Delta t_f) = (25, 25)ms$  and variable-frequency one utilizes  $(\Delta t_c, \Delta t_f) = (25, 100)ms$ . All MPCs use the prediction horizon  $K = 10$ , thus the fixed-frequency MPC does not cover the entire flight phase.

For each combination, we perform 23 jumps consisting of (i) 3 single jumps for each of five targets:  $x = 40$  cm,  $x = 50$  cm,  $x = 60$  cm,  $(x, z) = (60, 20)$  cm for jump on box, and  $x = 70$  cm, and (ii) 8 continuous jumps of 60 cm. This yields a total of 92 jumps for comparison. Fig. 6 shows the average final jump target errors of four different MPC combinations across different jumping tasks. Our approach outperforms the methods that adopt fixed-frequency or nominal models, reducing the jumping distance error up to 8 times. Our method demonstrates successful jumping on a box with landing angle errors  $< 15^\circ$ , while the fixed-frequency MPC or nominal model MPC fail, as shown in Fig. 7. Note that the box-jumping task is not executed during data collection, verifying the scalability of our method to generalize to unseen reference trajectories and unseen tasks. With our method, the robot also successfully jumps under model uncertainty, e.g. an unknown 5cm block (30% robot standing height) placed under the front feet (Fig. 7). This task uses a reference trajectory designed for flat ground and demonstrates an example where the robot significantly deviates from the reference trajectory.

We also studied the effect of uncertainty in the model by evaluating the jumping performance if we only utilize a joint PD controller to track the joint reference from full-body TO [1], via  $\tau_{pd} = \tau_{ff} + \mathbf{K}_p(\mathbf{q}_j^* - \mathbf{q}_j) + \mathbf{K}_d(\dot{\mathbf{q}}_j^* - \dot{\mathbf{q}}_j)$ . The full-order model can be conservative by assuming hard feet, point-to-surface contact, and hard ground. These assumptions, however, are not valid for the Unitree A1 robot equipped with deformable feet [36]. Uncertainties due to DC motors working at extreme conditions or motor deficiency are difficult to model and are normally ignored in the full-body model. These factors affect the model accuracy and prevent the robot from reaching the jumping target if we solely rely on the joint PD, as can be seen in Fig. 7.

We evaluate the target accuracy and robustness with our method for continuous jumps on flat ground. The results are presented in Fig. 8, showing some selected snapshots of the first and the last jump to the final target 2.4 m. Our method



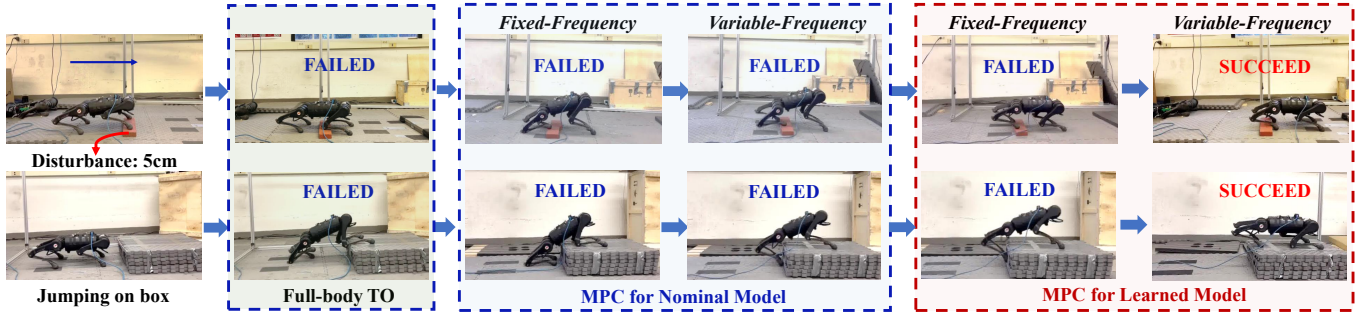


Fig. 7: **Jumping from uneven terrain and Jumping on a box:** Comparison among full-body TO, nominal-model MPC with fixed and variable frequency, and learned-model MPC with fixed and variable frequency. All MPCs have a horizon length of  $K = 10$ . An unknown 5 cm block (30% of the robot’s initial height) was introduced under the robot’s front feet (upper figures). With variable-frequency learned-model MPC (our method), the robot is able to compensate for the disturbance and perform successful jumps. Supplemental video: [https://youtu.be/yUqL\\_MBOC6Q](https://youtu.be/yUqL_MBOC6Q).

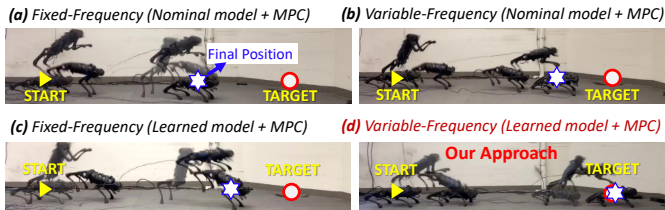


Fig. 8: **Hardware experiments:** We conduct 2 rounds with 4 jumps in each round for each method. The figures (a)-(d) select snapshots of the first and the last jumps when the robot executes four consecutive jumps with different MPC approaches. The target for each jump is 0.6 m.

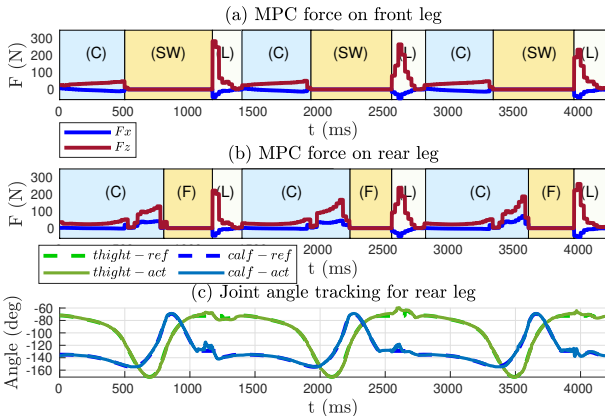


Fig. 9: **Continuous jumping on uneven terrain with A1 robot, shown in Fig. 1:** control forces on the front (a) and rear (b) legs, where the forces are 0 for swing legs; and (c) accurate joint tracking of thigh and calf of rear legs. Plots (a) and (b) show the force outputs satisfy  $|F_z| \leq F_{max} = 350$  N and the friction cone  $|F_x/F_z| \leq \mu = 0.6$  during all phases.

achieves the highest target accuracy, allowing the robot to traverse 2.47 m with an average distance error of only 1.75 cm per jump. During flight periods, the robot can jump up to  $4\times$  its standing height.

Note that computing residual matrices  $\mathbf{G}_{C,\theta}$ ,  $\mathbf{H}_{C,\theta}$ ,  $\mathbf{H}_{F,\theta}$  can consume significant time if we feed the neural network with  $(\mathbf{x}_0, \mathbf{u}_0), (\mathbf{x}_1^*, \mathbf{u}_1^*), \dots, (\mathbf{x}_K^*, \mathbf{u}_K^*)$  sequentially. To improve computational efficiency, we combined the current state and reference as a batch, then feed-forward the entire batch to the learned neural networks all at once, reducing the neural network query time to less than 1 ms.

We further evaluate MPC with different models learned with other timing choices of  $(\Delta t_c, \Delta t_f) = \{(50, 50), (25, 80)\}$  ms, in addition to the two timesteps

TABLE II: Distance errors (cm) when varying coarseness of timesteps  $(\Delta t_c, \Delta t_f)$  ms. Red colors indicate jumping failure as robot falls short of the box (positive values). The distance errors are measured by the motion capture system at the end of the predefined flight period (1200ms). Supplemental video: <https://youtu.be/2QjZkARs1mU>.

Task	(25, 25)	(50, 50)	(25, 80)	(25, 100)
Box-Jump	18.3 ± 3.7	10.8 ± 2.1	-1.8 ± 0.9	-2.0 ± 1.3

$(\Delta t_c, \Delta t_f) = \{(25, 25), (25, 100)\}$  we already discussed above. Three jumps are conducted separately for each timing choice of jumping on box task, and the experiments are summarized in Tab. II. We also achieved successful jumps using the variable-frequency scheme (25, 80) ms. We also observed that using higher timestep in contact to cover the whole flight phase even with fixed-frequency scheme, e.g., (50, 50) ms does not guarantee successful jumps. This confirms our rationale that due to the higher complexity of the robot dynamics during the contact phase, a coarse time step during this phase leads to higher accumulated prediction errors, even in the flight phase, causing task failures.

#### D. Continuous Jumping on Uneven Terrain

We tested the robustness and target accuracy of our learning-based MPC for continuous jumping on uneven terrains with a target of 60 cm for each jump, as illustrated in Fig. 1 and Fig. 9. The terrain consisted of multiple blocks with random heights between 2 – 4 cm, randomly placed on the ground. Since the robot legs can impact the ground early or late on the real robot, there is usually a mismatch between scheduled contact and actual contact states. Whenever both contacts happen (i.e., the landing phase  $L$  starts), we activate a separate landing MPC to make a transition between two jumps in a short period of  $\sim 200$  ms [28]. The landing MPC is designed with simplified dynamics,  $K = 10$ , and  $\Delta t = 25$  ms. This aims to track the body reference trajectory from TO for continuous jumping [28] and connects two jumps seamlessly. Figure 1 shows the actual trajectory of three continuous jumps, leaping around 175 cm in total that yields only 1.67 cm distance error ( $< 3\%$ ) for each jump on the uneven terrain. Compared to our prior work [28], we successfully achieve continuous jumping in hardware in this work. With our framework, we also achieve a target jumping error of less than 2 cm (3.5%) on 34 single and continuous jumps with different targets and uneven terrain.

## VII. DISCUSSION AND CONCLUSION

In this letter, we developed a learned-model MPC approach that enables both target accuracy and robustness for aggressive planar jumping motions. We learned a residual model from a limited dataset to account for the effect of leg dynamics and model mismatch with real hardware. Given the learned model, we designed a learning-based variable-frequency MPC that considers the jumping transitions, the entire flight phase, and the jumping target during the optimization process. We demonstrated the scalability of our approach to handle new jumping targets with unseen reference trajectories. While allowing real-time computation, our approach uses the reference trajectory to evaluate the neural network residual model and ensure a linear model for MPC, which sacrifices a certain degree of accuracy if the system does not operate around the reference under substantial disturbances. To overcome this limitation, our future work will explore the use of nonlinear MPC and address its associated higher computational cost. We will also generalize the model learning and control framework with varying contact schedules. Further, we will incorporate line-foot, rolling, and soft contact in learning aggressive legged locomotion maneuvers.

## REFERENCES

- [1] Q. Nguyen, M. J. Powell, B. Katz, J. D. Carlo, and S. Kim, "Optimized jumping on the MIT Cheetah 3 robot," in *International Conference on Robotics and Automation (ICRA)*, pp. 7448–7454, IEEE, 2019.
- [2] C. Nguyen and Q. Nguyen, "Contact-timing and trajectory optimization for 3D jumping on quadruped robots," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2022.
- [3] M. Chignoli, "Trajectory optimization for dynamic aerial motions of legged robots," *MIT Libraries*, 2021.
- [4] M. Chignoli and S. Kim, "Online trajectory optimization for dynamic aerial motions of a quadruped robot," in *IEEE International Conference on Robotics and Automation (ICRA)*, pp. 7693–7699, 2021.
- [5] Y. Ding, A. Pandala, C. Li, Y.-H. Shin, and H.-W. Park, "Representation-free model predictive control for dynamic motions in quadrupeds," *IEEE Transactions on Robotics*, vol. 37, no. 4, pp. 1154–1171, 2021.
- [6] J. Li, O. Kolt, and Q. Nguyen, "Continuous dynamic bipedal jumping via adaptive-model optimization," *arXiv:2404.11807*, 2024.
- [7] G. Garcia, R. Griffin, and J. Pratt, "Time-varying model predictive control for highly dynamic motions of quadrupedal robots," in *IEEE International Conference on Robotics and Automation (ICRA)*, pp. 7344–7349, 2021.
- [8] H.-W. Park, P. M. Wensing, and S. Kim, "High-speed bounding with the mit cheetah 2: Control design and experiments," *The International Journal of Robotics Research*, vol. 36, no. 2, pp. 167–192, 2017.
- [9] Z. Song, L. Yue, G. Sun, Y. Ling, H. Wei, L. Gui, and Y.-H. Liu, "An optimal motion planning framework for quadruped jumping," in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Japan, pp. 11366–11373, IEEE, 2022.
- [10] C. Mastalli, W. Merkt, G. Xin, J. Shim, M. Mistry, I. Havoutis, and S. Vijayakumar, "Agile maneuvers in legged robots: a predictive control approach," *arXiv preprint arXiv:2203.07554*, 2022.
- [11] Z. Zhuang, Z. Fu, J. Wang, C. Atkeson, S. Schwertfeger, C. Finn, and H. Zhao, "Robot parkour learning," in *Conference on Robot Learning (CoRL)*, 2023.
- [12] Z. Li, X. B. Peng, P. Abbeel, S. Levine, G. Berseth, and K. Sreenath, "Robust and versatile bipedal jumping control through reinforcement learning," *Robotics: Science and Systems*, 2023.
- [13] Y. Yang, G. Shi, X. Meng, W. Yu, T. Zhang, J. Tan, and B. Boots, "CAJun: Continuous adaptive jumping using a learned centroidal controller," *Conference on Robot Learning (CoRL)*, 2023.
- [14] V. Atanassov, J. Ding, J. Kober, I. Havoutis, and C. Della Santina, "Curriculum-based reinforcement learning for quadrupedal jumping: A reference-free design," *arXiv:2401.16337v2*, 2024.
- [15] G. Bellegarda, C. Nguyen, and Q. Nguyen, "Robust quadruped jumping via deep reinforcement learning," *Robotics and Autonomous Systems*, vol. 182, p. 104799, 2024.
- [16] L. Krishna, N. Sobanbabu, and Q. Nguyen, "OGMP: Oracle guided multimodal policies for agile and versatile robot control," *arXiv preprint arXiv:2403.04205*, 2024.
- [17] N. Csomay-Shanklin, V. D. Dorobantu, and A. D. Ames, "Nonlinear model predictive control of a 3D hopping robot: Leveraging Lie group integrators for dynamically stable behaviors," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2023.
- [18] C. Nguyen, L. Bao, and Q. Nguyen, "Mastering agile jumping skills from simple practices with iterative learning control," *arXiv preprint arXiv:2408.02619*, 2024.
- [19] L. Bauersfeld, E. Kaufmann, P. Foehn, S. Sun, and D. Scaramuzza, "NeuroBEM: Hybrid aerodynamic quadrotor model," in *Robotics: Science and Systems*, 2021.
- [20] L. Hewing, J. Kabzan, and M. N. Zeilinger, "Cautious model predictive control using gaussian process regression," *IEEE Transactions on Control Systems Technology*, vol. 28, no. 6, pp. 2736–2743, 2019.
- [21] A. Saviolo, G. Li, and G. Loianno, "Physics-inspired temporal learning of quadrotor dynamics for accurate model predictive trajectory tracking," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 10256–10263, 2022.
- [22] T. Duong, A. Altawaitan, J. Stanley, and N. Atanasov, "Port-Hamiltonian neural ODE networks on Lie groups for robot dynamics learning and control," *IEEE Transactions on Robotics*, 2024.
- [23] T. Salzmann, E. Kaufmann, J. Arrizabalaga, M. Pavone, D. Scaramuzza, and M. Ryll, "Real-time neural mpc: Deep learning model predictive control for quadrotors and agile robotic platforms," *IEEE Robotics and Automation Letters*, vol. 8, no. 4, pp. 2397–2404, 2023.
- [24] G. Cao, E. M.-K. Lai, and F. Alam, "Gaussian process model predictive control of an unmanned quadrotor," *Journal of Intelligent & Robotic Systems*, vol. 88, no. 1, 2017.
- [25] J. Pohlodek, B. Morabito, C. Schlauch, P. Zometa, and R. Findeisen, "Flexible development and evaluation of machine-learning-supported optimal control and estimation methods via HILO-MPC," in *arXiv preprint arXiv:2203.13671*, 2022.
- [26] Y. Sun, W. L. Ubellacker, W.-L. Ma, X. Zhang, C. Wang, N. V. Csomay-Shanklin, M. Tomizuka, K. Sreenath, and A. D. Ames, "Online learning of unknown dynamics for model-based controllers in legged locomotion," *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 8442–8449, 2021.
- [27] A. Pandala, R. T. Fawcett, U. Rosolia, A. D. Ames, and K. A. Hamed, "Robust predictive control for quadrupedal locomotion: Learning to close the gap between reduced- and full-order models," *IEEE Robotics and Automation Letters*, vol. 7, no. 3, pp. 6622–6629, 2022.
- [28] C. Nguyen, L. Bao, and Q. Nguyen, "Continuous jumping for legged robots on stepping stones via trajectory optimization and model predictive control," in *Conference on Decision and Control*, 2022.
- [29] Z. Zhou, B. Wingo, N. Boyd, S. Hutchinson, and YeZhao, "Momentum-aware trajectory optimization and control for agile quadrupedal locomotion," *IEEE Robotics and Automation Letters*, vol. 7, no. 3, pp. 7755–7762, 2022.
- [30] Z. He, J. Wu, J. Zhang, S. Zhang, Y. Shi, H. Liu, L. Sun, Y. Su, and X. Leng, "CDM-MPC: An integrated dynamic planning and control framework for bipedal robots jumping," *IEEE Robotics and Automation Letters*, 2024.
- [31] H. Li and P. M. Wensing, "Cafe-mpc: A cascaded-fidelity model predictive control framework with tuning-free whole-body control," *arXiv preprint arXiv:2403.03995*, 2024.
- [32] J. Norby, A. Tajbakhsh, Y. Yang, and A. M. Johnson, "Adaptive complexity model predictive control," *IEEE Transactions on Robotics*, vol. 40, pp. 4615–4634, 2024.
- [33] H. Li, R. J. Frei, and P. M. Wensing, "Model hierarchy predictive control of robotic systems," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 3373–3380, 2021.
- [34] J. D. Carlo, P. M. Wensing, B. Katz, G. Bledt, and S. Kim, "Dynamic locomotion in the MIT Cheetah 3 through convex model-predictive control," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018.
- [35] J. A. E. Andersson, J. Gillis, G. Horn, J. B. Rawlings, and M. Diehl, "CasADI – A software framework for nonlinear optimization and optimal control," *Mathematical Programming Computation*, vol. 11, no. 1, pp. 1–36, 2019.
- [36] S. Yang, H. Choset, and Z. Manchester, "Online kinematic calibration for legged robots," *IEEE Robotics and Automation Letters*, vol. 7, pp. 8178–8185, 2022.