

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

# First, Learn What You Don't Know: Active Information Gathering for Driving at the Limits of Handling

Alexander Davydov<sup>1,2</sup>, Franck Djeumou<sup>2,3</sup>, Marcus Greiff<sup>2</sup>,  
Makoto Suminaka<sup>2</sup>, Michael Thompson<sup>2</sup>, John Subosits<sup>2</sup>, and Thomas Lew<sup>2</sup>

**Abstract**—Combining data-driven models that adapt online and model predictive control (MPC) has enabled effective control of nonlinear systems. However, when deployed on unstable systems, online adaptation may not be fast enough to ensure reliable simultaneous learning and control. For example, a controller on a vehicle executing highly dynamic maneuvers—such as drifting to avoid an obstacle—may push the vehicle's tires to their friction limits, destabilizing the vehicle and allowing modeling errors to quickly compound and cause a loss of control. To address this challenge, we present an active information gathering framework for identifying vehicle dynamics as quickly as possible. We propose an expressive vehicle dynamics model that leverages Bayesian last-layer meta-learning to enable rapid online adaptation. The model's uncertainty estimates are used to guide informative data collection and quickly improve the model prior to deployment. Dynamic drifting experiments on a Toyota Supra show that (i) the framework enables reliable control of a vehicle at the edge of stability, (ii) online adaptation alone may not suffice for zero-shot control and can lead to undesirable transient errors or spin-outs, and (iii) active data collection helps achieve reliable performance.

**Index Terms**—Model Learning for Control, Intelligent Transportation Systems

## I. INTRODUCTION

CONTROLLING unstable nonlinear systems remains challenging [1]. For example, driving a vehicle through dynamic and unstable drifting maneuvers is a difficult task where slight decision-making mistakes can lead to spin outs or crashes. In recent years, advanced control methods have been developed to give autonomous vehicles the ability to leverage their full handling potential [2]–[6], with possible applications to the future design of advanced driver-assistance systems and autonomous driving safety systems.

In this setting, data-driven model predictive control (MPC) methods have gained wide popularity [7] due to their high performance, capability to use expressive learned models,

Manuscript received: April 8, 2025; Revised July 22, 2025; Accepted August 13, 2025.

This paper was recommended for publication by Editor Clément Gosselin upon evaluation of the Associate Editor and Reviewers' comments. This work was supported by Toyota Research Institute.

\* The work of A. Davydov was carried out at Toyota Research Institute.

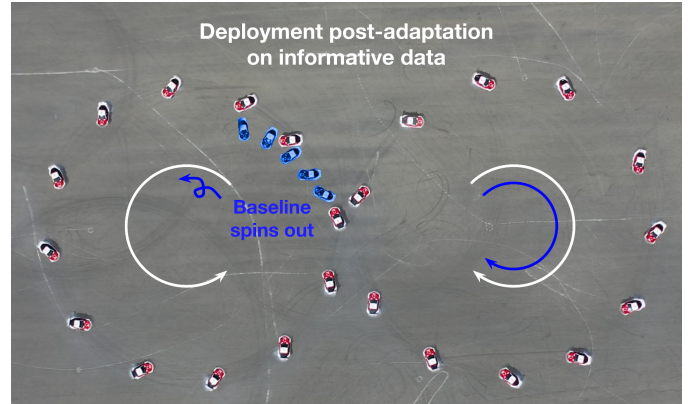
<sup>1</sup>Alexander Davydov is with the Department of Mechanical Engineering, Rice University, Houston, TX, USA [davydov@rice.edu](mailto:davydov@rice.edu)

<sup>2</sup>Alexander Davydov, Franck Djeumou, Marcus Greiff, Makoto Suminaka, Michael Thompson, John Subosits, and Thomas Lew are with Toyota Research Institute, Los Altos, CA, USA  [{marcus.greiff, makoto.suminaka, michael.thompson, john.subosits, thomas.lew}@tri.global](mailto:{marcus.greiff, makoto.suminaka, michael.thompson, john.subosits, thomas.lew}@tri.global)

<sup>3</sup>Franck Djeumou is with the Department of Mechanical, Aerospace, and Nuclear Engineering, Rensselaer Polytechnic Institute, Troy, NY, USA [djeumf2@rpi.edu](mailto:djeumf2@rpi.edu)

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

©2026 IEEE



**Fig. 1:** We propose an active information gathering framework that guides data collection to quickly learn a model in a controlled (i.e., safe) setting. Adapting on this information-rich data enables reliable control in dynamic drifting trajectories. In contrast, directly using the model without adapting prior to deployment may lead to unreliable control performance, such as spinning out.

and ability to account for constraints such as actuator limits and obstacle avoidance. In particular, data-driven models that adapt using online data can reduce model errors and sim-to-real discrepancies at deployment time [8]–[11]. However, simultaneously controlling and learning unstable systems is notoriously difficult. For instance (see Sec. V), successfully executing a drifting maneuver requires a sufficiently accurate initial model to precisely initiate the drift, and online adaptation alone may not be sufficient to execute the maneuver.

Autonomous driving at the limits presents a challenge for data-driven MPC with online adaptation due to its fast-paced nature and the inherent instability of the executed maneuvers. These difficulties raise the following questions: How can we effectively identify uncertain nonlinear dynamics to ensure that subsequent control is reliable? How should informative data be collected efficiently and safely to minimize time-intensive and potentially expensive testing on hardware?

**Contributions:** We present an active information gathering framework to quickly identify nonlinear dynamics for reliable subsequent deployment in MPC. Specifically:

- We present a data-driven dynamics model tailored for driving at the limits of handling. This model combines prior physics knowledge with learned expressive neural network features. By adapting only the last layer, this data-driven model enables rapid online adaptation, significantly improving upon non-adaptive models in controlling unstable systems where small initial errors can quickly amplify.

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

- To enable reliable deployment of the learned model in an MPC framework, we propose a pre-deployment active learning approach to identify an accurate model as fast as possible while satisfying constraints during data collection.
- We extensively validate the proposed approach on a Toyota Supra in two different dynamic drifting maneuvers. In particular, results show that online adaptation of the meta-trained model alone may be insufficient to initiate a drift reliably and can lead to large transients in tracking error. On the other hand, the proposed active information gathering approach enables rapid identification of the system's dynamics and yields improved tracking performance.

From a vehicle control perspective, our results demonstrate that executing drifting maneuvers is possible by adapting a model on informative non-drifting trajectories. Specifically, information-rich trajectories do not necessarily involve drifting, and adapting on such data significantly improves the execution of challenging drifting maneuvers.

## II. RELATED WORK

Data-driven models are increasingly popular for control due to their ability to leverage expressive models and adapt them online given measurements from the system, see [7] and [12] for recent reviews. For example, model predictive control (MPC) approaches leveraging neural network models have been widely used in recent years for dynamic driving at the limits of handling [4], [9]–[11]. These approaches adapt models online with the implicit assumption that the model will adapt fast enough during the control task to yield satisfactory performance, which may be difficult for systems where small initial errors quickly amplify.

*Online learning* and control algorithms that *actively* seek informative data to improve the model [13]–[20] often result in improved control performance compared to passively adapting a model while performing a task, and allow enforcing safety constraints during the data collection process [15], [20]. Leveraging such active learning approaches for controlling unstable nonlinear systems remains challenging. Indeed, the tradeoffs between seeking informative data (exploration) and performing the task (exploitation) are well-known; unstable systems are particularly difficult to simultaneously learn and control [21], [22]. These tradeoffs motivate the design of a framework where active learning is done *prior to deployment* in a controlled setting to learn a model for control with minimal experimentation.

Autonomous driving at the limits of handling exemplifies these tradeoffs. In this application, the vehicle may operate with saturated tire forces close to the road boundaries, and it is desirable to identify the vehicle dynamics as accurately as possible prior to deployment. Previous approaches to driving at the limits leverage MPC with physics-based [5], [6] and neural network-based [4], [9], [11] dynamics models. Recently, [10] introduced a neural SDE dynamics model for MPC where the parameters of the SDE are sampled from a diffusion model conditioned on online observations, enabling *passive* online adaptation during task execution. Our proposed active learning framework complements these works, enabling the

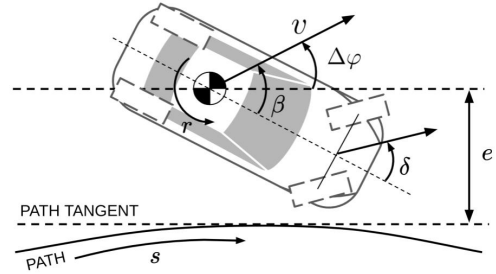


Fig. 2: Curvilinear coordinate system of the vehicle, from [11].

identification of uncertain dynamics as fast as possible prior to deployment in model-based control.

Active learning approaches to nonlinear system identification include methods that maximize information gain [16], [20], [23], [24]. Our approach relies on similar information-theoretic ideas to identify uncertain dynamics as quickly as possible. We leverage a Bayesian modeling approach [25] to last-layer meta-learning [25]–[29] that is conducive to information gathering [20]. Compared to driving at the limits of handling, experiments in [16], [20], [24] involve simulated or slow-paced robotic systems that do not face the same challenges of controlling and learning unstable dynamics. [23] also study active exploration for off-road driving, but use probabilistic ensemble neural networks that are computationally expensive compared to last-layer meta-trained neural network models, and penalize the vehicle sideslip which prevents executing drifting maneuvers. Additionally, as we show in Sec. V, online adaptation alone may not suffice to reliably execute drifting maneuvers, even using a meta-learned model capable of rapid adaptation.

## III. BAYESIAN VEHICLE DYNAMICS MODEL

Central to our approach is a vehicle dynamics model that is expressive, captures uncertainty, and enables rapid online adaptation. To this end, we combine a physics-based nominal bike model with a Bayesian last-layer meta-learning model. The former encodes prior knowledge about the vehicle, while the latter captures model mismatch and enables rapid adaptation over expressive neural network features. The vehicle state is expressed in curvilinear coordinates [5], [11], see Fig. 2. As in [5], [11], we do not use brake inputs in this work. The states and control inputs of the vehicle are

$$x = (r, v, \beta, \omega_r, e, \Delta\varphi) \in \mathbb{R}^n, \quad u = (\delta, \tau) \in \mathbb{R}^m, \quad (1)$$

where  $(n, m) = (6, 2)$  are the state and control dimensions,  $r$  is the yaw rate,  $v$  is the total planar velocity,  $\beta$  is the sideslip,  $\omega_r$  is the rear wheelspeed,  $e$  and  $\Delta\varphi$  are the lateral and angle deviations to the reference trajectory,  $\delta$  is the steering angle, and  $\tau$  is the engine torque, respectively.

The vehicle is modeled with discrete-time dynamics as

$$x_{k+1} = f_\xi(x_k, \tilde{u}_k) + \epsilon_k, \quad (2)$$

where  $k \in \mathbb{N}$  and  $\epsilon_k \sim \mathcal{N}(0, \text{diag}(\sigma_1^2, \dots, \sigma_n^2))$  are Gaussian-distributed disturbances that are independent over time  $k$  and state dimensions.  $\xi$  represents unknown parameters of the

**IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.**

dynamics (e.g., corresponding to unmodeled phenomena) and  $\tilde{u}_k = (u_k, \dot{u}_k)$  includes the control input rate  $\dot{u}_k = (\dot{\delta}_k, \dot{\tau}_k)$  at time  $k$ . Including  $\dot{u}_k$  in  $\tilde{u}_k$  and (2) accounts for the fact that MPC plans  $(u_0, u_1, \dots, u_N)$  are executed by a low-level controller via linear interpolation on the control horizon. That is,  $u(k\Delta t + \tau) = u_k + \tau\dot{u}_k$  with a time step  $\Delta t$  and  $\dot{u}_k = (u_{k+1} - u_k)/\Delta t$  for  $\tau \in [0, \Delta t)$ .

We approximate the dynamics (2) using the model

$$x_{k+1} = h(x_k, u_k) + g(z_k) + \epsilon_k, \quad (3)$$

where  $h$  is the  $\Delta t$ -time step Euler integration of a standard bike model (see Appendix A for details). The residual  $g$  in (3) is modeled by  $g(z) = (g_1, \dots, g_n)(z)$  with inputs

$$z_k = ((r_k, v_k, \beta_k, \omega_{r,k}), u_k, u_{k+1}),$$

and each dimension  $i$  of  $g(\cdot)$  is linear in the learned features  $\phi_i(z) \in \mathbb{R}^d$  for some parameters  $\theta_i \in \mathbb{R}^d$ , that is

$$g_i(z_k) := \theta_i^\top \phi_i(z_k), \quad i \in \{1, \dots, n\}. \quad (4)$$

We omit the variables  $e$  and  $\Delta\varphi$  since they do not affect the dynamics of the other state variables, see [5]. That is, given the state in (1), we let  $(g_5, g_6)(z) = 0$  for all  $z$ .

The features  $\phi_i(\cdot)$  are neural networks that are learned jointly with the parameters  $\theta_i$  via *last-layer Bayesian meta-learning* [20], [25], [30] to yield an expressive model capable of rapid adaptation with uncertainty estimates. Specifically, viewing  $g_i(z_k)$  as a neural network, the last-layer parameter  $\theta_i$  are parameterized using a Gaussian distribution

$$\theta_i \sim \mathcal{N}(\bar{\theta}_i, \sigma_i^2 \Lambda_i^{-1}), \quad i \in \{1, \dots, n\}, \quad (5)$$

with mean parameters  $\bar{\theta}_i$  and positive definite precision matrices  $\Lambda_i/\sigma_i^2$ . Following [20], we model independent Gaussian distributions of each  $\theta_i$ , such that the joint distribution for the last layer parameters  $(\theta_1, \dots, \theta_n)$  are Gaussian distributed.

*Computing the posterior:* The linear structure of  $g_i(z)$  enables the computation of the last-layer parameters  $\theta_i$  via Bayesian linear regression. Specifically, given prior distributions  $\mathcal{N}(\bar{\theta}_{i,0}, \sigma_i^2 \Lambda_{i,0}^{-1})$  for  $\theta_i$ , and a state-control trajectory  $\mathcal{D}_T = \{(z_k, x_{k+1})\}_{k=0}^T$ , the posterior  $\theta_{i,T} | \mathcal{D}_T$  is Gaussian and can be computed recursively from  $Q_{i,0} = \Lambda_{i,0} \bar{\theta}_{i,0}$  by

$$\Lambda_{i,k+1}^{-1} = \Lambda_{i,k}^{-1} - \frac{(\Lambda_{i,k}^{-1} \phi_i(z_k))(\Lambda_{i,k}^{-1} \phi_i(z_k))^\top}{1 + \phi_i(z_k)^\top \Lambda_{i,k}^{-1} \phi_i(z_k)}, \quad (6a)$$

$$Q_{i,k+1} = Q_{i,k} + (x_{i,k+1} - h_i(x_k, u_k)) \phi_i(z_k), \quad (6b)$$

$$\bar{\theta}_{i,k+1} = \Lambda_{i,k+1}^{-1} Q_{i,k+1}. \quad (6c)$$

Moreover, the one-step predictions are Gaussian-distributed, that is,  $x_{i,k+1} | x_{0:k}, \tilde{u}_{0:k} \sim \mathcal{N}(\mu_{i,k+1}, \Sigma_{i,k+1})$  with

$$\mu_{i,k+1} = h_i(x_k, u_k) + \bar{\theta}_{i,k}^\top \phi_i(z_k), \quad (7a)$$

$$\Sigma_{i,k+1} = \sigma_i^2 (1 + \phi_i(z_k)^\top \Lambda_{i,k}^{-1} \phi_i(z_k)). \quad (7b)$$

With the proposed model, the mismatch between the true dynamics  $f$  and the nominal model  $h$  through  $g$  is learned such that the one-step-ahead predictions are Gaussian by design, facilitating both meta-learning and information gathering.

*Offline Bayesian meta-training:* The prior parameters  $(\bar{\theta}_{i,0}, \Lambda_{i,0}^{-1})$ , features  $\phi_i$ , and noise covariance scalings,  $\sigma_i^2$ , are

pre-trained offline to obtain an expressive model that can be adapted online using the update rule (6). Specifically, given a dataset of  $J$  trajectories corresponding to different parameters  $\xi$  of the true system (2) and a meta-training horizon  $T$ , we minimize the negative posterior log-likelihood objective

$$\mathcal{L}_{\text{all}}(\bar{\theta}, \Lambda, \phi, \sigma) = \sum_{j=1}^J \sum_{k=1}^T \|x_k^j - \mu_k^j\|_{(\Sigma_k^j)^{-1}}^2 + \log(\det(\Sigma_k^j)),$$

where  $\|y\|_A^2 = y^\top A y$  and  $(\mu_k^j, \Sigma_k^j)$  denotes the one-step-ahead prediction computed by (6) at a time  $k-1$  in dataset  $j$ . The fact that the posterior parameters are obtained via the update rule is key to learning expressive features  $\phi$  that favor rapid online adaptation of the model. For further details on Bayesian meta-learning, we refer to [20], [25], [30].

*Implementation details:* We process vehicle state data at 10Hz and thus define the discrete-time dynamics (2) and the model (3) with  $\Delta t = 0.1$ s. We found that the bike model gives poor estimates of wheelspeed dynamics, and thus learn the full wheelspeed dynamics by setting  $h_4(x_k, u_k) = 0$ . We train the model using code from [30].

As a dataset, we train on a mix of successful and unsuccessful manual and autonomous drifting data totaling approximately 19 minutes of driving. This data corresponds to 35 different trajectories, i.e., realizations of  $f_\xi$ , and contains a total of 11382 data points. Following [11], we additionally augment our dataset with a mirrored set of data by negating the signs of the yaw rate, sideslip, and steering angle to simulate turning in the opposite direction. This doubles our dataset size.

For our architecture, we use a tanh-feedforward neural network for each  $\phi_i$  with two hidden layers of width 128, and the dimension of each  $\theta_i$  is 16. Following [20], we share weights and biases across each  $\phi_i$  except for their last layer weights and biases  $(W_{\text{out}}^i, b_{\text{out}}^i) \in \mathbb{R}^{16 \times 128} \times \mathbb{R}^{16}$ . In total, there are 27012 trainable parameters for the neural network, prior parameters,  $(\bar{\theta}_{i,0}, \Lambda_{i,0}^{-1})$ , and noise covariance scalings,  $\sigma_i^2$ . We use a meta-training horizon  $T = 10$  and train our model using Adam with an initial learning rate of  $10^{-3}$  and an exponential decay rate of 0.9975 over 1000 epochs.

#### IV. INFORMATION GATHERING MPC FRAMEWORK

We present an active information gathering framework, shown in Fig. 3, to identify nonlinear dynamics as fast as possible for reliable subsequent deployment within MPC. First, we meta-train a Bayesian vehicle model offline (Sec. III) using a dataset of vehicle trajectories. Second, we leverage the model's expressive feature space  $\phi$  and uncertainty representation encoded by  $(\bar{\theta}, \Lambda^{-1})$  to safely collect informative data and refine the model. After this information-guided adaptation, we leverage the refined vehicle model within an MPC framework to control the vehicle in challenging driving scenarios such as drifting. Next, we formulate the optimal control problems used for MPC and information gathering.

We define an MPC formulation that enables tracking dynamic trajectories such as drifting maneuvers parameterized by a reference trajectory  $x_{\text{ref}}$ . To favor smooth control inputs, we penalize fast changes in the control inputs and define

$$\ell_k(x, u, w) = \|x - x_{\text{ref},k}\|_Q^2 + \|u - w\|_R^2,$$

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

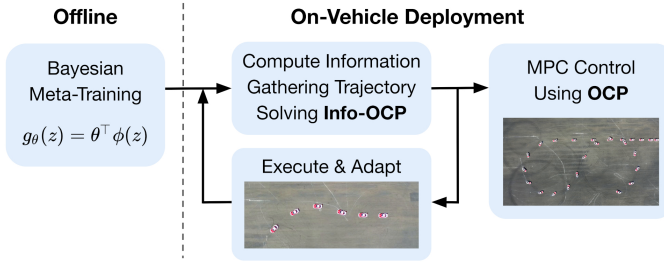


Fig. 3: Proposed active information gathering framework.

where  $(Q, R)$  are positive semidefinite diagonal matrices. We impose constraints to ensure that the vehicle remains on the track and that actuators operate within hardware constraints. To plan over a horizon  $N$  from an initial state  $x_{\text{init}}$ , we formulate an optimal control problem (**OCP**)

$$\min_{x,u} \sum_{k=0}^{N-1} \ell_k(x_k, u_k, u_{k+1}) \quad \boxed{\text{OCP}} \quad (8a)$$

$$\text{s.t. } x_{k+1} = f_{\bar{\theta}}(x_k, u_k, u_{k+1}), \quad k \in \{0, \dots, N-1\}, \quad (8b)$$

$$x_0 = x_{\text{init}}, \quad (8c)$$

$$u_{\min} \leq u_k \leq u_{\max}, \quad k \in \{1, \dots, N\}, \quad (8d)$$

$$e_{\min,k} \leq e_k \leq e_{\max,k}, \quad k \in \{1, \dots, N\}. \quad (8e)$$

The quality of solutions to **OCP** depend on the accuracy of the model  $f_{\bar{\theta}}$  parameterized by the last-layer parameters  $\bar{\theta}$ :

$$f_{\bar{\theta}}(x_k, u_k, u_{k+1}) = h(x_k, u_k) + g(z_k; \bar{\theta}),$$

where  $g_i(z_k; \bar{\theta}) = \bar{\theta}_i^\top \phi_i(z_k)$  for each  $i = 1, \dots, n$ .

Directly deploying an MPC controller that recursively solves **OCP** using the prior model parameters  $\bar{\theta}_{i,0}$  after offline meta-training may lead to poor performance in practice, even if one adapts the model online during the task, see Sec. V. These observations motivate an active data collection approach to refine the model prior to deployment in challenging applications. Since data collection on a vehicle may be time-consuming and expensive, this raises the question: *What data should we collect to best identify the unknown dynamics?*

The key insight in answering this question is that learning the model as fast as possible amounts to maximizing the *information gain* from future observations, i.e., to maximizing the *mutual information* between observations and the true system (see [31, Chapter 2] and [32, Chapter 8]). The mutual information between two random variables characterizes their mutual dependence and provides a quantitative measure for selecting data that is most informative in identifying the true dynamics (2). Thus, to steer the system towards regions of the state-input space with high model uncertainty, we define the information gathering objective

$$\ell_{\text{info}}(x, u) := \frac{1}{2} \sum_{k=0}^{N-1} \sum_{i=1}^n \log(1 + \phi_{i,k}^\top \Lambda_{i,0}^{-1} \phi_{i,k}), \quad (9)$$

where  $\phi_{i,k} := \phi_i(x_k, u_k, u_{k+1})$ . Using  $\ell_{\text{info}}$ , we define the information gathering optimal control problem

$$\min_{x,u} -\alpha \ell_{\text{info}}(x, u) + \sum_{k=0}^{N-1} \ell_k(x_k, u_k, u_{k+1}) \quad \boxed{\text{Info-OCP}} \\ \text{s.t. } (8b) - (8e), \quad (10)$$

where  $\alpha > 0$  weighs the information gain compared to the nominal MPC objective. The nominal costs  $\ell_k$  encourage smooth control inputs, and the constraints (8b)-(8e) ensure that the vehicle remains on the track during data collection.

The information gathering objective (9) is derived in [20] by leveraging the last-layer uncertainty representation of the model,  $g$ , which yields a closed-form expression of the one-step mutual information between an observation and the dynamics. Similar objectives are used for active learning in [16], [33], [34]. As discussed in [20],  $\ell_{\text{info}}$  is an approximation of the true total information gain over a trajectory, which is not exactly the sum of the expected information gains per timestep. We found that incorporating the linear regression updates (6) in the objective (by defining the cost  $\bar{\ell}_{\text{info}} = \frac{1}{2} \sum_{k=0}^{N-1} \sum_{i=1}^n \log(1 + \phi_{i,k}^\top \Lambda_{i,k}^{-1} \phi_{i,k})$  instead) does not result in a substantial difference in computed trajectories, while resulting in a more challenging numerical resolution of the resulting information gathering problem.

The information gathering objective (9) guides data collection toward regions of the feature space that enable rapid adaptation. As shown in Fig. 3, the procedure of generating information gathering trajectories and adapting the model on them can be iterated as many times as needed to lower estimates of uncertainty. In our experiments, we found that adaptation on only one or two information gathering trajectories is sufficient for performing reliable drifting maneuvers, as shown in the next section.

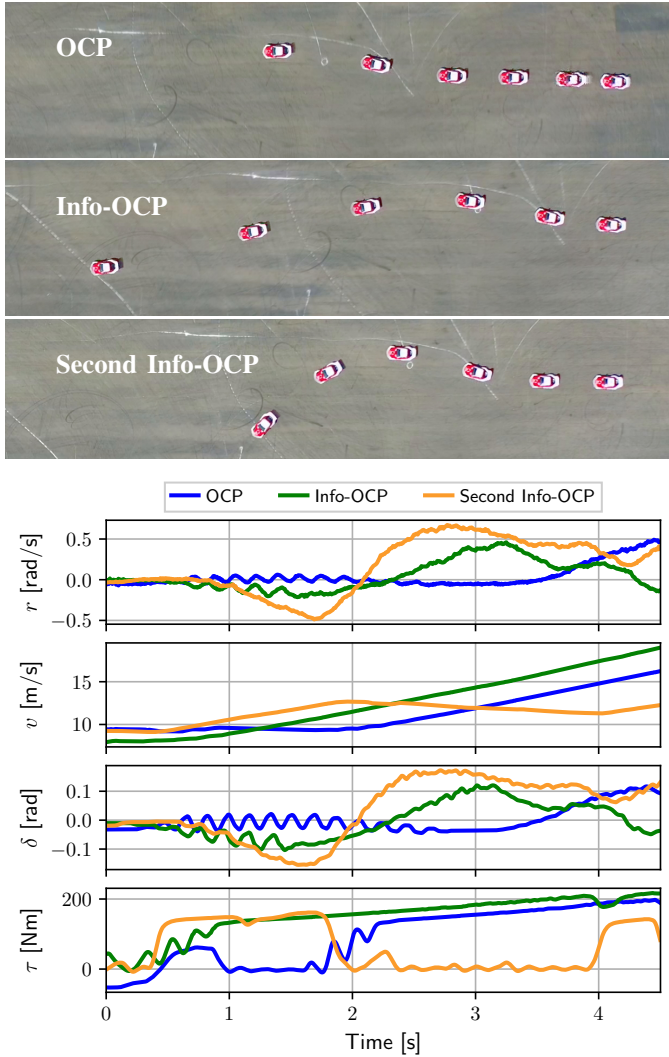
## V. RESULTS

We validate the proposed active information gathering framework on a Toyota Supra. The states are estimated at a rate of 100Hz by an OxTS Inertial Navigation System [35], and the MPC problems are solved at approximately 100Hz on an on-board computer equipped with a 3.30GHz Intel Xeon E-2278GE CPU using a standard line search sequential quadratic programming method [36, Chapter 18] where quadratic programming subproblems are solved with the OSQP solver [37]. Reference trajectories for drifting are generated by the method in [38]. All videos of our experiments are available at <https://tinyurl.com/d4pnh8kc>.

### A. What do information-rich trajectories look like?

We compare solutions to **OCP** with those to **Info-OCP**. We define the problems with  $x_{\text{init}} = (0, 7.5, 0, 7.5, 0, 0)$ , a straight-line reference trajectory  $x_{\text{ref}}$  from  $x_{\text{init}}$  to  $x_{\text{ref},N} = (0, 20, 0, 20, 0, 0)$ , and a planning horizon  $N = 45$ , corresponding to 4.5s-long trajectories. Solutions to **OCP** and **Info-OCP** are executed on the vehicle using a low-level tracking controller. We record the trajectories and use them to adapt the Bayesian vehicle model using (6). We refer to (i) the posterior model adapted on the nominal trajectory from **OCP** as Once-OCP-Model and (ii) the posterior model adapted on the information gathering trajectory from **Info-OCP** as Once-Info-Model. We then use Once-Info-Model to solve **Info-OCP** again using the adapted model parameters to generate a second information gathering trajectory. We refer to the posterior model adapted on both the first and second

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.



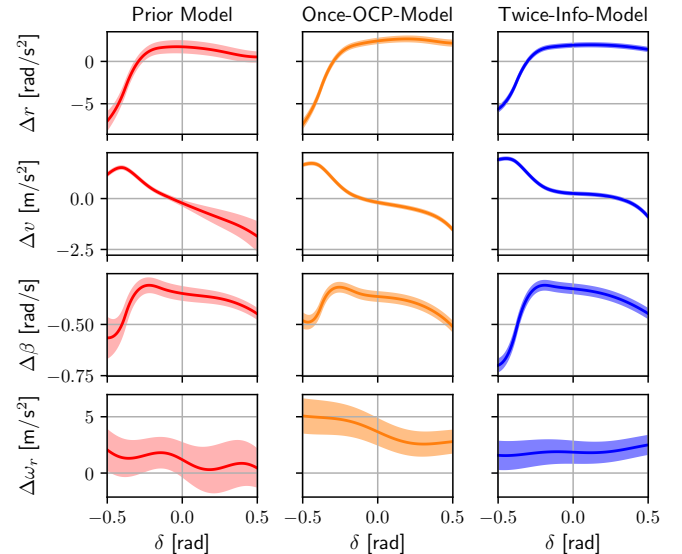
**Fig. 4:** Nominal and information gathering trajectories. (Top) Overhead view of the trajectories executed on the vehicle and (Bottom) corresponding yaw rate, velocity, and control inputs. The duration of all executed trajectories is 4.5 seconds.

**TABLE I:** Covariance norms  $\|\Lambda_i^{-1}\|_2$  for the prior model, Once-OCP-Model (1-OCP-Model), Once-Info-Model (1-Info-Model), and Twice-Info-Model (2-Info-Model).

$\ \Lambda_i^{-1}\ _2$	Prior	1-OCP-Model	1-Info-Model	2-Info-Model
$r$	3.05	1.43	0.81	0.49
$v$	27.47	1.36	1.56	0.52
$\beta$	1.89	0.56	0.44	0.31
$\omega_r$	1.57	0.43	0.47	0.39
Total	33.98	3.79	3.24	1.71

information gathering trajectories as Twice-Info-Model. Vehicle trajectories are shown in Figs. 4. Table I shows the model covariance values  $\|\Lambda_i^{-1}\|_2$  prior to adaptation and after adaptation on each trajectory, and Fig. 5 shows the models' predictions for a relevant region of the statespace.

In Fig. 4, we see that although the model is meta-trained only on drifting data, solutions to **Info-OCP** are not exactly drifting trajectories. Solutions to **OCP** roughly follow the straight-line reference trajectory, whereas the solution to



**Fig. 5:** Predictions of the model pre- and post-adaptation on the information gathering trajectories (see Fig. 4). Mean predictions  $\pm 3$  standard deviations for different steering angles  $\delta_k = \delta_{k+1}$  at  $(r_k, v_k, \beta_k, \omega_{r,k}, \tau_k, \tau_{k+1}) = (0.9, 14.5, -0.45, 21, 200, 200)$ .  $\Delta r$  denotes  $(r_{k+1} - r_k)/\Delta t$  and similarly for other state variables.

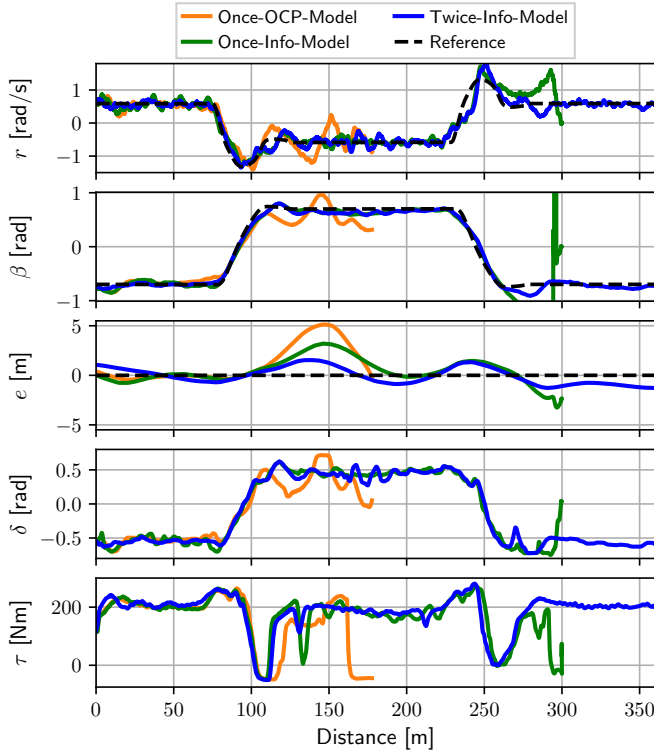
**Info-OCP** applies larger engine torques and steers in each direction to excite the dynamics and provide informative data: Table I shows that following the solution to **Info-OCP** helps identify dynamics quicker than straight-line driving. Finally, after adapting on the first informative trajectory, the solution to **Info-OCP** applies even larger engine torques and steers the vehicle in each direction to reach larger yaw rates (see Fig. 4). This helps to reduce uncertainty over the yaw rate and sideslip dynamics (see Table I). Intuitively, reaching higher yaw rates may help saturate tire forces and identify uncertain tire properties. Although solutions to **Info-OCP** are task-agnostic, identifying tire dynamics may be key to subsequent deployment in dynamic drifting tasks.

### B. Information gathering enables reliable drifting

Next, we consider the task of drifting a Figure-8 trajectory, see Fig. 1. This maneuver is challenging due to the fast and unstable transients during transitions between the circular arcs of the trajectory, shown in grey in Fig. 6.

We evaluate the performance of three models for MPC: Once-OCP-Model, Once-Info-Model, and Twice-Info-Model. To assess the value of the proposed information gathering procedure in isolation, we do not perform online adaptation in this experiment. Each model is given three attempts to drift along the Figure-8, and we present the median tracking performance of each model in Fig. 6 (additional statistics are reported in Appendix B). Tracking performance is measured in terms of the root mean squared lateral error along the reference.

We observe that Once-OCP-Model is unable to complete the drift and deviates too far from the reference trajectory after the first transition. This undesirable large deviation is a result of the car transitioning from slipping to gripping after the first transition and then being unable to re-initiate drifting. In contrast, Once-Info-Model successfully completes



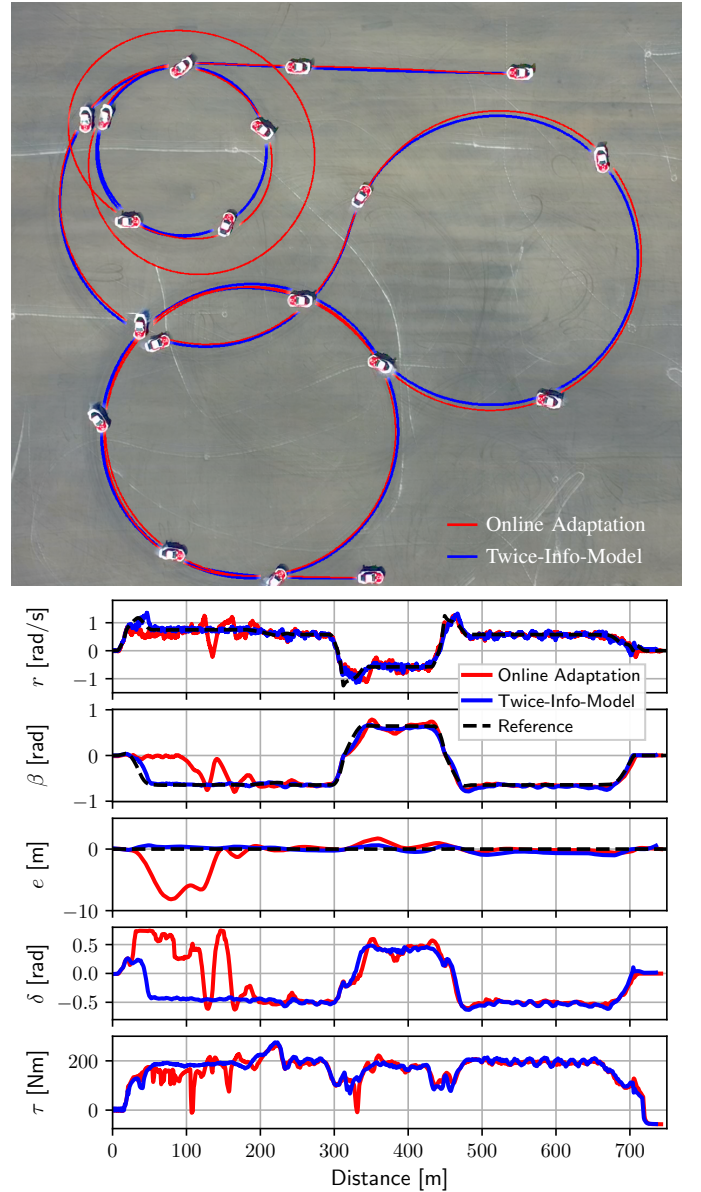
**Fig. 6:** Median Figure-8 drifting performance for the prior model after adapting on the nominal and informative trajectories (Fig. 4). Online adaptation is disabled for all models. Transitions are highlighted in grey. Only Twice-Info-Model successfully completes the full maneuver, also shown overhead in Fig. 1.

the first transition but spins out after the second transition, whereas Twice-Info-Model is able to complete the full Figure-8 maneuver, succeeding every time.

Performing at least one round of information gathering before deployment helps in completing the first transition. We hypothesize that this is because adapting on trajectories from **Info-OCP** results in better estimates of yaw rate dynamics compared to adapting on the nominal **OCP** trajectory. Indeed, Fig. 4 shows that the **Info-OCP** trajectories excite yaw rate dynamics more significantly than the **OCP** trajectory does. Moreover, Table I shows that adapting on **Info-OCP** trajectories results in smaller yaw rate covariance norms compared to adapting on the nominal **OCP** trajectory, suggesting more accurate yaw rate dynamics estimates after adapting on **Info-OCP** trajectories.

### C. Online adaptation may not be enough for reliable control

Having established the utility of the information gathering and offline adaptation on trajectories generated via **Info-OCP**, we now evaluate the efficacy of online adaptation of the model through the last-layer update rule in (6) on a loops maneuver shown in Fig. 7. This maneuver is challenging as the drift is initiated by the MPC controller, in contrast to the hand-brake initiation done in the Figure-8 maneuver. Thus, to successfully initiate the drift, it is essential that the model precisely describes the challenging initial transition between gripping and sliding.



**Fig. 7:** Best performance in the loops maneuver for the prior model with online adaptation and the model adapted on informative trajectories. (Top) Overhead view and (Bottom) time-series data of key signals. Online adaptation is disabled in the Twice-Info-Model.

We compare the performance of the prior model adapting online to the model previously adapted on the two information gathering trajectories from Sec. V-A. We test each of these models twice and report the best performance of each model in Fig. 7.

Fig. 7 shows that the model with online adaptation enables accurate tracking of this challenging trajectory. However, while it successfully completes the maneuver and eventually yields accurate tracking, MPC with online adaptation alone leads to large initial errors, requiring upward of 150 meters to initiate the drift properly. This poor transient performance may be problematic for safety-critical applications of controllers for driving at the limits, e.g., for emergency obstacle avoidance [39].

In contrast, using the model that has adapted offline on the information gathering trajectories enables seamlessly drifting

**IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.**

across all transitions, as shown in the in the plots of the tracking errors in position, yaw rate, and sideslip angle in Fig. 7. We postulate that the prior model does not capture the transition between gripping and sliding accurately enough to properly initiate the drift. Although the model is meta-trained to be capable of rapid adaptation with little data, results show that online adaptation alone from the prior model does not suffice to initiate the drift seamlessly and leads to a large transient response. Information-rich data collection prior to deployment clearly helps in reliably performing this difficult maneuver while minimizing the transient response.

## VI. CONCLUSION

We present an information gathering framework to actively and efficiently identify nonlinear dynamics for reliable subsequent deployment in MPC. Our approach hinges on an expressive vehicle dynamics model leveraging meta-learned features and Bayesian last-layer weight updates for rapid adaptation. Our results show that collecting informative data first unlocks reliable few-shot deployment within an MPC framework, and we expect that these findings translate to other challenging tasks that involve controlling unstable systems where small model errors may result in large tracking errors. This methodology is especially valuable in applications where passive online adaptation may not be fast enough to correct the initial model mismatch to complete the task successfully, precisely because unstable systems are difficult to simultaneously learn and control.

In future work, the uncertainty estimates encoded in the Bayesian model could be used in the MPC formulation to robustify constraints during exploration and task deployment [20]. These uncertainty estimates could also be used to identify when a sufficient amount of data is collected or to detect out-of-distribution behavior [40]. Finally, although the proposed method is task-agnostic and already demonstrates reliable performance in challenging drifting tasks, leveraging control-oriented identification tools [14], [28], [41], [42] may further enhance the performance of the method and further decrease the data requirements for information gathering.

## APPENDIX A

## SINGLE-TRACK BICYCLE DYNAMICS

The nominal dynamics model in (3) consists of a single-track bicycle model (see e.g. [5], [11]) given by

$$\begin{bmatrix} \dot{r} \\ \dot{v} \\ \dot{\beta} \\ \dot{\omega}_r \\ \dot{\Delta\varphi} \end{bmatrix} = \begin{bmatrix} (aF_{yf} \cos(\delta) + aF_{xf} \sin(\delta) - bF_{yr})/I_z \\ (-F_{yf} \sin(\delta - \beta) + F_{xf} \cos(\delta - \beta) + F_{yr} \sin(\beta) + F_{xr} \cos(\beta))/m \\ -r + (F_{yf} \cos(\delta - \beta) + F_{xf} \sin(\delta - \beta) + F_{yr} \cos(\beta) - F_{xr} \sin(\beta))/(mv) \\ (G\tau - F_{xr}R_w)/I_w \\ v \sin(\Delta\varphi) \\ \dot{\beta} + r - \kappa_{\text{ref}}v \cos(\Delta\varphi)/(1 - \kappa_{\text{ref}}e) \end{bmatrix},$$

where  $(a, b)$  are the distances from the center of gravity to the front and rear axles, respectively,  $(I_z, m)$  are the vehicle inertia and mass, respectively,  $(R_w, I_w)$  are the wheel radius and the rear axle inertia, respectively,  $G$  is the gear ratio from engine

**TABLE II:** Root mean squared (RMS) lateral deviation for Once-OCP-Model, Once-Info-Model, and Twice-Info-Model over three attempts of drifting the Figure-8 trajectory. Units are [m].

RMS( $e$ )	Once-OCP-Model	Once-Info-Model	Twice-Info-Model
Best	2.31	1.18	<b>0.61</b>
Median	3.19	1.31	<b>0.83</b>
Worst	7.35	3.64	<b>1.29</b>

**TABLE III:** Root mean squared (RMS) sideslip angle error  $e_\beta = \beta - \beta_{\text{ref}}$  for Once-OCP-Model, Once-Info-Model, and Twice-Info-Model over three attempts of drifting the Figure-8 trajectory. Units are [rad].

RMS( $e_\beta$ )	Once-OCP-Model	Once-Info-Model	Twice-Info-Model
Best	0.137	0.105	<b>0.068</b>
Median	0.261	0.348	<b>0.070</b>
Worst	0.384	0.526	<b>0.084</b>

to rear axle,  $\kappa_{\text{ref}}$  is the curvature of the reference trajectory, and  $(F_{yf}, F_{xf}, F_{yr}, F_{xr})$  denote tire forces and are modeled using an isotropic coupled slip brush Fiala model [43]. Following [5], since we test on a rear-wheel drive vehicle, we set  $F_{xf} = 0$ .

## APPENDIX B

## ADDITIONAL STATISTICS ON FIGURE-8 TRAJECTORY

We report additional statistics to compare the performance of Once-OCP-Model, Once-Info-Model, and Twice-Info-Model for the task of drifting the Figure-8 trajectory from Sec. V-B. In Table II, we present the root mean square (RMS) lateral deviation  $e$  over three trials for each of the three models. We also report the RMS sideslip angle error,  $e_\beta = \beta - \beta_{\text{ref}}$ , in Table III. Since neither Once-OCP-Model nor Once-Info-Model successfully completed the entirety of the Figure-8, we truncated the RMS computation before the spin out or the manual takeover for these models. Twice-Info-Model always achieves the lowest RMS errors and is also the only model that completes the entirety of the Figure-8 trajectory without ever spinning out.

## ACKNOWLEDGMENTS

We thank Phung Nguyen, Steven Goldine, and William Kettle for their support with the test platform and experiments.

## REFERENCES

- [1] G. Stein, "Respect the unstable," *IEEE Control Systems Magazine*, vol. 23, no. 4, pp. 12–25, 2003.
- [2] E. Velenis, D. Katzourakis, E. Frazzoli, P. Tsiotras, and R. Happee, "Steady-state drifting stabilization of RWD vehicles," *Control Engineering Practice*, vol. 19, no. 11, pp. 1363–1376, 2011.
- [3] M. Cutler and J. P. How, "Autonomous drifting using simulation-aided reinforcement learning," in *Proc. IEEE Conf. on Robotics and Automation*, pp. 5442–5448, 2016.
- [4] F. Djeumou, J. Y. M. Goh, U. Topcu, and A. Balachandran, "Autonomous drifting with 3 minutes of data via learned tire models," in *Proc. IEEE Conf. on Robotics and Automation*, pp. 968–974, 2023.
- [5] J. Y. M. Goh, M. Thompson, J. Dallas, and A. Balachandran, "Beyond the stable handling limits: Nonlinear model predictive control for highly transient autonomous drifting," *Vehicle System Dynamics*, pp. 1–24, 2024.

**IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.**

- [6] T. P. Weber and J. C. Gerdes, "Modeling and control for dynamic drifting trajectories," *IEEE Transactions on Intelligent Vehicles*, vol. 9, no. 2, pp. 3731–3741, 2024.
- [7] L. Hewing, K. P. Wabersich, M. Menner, and M. N. Zeilinger, "Learning-based model predictive control: Toward safe learning in control," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 3, no. 1, pp. 269–296, 2020.
- [8] 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, no. 4, pp. 3363–3370, 2019.
- [9] N. A. Spielberg, M. Brown, and J. C. Gerdes, "Neural network model predictive motion control applied to automated driving with unknown friction," *IEEE Transactions on Control Systems Technology*, vol. 30, no. 5, pp. 1934–1945, 2021.
- [10] F. Djeumou, T. Lew, N. Ding, M. Thompson, M. Suminaka, M. Greiff, and J. Subosits, "One model to drift them all: Physics-informed conditional diffusion model for driving at the limits," in *Conf. on Robot Learning*, 2024.
- [11] N. Ding, M. Thompson, J. Dallas, J. Y. M. Goh, and J. Subosits, "Drifting with unknown tires: Learning vehicle models online with neural networks and model predictive control," in *IEEE Intelligent Vehicles Symposium*, pp. 2545–2552, 2024.
- [12] L. Brunke, M. Greeff, A. W. Hall, Z. Yuan, S. Zhou, J. Panerati, and A. P. Schoellig, "Safe learning in robotics: From learning-based control to safe reinforcement learning," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 5, no. 1, pp. 411–444, 2022.
- [13] Y. Abbasi-Yadkori, D. Pal, and C. Szepesvári, "Improved algorithms for linear stochastic bandits," in *Conf. on Neural Information Processing Systems*, 2011.
- [14] A. Wagenmaker, G. Shi, and K. G. Jamieson, "Optimal exploration for model-based RL in nonlinear systems," *Conf. on Neural Information Processing Systems*, 2024.
- [15] T. Koller, M. T. F. Berkenkamp, and A. Krause, "Learning-based model predictive control for safe exploration," in *Proc. IEEE Conf. on Decision and Control*, pp. 6059–6066, 2018.
- [16] M. Buisson-Fenet, F. Solowjow, and S. Trimpe, "Actively learning Gaussian process dynamics," in *Conf. on Learning for Dynamics and Control*, pp. 5–15, 2020.
- [17] H. Mania, M. I. Jordan, and B. Recht, "Active learning for nonlinear system identification with guarantees," *Journal of Machine Learning Research*, vol. 23, no. 32, pp. 1–30, 2022.
- [18] A. Saviolo, J. Frey, A. Rathod, M. Diehl, and G. Loianno, "Active learning of discrete-time dynamics for uncertainty-aware model predictive control," *IEEE Transactions on Robotics*, vol. 40, pp. 1273–1291, 2024.
- [19] W. Zhang, M. Tognon, L. Ott, R. Siegart, and J. Nieto, "Active model learning using informative trajectories for improved closed-loop control on real robots," in *Proc. IEEE Conf. on Robotics and Automation*, pp. 4467–4473, 2021.
- [20] T. Lew, A. Sharma, J. Harrison, A. Bylard, and M. Pavone, "Safe active dynamics learning and control: A sequential exploration–exploitation framework," *IEEE Transactions on Robotics*, vol. 38, no. 5, pp. 2888–2907, 2022.
- [21] S. Dean, H. Mania, N. Matni, B. Recht, and S. Tu, "On the sample complexity of the linear quadratic regulator," *Foundations of Computational Mathematics*, vol. 20, no. 4, pp. 633–679, 2019.
- [22] H. Mania, S. Tu, and B. Recht, "Certainty equivalence is efficient for linear quadratic control," in *Conf. on Neural Information Processing Systems*, 2019.
- [23] T. Kim, J. Mun, J. Seo, B. Kim, and S. Hong, "Bridging active exploration and uncertainty-aware deployment using probabilistic ensemble neural network dynamics," in *Robotics: Science and Systems*, 2023.
- [24] B. Sukhija, L. Treven, C. Sancaktar, S. Blaes, S. Coros, and A. Krause, "Optimistic active exploration of dynamical systems," *Advances in Neural Information Processing Systems*, pp. 38122–38153, 2024.
- [25] J. Harrison, A. Sharma, and M. Pavone, "Meta-learning priors for efficient online Bayesian regression," in *Algorithmic Foundations of Robotics XIII: Proceedings of the 13th Workshop on the Algorithmic Foundations of Robotics*, pp. 318–337, 2020.
- [26] K. Lee, S. Maji, A. Ravichandran, and S. Soatto, "Meta-learning with differentiable convex optimization," in *IEEE/CVF Conf. on Computer Vision and Pattern Recognition*, pp. 10657–10665, 2019.
- [27] L. Bertinetto, J. F. Henriques, P. H. S. Torr, and A. A. Vedaldi, "Meta-learning with differentiable closed-form solvers," in *International Conference on Learning Representations*, 2019.
- [28] S. M. Richards, N. Azizan, J.-J. Slotine, and M. Pavone, "Control-oriented meta-learning," *Int. Journal of Robotics Research*, vol. 42, no. 10, pp. 777–797, 2023.
- [29] M. O’Connell, G. Shi, X. Shi, K. Azizzadenesheli, A. Anandkumar, Y. Yue, and S.-J. Chung, "Neural-fly enables rapid learning for agile flight in strong winds," *Science Robotics*, vol. 7, no. 66, pp. 1–15, 2022.
- [30] S. Banerjee, J. Harrison, P. M. Furlong, and M. Pavone, "Adaptive meta-learning for identification of rover-terrain dynamics," in *Int. Symp. on Artificial Intelligence, Robotics and Automation in Space*, 2020.
- [31] T. M. Cover, *Elements of Information Theory*. John Wiley & Sons, 1999.
- [32] D. J. C. MacKay, *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, 2003.
- [33] A. Krause and C. Guestrin, "Nonmyopic active learning of Gaussian processes," in *Int. Conf. on Machine Learning*, 2007.
- [34] C. Zimmer, M. Meister, and D. Nguyen-Tuong, "Safe active learning for time-series modeling with Gaussian processes," in *Conf. on Neural Information Processing Systems*, 2018.
- [35] OxTS, "User manual: NCOM data format for the efficient communication of navigation measurements," 2022.
- [36] J. Nocedal and S. J. Wright, *Numerical Optimization*. Springer, 1999.
- [37] B. Stellato, G. Banjac, P. Goulart, A. Bemporad, and S. Boyd, "OSQP: An operator splitting solver for quadratic programs," *Mathematical Programming Computation*, vol. 12, no. 4, pp. 637–672, 2020.
- [38] J. Y. Goh and J. C. Gerdes, "Simultaneous stabilization and tracking of basic automobile drifting trajectories," in *IEEE Intelligent Vehicles Symposium*, pp. 597–602, 2016.
- [39] T. Zhao, E. Yurtsever, and G. Rizzoni, "Justifying emergency drift control for automated vehicles," in *IFAC Symposium on Advances in Automotive Control*, vol. 55, pp. 141–148, 2022.
- [40] R. Sinha, A. Sharma, S. Banerjee, T. Lew, R. Luo, S. M. Richards, Y. Sun, E. Schmerling, and M. Pavone, "A system-level view on out-of-distribution data in robotics," *arXiv preprint arXiv:2212.14020*, 2022.
- [41] S. Anderson and J. P. Hespanha, "Control-oriented identification via stochastic optimization," *IEEE Control Systems Letters*, vol. 8, pp. 1865–1870, 2024.
- [42] B. D. Lee, I. Ziemann, G. J. Pappas, and N. Matni, "Active learning for control-oriented identification of nonlinear systems," *arXiv preprint arXiv:2404.09030*, 2024.
- [43] J. Svendenius, *Tire modeling and friction estimation*. PhD thesis, Lund University, 2007.