

# A Multi-step Dynamics Modeling Framework For Autonomous Driving In Multiple Environments

Jason Gibson<sup>1</sup>, Bogdan Vlahov<sup>1</sup>, David Fan<sup>2</sup>, Patrick Spieler<sup>2</sup>, Daniel Pastor<sup>2</sup>,  
Ali-akbar Agha-mohammadi<sup>2</sup>, Evangelos A. Theodorou<sup>1</sup>

**Abstract**—Modeling dynamics is often the first step to making a vehicle autonomous. While on-road autonomous vehicles have been extensively studied, off-road vehicles pose many challenging modeling problems. An off-road vehicle encounters highly complex and difficult-to-model terrain/vehicle interactions, as well as having complex vehicle dynamics of its own. These complexities can create challenges for effective high-speed control and planning. In this paper, we introduce a framework for multistep dynamics prediction that explicitly handles the accumulation of modeling error and remains scalable for sampling-based controllers. Our method uses a specially-initialized Long Short-Term Memory (LSTM) over a limited time horizon as the learned component in a hybrid model to predict the dynamics of a 4-person seating all-terrain vehicle (Polaris S4 1000 RZR) in two distinct environments. By only having the LSTM predict over a fixed time horizon, we negate the need for long term stability that is often a challenge when training recurrent neural networks. Our framework is flexible as it only requires odometry information for labels. Through extensive experimentation, we show that our method is able to predict millions of possible trajectories in real-time, with a time horizon of five seconds in challenging off road driving scenarios.

## I. INTRODUCTION

Model-Predictive Control (MPC) methods have been used for real-time controls applications for many years [1]–[4] in real-time control of complex systems. These methods are able to handle noisy inputs and uncertainty through fast replanning. However, one of the main limiting factors of these approaches is fast, accurate dynamics models.

The most common dynamics model historically has been linear approximations obtained through system identification [5]. The inherent nonlinear nature of the world pushed the models towards nonlinear or linearized approximations focused around the core principal of physics-based approximations. Unfortunately, these equations are prohibitively difficult to compute, requiring simplifying assumptions [1], or rely on parameters that are costly to determine with any certainty [6]. Therefore, we often have first principal models where nonlinear effects are ignored or greatly simplified [7]. Even with this limitation, MPC methods are able to accomplish challenging real world tasks. With fast replanning, even significant model error can be mitigated.

<sup>1</sup>Autonomous Control and Decision Systems Lab, Georgia Institute of Technology, Atlanta GA 30313 USA (Corresponding author: jgibson37@gatech.edu)

<sup>2</sup>NASA Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA

This paper has a supplementary video available to view at <https://youtu.be/bzyn-qpRJIk>

Fast replanning induces computational restrictions that are magnified when considering sampling-based controllers, where the number of forward passes on the prediction can exceed millions of computation a second. Still, a desire for increased accuracy of dynamics without sacrificing computational efficiency has led to a prevalence of black-box learning methods in dynamics modeling. The direction of learned components can be broadly categorized by the following methods: Gaussian Processes (GPs) [2], [8], Neural Networks [4], [9], and others outlined in this survey paper on dynamics model learning [5]. Each approach has potential drawbacks for hardware systems in the real world. The balance of computation and predictive accuracy is the main limitation when it comes to selecting models for MPC. Using physics-based equations to integrate predicted accelerations or velocities can simplify the learning problem and improve generalization in various systems [10] [11]. Models using this structure are often called hybrid models.

While there are various architectures and frameworks for the dynamics model in MPC, they tend to be used in a very similar fashion, and suffer from two main drawbacks. First, trajectories are often computed recursively, using the previous steps' output as the input to the model in the next step. This inherent recursive nature can lead to significantly decreased accuracy as the time horizon increases, and a small overestimate can quickly blow up a state value and lead to nonsensical trajectories. Second, a more nuanced drawback of single-step learning is that lower frequency phenomena are effectively treated as random noise in the learning problem. Indeed, prior work has shown the sensitivity of selecting  $\Delta t$  for different systems [12] in single-step prediction. Therefore, a good solution to both of these problems is to consider the trajectory prediction problem as a multi-step prediction problem rather than a single-step one.

The recursive nature of how these models are used lends itself toward a model structure that is designed for recurrence. In this work, we will look primarily at the Long Short-Term Memory (LSTM) architecture [13], [14] due to its ability to avoid the vanishing gradient problem often seen in other recurrent structures. We will follow the example of [15] and use another network to initialize the hidden and cell state of the LSTM. Our LSTM prediction models will only need to be run for a fixed time horizon, so training them for continual stability is not required.

In this work, we demonstrate a hybridization architecture that is successfully able to learn the dynamics of the same vehicle in two diverse environments. We combine initialized

LSTMs [15] with the integrated loss function [16] while keeping the networks small enough to run in real-time on hardware systems. Through the use of the integrated loss function, we are able to construct a modeling architecture that is robust to changing environments, while only requiring relatively smooth odometry estimates for training. This allows seamless integration of human driving data and autonomously controlled data into the training process. We emphasize that our specific hybrid setup should be viewed as a hyperparameter with extensive room for exploration. The key is the integrated loss function allowing for predictions with arbitrary transformations and integration. To our knowledge, this is the first work to combine initialized LSTMs with the integrated loss function. The results are validated on a real vehicle using 650,000 trajectories of driving data.

## II. RELATED WORKS

In the world of autonomous racing, accurate dynamics modeling is crucial to achieving performance at the limits of handling. [1] uses a kino-dynamic model that improves upon the nominal bicycle model by incorporating drag, road slop effects, understeer/oversteer characteristics, and pose relative to an optimal race line. They achieve real-time performance with this model in an iterative MPC solver. However, this model assumes there is no tire slip, which is not applicable in off-road conditions. [17] demonstrates a MPC approach using a model that learns tire forces as deviations from a library of pre-selected tire models. This allows for control over snow-covered terrain but requires a large body of domain-specific knowledge to generate the library of tire models. Our approach allows for a relaxation on the amount of knowledge required to build the model.

Using learned models for MPC to achieve an acceptable balance between accuracy and computational complexity has been investigated for many years. In 1995, [4] used a feed-forward Neural Network (FNN) to learn the dynamics of the pH in a plant with pump flow rates as the control. They successfully showed that a model trained using a single-step prediction loss and combined with MPC could outperform a PI controller for achieving and maintaining desired pH setpoints. Since then, [18] has used neural networks trained on single-step prediction loss functions in sampling-based MPC controllers to capture nonlinear behavior, such as vehicle drift while turning, leading to improved performance.

One approach to prevent this recursive error magnification while learning a model is to use GPs [2], [8], whose errors remain bounded when used for predictions which lie near previously collected data points. [2] uses Sparse Gaussian Processes (SGPs) to learn the difference between a nominal bicycle dynamics model and real data gathered from an electric race car. This combination allows them to run an iterative MPC solver in real-time and update the model in real-time to new environment information. However, this sort of model is restricted to problems which allow for traversing the same location repeatedly, such as on a race track. [11] [19] uses LSTM models for prediction of dynamical system inside hybrid models, but none applies the integrated loss

function which prevents our hybrid model architecture from being applied.

## III. MODEL LEARNING FRAMEWORK

We will be using the Model-Predictive Path Integral Control (MPPI) algorithm as a baseline MPC approach for computational comparisons. Ultimately, we want models that can compute 18,432 samples at around 15 Hz. The target time horizon of the prediction is  $T = 5.0$  with a  $\Delta t = 0.02$ . This results in  $\approx 69,000,000$  forward passes on the dynamics a second. Our specific implementation of MPPI will be able to leverage GPUs through CUDA, as with prior publications [20]. We specifically target a Nvidia RTX 3080, but these networks can run in real-time on smaller hardware systems as well. We will follow prior work on hybrid dynamics models [5], and predict the derivative of the state we care about and integrate using the relevant kinematic equations. Our overall architecture can be summarized in Fig. 1.

**Initialized LSTM:** We now briefly review [21] where a new method of initializing LSTMs was proposed. LSTM training often has to deal with how to initialize the hidden and cell states of the LSTM. The most common approach with recurrent networks is to define a washout period, where the network is run on prior data for a predefined time period [22]. Prior work has treated the washout period as a hyperparameter, and effectively loses the data used during the washout period since no loss can be computed on it. Furthermore, washout can also lead to instability during training [23]. [21] treats the problem of initializing the hidden/cell states as part of the training process. This is done by having a network predict the initial hidden/cell state of another network using some predefined buffer over history denoted by  $\tau$ .

The followup paper [15] primarily focuses on exploring what type of initialization network is best. It does a robust comparison of washout, a FNN, and a Recurrent Neural Network (RNN). Their extensive results with very large networks indicate that RNNs are the best suited architecture. They also investigate and demonstrate that hybrid architectures are better able to predict farther into the future with fewer weights. We will continue this analysis for the autonomous driving setting in off-road environments. The general initialized network architecture can be shown in the following equations.

$$[\mathbf{c}_{t-\tau}^I \quad \mathbf{h}_{t-\tau}^I]^T = \kappa_I(\mathbf{x}_{t-\tau}, \mathbf{0}, \mathbf{0}) \quad (1)$$

⋮

$$[\mathbf{c}_t^I \quad \mathbf{h}_t^I]^T = \kappa_I(\mathbf{x}_t, \mathbf{c}_{t-1}^I, \mathbf{h}_{t-1}^I) \quad (2)$$

$$[\mathbf{c}_t^P \quad \mathbf{h}_t^P]^T = \vartheta_I(\mathbf{h}_t^I, \mathbf{x}_t) \quad (3)$$

$$[\mathbf{c}_{t+1}^P \quad \mathbf{h}_{t+1}^P]^T = \kappa_P(\mathbf{x}_t, \mathbf{c}_t^P, \mathbf{h}_t^P) \quad (4)$$

$$\zeta_t = \vartheta_P(\mathbf{h}_{t+1}^P, \mathbf{x}_t) \quad (5)$$

Where  $\kappa_I$  represents the LSTM initializer network,  $\vartheta_I$  is the FNN initializer output network,  $\kappa_P$  represents the LSTM

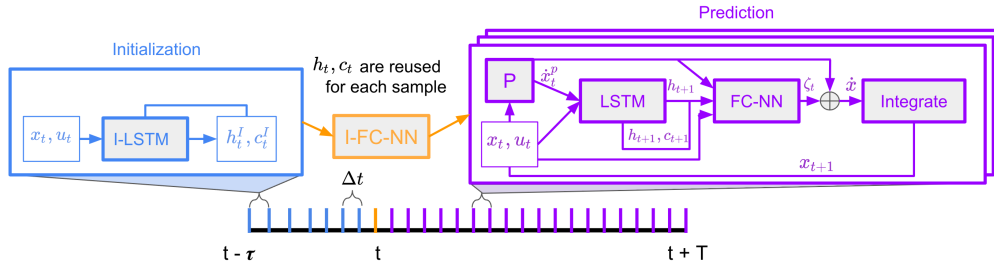


Fig. 1. Figure showing overall architecture. Note the initialization network is run a single time and  $h_t, c_t$  are reused for each trajectory sampled using the prediction step.

prediction network,  $\vartheta_P$  represents the FNN prediction output network, and  $\mathbf{h}_t^I, \mathbf{c}_t^I, \mathbf{h}_t^P, \mathbf{c}_t^P, \mathbf{x}_t, \zeta_t$ , are the initializer hidden and cell vectors, prediction hidden and cell vectors, state vector, and output vector respectively at time  $t$ . A key note is that the initializer network only runs on historical data and has no information about the future controls of the vehicle. This means sampling-based approaches can reuse the same initialization output for all samples. Furthermore, there is no restriction on the input to the initializer networks; it can use a different input vector than the predictor network.

**Integrated Loss:** The integrated loss function penalizes the deviation from the true state trajectory in position mainly [16]. In essence this penalizes the accuracy of the integration as well as the networks prediction. The main benefits of [16] can be summarized by the following points. First, we only require smooth position estimates, so there is no need to differentiate noisy odometry readings to get accelerations or approximate outputs that are difficult to compute. The input to the loss function is only position estimates so learned models can predict arbitrary values which would be challenging to determine labels for. We use this in our hybrid architecture to predict drag forces on wheels in the wheel frame. Second, we explicitly penalize accumulated error which allows for stable long term predictions far into the future that can account for slowly evolving changes to the dynamics. The two main drawbacks of this method is that we can get non-smooth outputs with low loss in the first couple predictions and increased loss function complexity during learning. The first is an ongoing problem in this work, and the second is mitigated by using modern libraries for backpropagation [24].

#### IV. HYBRID MODEL SETUP

The vehicle dynamics that we learn are for the Polaris S4 1000 RZR. It is fit with a drive-by-wire kit that enables autonomous driving as well as logging of control inputs from a human driver. The control inputs for the vehicle are desired throttle, brake, and steering wheel position. During autonomous operation, a physical actuator must turn the steering wheel and actuate the brake, which introduces significant delay. The delay on throttle actuation is very small and can safely be ignored. These delays are not present in the human driving data, and the human is able to actuate these controls far faster than the autonomous system. This

poses a problem as human driving data is cheaper to attain, and the vehicle can be driven more aggressively. In order to make use of this data, we create a model architecture that allows for the decoupling of the steering wheel angle and the brake commanded from the prediction of the body accelerations. The inputs to our models consist of position and velocity coming from odometry and an approximation of roll and pitch derived from an elevation map. Our model is broken into 3 main components summarized below.

The first two components predict the brake and steering actuation delay. These are trained only on autonomous data and use ground-truth odometry and control commands as inputs. Both systems can be modeled parametrically by a rate-limited first-order lag in addition to a neural network as outlined below,

$$p_b = \min \left( \max \left( (u_t^b - b_t^u) C_B, -\dot{b}_{lim}^u \right), \dot{b}_{lim}^u \right) \quad (6)$$

$$\zeta_t^b = \vartheta_P^b (b_t^u, u_t^b, p_b) \quad (7)$$

$$b_{t+1}^u = b_t^u + (p_b + \zeta_t^b) \Delta t \quad (8)$$

Where  $p_b$  is the parametric approximation of the model with rate limit  $\dot{b}_{lim}^u$  and constant  $C_B$ ,  $u_t^b$  is the commanded brake, and  $b_t^u$  is the current brake state.  $\zeta_t^b$  is the output of the network, the variables after that denote the inputs. The steering delay is modeled with the following equations,

$$p_\delta = \min \left( \max \left( (u_t^\delta - \delta_t) C_\delta, -\dot{\delta}_{lim}^u \right), \dot{\delta}_{lim}^u \right) \quad (9)$$

$$\zeta_t^\delta = \vartheta_P^\delta (v_t^x, u_t^\delta, \delta_t, \dot{\delta}_t, p_\delta) \quad (10)$$

$$\delta_{t+1} = \delta_t + (p_\delta + \zeta_t^\delta) \Delta t \quad (11)$$

where the notation matches the above but with  $\delta$  as the steering angle position and  $v_t^x$  denotes the body frame forward velocity of the vehicle. The parameters used in the above equations are fit using the Adam optimizer [25] on the purely parametric version. The loss function is mean squared error on the state  $(\delta, b_t^u)$ .

The third component is made up of two different networks. An engine model predicts the force imparted by the engine on the wheels,  $\zeta_t^e$ , and a terradynamics model predicts an environment-dependent drag force,  $\zeta_t^t$ . The coordinate frames are pictured in Fig. 2. We normalize by the mass of the vehicle as it is constant between every equation. The outputs of the networks are scaled by 10, which was chosen to be the maximum expected acceleration value output from the



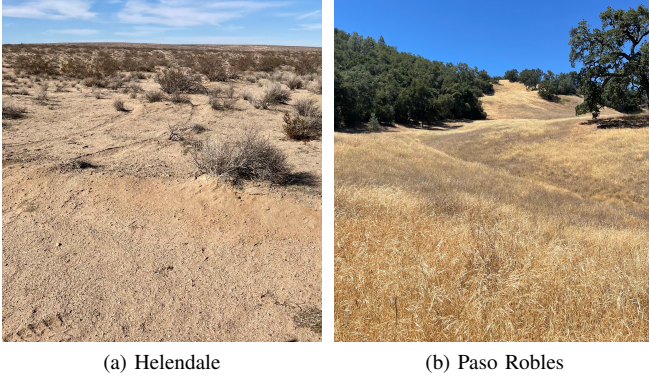


Fig. 3. Pictures from the two environments where the datasets were collected.

loose sand. The dataset includes  $\approx 300,000$  trajectories of human driving data and  $\approx 100,000$  trajectories of autonomous driving data.

**Paso Robles:** The second dataset is from trails and off-road driving near Paso Robles, CA. It is a primarily consists of dry grassland and oak woodland, denoted as the chaparral environment (Fig. 3b). This subsection of data has large, steep slopes of up to  $40^\circ$  inclination and more than 100 m tall. The ground is entirely covered with dense, dry grass where the vehicle easily loses traction. Our dataset includes  $\approx 100,000$  trajectories of human driving and  $\approx 150,000$  trajectories of autonomous driving data.

## VI. RESULTS

Throughout this results section, we will be covering results on the RZR vehicle trained with a time horizon of  $T = 5s$  and a  $\Delta t = 0.02s$ . The main results are summarized in Table I. This compares 5 different architectures of models on the total combined dataset. The overall training set was approximately  $\approx 500,000$  trajectories, and a test set of  $\approx 150,000$  trajectories. There is potential overlap within the same trajectory since a trajectory going from  $[t_0, t_T]$  could include values from another trajectory from  $[t_1, t_{T+1}]$ . Overlap is constrained within the validation or test set. These two sets are pulled from the same day of testing, but are otherwise independent.

$H$  represents the hybrid model outlined in Section IV, whereas  $K$  represents the kinematic only model, and  $I$  denotes use of an initializer network for prediction, such that  $HI-LSTM$  is an hybrid architecture with an initialization network. If there is no initializer network the initial hidden/cell state of the predictor network is set to zero. The  $K$  networks directly predict the outputs  $\dot{v}_t^x, \dot{v}_t^y, \dot{\psi}$  without the hybrid architecture.  $K-LSTM$  is a LSTM with no initializer network and only uses input from odometry and elevation map, but still uses the kinematic Eqs. (19) to (23) for integration. This restricts  $K-LSTM$  to using  $(v_t^x, v_t^y, \dot{\psi}_t, \delta_t, \dot{\delta}_t, \phi_t, \theta_t, u_t^t, u_t^b, \eta_t)$  as inputs, where  $\eta_t$  is -1 for Paso Robles and 1 for Helendale. This means that the  $K$ -networks do have access to what environment they are being run on. The initialization and prediction networks take identical input values.

Our networks are implemented in PyTorch [24] for training and comparisons on data. The real-time results are in a custom implementation of an LSTM network written in C++/CUDA. All networks are trained for 10 epochs, with the last 3 epochs allowing for specialization of the terradynamics network to each environment. The network size for each architecture is given in the following notation:

$$[\kappa_I \text{ input dim}][\kappa_I \text{ hidden dim}] \times [\vartheta_I \text{ architecture}] - [\kappa_P \text{ input dim}][\kappa_P \text{ hidden dim}] \times [\vartheta_P \text{ architecture}].$$

Where the notation follows that used in Eqs. (1) to (5). For the output networks,  $\vartheta$ , the numbers represent the size of the layers. For example 5, 10, 3 would mean a FNN that takes in 5 inputs, has a middle layer with 10 neurons with a tanh activation, and outputs 3 values.

- $K-LSTM$  and  $KI-LSTM$ :
  - Delay:  $[2][60] \times [62, 100, 10] - [2][5] \times [7, 10, 1]$
  - Steering:  $[4][60] \times [64, 100, 10] - [4][5] \times [9, 5, 1]$
  - Engine/Terra Combo:  $[10][60] \times [70, 100, 60] - [10][30] \times [40, 40, 3]$
- $H-LSTM$  and  $HI-LSTM$ :
  - Delay:  $[2][60] \times [62, 100, 10] - [3][5] \times [8, 10, 1]$
  - Steering:  $[4][60] \times [64, 100, 10] - [5][5] \times [10, 5, 1]$
  - Engine:  $[3][60] \times [63, 100, 10] - [3][5] \times [8, 20, 1]$
  - Terra:  $[10][60] \times [70, 100, 20] - [10][10] \times [20, 20, 3]$

Notice that the combined engine and terradynamics network in the  $K$  versions has significantly more weights  $\approx 3X$ . The increase in weights was to compensate for the lack of a parametric approximation to predict around.

A key part of optimizing performance in CUDA code is reducing access to memory in the parallel threads. Some additional memory beyond the weights must be used in order to parallelize the computation of the network. As the initializer network only needs to be run once for all samples, it can remain on the CPU. Only the predictor network needs to run in parallel on the GPU. The memory usage and runtimes of the models in CUDA are outlined in Table II. Notice that the amount of memory required for computing the kinematic networks greatly exceeds that needed by the hybrid architecture while giving similar results.

The parametric model used as a baseline comparison is a linear approximation of the engine and terradynamics model combined with the parametric components outlined in Eqs. (19) to (23), using Eq. (13) for the yaw rate. This model does not have any  $v_t^y$ , just a linear approximation of  $\dot{v}_t^x$ ,

$$\dot{v}_t^x = C_T u_t^t + C_B b_t^u - C_V v_t^x \quad (24)$$

where  $C_T = 4.0, C_B = 10, C_V = 0.2$ . This was the primary dynamics model used during collection of the autonomous driving data, and performs reasonably well with MPPI.

**Engine/Terradynamics Performance:** For a comparison independent of the learned actuator delays, we compared our models with ground truth values for the brake state  $b_t^u$ , steering angle  $\delta$  and steering rate  $\dot{\delta}$ . The models were run on combined data from autonomous driving and human

TABLE I  
MEAN ERROR AND FIRST STANDARD DEVIATION OF ERROR FOR COMBINED AND COMBINED AUTONOMOUS

Data Type	Network	Distance Error (m) $t = 1$	Yaw Error (rad) $t = 1$	Distance Error (m) $t = 5$	Yaw Error (rad) $t = 5$	Distance Error (m) $t = 10$	Yaw Error (rad) $t = 10$
C	P	$0.377 \pm 0.127$	$0.036 \pm 0.002$	$4.203 \pm 10.461$	$0.175 \pm 0.054$	$11.153 \pm 77.551$	$0.339 \pm 0.184$
	K-LSTM	<b><math>0.270 \pm 0.088</math></b>	<b><math>0.033 \pm 0.001</math></b>	$1.845 \pm 2.309$	$0.117 \pm 0.013$	$5.023 \pm 18.833$	$0.208 \pm 0.046$
	KI-LSTM	$0.277 \pm 0.087$	$0.036 \pm 0.001$	$1.975 \pm 2.476$	<b><math>0.113 \pm 0.014</math></b>	<b><math>4.931 \pm 18.095</math></b>	<b><math>0.195 \pm 0.045</math></b>
	H-LSTM	$0.284 \pm 0.096$	$0.042 \pm 0.001$	$1.995 \pm 2.708$	$0.138 \pm 0.021$	$5.459 \pm 21.833$	$0.249 \pm 0.074$
	HI-LSTM	$0.278 \pm 0.078$	$0.049 \pm 0.002$	<b><math>1.765 \pm 2.021</math></b>	$0.132 \pm 0.021$	$4.981 \pm 17.143$	$0.229 \pm 0.065$
CA	K-LSTM	$0.539 \pm 0.136$	$0.239 \pm 0.026$	$7.259 \pm 28.837$	$0.775 \pm 0.307$	$18.159 \pm 202.346$	$1.055 \pm 0.520$
	KI-LSTM	$0.472 \pm 0.109$	$0.177 \pm 0.015$	$7.040 \pm 21.287$	$0.813 \pm 0.0307$	$21.015 \pm 155.293$	$1.431 \pm 0.688$
	H-LSTM	$0.376 \pm 0.071$	<b><math>0.036 \pm 0.001</math></b>	$2.465 \pm 3.737$	$0.149 \pm 0.016$	$6.239 \pm 26.865$	$0.276 \pm 0.049$
	HI-LSTM	<b><math>0.272 \pm 0.042</math></b>	$0.040 \pm 0.001$	<b><math>2.089 \pm 3.463</math></b>	<b><math>0.145 \pm 0.016</math></b>	<b><math>5.722 \pm 22.936</math></b>	<b><math>0.268 \pm 0.051</math></b>

TABLE II  
AVERAGE CUDA RUNTIMES OF 1 MPPI ITERATION  
WITH 18432 SAMPLES ON A NVIDIA RTX 3080

Architecture	CUDA Memory (kB)	CUDA Runtime (ms)
P	0.024	23.327
K-LSTM and KI-LSTM	62.468	83.708
H-LSTM and HI-LSTM	38.08	59.346

driving. Results are in Table I under Combined. All models perform relatively similarly with ground truth values, most resulting in positional errors around 5 meters with a 10 second trajectory. This shows that the LSTM networks are able to predict stable trajectories twice their training horizon. For engine/terradyamics modeling, the hybrid architecture did not seem to improve performance significantly over the learned networks, indicating that the hybrid architecture was likely not assistive with the learning. The design space of the hybrid architecture and integration is broad, a different setup could result in improved performance. Furthermore, the initializer network had a limited effect on accuracy. Increasing the size of the initialization network would likely improve accuracy, a direction for future work would be a detailed analysis of how varying the size effects the accuracy.

**Entire System Performance:** In this section, we reuse the same networks as previously, but include matching architectures for the brake and steering delay models. Therefore, the  $K$  networks will have an LSTM directly predict  $\dot{b}_t^u$  and  $\dot{\delta}$  without any parametric assistance. This focuses only on the autonomy dataset, a subset of  $\approx 50,000$  trajectories for a test set. Here we see a stark contrast between the hybrid approach and the kinematic approach. The brake and steering networks perform worse than their parametric counterparts in these cases where a black box function can be hard to predict. On average, they have about 2X the average error in steering and very similar prediction accuracy for the brake models. The brake prediction error is deceptive though, since most trajectories have a constant brake state of zero and no brake applied at all. Looking at the individual state trajectories, we see slight perturbations from the network that can impact performance significantly.

The worse steering and brake prediction performance leads to a large increase in the yaw error over time, giving very large distance errors at higher speeds. Notably, the outputs

tend to be more jagged, when previously they were smooth. It is likely that the network was relying on the perfect values for  $\delta, \dot{\delta}, b_t^u$ . Once these values started to become inaccurate due to the modeling errors, the network entered sequences of inputs that were not encountered during training and gave poor results. The hybrid architecture reduces the reliance by providing a reasonable estimate that is independent of the internal state of the LSTM.

The network tends to oscillate the predictions from large positive values to large negative values in a single time step on all values.  $\dot{\psi}$  and  $\dot{v}_x$  are prone to larger oscillations, likely due to more noise in labels from odometry. The methodology is sensitive to drift and the small jumps in odometry that come from the constraint of delayed LIO optimization with IMU preintegration. Logically, the integrated loss function will not penalize discontinuous looking state trajectories that match well with position from odometry. drift is treated as potential movement the network needs to predict, therefore we often see short dynamically infeasible trajectories at velocities close to zero. This was somewhat reduced in the trajectory through adding additional losses on velocity, but further investigation into the effect of odometry noise is needed. In general the models show great ability to predict stable outputs into the future but tend to over smooth the outputs. Since odometry noise in velocities can result in incorrect positions if just directly integrated, we only penalize the loss of velocity when the error is large.

## VII. CONCLUSION

In this work, we propose a framework for dynamics model learning in challenging off-road environments. Our method combines specially initialized LSTM networks with an integrated loss function to allows for flexible learned outputs. Through hundreds of thousands of trajectories we are able to show generalization across two distinct off-road environments. Our method is computationally efficient and can run in real time on our hardware in a sampling based MPC approach.

## ACKNOWLEDGMENTS

The research was partially carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (80NM0018D0004).

## REFERENCES

- [1] E. Pagot, M. Piccinini, and F. Biral, "Real-time optimal control of an autonomous RC car with minimum-time maneuvers and a novel kineto-dynamical model," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, (Las Vegas, NV, USA), pp. 2390–2396, IEEE, Oct. 2020. 1, 2
- [2] J. Kabzan, L. Hewing, A. Liniger, and M. N. Zeilinger, "Learning-Based Model Predictive Control for Autonomous Racing," *IEEE Robotics and Automation Letters*, vol. 4, pp. 3363–3370, Oct. 2019. 1, 2
- [3] 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 Transactions on Robotics*, vol. 34, no. 6, pp. 1603–1622, 2018. 1
- [4] A. Draeger, S. Engell, and H. Ranke, "Model predictive control using neural networks," *IEEE Control Systems Magazine*, vol. 15, pp. 61–66, Oct. 1995. 1, 2
- [5] D. Nguyen-Tuong and J. Peters, "Model learning for robot control: A survey," *Cognitive processing*, vol. 12, pp. 319–40, 04 2011. 1, 2
- [6] H. Hjalmarsson, "System identification of complex and structured systems," *European journal of control*, vol. 15, no. 3-4, pp. 275–310, 2009. 1
- [7] A. Freddi, A. Lanzon, and S. Longhi, "A feedback linearization approach to fault tolerance in quadrotor vehicles," *IFAC Proceedings Volumes (IFAC-PapersOnline)*, vol. 44, no. 1 PART 1, pp. 5413–5418, 2011. 1
- [8] 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, 2017. 1, 2
- [9] G. R. Williams, B. Goldfain, K. Lee, J. Gibson, J. M. Rehg, and E. A. Theodorou, "Locally weighted regression pseudo-rehearsal for adaptive model predictive control," in *Conference on Robot Learning*, pp. 969–978, 2020. 1
- [10] Y. Yu, H. Yao, and Y. Liu, "Aircraft dynamics simulation using a novel physics-based learning method," *Aerospace Science and Technology*, vol. 87, pp. 254–264, 2019. 1
- [11] S. K. Singh, R. Yang, A. Behjat, R. Rai, S. Chowdhury, and I. Matei, "Pi-ilstm: Physics-infused long short-term memory network," in *2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA)*, pp. 34–41, 2019. 1, 2
- [12] Y. Liu, J. N. Kutz, and S. L. Brunton, "Hierarchical deep learning of multiscale differential equation time-steppers," *Philosophical Transactions of the Royal Society A*, pp. 1–22, 2020. 1
- [13] T. Zia and U. Zahid, "Long short-term memory recurrent neural network architectures for Urdu acoustic modeling," *International Journal of Speech Technology*, vol. 22, no. 1, pp. 21–30, 2019. 1
- [14] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, pp. 1735–1780, 1997. 1
- [15] N. Mohajerin and S. L. Waslander, "Multistep Prediction of Dynamic Systems with Recurrent Neural Networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 11, pp. 3370–3383, 2019. 1, 2
- [16] N. Seegmiller, F. Rogers-Marcovitz, G. Miller, and A. Kelly, "Vehicle model identification by integrated prediction error minimization," *International Journal of Robotics Research*, vol. 32, no. 8, pp. 912–931, 2013. 2, 3
- [17] K. Berntorp, R. Quirynen, T. Uno, and S. Di Cairano, "Trajectory tracking for autonomous vehicles on varying road surfaces by friction-adaptive nonlinear model predictive control," *Vehicle System Dynamics*, vol. 58, pp. 705–725, May 2020. 2
- [18] G. Williams, P. Drews, B. Goldfain, J. M. Rehg, and E. A. Theodorou, "Aggressive driving with model predictive path integral control," in *International Conference on Robotics and Automation (ICRA)*, pp. 1433–1440, IEEE, 2016. 2
- [19] Q. Teng and L. Zhang, "Data driven nonlinear dynamical systems identification using multi-step cldnn," *AIP Advances*, vol. 9, p. 085311, 08 2019. 2
- [20] M. S. Gandhi, B. Vlahov, J. Gibson, G. Williams, and E. A. Theodorou, "Robust model predictive path integral control: Analysis and performance guarantees," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 1423–1430, 2021. 2
- [21] N. Mohajerin, M. Mozifian, and S. Waslander, "Deep learning a quadrotor dynamic model for multi-step prediction," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 2454–2459, 2018. 2
- [22] H. Jaeger, "Tutorial on training recurrent neural networks, covering bppt, rtrl, ekf and the "echo state network" approach," 2002. 2
- [23] H.-G. Zimmermann, C. Tietz, and R. Grothmann, "Forecasting with Recurrent Neural Networks: 12 Tricks," in *Neural Networks: Tricks of the Trade: Second Edition* (G. Montavon, G. B. Orr, and K.-R. Müller, eds.), pp. 687–707, Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. 2
- [24] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, "Pytorch: An imperative style, high-performance deep learning library," in *Advances in Neural Information Processing Systems 32* (H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, eds.), pp. 8024–8035, Curran Associates, Inc., 2019. 3, 5
- [25] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014. 3
- [26] S. Fakoorian, K. Otsu, S. Khattak, M. Palieri, and A. A. Aghamohammadi, "ROSE: Robust State Estimation via Online Covariance Adaptation," *International Foundation of Robotics Research (ISRR)*, 2022. 4
- [27] T. Shan, B. Englot, D. Meyers, W. Wang, C. Ratti, and R. Daniela, "Lio-sam: Tightly-coupled lidar inertial odometry via smoothing and mapping," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 5135–5142, IEEE, 2020. 4
- [28] D. D. Fan, K. Otsu, Y. Kubo, A. Dixit, J. Burdick, and A.-A. Aghamohammadi, "Step: Stochastic traversability evaluation and planning for risk-aware off-road navigation," in *Robotics: Science and Systems*, pp. 1–21, RSS Foundation, 2021. 4
- [29] T. Overbye and S. Saripalli, "G-vom: A gpu accelerated voxel off-road mapping system," in *2022 IEEE Intelligent Vehicles Symposium (IV)*, pp. 1480–1486, 2022. 4