Augmented Incremental Nonlinear Dynamic Inversion for Agile Flight of Quadcopter

Mike Timmerman
Department of Aeronautics and Astronautics
Stanford University
mtimmerm@stanford.edu

Abstract

This paper presents an augmentation to the incremental nonlinear dynamic inversion control law for trajectory tracking of quadcopter drones. The conventional control law, derived by neglecting higher-order terms and external forces in the Taylor expansion, is augmented by incorporating a deep neural network policy. This policy, trained using reinforcement learning, approximates these neglected terms, which primarily arise from aerodynamic forces and nonlinearities. The proposed augmentation is shown to enhance trajectory tracking performance through simulations and experiments with a 50% decrease in positional error at a target velocity of $9m/s$. Additionally, disturbance rejection performance is assessed under wind conditions. Results demonstrate a 15% improvement in disturbance rejection compared to the conventional control law. These findings highlight the effectiveness of leveraging neural network policies to address complex nonlinear dynamics in autonomous drone control.

Introduction

Aerial vehicles encompass a variety of types, with the quadcopter arguably standing out as one of the most agile variants due to its four actuators, enabling it to execute high-speed maneuvers. This agility has led to its integration into drone racing, where participants achieve remarkable lap times by skillfully navigating through obstacles, maximizing the quadcopter's capabilities. To replicate such performance with a fully autonomous drone, precise trajectory tracking control is imperative, leveraging the drone's nonlinear dynamics to execute agile maneuvers at high speeds. At elevated velocities, aerodynamic forces exert significant influence, necessitating the controller's ability to accommodate both the quadcopter's nonlinear dynamics and the impact of external forces such as aerodynamics.

Previous work has focused on extending nonlinear control techniques to take advantage of the quadcopter's full potential. In particular, (Richards et al., 2021) developed an adaptive-control-oriented meta-learning approach in which they extend adaptive control to incorporate meta-learning to model uncertain dynamics. (Zhang and Ran, 2023) also uses meta-learning to extend upon incremental nonlinear dynamic inversion by learning the control effectiveness matrix directly online. (Tal and Karaman, 2021) utilizes incremental nonlinear dynamic inversion for accurate tracking of aggressive trajectories by incorporating trajectory snap and jerk tracking into the control architecture.

This paper proposes an augmentation directly to the incremental nonlinear dynamic inversion control law. In the conventional control law, the Taylor expansion is used to linearize the linear acceleration dynamics, in which higher-order terms and any external forces are neglected, to then inverse the dynamics to arrive at the control law. The control law takes a desired linear acceleration and outputs a command in roll, pitch and thrust magnitude. The augmentation proposed in the paper approximates the neglected terms in the Taylor expansion by training a deep neural network policy. Adding this approximation to the control law corrects the commanded roll, pitch and thrust magnitude such that the commanded linear acceleration is achieved. The other parts in the control architecture are adopted from previous literature which describe using incremental nonlinear dynamic inversion for attitude control (Acquatella et al., 2012).

This approach differs in that it is a simple, yet effective, one term augmentation to the control law. The nonlinearities are represented by a deep neural network which are perfectly suited for this task. Furthermore, since the policy actions only affect the augmentation, the physical drone can safely be flown while the policy is trained without risking the drone becoming unstable. This allows the neural network to accurately represent the nonlinearities and neglected force contributions. Since the approach uses a feed-forward reference acceleration, it only requires a three times differentiable trajectory.

The remainder of the article is structured as follows. First, the quadcopter model and the overall trajectory control architecture is covered, as well as a derivation of the control law used to convert a desired linear acceleration into roll, pitch and thrust magnitude commands. Then, the augmentation to this control law is discussed as well as the specification of the policy which learns the augmentation. Finally, an evaluation with experimental results and comparison between the augmented and conventional control law are given.

Copyright © 2024, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

Preliminaries

In this section, the quadcopter dynamics model is described as well as the trajectory tracking controller which uses a conventional incremental nonlinear dynamic inversion based control law.

Quadcopter Model

For the quadcopter model, we consider the six-degree-of-freedom dynamics which consists of 12 states; position $\boldsymbol{p}$, linear velocity $\boldsymbol{v}$, attitude $\boldsymbol{\eta}$, angular velocity $\boldsymbol{\Omega}$ and 4 inputs; four rotational speeds of the propellers $\boldsymbol{\omega}_i$. The inertial reference frame is taken as the local tangent plane with north-east-down coordinate convention. The body-fixed reference frame has its origin at the quadcopter's center of mass. The reference frame definitions are shown in Figure 1.

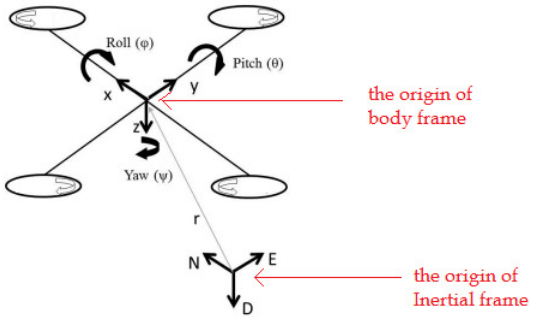

Figure 1: Quadcopter with body-fixed reference system and inertial reference system definitions.

The quadcopter's translational dynamics are given by

$$\ddot{\boldsymbol{\xi}} = \boldsymbol{g} + \frac{1}{m}\vec{T}(\boldsymbol{\eta}, T) + \frac{1}{m}\vec{F}(\dot{\boldsymbol{\xi}}, \boldsymbol{\eta}, \boldsymbol{w}) \tag{1}$$

where $\ddot{\boldsymbol{\xi}}$ is the linear acceleration in the inertial frame and m the quadcopter's mass. The linear acceleration is modeled to have contributions from three terms; gravitational acceleration $\boldsymbol{g}$, acceleration due to thrust $\vec{T}$ produced by the propellers and acceleration experienced due to aerodynamic forces $\vec{F}$. The thrust force is a function of the quadcopter's attitude $\boldsymbol{\eta}$ and thrust magnitude T. The aerodyanmic force is a function of the quadcopter's velocity $\dot{\boldsymbol{\xi}}$, attitude $\boldsymbol{\eta}$, and the wind velocity $\boldsymbol{w}$.

The rotational dynamics are given by

$$\dot{\boldsymbol{\Omega}} = \boldsymbol{J}^{-1}\left(\vec{M} - \boldsymbol{\Omega} \times \boldsymbol{J}\boldsymbol{\Omega}\right) \tag{2}$$

where $\boldsymbol{\Omega}$ is the angular velocity in body-fixed reference frame and $\boldsymbol{J}$ is the inertia tensor. The external moment $\vec{M}$ is modeled to only consist of the control moments $\vec{M}_c$ induced by the propeller thrust.

Trajectory Tracking

The control design is split up in several components, as shown in Figure 2. The trajectory generation provides position $\boldsymbol{p}$, velocity $\boldsymbol{v}$ and acceleration $\boldsymbol{a}$ references according to a reference trajectory. The trajectory regularization uses proportional feedback on the position, velocity and acceleration to generate an adjusted desired acceleration. The acceleration control uses an incremental nonlinear dynamic inversion control law to command roll ϕ, pitch θ and thrust magnitude T. The attitude control uses nonlinear dynamic inversion to convert the attitude commands into angular velocity commands which are then converted into the control moment $\vec{M}_c$. Control allocation is performed to set the propeller rotational speeds $\boldsymbol{\omega}$ to achieve the desired thrust magnitude and control moments.

The augmentation to the incremental nonlinear dynamic inversion control law is performed in the acceleration control. The derivation of the control law, based on (Smeur et al., 2018), starts from Equation 1 by taking the Taylor expansion around the current state, denoted by subscript 0, and neglecting the higher-order terms.

$$\begin{aligned}
\ddot{\boldsymbol{\xi}} \approx & \boldsymbol{g} + \frac{1}{m}\vec{T}(\boldsymbol{\eta}_0, T_0) + \frac{1}{m}\vec{F}(\dot{\boldsymbol{\xi}}_0, \boldsymbol{\eta}_0, \boldsymbol{w}_0) \\
& + \frac{\partial}{\partial T}\frac{1}{m}\vec{T}(\boldsymbol{\eta}_0, T)|_{T=T_0}(T - T_0) \\
& + \frac{\partial}{\partial \boldsymbol{\eta}}\frac{1}{m}\vec{T}(\boldsymbol{\eta}, T_0)|_{\boldsymbol{\eta}=\boldsymbol{\eta}_0}(\boldsymbol{\eta} - \boldsymbol{\eta}_0) \\
& + \frac{\partial}{\partial \dot{\boldsymbol{\xi}}}\frac{1}{m}\vec{F}(\dot{\boldsymbol{\xi}}, \boldsymbol{\eta}_0, \boldsymbol{w}_0)|_{\dot{\boldsymbol{\xi}}=\dot{\boldsymbol{\xi}}_0}(\dot{\boldsymbol{\xi}} - \dot{\boldsymbol{\xi}}_0) \\
& + \frac{\partial}{\partial \boldsymbol{\eta}}\frac{1}{m}\vec{F}(\dot{\boldsymbol{\xi}}_0, \boldsymbol{\eta}, \boldsymbol{w}_0)|_{\boldsymbol{\eta}=\boldsymbol{\eta}_0}(\boldsymbol{\eta} - \boldsymbol{\eta}_0) \\
& + \frac{\partial}{\partial \boldsymbol{w}}\frac{1}{m}\vec{F}(\dot{\boldsymbol{\xi}}_0, \boldsymbol{\eta}_0, \boldsymbol{w})|_{\boldsymbol{w}=\boldsymbol{w}_0}(\boldsymbol{w} - \boldsymbol{w}_0)
\end{aligned} \tag{3}$$

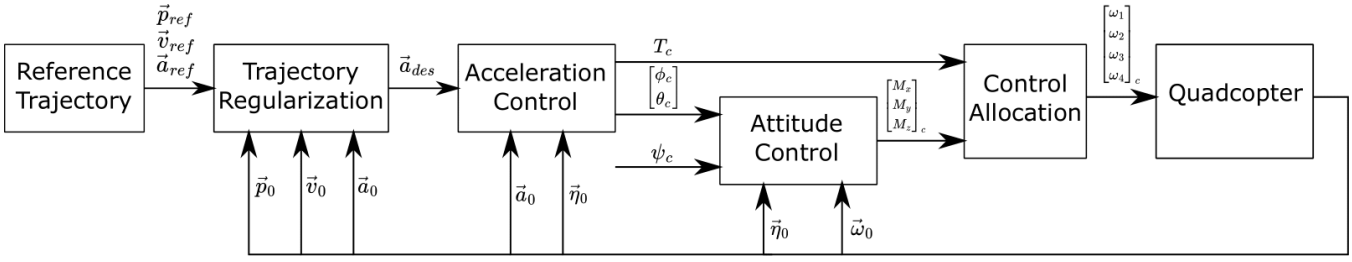

Figure 2: Trajectory tracking control architecture.

The first-order Taylor expansion is further simplified by omitting the derivatives with respect to the aerodynamic forces. The authors make the argument that there is no good estimate for a model of the aerodynamic drag of the airframe and that it is difficult to predict changes in wind. Equation 3 then becomes:

$$\begin{aligned}\ddot{\boldsymbol{\xi}} \approx & \boldsymbol{g} + \frac{1}{m}\vec{T}(\boldsymbol{\eta}_0, T_0) + \frac{1}{m}\vec{F}(\dot{\boldsymbol{\xi}}_0, \boldsymbol{\eta}_0, \boldsymbol{w}_0) \\ & + \frac{\partial}{\partial\phi}\frac{1}{m}\vec{T}(\boldsymbol{\eta}, T_0)|_{\phi=\phi_0}(\phi - \phi_0) \\ & + \frac{\partial}{\partial\theta}\frac{1}{m}\vec{T}(\boldsymbol{\eta}, T_0)|_{\theta=\theta_0}(\theta - \theta_0) \\ & + \frac{\partial}{\partial T}\frac{1}{m}\vec{T}(\boldsymbol{\eta}_0, T)|_{T=T_0}(T - T_0) \\ = & \ddot{\boldsymbol{\xi}}_0 + \boldsymbol{G}(\boldsymbol{u} - \boldsymbol{u}_0)\end{aligned} \tag{4}$$

where $\boldsymbol{u} = [\phi \quad \theta \quad T]$, $\ddot{\xi}_0$ is the current linear acceleration obtained from measurements and $\boldsymbol{G}$ is the control effectiveness matrix:

$$\boldsymbol{G} = \frac{1}{m}\left[\frac{\partial}{\partial\phi}\vec{T}(\boldsymbol{\eta}, T) \quad \frac{\partial}{\partial\theta}\vec{T}(\boldsymbol{\eta}, T) \quad \frac{\partial}{\partial T}\vec{T}(\boldsymbol{\eta}, T)\right] \tag{5}$$

Equation 4 represents an approximation of the linear acceleration increment based on an increment in $\boldsymbol{u}$. Inverting the equation results in a control law based on the difference between the current linear acceleration and a commanded linear acceleration, incremented from the current $\boldsymbol{u}$. The control law then becomes:

$$\boldsymbol{u}_c = \boldsymbol{G}^{-1}(\ddot{\boldsymbol{\xi}}_c - \ddot{\boldsymbol{\xi}}_0) + \boldsymbol{u}_0 \tag{6}$$

This control law takes a commanded linear acceleration $\ddot{\boldsymbol{\xi}}_c$ from the trajectory regularization, current linear acceleration $\ddot{\boldsymbol{\xi}}_0$, attitude and thrust magnitude from measurements and returns a commanded roll angle ϕ_c, pitch angle θ_c and thrust magnitude T_c.

Control Law Augmentation

In the previous section, the control law was derived for acceleration control. To simplify the control law, the terms related to the change in linear acceleration due to aerodynamic forces were omitted as well as any higher order terms. To analyze the implication of neglecting these terms, we can write out the error between the desired commanded acceleration and the achieved acceleration.

$$\ddot{\boldsymbol{\xi}}_{i+1} = \ddot{\boldsymbol{\xi}}_i + \boldsymbol{G}(\boldsymbol{u}_c - \boldsymbol{u}_i) + \boldsymbol{\epsilon} \tag{7}$$

where $\boldsymbol{\epsilon}$ represents all terms neglected to derive the control law from Equation 6. Substituting the control law results in the following error.

$$\ddot{\boldsymbol{\xi}}_{i+1} - \ddot{\boldsymbol{\xi}}_{c_i} = \boldsymbol{\epsilon} \tag{8}$$

This result indicates that the control law is effective as long as the omitted terms remain small during flight. Nevertheless, in high-speed agile flight paths, these terms may become substantial factors, leading to a notable increase in error and ultimately jeopardizing the stability of the quadcopter.

This work proposes an augmentation of the control law by an additional term β.

$$\boldsymbol{u}_c = \boldsymbol{G}^{-1}(\ddot{\boldsymbol{\xi}}_c - \ddot{\boldsymbol{\xi}}_0 - \boldsymbol{\beta}) + \boldsymbol{u}_0 \tag{9}$$

The resulting error between the commanded linear acceleration and achieved linear acceleration for the augmented control law now becomes

$$\ddot{\boldsymbol{\xi}}_{i+1} - \ddot{\boldsymbol{\xi}}_{c_i} = (\boldsymbol{\epsilon} - \boldsymbol{\beta}) \tag{10}$$

Therefore, if $\beta = \epsilon$, a correction in the commanded control input is made to account for the omitted terms, resulting in the commanded acceleration being achieved at the next timestep. Since these neglected terms consist of nonlinear and complex aerodynamic behaviour, β is chosen to be the output of a deep neural network. The neural network is trained using deep reinforcement learning, where the neural network represents the policy which outputs β. Through training the policy on simulated flights, it is able to learn the underlying representation of these neglected terms.

In order for the policy to achieve its objective, which is to approximate the neglected terms $\boldsymbol{\epsilon}$, it is important to choose an appropriate reward function. From Equation 10, it can be seen that by choosing $r = -||\ddot{\boldsymbol{\xi}}_{i+1} - \ddot{\boldsymbol{\xi}}_{c_i}||_2$, then as $r \to 0$, implies that $\beta \to \epsilon$, where 0 is the maximum achievable reward.

Therefore, the policy specification is as follows:

1. Observation: the policy takes the quadcopter's position error $\boldsymbol{e}_p$, velocity error $\boldsymbol{e}_v$ and attitude error $\boldsymbol{e}_\eta$.
2. Action: the action β produced by the policy is continuous and restricted to a range of $[-4, 4]$.
3. Transitions are dictated by the quadcopter's simulated dynamics model
4. Reward: the reward function is the 2-norm of the error between achieved linear acceleration and commanded acceleration $||\ddot{\boldsymbol{\xi}}_{i+1} - \ddot{\boldsymbol{\xi}}_{c_i}||_2$.

This policy is trained in an environment for a drone with specified properties and learns the drone's nonlinear behaviour through interacting with the environment. For policy optimization, a well-known policy optimization algorithm such as PPO may be used.

A diagram of the augmented control law is shown in Figure 3, where the policy network is shown schematically.

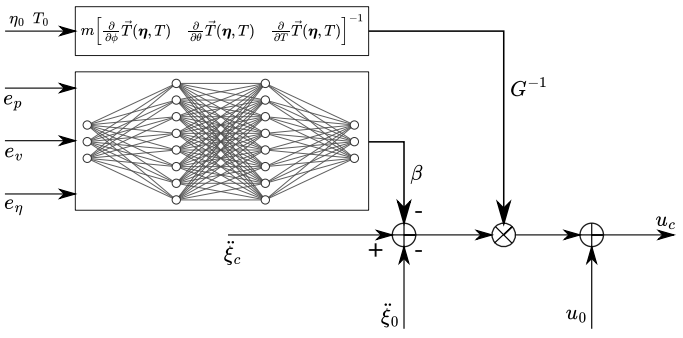

Figure 3: Diagram of augmented control law

Experimental Results

This section first described the details of the experimental setup, such as the environment, the drone characteristics and training of the policy. Then, the evaluation of the proposed augmented control law is given.

Experimental Setup

The policy is trained in a *gymnasium* (Brockman et al., 2016) environment which handles the observations, actions, rewards and environment steps and resets. For the purpose of training the policy, a figure-8 trajectory in 3d-space has been created, as shown in Figure 4. The trajectory is based on a provided set of waypoints between which cubic splines with clamped boundary conditions are fit. These splines provide a smooth reference trajectory between waypoints resulting in a piece-wise linear reference acceleration. The figure-8 trajectory extends into each dimension in order to excite all dynamics of the quadcopter during training.

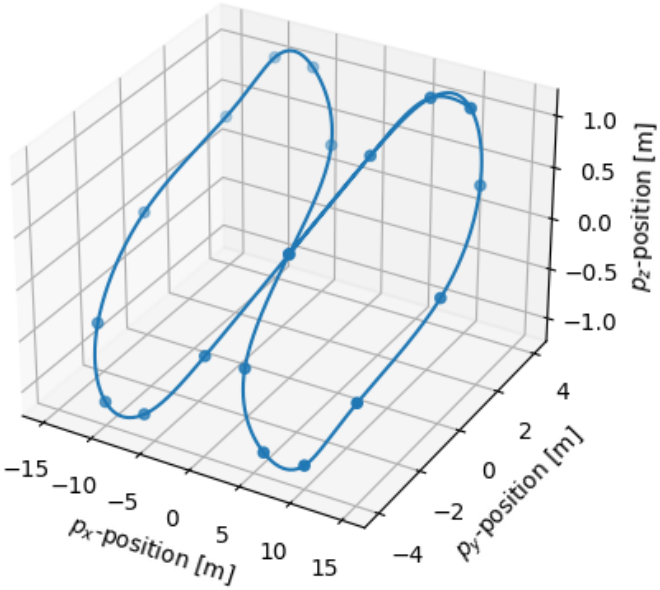

Figure 4: Figure-8 trajectory on which policy is trained.

The quadcopter under consideration is modeled as two perpendicular rigid beams crossing each other midway with propellers at each end of the beams. The center of mass is at the crossing of these beams and the moments of inertia is symmetric about the x- and y-axis. Table 1 contains all properties of the drone under consideration, which are based on the *iFlight Chimera7 Pro* Racing Drone.

Table 1: Quadcopter properties used in simulation. These values are based on the *iFlight Chimera7 Pro* Racing Drone

Property	Value		Property	Value	
Mass	0.695	kg	Mass Inertia I_x	3.56×10^{-4}	kgm^2
Half-Length	0.129	m	Mass Inertia I_y	2.22×10^{-3}	kgm^2
Half-Width	0.100	m	Mass Inertia I_z	6.22×10^{-3}	kgm^2
Effective Area	0.026	m^2	Drag coefficient	0.3	–

The policy is optimized using *Proximal Policy Optimization* (Schulman et al., 2017). This is a policy gradient method directly optimizing the policy by alternating between sampling data through interaction with the environment, and optimizing a surrogate objective function using stochastic gradient ascent. The policy is trained over a total of 1.2×10^6 steps, which corresponds to in-simulation timesteps. Since one timestep is $0.01s$, this corresponds to a total of $3hrs$ and $20mins$ of real-time flight. For each policy update iteration, 2048 steps are made to make the update. The discount factor is set at $\gamma = 0.99$ and the learning rate at $\eta = 3 \times 10^{-4}$.

The learning curve for the policy training is shown in Figure 5. Since the reward $r = -||\ddot{\boldsymbol{\xi}}_{i+1} - \ddot{\boldsymbol{\xi}}_{c_i}||_2$, the theoretical reward ceiling is set at 0. Figure 5 also shows the mean episode root mean square error in position. The position error decreases as the reward increases. This indicates the relation between decreasing the error in achieved acceleration and commanded acceleration (i.e. maximizing the reward) and obtaining an improved trajectory tracking performance.

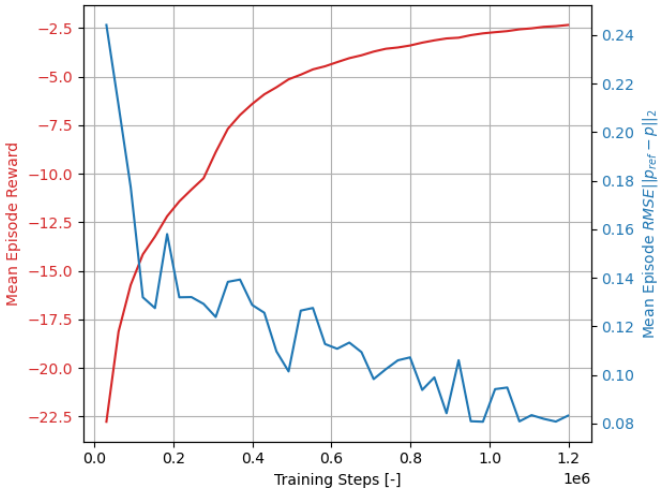

Figure 5: Mean episode reward and root mean square error in position over the policy training cycle.

The evaluation of the augmented control law is performed on a set of randomly generated waypoints. In order to ensure a feasible reference trajectory can be generated, these waypoints are generated according to several restrictions. Firstly, the waypoints are constrained to be within an area of $100m \times 100m \times 30m$. Secondly, the minimum distance between two consecutive waypoints is set to be $25m$ and maximum distance is set to be $50m$. Time stamps are assigned to each waypoint based on a provided target velocity, which specifies how fast the drone is set to track these waypoints. Given the set of waypoints with corresponding time stamps, a cubic spline with clamped boundary conditions is fit between consecutive waypoints. This spline provides a reference position, the first derivative provides a reference velocity and the second derivative provides the reference acceleration. An example of such randomly generate trajectory is provided in Figure 6.

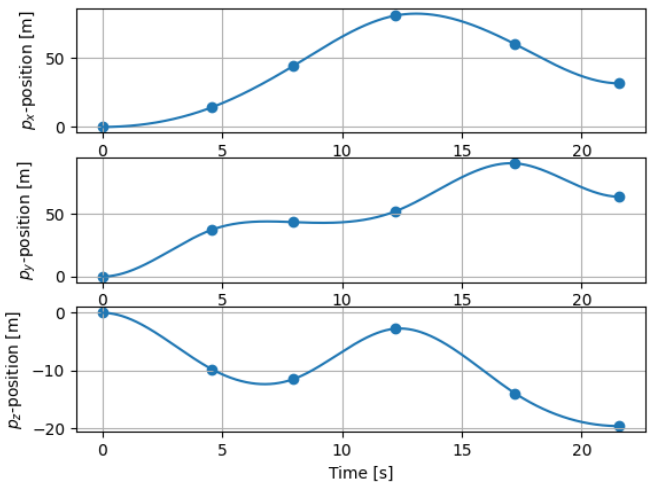

Figure 6: Random waypoint-spline trajectory for evaluation.

Evaluation of Control Law

This section provides an evaluation of the proposed augmented control law. The evaluation is performed by comparing the tracking performance of the drone using the augmented control law against the conventional control law over several experiments.

The first experiment simulates the drone tracking 100 randomly generated evaluation trajectories at different target velocities of $3m/s$, $6m/s$ and $9m/s$. For each trajectory tracking task, the root mean square error, maximum error in position, root mean square velocity and maximum velocity are logged. These are averaged over the 100 random trajectories. The results for the random seed 10 are shown in Table 2, Table 3 and Table 4 for $3m/s$, $6m/s$ and $9m/s$ respectively. These results indicate that the augmented control law is capable of achieving an improvement in position tracking performance with a 50% decrease in the average and maximum tracking error for a target velocity of $9m/s$.

Table 2: Evaluation results over 100 randomly generated evaluation trajectories with target velocity of $3m/s$.

Metric		**Conventional**	**Augmented**
$RMSE\|\|\boldsymbol{p}-\boldsymbol{p}_{ref}\|\|_2$	$[cm]$	4.98	4.64
$\max\|\|\boldsymbol{p}-\boldsymbol{p}_{ref}\|\|_2$	$[cm]$	8.60	6.69
$RMS\|\|\boldsymbol{v}\|\|_2$	$[m/s]$	3.33	3.33
$\max\|\|\boldsymbol{v}\|\|_2$	$[m/s]$	4.32	4.32

Table 3: Evaluation results over 100 randomly generated evaluation trajectories with target velocity of $6m/s$.

Metric		**Conventional**	**Augmented**
$RMSE\|\|\boldsymbol{p}-\boldsymbol{p}_{ref}\|\|_2$	$[cm]$	22.0	14.9
$\max\|\|\boldsymbol{p}-\boldsymbol{p}_{ref}\|\|_2$	$[cm]$	42.3	26.5
$RMS\|\|\boldsymbol{v}\|\|_2$	$[m/s]$	6.69	6.67
$\max\|\|\boldsymbol{v}\|\|_2$	$[m/s]$	8.69	8.64

Table 4: Evaluation results over 100 randomly generated evaluation trajectories with target velocity of $9m/s$.

Metric		**Conventional**	**Augmented**
$RMSE\|\|\boldsymbol{p}-\boldsymbol{p}_{ref}\|\|_2$	$[cm]$	126	62.3
$\max\|\|\boldsymbol{p}-\boldsymbol{p}_{ref}\|\|_2$	$[cm]$	298	145
$RMS\|\|\boldsymbol{v}\|\|_2$	$[m/s]$	10.3	10.1
$\max\|\|\boldsymbol{v}\|\|_2$	$[m/s]$	14.1	13.6

To further analyze the impact of the augmentation, we compare the error in achieved linear acceleration and commanded linear acceleration for the conventional control law against the augmented control law. The results for a simulated tracking task on a random evaluation trajectory with target velocity of $6m/s$ are shown in Figure 3. Here, the augmentation policy action β is shown in blue, the orange curve is the acceleration error using the conventional control law, the blue curve is the acceleration error using the augmented control law. As can be seen, the control law augmentation achieves a smaller error as compared to the conventional control law. This shows that the policy successfully learned to represent the neglected terms in the Taylor expansion from the control law derivation, thus achieving enhanced trajectory tracking performance.

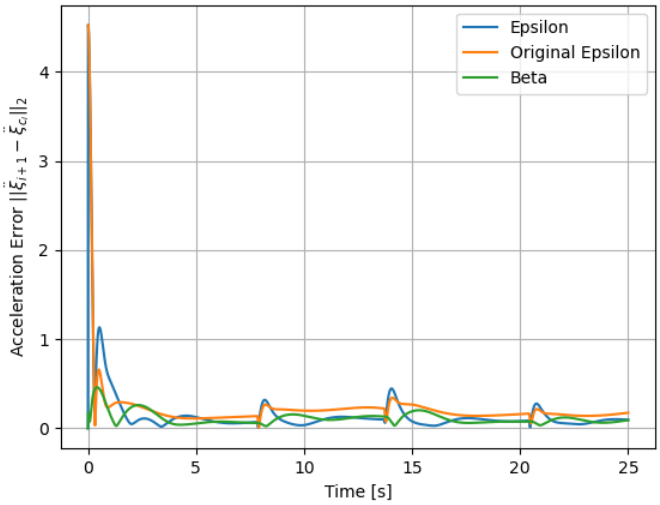

Figure 7: Improvemen in acceleration error through control law augmentation. The augmentation policy action β is the blue curve, the orange curve is the acceleration error using the conventional control law, the blue curve is the acceleration error using the augmented control law.

To further evaluate the capabilities of the proposed augmented control law, the disturbance rejection performance is evaluated. This is done by simulating the quadcopter in hover mode and subjecting it to a disturbance. At the beginning of the simulation, the quadcopter experiences a wind disturbance of $3m/s$, decomposed over each axis as $\boldsymbol{v}_{dist} = \begin{bmatrix}\sqrt{3} & \sqrt{3} & \sqrt{3}\end{bmatrix} m/s$.

The result of the wind disturbance rejection experiment is shown in Figure 8. This figure shows the response of the quadcopter after receiving a disturbance in velocity. Both the response using the conventional and augmented control law are shown. It can be seen that the augmented control law achieves a smaller error in position than the conventional control law, as well as a faster convergence to the hover position. A quantitative comparison is given in Table 5.

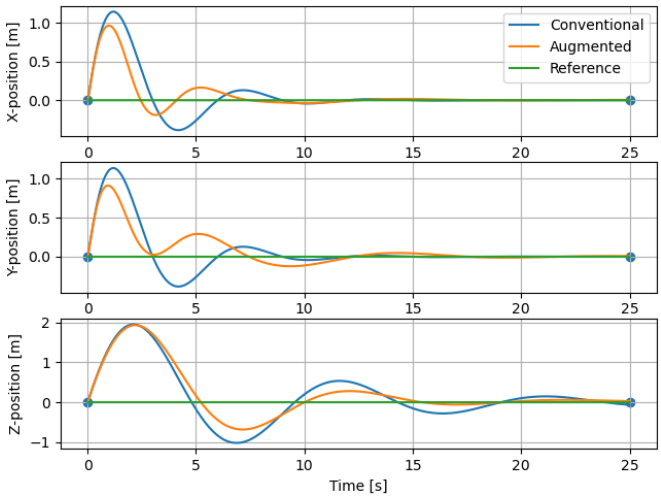

Figure 8: Quadcopter positional response against velocity disturbance $\boldsymbol{v}_{dist} = \begin{bmatrix}\sqrt{3} & \sqrt{3} & \sqrt{3}\end{bmatrix} m/s$ for the conventional and augmented control law.

Table 5: Comparison in disturbance rejection performance between conventional and augmented control law.

Metric		Conventional	Augmented
$RMSE\|\|\boldsymbol{p} - \boldsymbol{p}_{ref}\|\|_2$	$[cm]$	81.4	72.4
$\max \|\|\boldsymbol{p} - \boldsymbol{p}_{ref}\|\|_2$	$[cm]$	235	202
$RMS\|\|\boldsymbol{v}\|\|_2$	$[m/s]$	0.66	0.57
$\max \|\|\boldsymbol{v}\|\|_2$	$[m/s]$	3.00	3.00

While the experiments show that the control law augmentation performs better than the conventional control law, there are limitations to the approach. These mainly come from the assumption that the error between achieved linear acceleration and commanded linear acceleration are due to the neglected terms in the Taylor expansion. However, another assumption in incremental nonlinear dynamic inversion is the principle of time scale separation. This principle assumes that the inner loop dynamics, in this case the attitude dynamics, are sufficiently fast such that the commanded attitude is achieved instantly relative to the slower outer loop dynamics, the linear acceleration dynamics. Hence, if for a given quadcopter the time it takes to achieve a commanded attitude command becomes too large, than part of the error between the achieved linear acceleration and commanded linear acceleration is attributed to the slower attitude dynamics. In this case, augmenting the control law with an additional term will not be effective since the quadcopter's agility is inherently limited by its physical capabilities.

Conclusion

This paper introduces a novel approach to augment the incremental nonlinear dynamic inversion control law for agile trajectory tracking for a quadcopter at high speeds. This augmentation allows the quadcopter to follow trajectories that are three-times differentiable, reducing positional error by 50% compared to the conventional control law at a target velocity of $9m/s$. Additionally, it improves the quadcopter's ability to reject disturbances, showing a 15% decrease in positional error and quicker recovery from a wind disturbance of $3m/s$. These findings demonstrate the efficacy of the augmentation in enabling high-speed agile flight, particularly in scenarios with significant aerodynamic and nonlinear effects. However, one limitation is the assumption that attitude dynamics respond quickly enough to achieve commanded roll and pitch, thereby limiting their contribution to the error between achieved and commanded linear acceleration.

Subsequent research could explore employing alternative methods to train the control augmentation, in a more general context. This could involve generalizing across various drone configurations and extending to a wider spectrum of velocities. Another avenue for investigation is integrating the error stemming from slower attitude dynamics into the policy learning process.

Acknowledgements

This project was completed as part of the Stanford University course on Advanced Topics in Sequential Decision Making; AA229.

References

[Acquatella et al., 2012] Acquatella, P., Falkena, W., Van Kampen, E.-J., and Chu, Q. (2012). Robust nonlinear spacecraft attitude control using incremental nonlinear dynamic inversion.

[Brockman et al., 2016] Brockman, G., Cheung, V., Pettersson, L., Schneider, J., Schulman, J., Tang, J., and Zaremba, W. (2016). Openai gym.

[Richards et al., 2021] Richards, S. M., Azizan, N., Slotine, J.-J., and Pavone, M. (2021). Adaptive-control-oriented meta-learning for nonlinear systems.

[Schulman et al., 2017] Schulman, J., Wolski, F., Dhariwal, P., Radford, A., and Klimov, O. (2017). Proximal policy optimization algorithms.

[Smeur et al., 2018] Smeur, E., de Croon, G., and Chu, Q. (2018). Cascaded incremental nonlinear dynamic inversion for mav disturbance rejection. *Control Engineering Practice*, 73:79–90.

[Tal and Karaman, 2021] Tal, E. and Karaman, S. (2021). Accurate tracking of aggressive quadrotor trajectories using incremental nonlinear dynamic inversion and differential flatness. *IEEE Transactions on Control Systems Technology*, 29(3):1203–1218.

[Zhang and Ran, 2023] Zhang, X. and Ran, M. (2023). Meta-learning-based incremental nonlinear dynamic inversion control for quadrotors with disturbances. *Applied Sciences*, 13:11844.