

Learning dynamics for improving control of overactuated flying systems

Weixuan Zhang, Maximilian Brunner, Lionel Ott, Mina Kamel, Roland Siegwart, and Juan Nieto

Abstract—Overactuated omnidirectional flying vehicles are capable of generating force and torque in any direction, which is important for applications such as contact-based industrial inspection. This comes at the price of an increase in model complexity. These vehicles usually have non-negligible, repetitive dynamics that are hard to model, such as the aerodynamic interference between the propellers. This makes it difficult for high-performance trajectory tracking using a model-based controller. This paper presents an approach that combines a data-driven and a first-principle model for the system actuation and uses it to improve the controller. In a first step, the first-principle model errors are learned offline using a Gaussian Process (GP) regressor. At runtime, the first-principle model and the GP regressor are used jointly to obtain control commands. This is formulated as an optimization problem, which avoids ambiguous solutions present in a standard inverse model in overactuated systems, by only using forward models. The approach is validated using a tilt-arm overactuated omnidirectional flying vehicle performing attitude trajectory tracking. The results show that with our proposed method, the attitude trajectory error is reduced by 32% on average as compared to a nominal PID controller.

I. INTRODUCTION

Omnidirectional flying vehicles ([1]–[6]) are suitable for industrial inspection or interaction applications due to their decoupled translational and rotational dynamics (e.g. being able to hover at arbitrary attitudes) and the ability to exert forces and torques in arbitrary directions.

These vehicles are overactuated, resulting in a mechanical redundancy that provides the system with the necessary control authority to achieve omnidirectionality and increases their maneuverability and force/torque margin. However, the increased mechanical complexity also poses some challenges. Unlike standard quadrotors or hexacopters, for which a simple and accurate actuation model is available for low speed maneuvers [7], only an inaccurate and simplified actuation model can be obtained for omnidirectional vehicles from first principles. The actuation model is the mapping from the individual actuator commands to the total wrench, i.e., force and torque. To achieve omnidirectionality, vehicle designs often cause complex aerodynamic effects that are hard-to-model and non-negligible, such as the aerodynamic interference between the rotors. This is disadvantageous from a control perspective. A standard approach for overactuated flying vehicles is that a controller first computes a desired wrench, which, based on the inverse of the actuation model, is transformed into individual actuator commands. Due to unmodeled effects, this typically results in a different actual wrench than the intended one. This leads to degraded control performance and reduces the ability to accurately track desired trajectories.

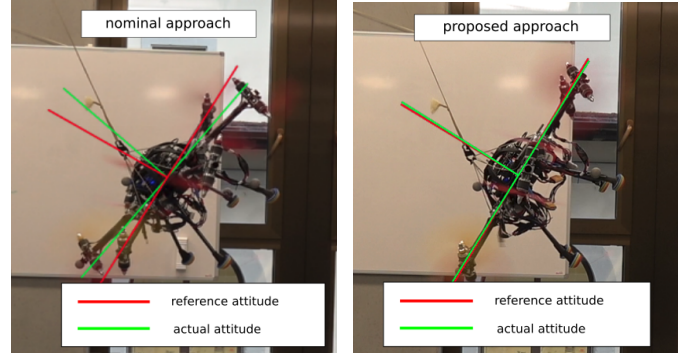


Fig. 1: Attitude tracking performance of an overactuated flying vehicle: (left) without model learning, and (right) using our proposed learning based approach. For perfect tracking the green frame should coincide with the red frame.

To solve the problem mentioned above, adding only integral actions in the feedback controller is often not enough, as they typically decrease the damping or stability of the system and are not able to react to fast-changing modeling errors. Another solution is to treat the unmodelled dynamics as disturbances [8] [9] and employ a disturbance observer. While being computationally efficient, this reactive strategy introduces delay in tracking. One alternative solution is to identify these mismatched dynamics and compute additional feedforward signals in an attempt to cancel out the effect of modeling errors. One way to model this mismatch is using data-driven machine learning models. Such models have recently been shown to be viable due to their ability to learn complex and nonlinear dynamic functions [10]. Directly learning an inverse mapping of the actuation, that is, a mapping from the wrench to the individual actuator commands, is a commonly used method. However, this approach is problematic for an overactuated system due to its one-to-many mapping between wrench and actuator inputs. The multi-modal nature of this mapping has the risk of producing invalid results when averaging over multiple distinct modes with standard inverse model approaches.

This paper presents an approach aiming to improve the control of overactuated flying systems. First, the actuation modeling error is learned offline using a Gaussian process (GP) regressor. To overcome the challenge arising due to the above-mentioned one-to-many mapping, we formulate the task of finding a suitable wrench command as an iterative optimization problem. We obtain a wrench by optimizing the commanded wrench using the analytical forward model and its learned error model. This yields in expectation the desired

behaviour of the system, necessitating no inverse models.

This paper makes the following contributions:

- Proposal of a GP-based model to capture the model-plant mismatch common among overactuated omnidirectional flying vehicles.
- Introduction of an optimization-based method using a first-principles model and a learned GP error model to select a signal correcting for the model-plant mismatch, necessitating only forward models.
- Experimental validation of the proposed approach on a tilt-arm overactuated flying vehicle called VoliroX [3] (Fig. 1), with a reduction of the attitude tracking error of 32% on average.

II. RELATED WORK

Optimal control such as the model predictive controller is shown to work well for standard multicopters [11]. However, it is shown in [12] that the model predictive controller has comparable performance to a standard PD controller in case of insufficient model knowledge.

In this section, we focus mainly on the review of approaches that learn the nonlinear, non-negligible model-plant-mismatch, which is then used to improve control performance. A broader review of model learning for control can be found in [13]. We categorize the related work by the types of model being used: forward models and inverse models. A forward model could be a static mapping (e.g. an actuator mapping from individual actuator commands to the resulting wrench in case of neglecting its dynamics) or a dynamics model (the mapping from the state and the control inputs to the derivative of the state), which are uniquely defined. The inverse model is the inverse of the forward model.

1) *Inverse model approach:* The inverse model approach aims to find a mapping that maps from the robot state to a feedforward control command. Several approaches [14]–[18] use a GP regression approach to model the inverse dynamics. The implied assumption of these approaches is that there is a unique mapping between the input and output data so that a direct regression is well-defined. Unfortunately, for an overactuated platform this is not the case.

2) *Forward model approach:* Forward model learning has been combined with iterative linear quadratic Gaussian algorithm [19], or model predictive control [20]–[22]. These methods use rollouts of the learned dynamics in an optimization-based framework to find a sequence of control actions in real-time. In [20]–[22], the model-plant mismatches are modeled using GPs and embedded into a model predictive controller. These methods are computationally expensive and thus they typically use approximation techniques to propagate the state distribution in the rollout. They experimentally validated these approaches ([20], [21]) on autonomous mobile vehicles. In [22] the authors demonstrated the proposed approach in the simulation of a quadcopter. One of the aims of our work is to validate the proposed approach on a real system.

Model-based reinforcement learning control may also be employed with a learned forward model. Among these ap-

proaches, the control policy is either found through experiments [23], which needs to make sure that the system does not damage the robots, or through simulation [10], [24], [25]. More specifically, the complete actuator model [10] or the residual dynamics ([24], [25]) are first learned using either a neural network or a GP and then embedded into the simulation. A high-performance control policy is then obtained in the simulation through reinforcement learning. These approaches typically require a large amount of training data.

Another approach is given a desired state, the control inputs are solved in an equation that includes the forward model. In [26], in order to improve the tracking performance of a quadcopter, an acceleration error model is learned using the linear regression. It is then embedded into the acceleration model equation, in which the body angular acceleration command is solved for in real-time as the control inputs. It does not take uncertainty into consideration.

III. PLATFORM DESCRIPTION

VoliroX (Fig. 1) is an overactuated omnidirectional flying vehicle with six tiltable arms in a hexagonal arrangement. A coaxial rotor configuration is rigidly attached to the end of each arm. The propellers spin in the opposite direction, which helps to reduce the gyroscopic effects introduced by the rotation of the arm and thus the tilt motor efforts. The rotation of each arm can be actively controlled by a servo motor, which results in a total of 18 actuators. This setup allows for force- and pose omnidirectionality for almost all configurations and an improved hover efficiency when compared to fixed tilt omnidirectional vehicles.

IV. MODELING

This section describes the dynamics of the flying vehicle and presents its actuation model.

A. Dynamic model

The position of the vehicle's center of mass with respect to a point fixed in the inertial coordinate system is denoted as $\mathbf{p} \in \mathbb{R}^3$. Boldface symbols are used throughout the paper to denote vectors. Two types of coordinate systems are used for the modeling: an inertial coordinate system E , a body-fixed coordinate system B . A vector expressed in a specific coordinate system is indicated by a superscript, for example \mathbf{p}^E expresses \mathbf{p} in coordinate system E . The vehicle rotates at an angular velocity $\boldsymbol{\omega}_{BE}$. The subscript BE in $\boldsymbol{\omega}_{BE}$ denotes the relative velocity of coordinate system B with respect to E .

Using a simplified physical model, the dynamics model can be written as follows [3]:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{v} \quad (1)$$

where $\mathbf{x} = [\dot{\mathbf{p}}^E, \boldsymbol{\omega}_{BE}^B]$ and $\mathbf{v} = [\mathbf{F}^E/m, (\mathbf{I}_B^B)^{-1}\mathbf{M}^B]$ with m being the vehicle mass and \mathbf{I}_B^B being the vehicle inertia matrix. The translational drag forces and rotational drag torques are neglected, as it is assumed that the vehicle travels at low translational and angular velocities. We assume no model-plant

mismatches in (1) as the rigid body dynamics are typically known and the parameters such as mass or inertia can be precisely obtained through CAD model.

B. Approximate actuation model

The design of the VoliroX leads to complicated aerodynamic effects (e.g. coaxial rotor configurations and interaction between the adjacent propeller flows), which are hard to model using first-principles. Therefore, the following assumptions are made to approximate the actuation model: 1). It is assumed that the force and torque produced by a stationary propeller are proportional to its angular velocity squared, independent of the vehicle's translational and angular velocity. The resulting force and torque vectors are also assumed to be parallel with each other and perpendicular to the rotor disk. 2). Every rotor has the same torque and force coefficient. 3). The rotors do not interfere with each other, therefore the total force and torque produced by the rotors are the sum of individual force and torque vectors. 4) Body drag forces and torques are neglected due to the low translational and angular velocity of the flying vehicle. 5). It is assumed that the actuators have no temporal dynamics, that is, the past wrench commands do not affect the current wrench.

Based on the above assumptions, a nonlinear approximate mapping $h(\cdot) : \mathbb{R}^{18} \rightarrow \mathbb{R}^6$ between the individual actuator commands and the total wrench command $\mathbf{W}_{\text{cmd}}^B = [\mathbf{F}_{\text{cmd}}^B, \mathbf{M}_{\text{cmd}}^B]$ expressed in the body coordinate system B may be established [3]. That is

$$\mathbf{W}_{\text{cmd}}^B = h(\mathbf{u}_{\text{cmd}}) \quad (2)$$

where $\mathbf{u}_{\text{cmd}} \in \mathbb{R}^{18}$ contains the twelve propeller thrusts and the six tilt arm angles. For the sake of brevity, the mapping $h(\cdot)$ is not described in detail as it does not affect the subsequent derivation.

This approximate modeling of the actuation introduces modeling errors, this results in a \mathbf{W}^B which differs from the intended $\mathbf{W}_{\text{cmd}}^B$:

$$\mathbf{W}^B = h^*(\mathbf{u}_{\text{cmd}}) = h(\mathbf{u}_{\text{cmd}}) + \eta(\mathbf{u}_{\text{cmd}}) \quad (3)$$

where $h^*(\cdot)$ is the true unknown actuation function and $\eta(\cdot)$ is an additive nonlinear modeling error function.

Note that this additive error modeling implies that $\eta(\cdot)$ and $h(\cdot)$ do not need to have the same parametric structure, which allows for more flexibility in the modeling. On the other hand, misidentification of $\eta(\cdot)$ can result in a physically implausible prediction of the generated wrench.

In addition, the functions $h(\cdot)$, $h^*(\cdot)$, and $\eta(\cdot)$ are surjective, that is, for each achievable wrench there are possibly more than one set of feasible inputs.

V. APPROACH

The problem we are addressing in this report is formulated as follows: Given the system dynamics (1), the actuation model (2), the unknown modeling error $\eta(\mathbf{u}_{\text{cmd}})$, and a given desired wrench $\mathbf{W}_{\text{des}}^B$ which is the output of a given model-based nominal controller, find the actuator commands \mathbf{u}_{cmd} such that

the $\mathbf{W}_{\text{des}}^B$ is achieved. We distinguish between $\mathbf{W}_{\text{des}}^B$, which is the output of the controller, and $\mathbf{W}_{\text{cmd}}^B$, which is the input to the allocation.

The proposed solution aims to find a compensation signal at the wrench level which, when applied to the nominal controller's command, results in the desired wrench being executed by the system. This reduces the problem to six dimensions since we do not search compensation signals for individual actuator commands \mathbf{u}_{cmd} , as they are typically computed on a low-level flight computer that is computationally limited. Furthermore, more training data would be needed for a higher dimensional input space. The approach is summarized as follows and illustrated in Fig. 2: the nominal controller computes a desired wrench $\mathbf{W}_{\text{des}}^B$ according to the reference trajectory and the state estimate. The discrepancy between the model and the plant is modeled using GPs. The regressed model is input into an optimization framework to obtain an additive compensation signal to the desired wrench, which in case of high prediction uncertainty is filtered towards zero. The resulting command wrench $\mathbf{W}_{\text{cmd}}^B$ is then sent to the control allocation, which outputs individual actuator commands \mathbf{u}_{cmd} .

This approach has the advantage that the GP and optimization block is modular and therefore can be added to the nominal controller. Furthermore, an important point in our proposed solution is that the uncertainty information provided by the GP is used to prevent uncertain predictions from destabilizing the system. This is a key for safely deploying the proposed controller in a physical system.

A. Nominal controller and control allocation

Let \mathbf{e}_p denote the position error of the vehicle (i.e. the difference between the actual position \mathbf{p} and the desired position \mathbf{p}_{des}) and $\dot{\mathbf{e}}_p$ its time-derivative. Let D denote the reference body frame and define the attitude error terms as follows:

$$\mathbf{e}_R = \frac{1}{2}(\mathbf{R}^{DB} - \mathbf{R}^{BD})^\vee \quad (4)$$

$$\mathbf{e}_\omega = \boldsymbol{\omega}_{BD}^B \quad (5)$$

where \mathbf{R}^{DB} is the coordinate transformation matrix from the coordinate system D to B . The mapping \vee maps a skew-symmetric matrix to its \mathbb{R}^3 space.

Similar to the position control in [27], the desired wrench $\mathbf{W}_{\text{des}}^B = [\mathbf{F}_{\text{des}}^B, \mathbf{M}_{\text{des}}^B]$ is calculated such that

$$\ddot{\mathbf{e}}_p^E = -2\zeta_p\omega_{n,p}\dot{\mathbf{e}}_p^E - \omega_{n,p}^2\mathbf{e}_p^E \quad (6)$$

$$\dot{\mathbf{e}}_\omega^B = -2\zeta_a\omega_{n,a}\mathbf{e}_\omega^B - \omega_{n,a}^2\mathbf{e}_R^B \quad (7)$$

where $(\zeta_p, \omega_{n,p})$ and $(\zeta_a, \omega_{n,a})$ are a set of damping ratios and natural frequencies for the translational dynamics and attitude dynamics, respectively.

From (4) and (5) we can see that to first order, \mathbf{e}_ω is the time-derivative of \mathbf{e}_R . This means that if the desired wrench is tracked perfectly, the translation deviation \mathbf{e}_p and the rotational deviation \mathbf{e}_R will behave like a damped second order system. This makes the parameter tuning more intuitive.

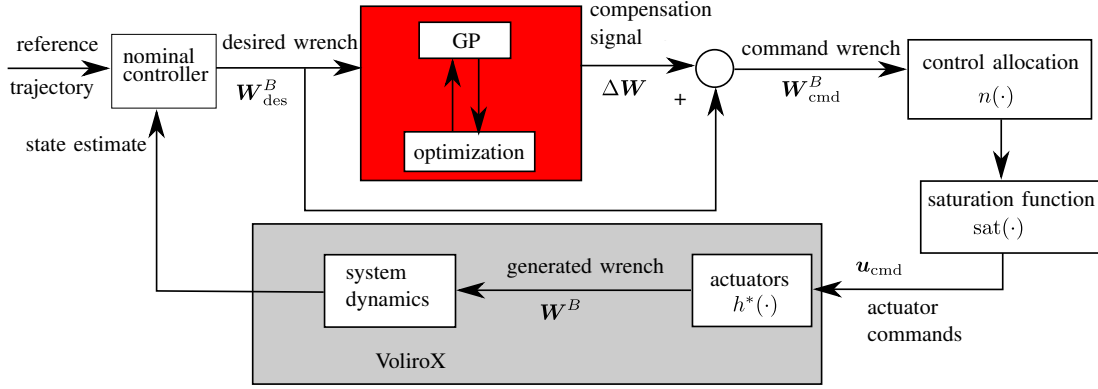


Fig. 2: A block diagram of the proposed solution. The model-plant mismatch is modeled using GPs and regressed offline using flight data. The model is then fed into an optimization to find a compensation signal ΔW so that the generated wrench W^B is equal to W_{des}^B .

Recall that the actuator dynamics are neglected, thus the wrench and the individual actuators are a static mapping. The wrench command then gets allocated through a chosen mapping $n(\cdot) : \mathbb{R}^6 \rightarrow \mathbb{R}^{18}$ [3] and a saturation function $\text{sat}(\cdot) : \mathbb{R}^{18} \rightarrow \mathbb{R}^{18}$ in the inner loop

$$u_{cmd} = \text{sat}(n(W_{cmd}^B)) \quad (8)$$

Again, we omit the details of the allocation for brevity. This allocation is essentially a pseudo-inverse mapping of $h(\cdot)$. The mapping $n(\cdot)$ is injective, that is, each set of actuation has at most one corresponding wrench.

B. Learning model-plant mismatch

We first derive the model-plant mismatch as a function of the wrench command. Substituting (2) and (8) into (3) yields

$$W^B = h(u_{cmd}) + \eta(\text{sat}(n(W_{cmd}^B))) \quad (9)$$

$$= W_{cmd}^B + g(W_{cmd}^B), \quad (10)$$

where $g(\cdot) : \mathbb{R}^6 \rightarrow \mathbb{R}^6$ is the composition of the injective mapping $n(\cdot)$, $\text{sat}(\cdot)$ and the surjective mapping $\eta(\cdot)$. It may therefore be a surjective mapping. This means that for each W^B there may be more than one set of possible W_{cmd}^B that achieves it. This leads to the problem mentioned in Section I: an inverse mapping from W^B to W_{cmd}^B is a one-to-many mapping and may not directly be learned.

The unknown error function $g(\cdot)$ is modeled as six independent GPs. We make the assumption that the outputs of these GPs are uncorrelated. A GP is a collection of random variables, any finite number of which have a joint Gaussian distribution [28], and can be viewed as a distribution over functions. It is characterized by a mean function $m(\cdot)$ and a covariance function $k(\cdot, \cdot)$, which is a positive-definite kernel function parametrized by a set of hyperparameters.

A GP-based representation was chosen because it handles stochasticity (e.g. measurement noise) naturally, its nonparametric property offers flexibility in the modeling, and gives us information about the uncertainty of predictions made. On the other hand, it is known that the GP prediction does not scale well with data (it has a computational complexity of $\mathcal{O}(N^3)$, where N is the number of data samples). Reducing

the computational complexity is beyond the scope of this work. We are currently exploring the use of local GPs to handle the large state space.

We may apply Gaussian process regression to approximate the error function $g_l(W_{cmd}^B)$ according to (10), with the subscript l denoting the l -th entry of the function $g(\cdot)$:

$$z = g_l(\xi) + \epsilon \quad (11)$$

where $z \in \mathbb{R}$ denotes the observation of the l -th entry of $W^B - W_{cmd}^B$, $\epsilon \in \mathbb{R}$ is the measurement noise with Gaussian distribution $\mathcal{N}(0, \sigma^2)$ with σ^2 being the variance, and $\xi \in \mathbb{R}^6$ denotes W_{cmd}^B for brevity for the remaining of this section.

Let (ξ_i, z_i) denote one observed data point and $X \in \mathbb{R}^{(N+1) \times 6}$, $Z \in \mathbb{R}^{N+1}$ denote the stacked version of (ξ_0, \dots, ξ_N) , (z_0, \dots, z_N) from $N+1$ data points. Conditioned on this data set and a query input ξ^* , the expected value and variance of the l -th GP is:

$$\mu_l(\xi^*) = m(\xi^*) + k(\xi^*, X)(K + \sigma^2 I)^{-1} Z \quad (12)$$

$$\sigma_l^2(\xi^*) = k(\xi^*, \xi^*) + \sigma^2 - k(\xi^*, X)(K + \sigma^2 I)^{-1} k(X, \xi^*) \quad (13)$$

where K is a $(N+1) \times (N+1)$ kernel matrix with $K_{ij} = k(\xi_i, \xi_j)$.

In this work we use the squared exponential covariance as the kernel

$$k(\xi_p, \xi_q) = \sigma_f^2 \exp\left(-\frac{1}{2}(\xi_p - \xi_q)^T \Sigma (\xi_p - \xi_q)\right). \quad (14)$$

where ξ_p and ξ_q are any two inputs, Σ is the lengthscale matrix which indicates the relevance between two data points and σ_f^2 is the signal variance which is a scale factor. The hyperparameters $(\sigma, \sigma_f, \Sigma)$ for a particular data set can be selected by maximum likelihood estimation, which maximizes the likelihood of the observed outputs given hyperparameters [28].

C. Finding compensation signal using optimization

From (10) it can be seen that if W_{cmd}^B is set equal to the controller output W_{des}^B , the actual resulting wrench W^B deviates from W_{des}^B by the error function $g(W_{des}^B)$.

In order to achieve the desired wrench $\mathbf{W}_{\text{des}}^B$, we propose to add a compensation element $\Delta \mathbf{W} \in \mathbb{R}^6$ to the $\mathbf{W}_{\text{des}}^B$ when setting $\mathbf{W}_{\text{cmd}}^B$, that is,

$$\Delta \mathbf{W} := \mathbf{W}_{\text{cmd}}^B - \mathbf{W}_{\text{des}}^B \quad (15)$$

A $\Delta \mathbf{W}$ is to be found such that when (15) is substituted into (10), the produced wrench \mathbf{W}^B is equal to the desired wrench $\mathbf{W}_{\text{des}}^B$:

$$\mathbf{W}_{\text{des}}^B \stackrel{!}{=} \mathbf{W}^B = (\mathbf{W}_{\text{des}}^B + \Delta \mathbf{W}) + g(\mathbf{W}_{\text{des}}^B + \Delta \mathbf{W}). \quad (16)$$

With $g(\cdot)$ modeled as GPs, we obtain

$$\Delta \mathbf{W} + \mu(\mathbf{W}_{\text{des}}^B + \Delta \mathbf{W}) = 0 \quad (17)$$

with $\mu(\cdot)$ being the mean vector function of the GPs.

An online optimization problem is formulated to approximately solve (17):

$$\underset{\Delta \mathbf{W}}{\text{minimize}} \quad \left\| \mu(\mathbf{W}_{\text{des}}^B + \Delta \mathbf{W}) + \Delta \mathbf{W} \right\|_2^2 \quad (18)$$

We optimize this in real time using gradient-based optimization. If the optimal function value is smaller than 1e-4, then we approximately find the compensation according to the mean function of the GP model. Otherwise, we compare the optimal value with the cost function value with $\Delta \mathbf{W} = 0$, and pick the corresponding signal that yields a smaller cost.

Remark: Formulating the problem as an optimization problem has several advantages. Firstly, assuming the update rate is high enough, the previous solution is close to the current iteration solution, which is ideally suited as a warm start for the next iteration. Secondly, the search is local, therefore avoiding the jump between feasible solutions that are far away from each other. Due to the saturation function it is possible that the cost function has locally zero gradients. For this, a regularization term $\|\mathbf{W}_{\text{cmd}}\|_2^2$ can be added to the cost function. Finally, the cost function is easy to manipulate and tune. We can add wrench command constraints to the optimization problem or add costs that smoothen the adjacent wrench commands.

D. Utilize posterior uncertainty prediction

Since we are using a nonparametric method to learn the dynamics, this might lead to incorrect predictions for inputs that are far outside of the training data and thus destabilize the system. To address this issue we make use of the uncertainty information provided by the GP. After we obtained the optimal solution $\Delta \mathbf{W}^*$ from the optimization, the compensation signal is filtered towards zero in case of high uncertainty of the prediction

$$\Delta \mathbf{W} \leftarrow a\beta \Delta \mathbf{W}_{\text{prev}} + (1 - \beta) \Delta \mathbf{W}^* \quad (19)$$

where $a \in (0, 1)$, $\Delta \mathbf{W}_{\text{prev}}$ is the previously computed compensation signal,

$$\beta = \frac{1}{1 + e^{-\kappa(\sigma - \sigma_{\text{th}})}} \quad (20)$$

where $\sigma = \max \{\sigma_l(\mathbf{W}_{\text{des}}^B + \Delta \mathbf{W}^*), l = 0, \dots, 5\}$ and with some constants $\sigma_{\text{th}} \in \mathbb{R}$ and $\kappa \in \mathbb{R}$.

For small prediction uncertainty β is near zero and thus the compensation signal directly takes in the optimal value $\Delta \mathbf{W}^*$. For σ larger than the threshold σ_{th} , β is near one. a determines how fast the compensation signal $\Delta \mathbf{W}$ should be filtered towards zero. κ determines the rapidness of changing of β near σ_{th} .

Once the compensation signal is obtained, it is added to the output of the controller \mathbf{W}_{des} . The resulting actuator command \mathbf{u}_{cmd} is obtained through the allocation mapping $n(\cdot)$.

E. Comparison with disturbance observer approaches

An alternative approach is to employ an observer to estimate the unmodelled dynamics in real time and inform the controller to reject it. This approach is computationally efficient and has been shown to work well in practice. However, the approach introduces delay into tracking due to its nature of being a filter and furthermore requires expert knowledge to handcraft a disturbance model. The presented approach does not introduce delay and is less dependent on expert knowledge as it uses a data driven approach. The trade-off for this is the increased computational complexity and inability to compensate for unobserved disturbances. A good comparison of the properties of these two methods is presented in [29].

VI. EXPERIMENTAL RESULTS

The experiments are carried out at an indoor aerial vehicle testbed at Autonomous System Lab, ETH Zurich. A motion capture system provides pose estimates at 100 Hz. The experimental vehicle platform VoliroX has a diameter of about 80 cm and a weight of 4 kg. The vehicle is also equipped with an onboard NUC i7 computer, which runs computationally expensive modules, and a PixHawk flight controller that takes care of the low-level, low latency tasks. Using this configuration the vehicle is able to run all the necessary algorithms onboard. For a more complete description see [3].

The proposed framework is implemented as a ROS node and evaluated by the performance of attitude trajectory tracking both in simulation and on a real platform. In simulation, the framework is evaluated on a figure-8 trajectory with the vehicle tilting up to 63 degrees in the RotorS Gazebo simulator [30]. For the real experiments, the vehicle follows a smooth feasible reference pitching trajectory from 0 to 60 degrees and back, while remaining stationary with zero roll and yaw. The maximal reference angular acceleration is 1 rad s⁻². These trajectories were chosen as from experience the most prominent model-plant mismatch occurs on the torque level, especially when the produced force vector is close to the body xy plane, corresponding to the vehicle attitude with high roll or pitch.

A. GP model-plant mismatch modeling and learning

This section shows the details of the model-plant mismatch using GP in real experiments and demonstrates its prediction performance on a validation trajectory. As mentioned before, the most prominent model-plant mismatch occurs on the

torque level, three single-output GPs are therefore used to model the torque model-plant-mismatch in the experiments and the **force model-plant-mismatch is neglected**. For the GP regression, the inputs are the wrench command $\mathbf{W}_{\text{cmd}}^B$ and the outputs are the components of $(\mathbf{M}_{\text{meas}}^B - \mathbf{M}_{\text{cmd}}^B)$, where $\mathbf{M}_{\text{meas}}^B$ is obtained from a high-quality ADIS16448 IMU sensor [31]. The **training data is collected while the vehicle tracks a test trajectory with a nominal controller. Sinusoidal excitation is added to the reference trajectory to collect diverse training data**. This increases the probability of the learned model covering the output space during validation flight with the proposed strategy. Data points are **subsampled** from the experimental data offline using the **k-medoids algorithm** where the Euclidean squared distance between the inputs is used as the distance metric. Through empirical validation we found a hundred subsampled data points are sufficient for this trajectory. The **hyperparameters** of the GPs are fixed after being estimated using the **maximum likelihood methods**, as we assume the model-plant-mismatch characteristics along the reference trajectory are repeatable. The GPs are implemented using **GPy** [32], the optimization uses L-BFGS implemented from **nlopt** [33], and the **k-medoids algorithm** is taken from the **PyClustering** [34]. The learned model is then embedded into the online optimization framework, whose performance is then evaluated and presented in the following.

Fig. 3 shows the command torque \mathbf{M}_{cmd} , the measured torque \mathbf{M}_{meas} , and the predicted torque \mathbf{M}_{pred} of a 10 seconds trajectory. \mathbf{M}_{pred} is $\mu_{\mathbf{M}}(\mathbf{W}_{\text{cmd}}) + \mathbf{M}_{\text{cmd}}$, where $\mu_{\mathbf{M}}(\cdot)$ denotes the torque outputs of the GPs. Table I shows the mean and standard deviation of the absolute differences between the prediction and the measurement on the same trajectory (that is, $|\mathbf{M}_{\text{cmd}} - \mathbf{M}_{\text{meas}}|$ and $|\mathbf{M}_{\text{pred}} - \mathbf{M}_{\text{meas}}|$). It can be observed from Table I that the prediction error on the torque has been reduced by 62%, 70%, and 74% on the body x , y and z axis, respectively. In addition, the mean of the prediction error using the learned model on the body x and y axis are comparable (0.08 and 0.067 Nm), while the **error on the body z axis is much larger** (0.13 Nm). This is expected as the torque around the body z -axis is mainly **generated by tilting the motor arms** (especially during hover), while the torque around the body x - and y -axis is mainly induced by the sum of the propeller thrust times the motor arm length. While **both servo dynamics and motor dynamics are neglected during the model learning, the prediction around the body z -axis is more affected since the servo dynamics are much slower than the motor dynamics**.

TABLE I: Prediction performance

[Nm]	body x error	body y error	body z error
nominal mean \pm std	0.21 \pm 0.16	0.23 \pm 0.16	0.50 \pm 0.17
learned mean \pm std	0.08 \pm 0.06	0.07 \pm 0.05	0.13 \pm 0.09

B. Trajectory tracking

1) *Real experiments*: The experiments demonstrated in this section compare the trajectory tracking performance between

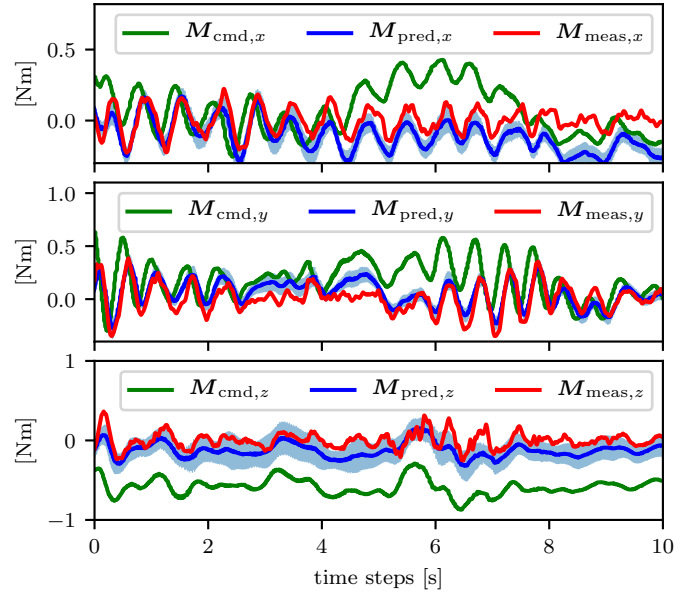


Fig. 3: Measurement, prediction (our model), and command of the torque vector (nominal model) of a 10 s trajectory. The 3 sigma confidence region of the GP prediction is shaded in blue.

the cases with and without compensation signals. Fig. 4 shows a box plot of the absolute values of the attitude tracking errors $|\mathbf{e}_{\mathbf{R}}|$ from two experiments. In the first experiment the vehicle tracks the pitch trajectory 10 times consecutively using the control strategy without the compensation, whereas the second experiment repeats the same trajectory using the proposed framework. Running onboard a NUC i7 computer, it takes on **average 7.5 ms and maximally 24.0 ms for solving one optimization problem in the proposed framework**.

In both cases, an integral term is added to the attitude controller (7). The **gains of the nominal controller are tuned such that both approaches result in a stable flight in both experiments for the sake of comparison**. They are $(\zeta_p, \omega_{n,p}) = (0.707, 3.5)$ and $(\zeta_a, \omega_{n,a}) = (1.3, 3.5)$ and $(0.74, 3.5)$, where the former gain is tuned for roll and pitch axis and the latter gain tuned for the yaw axis. The integral attitude gain is 0.3.

It can be seen that the medians of the absolute attitude tracking errors around the body x and y axis have reduced from 0.07, 0.087 to 0.04, 0.03, respectively, corresponding to a reduction of 43% and 65%. Their box ranges are also reduced from 0.114 and 0.10 to 0.050 to 0.040. On the other hand, there are no improvements shown on the body z axis. The median increases from 0.051 to 0.058 and the box range decreases from 0.066 to 0.054. **The tracking in body z -axis has not been improved** mainly due to the following two reasons: firstly, it can be seen in Fig. 3 that the model-plant mismatch is mostly a **constant offset**. In the nominal controller, this offset is already mitigated by the **integrator term** around the body- z axis. Secondly, as mentioned in the previous section, **the servo dynamics and the motor dynamics are not learned** using the GP. To further improve the tracking performance in body z axis, the **input space of the GP could be augmented** to include

the past history of the wrench commands.

2) *Simulation*: A simulation is conducted to demonstrate the generalizability of the learned model. The training data is obtained by tilting the flying vehicle along various body axes up to 70 degrees (with sinusoidal excitation around all three body axes). 200 data points are subsampled to train the GP model. A figure 8 trajectory with the vehicle tilting up to 63 degrees is then flown to evaluate the approach. A comparison of tracking performance with and without compensation signal is shown in Fig. 5. The controller gains are again the same in both cases. Although this time the integral gains are set to zero. The tracking errors along 3 body axes have been reduced by 94.5%, 87.9%, and 83.7%, respectively. Similar to the real experiments, the key to having a good model is to excite all body axes during the flight. This ensures that the model can provide informative predictions for all required states.

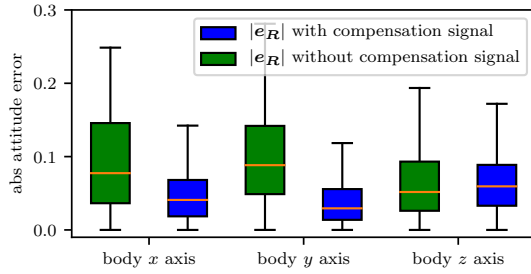


Fig. 4: A box plot comparison of the tracking performance for a pitching trajectory with and without the compensation signal. In both experiments, the same reference trajectory is executed consecutively 10 times. Note that by small angle approximation e_R corresponds to the Euler angles in radian.

C. Incorporation of prediction uncertainty

The experiment presented in this section aims to demonstrate that even in face of highly uncertain predictions, the system is still robust thanks to the uncertainty provided by the model. Two experiments using the learned model are conducted, one with the uncertainty check (that is, with performing (19)) and the other one without. In both cases the vehicle in hover is given a step reference of 40 degrees roll. Since the training data consists of smooth pitching trajectories, the wrench command for a step input in roll is far away from the training data. Fig. 6 shows the experiment with the controller with uncertainty check. The standard deviation of the prediction uncertainty in body z direction increases from 0.27 to 0.5. The compensation signal is then filtered towards zero and the flying vehicle is able to remain stable. The parameters Σ_{th} , κ , a are heuristically chosen. In another experiment with the controller without the uncertainty check, the system becomes unstable due to the unreliable compensation signal¹. In this case, it can be expected that the tracking performance of the approach with the uncertainty check cannot be better than the approach without compensation. GP will not predict well in region where no data is available.

¹For a visualization of the behaviors, please see the attached video

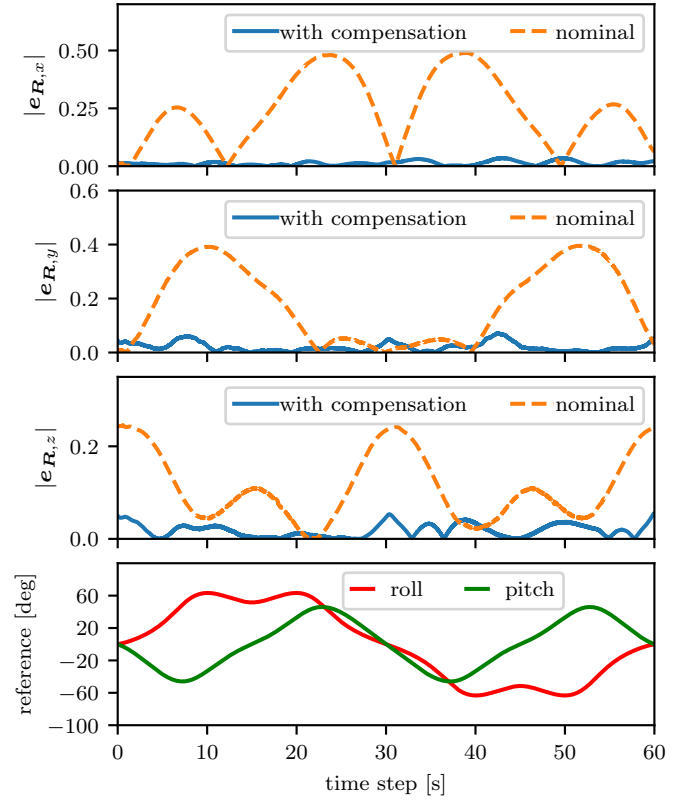


Fig. 5: VoliroX tracking a figure 8 reference trajectory with and without compensation signals in the simulation. Note that by small angle approximation e_R corresponds to the Euler angles in radian.

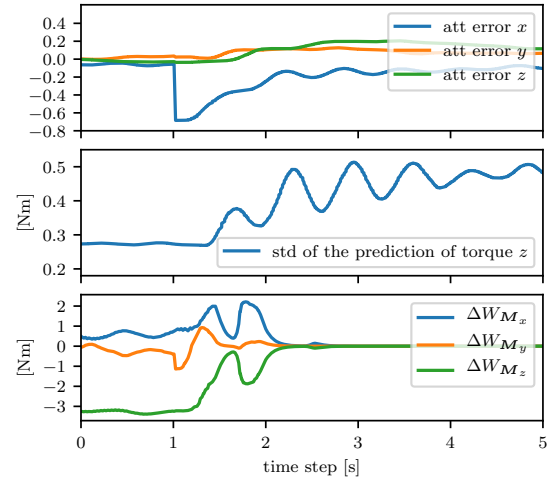


Fig. 6: This figure shows that for a step input (around 1.0 s) around body- x direction. The desired wrench command is far away from the training data and GP prediction is uncertain at this query input. Therefore the compensation signals (ΔW_{M_x} , ΔW_{M_y} , ΔW_{M_z}) get filtered towards zero according to (19) and the vehicle remains stable. Note that by small angle approximation e_R corresponds to the Euler angles (roll, pitch, yaw from the reference frame to the body frame) in radian.

VII. CONCLUSION AND OUTLOOK

This paper presented an approach that addresses the control challenges caused by the model-plant mismatches of an

overactuated flying vehicle. Specifically, our approach learns a well-defined forward model using a GP and avoids the multi-valued mapping of the inverse model of the overactuated systems. It finds a compensation signal through an optimization, which then cancels out the effect caused by the model-plant mismatch. In addition, the uncertainty prediction from the GP is exploited to prevent uncertain predictions from destabilizing the flying vehicle. Experiments on a real platform show that the proposed approach reduces the tracking error by 43% and 65% around body x - and y -axis and could prevent destabilizing the system in case of highly uncertain prediction of the model.

ACKNOWLEDGEMENTS

The authors would like to thank Nicholas Lawrance for thoroughly going through the manuscript. This work was partly supported by funding from NCCR Digital Fabrications, NCCR Robotics and Armasuisse.

REFERENCES

- [1] M. Ryll, H. H. Bühlhoff, and P. R. Giordano, "Modeling and control of a quadrotor uav with tilting propellers," in *IEEE International Conference on Robotics and Automation*, 2012.
- [2] D. Brescianini and R. D'Andrea, "An omni-directional multirotor vehicle," *Mechatronics*, 2018.
- [3] K. Bodie, M. Brunner, M. Pantic, S. Walser, P. Pfändler, U. Angst, R. Siegwart, and J. Nieto, "An omnidirectional aerial manipulation platform for contact-based inspection," *arXiv preprint arXiv:1905.03502*, 2019.
- [4] S. Park, J. Lee, J. Ahn, M. Kim, J. Her, G.-H. Yang, and D. Lee, "Odar: Aerial manipulation platform enabling omnidirectional wrench generation," *IEEE/ASME Transactions on mechatronics*, 2018.
- [5] N. Staub, D. Bicego, Q. Sablé, V. Arellano, S. Mishra, and A. Franchi, "Towards a flying assistant paradigm: The othex," in *IEEE International Conference on Robotics and Automation*, 2018.
- [6] M. Kamel, S. Verling, O. Elkhatib, C. Sprecher, P. Wulkop, Z. Taylor, R. Siegwart, and I. Gilitschenski, "Voliro: An omnidirectional hexacopter with tiltable rotors," *arXiv preprint arXiv:1801.04581*, 2018.
- [7] R. Mahony, V. Kumar, and P. Corke, "Multirotor aerial vehicles: Modeling, estimation, and control of quadrotor," *IEEE robotics & automation magazine*, 2012.
- [8] D. Hentzen, T. Stastny, R. Siegwart, and R. Brockers, "Disturbance estimation and rejection for high-precision multirotor position control," *arXiv preprint arXiv:1908.03166*, 2019.
- [9] B. Yüksel, C. Secchi, H. H. Bühlhoff, and A. Franchi, "A nonlinear force observer for quadrotors and application to physical interactive tasks," in *IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, 2014.
- [10] J. Hwangbo, J. Lee, A. Dosovitskiy, D. Bellicoso, V. Tsounis, V. Koltun, and M. Hutter, "Learning agile and dynamic motor skills for legged robots," *Science Robotics*, 2019.
- [11] M. Kamel, M. Burri, and R. Siegwart, "Linear vs nonlinear mpc for trajectory tracking applied to rotary wing micro aerial vehicles," *IFAC-PapersOnLine*, 2017.
- [12] M. Brunner, K. Bodie, M. Kamel, W. Zhang, J. Nieto, and R. Siegwart, "Trajectory tracking nonlinear model predictive control for an overactuated mav," in *IEEE International Conference on Robotics and Automation*, 2020.
- [13] D. Nguyen-Tuong and J. Peters, "Model learning for robot control: a survey," *Cognitive processing*, 2011.
- [14] G. Chowdhary, H. A. Kingravi, J. P. How, and P. A. Vela, "Bayesian nonparametric adaptive control using gaussian processes," *IEEE transactions on neural networks and learning systems*, 2014.
- [15] D. Nguyen-Tuong, J. R. Peters, and M. Seeger, "Local gaussian process regression for real time online model learning," in *Advances in Neural Information Processing Systems*, 2009.
- [16] F. Meier, D. Kappler, N. Ratliff, and S. Schaal, "Towards robust online inverse dynamics learning," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2016.
- [17] T. Beckers, J. Umlauf, D. Kulic, and S. Hirche, "Stable gaussian process based tracking control of lagrangian systems," in *IEEE Conference on Decision and Control*, 2017.
- [18] M. K. Helwa, A. Heins, and A. P. Schoellig, "Provably robust learning-based approach for high-accuracy tracking control of lagrangian systems," *IEEE Robotics and Automation Letters*, 2019.
- [19] D. Mitrovic, S. Klanke, and S. Vijayakumar, "Adaptive optimal feedback control with learned internal dynamics models," in *From Motor Learning to Interaction Learning in Robots*, 2010.
- [20] J. Kabzan, L. Hewing, A. Liniger, and M. N. Zeilinger, "Learning-based model predictive control for autonomous racing," *IEEE Robotics and Automation Letters*, 2019.
- [21] C. J. Ostafew, A. P. Schoellig, and T. D. Barfoot, "Learning-based nonlinear model predictive control to improve vision-based mobile robot path-tracking in challenging outdoor environments," in *IEEE International Conference on Robotics and Automation*, 2014.
- [22] G. Cao, E. M.-K. Lai, and F. Alam, "Gaussian process model predictive control of an unmanned quadrotor," *Journal of Intelligent & Robotic Systems*, 2017.
- [23] M. Deisenroth and C. E. Rasmussen, "Pilco: A model-based and data-efficient approach to policy search," in *International Conference on machine learning*, 2011.
- [24] M. Saveriano, Y. Yin, P. Falco, and D. Lee, "Data-efficient control policy search using residual dynamics learning," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2017.
- [25] J. Ko, D. J. Klein, D. Fox, and D. Haehnel, "Gaussian processes and reinforcement learning for identification and control of an autonomous blimp," in *IEEE International Conference on Robotics and Automation*, 2007.
- [26] A. Spitzer and N. Michael, "Inverting learned dynamics models for aggressive multirotor control," *arXiv preprint arXiv:1905.13441*, 2019.
- [27] W. Zhang, M. W. Mueller, and R. D'Andrea, "A controllable flying vehicle with a single moving part," in *IEEE International Conference on Robotics and Automation*, 2016.
- [28] C. E. Rasmussen, "Gaussian processes in machine learning." Springer, 2003.
- [29] V. R. Desaraju, A. Spitzer, and N. Michael, "Experience-driven predictive control with robust constraint satisfaction under time-varying state uncertainty," in *Robotics: Science and Systems*, 2017.
- [30] F. Furrer, M. Burri, M. Achtelik, and R. Siegwart, *Robot Operating System (ROS): The Complete Reference (Volume 1)*. Cham: Springer International Publishing, 2016, ch. RotorS—A Modular Gazebo MAV Simulator Framework, pp. 595–625. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-26054-9_23
- [31] "Adis16448 fact sheet," (Date last accessed 5-May-2020). [Online]. Available: <https://www.analog.com/media/en/technical-documentation/data-sheets/ADIS16448.pdf>
- [32] GPy, "GPy: A gaussian process framework in python," <http://github.com/SheffieldML/GPy>, since 2012.
- [33] S. G. Johnson, "The nlopt nonlinear-optimization package," (Date last accessed 10-Sept-2019). [Online]. Available: <http://github.com/stevengj/nlopt>
- [34] A. Novikov, "PyClustering: Data mining library," *Journal of Open Source Software*, 2019.