



Addressing incremental backstepping control limitations with direct online Gaussian process adaptation

Dmitry Ignatyev^{*}, Antonios Tsourdos

School of Aerospace, Transport and Manufacturing, Cranfield University, Cranfield, UK

ARTICLE INFO

Keywords:

Gaussian processes
Data-driven
Sensor based control
Incremental backstepping
stability
Nonlinear control systems
Uncertainty

ABSTRACT

Sensor-based incremental control is a recently developed technique that reduces dependency on precise model knowledge. This approach uses measurements or estimates of current state derivatives and actuator states to linearise the dynamics with respect to the previous time instant, albeit at the expense of increased sensitivity to measurement quality. Unknown system behaviour due to unforeseen malfunctions in measurement or actuation systems can lead to significant performance degradation. The novelty of this paper lies in a new method that combines the reduced model dependency of sensor-based Incremental Backstepping (IBKS) control with the adaptive capabilities of data-driven Gaussian Processes (GPs). The resulting controller exhibits significantly reduced sensitivity to model and measurement uncertainties. A theoretical proof of the global uniform ultimate boundedness of the IBKS tracking error is provided. The direct GP-based adaptation reduces the error bound and offers both long-term dependency learning and noise filtering capabilities. The effectiveness of the proposed approach is demonstrated through a missile flight control example.

1. Introduction

Ensuring superior control performance while maintaining system stability is essential in the design of safety-critical systems. Many real-world systems exhibit nonlinear dynamics, which makes their control particularly challenging. Gain-scheduling of linear feedback controllers is a common solution in commercial applications, providing stabilisation and satisfactory tracking across a range of conditions [1]. However, these controllers tend to underperform in scenarios involving severe and unpredictable changes in system behaviour, where their functionality becomes limited. Nonlinear Dynamics Inversion (NDI) and Backstepping (BS) techniques have gained popularity for their ability to globally linearize system dynamics and decouple control efforts [2–4]. NDI is a feedback linearization technique [5] that aims to invert the original dynamics and replace them with the desired linear dynamics [6]. BS is a control method for strict-feedback systems [3] and can be considered a particular case of the block-controllable form, which involves the iterative application of the regular form [7–9]. To enhance the robustness and fault tolerance of NDI and BS controls, incremental sensor-based approaches have been developed [10–12]. These methods exploit the idea that nonlinear system dynamics can be linearised incrementally based on the system states from the previous time step, thereby allowing real-time measurements or estimates of state vector derivatives to partially replace detailed model knowledge.

However, the incremental linearisation introduces error dynamics, which can degrade control stability and performance. Despite

^{*} Corresponding author.

E-mail address: d.ignatyev@cranfield.ac.uk (D. Ignatyev).

their robustness to certain failures and model uncertainties, incremental controllers remain sensitive to the quality of sensor information [13–16]. Noisy measurements can degrade control system performance, necessitating filtering techniques such as Kalman filtering or Gaussian Process methods [17,18]. However, filtering highly noisy signals can introduce significant delays in the measurement and control loops, which compromises control performance [14,19–22]. Input Scaling Gain (ISG) [23] has been proposed to mitigate the effects of high-frequency noise and feedback loop delays; however, it reduces the closed-loop response bandwidth, rendering the controller more conservative. Furthermore, sensor or actuator failures can lead to devastating consequences [14].

Adaptive augmentation strategies have been proposed to compensate for uncertainties due to unknown control efficiency matrices [14,20,24,25]. In cases involving unconventional control actuation, such as morphing wings [26] or plasma actuators [27], the system dynamics become significantly more complicated, and control efficiency depends on the states, losing the input-affine property. However, most techniques address only one of these problems—either sensor noise, uncertainty in the control efficiency matrix, or actuator model inaccuracies. A generalised adaptation strategy capable of tackling a broader range of issues could provide a universal remedy.

Neural networks, including Radial Basis Functions (RBFs), are popular for online identification and adaptive control due to their universal approximation capabilities (for example, see [28,29]). RBFs, being linear in the parameters, offer advantages over multilayer perceptron neural networks. However, their performance heavily depends on the selection of RBF centres, which are typically pre-allocated over a presumed domain [28,30]. If the system states deviate from these centres, the RBFs fail to capture the uncertainties.

To address these challenges, we propose using Gaussian Processes (GPs) to implement an adaptive control loop that augments the baseline IBKS controller. GPs introduce a Bayesian paradigm to adaptive control, treating estimation as a statistical problem [31]. Unlike previous approaches [32] that collected data prior to control design, limiting controller performance, our method utilises GPs to represent uncertainties online, as a distribution over functions. GPs dynamically choose new kernel locations to ensure domain coverage without prior assumptions. Measurement noise is explicitly handled, and there is no need for preallocating RBF centres. GP-based Bayesian inference overcomes standard gradient-based parameter update shortcomings, such as lack of convergence guarantees and potential instability under noisy conditions [33,34]. This method was applied in the design of direct Model Reference Adaptive Control (GP-MRAC) [35], where GPs were used to match uncertainties and produce compensating control commands. Additionally, GP-based indirect adaptive augmentation to IBKS [20] addressed actuation system failures in aircraft, using sparse GP adaptation to account for flight control uncertainties. Sparse GP adaptation was used to take into account other possible uncertainties in flight control, for example, change of the inertia due to slung payload motion [36,37].

Stability analysis of the system dynamics is essential for control clearance [38–40]. In this work, we analyse the IBKS control strategy and theoretically prove that a controller designed for incremental dynamics can stabilise the original nonlinear dynamics. We derive a theoretical estimate of the tracking error dynamics and demonstrate that it is bounded. To compensate for the error dynamics, we develop a direct GP-based adaptation of the incremental controller, capable of addressing a wide range of uncertainties, including linearisation errors, measurement inaccuracies, and actuation failures. This approach differs from the GP-based indirect adaptive augmentation to IBKS [20], which was developed to tackle only a specific type of error arising in the control efficiency matrix. Our detailed analysis of the indirect and direct GP adaptations to IBKS, provided in this paper, shows that direct adaptation can handle more general types of uncertainties. These uncertainties could arise from imperfect knowledge of the model cascade coupling matrices, linearisation of the system, failed estimation or measurement of state vector derivatives, and faulty actuation. Our approach combines the reduced model dependency of sensor-based IBKS control with the adaptive capabilities of GPs. The resulting tracking error is globally uniformly ultimately bounded, with the ultimate bound reducible by increasing the sampling rate and control gains, minimising the residual vector from the projection in the Reproducing Kernel Hilbert Space (RKHS), or collecting new measurements at points of high uncertainty.

The paper is organised as follows. The Introduction is followed by Section II, which considers the incremental form of the system dynamics. Section III describes the design of the incremental control law. The stability analysis of the IBKS is presented in Section IV. Section V discusses Sparse Online GPs, while Section VI addresses the convergence analysis of the proposed framework. Section VII considers an example dynamical system, specifically the longitudinal flight dynamics of a missile with the corresponding incremental controller. Section VIII presents the simulation results, and the paper is summarised in the Conclusion.

2. Incremental form of system dynamics

We consider nonlinear strict-feedback systems that are affine in the control input [3], which can be represented in the following form:

$$\dot{\mathbf{x}}_i = \mathbf{F}_i(\mathbf{w}_i) + \mathbf{G}_i(\mathbf{w}_i)\mathbf{x}_{i+1}, i = 1, \dots, h-1, \dot{\mathbf{x}}_h = \mathbf{F}_h(\mathbf{w}_h) + \mathbf{G}_u \mathbf{u}, \quad (1)$$

where $\mathbf{x}_i \in \mathcal{X}_{\mathbf{x}_i} \subseteq \mathbb{R}^n$ are the system states, $\mathbf{u} \in \mathcal{U} \subseteq \mathbb{R}^m$ is the control input, $\mathbf{w}_i = [\mathbf{x}_1^T, \dots, \mathbf{x}_i^T]^T$, $i = 1, \dots, h$ represents concatenations introduced for notation purposes. Similarly, the notation $\mathcal{W}_i = \mathcal{X}_{\mathbf{x}} \times \mathcal{X}_{\xi_1} \times \mathcal{X}_{\xi_2} \times \dots \times \mathcal{X}_{\xi_i}$ is used to denote concatenations of subspaces of the state spaces. For brevity, unless stated otherwise, whenever the subscript i is utilized, the full set, i.e., $i = 1, \dots, h$, is referred to. $\mathbf{x}_{i+1}$ is considered as a control input for $\mathbf{x}_i$ sub-system.

Assumption 1. The nonlinear functions $\mathbf{F}_i$ represent the unknown system dynamics, whereas the nonlinear functions $\mathbf{G}_i$, $i = 1, \dots, h-1$ and $\mathbf{G}_u$ are assumed to be known. $\mathbf{F}_i$, $\mathbf{G}_i$, $i = 1, \dots, h-1$ and $\mathbf{G}_u$ are assumed to be Lipschitz continuous:

$$|F_i(\mathbf{y}_1) - F_i(\mathbf{y}_2)| < \hat{\Gamma}_{F_i} |\mathbf{y}_1 - \mathbf{y}_2|, \quad i = 1 \dots n,$$

$$|G_i(\mathbf{y}_1) - G_i(\mathbf{y}_2)| < \hat{\Gamma}_{G_i} |\mathbf{y}_1 - \mathbf{y}_2|, \quad i = 1 \dots n.$$

$\mathbf{x}_i$ are continuous and bounded on the compact $[0; \tau]$, where τ denotes the sampling time; therefore, $\mathbf{x}_i$ are also Lipschitz continuous functions. Furthermore, $\mathbf{x}_i$ are bounded $|\mathbf{x}_i - \mathbf{x}_{i_0}| \leq \delta_i \tau$, where $\delta_i \geq 0$. Therefore, the following inequalities are true

$$|F_i(\mathbf{w}_{i_0}) - F_i(\mathbf{w}_i)| < \Gamma_{F_i} \tau, \quad i = 1 \dots h,$$

$$|G_i(\mathbf{w}_{i_0}) - G_i(\mathbf{w}_i)| < \Gamma_{G_i} \tau, \quad i = 1 \dots h. \quad (2)$$

This assumption is commonly made and is applicable to a wide range of real-world systems.

The IBKS is based on the concept that system dynamics (1) can be represented incrementally. This is achieved by employing sensor and actuator measurements for feedback, significantly reducing the dependence on precise knowledge of the plant dynamics.

By expanding Eq. (1) into a Taylor series around $(\mathbf{x}_0, \mathbf{u}_0)$ corresponding to the previous time instant t_0 , the dynamics in Eq. (1) can be expressed in the following form:

$$\begin{aligned} \dot{\mathbf{x}}_i &= \dot{\mathbf{x}}_{i_0} + \mathbf{A}_{x_i} \Delta \mathbf{x}_i + \mathbf{B}_{x_i} \Delta \mathbf{x}_{i+1} + \mathcal{O}(\Delta \mathbf{x}_i^2), \quad i = 1, \dots, h-1, \\ \dot{\mathbf{x}}_h &= \dot{\mathbf{x}}_{h_0} + \mathbf{A}_{x_h} \Delta \mathbf{x}_h + \mathbf{B}_u \Delta \mathbf{u} + \mathcal{O}(\Delta \mathbf{x}_h^2). \end{aligned} \quad (3)$$

where

$$\begin{aligned} \mathbf{A}_{x_i} &= \left. \frac{\partial(F_i(\mathbf{w}_i) + \mathbf{G}_i(\mathbf{w}_i)\mathbf{x}_{i+1})}{\partial \mathbf{x}_i} \right|_{\mathbf{w}_i=\mathbf{w}_{i_0}} \\ \mathbf{B}_{x_i} &= \left. \frac{\partial(F_i(\mathbf{w}_i) + \mathbf{G}_i(\mathbf{w}_i)\mathbf{x}_{i+1})}{\partial \mathbf{x}_{i+1}} \right|_{\mathbf{w}_i=\mathbf{w}_{i_0}} = \mathbf{G}_i(\mathbf{w}_{i_0}) \\ \mathbf{B}_u &= \left. \frac{\partial(f_h(\mathbf{w}_h) + \mathbf{G}_h(\mathbf{w}_h)\mathbf{u})}{\partial \mathbf{u}} \right|_{\mathbf{w}_h=\mathbf{w}_{h_0}} = \mathbf{G}_h(\mathbf{w}_{h_0}) \end{aligned} \quad (4)$$

are partial derivatives with respect to the state vector parameters and the control input; $\Delta \mathbf{x}_i = \mathbf{x}_i - \mathbf{x}_{i_0}$, $\Delta \mathbf{u} = \mathbf{u} - \mathbf{u}_0$ are increments in the state derivatives and the control input, $\mathbf{x}_{i_0}$, $\mathbf{u}_0$ can be measured by sensors. This means an approximate linearization about the previously measured signals is performed incrementally; $\mathcal{O}(\Delta \mathbf{x}_i^2)$ are high order terms in the Taylor series. The system dynamics in Eq. (3) are updated at frequency ν , corresponding to the sampling time $\tau = \frac{1}{\nu}$.

Assumption 2. Due to the timescale separation principle, which holds well for some real systems ([11,41]), the following set of assumptions can be made when the sampling frequency is sufficiently high:

$$\Delta \mathbf{x}_i \ll \Delta \mathbf{x}_{i+1}; \Delta \mathbf{x}_h \ll \Delta \mathbf{u}. \quad (5)$$

Such assumptions become feasible for a real plant because, over small time increments, the control inputs $\mathbf{x}_{i+1}$, $\mathbf{u}$ directly influence the state derivatives $\dot{\mathbf{x}}_i$, while changes in the state vector itself occur more slowly, as they evolve only through the integration of these state derivatives.

Assumption 3. The actuators are presumed to be sufficiently fast to achieve the required input increments within the small sampling interval. This assumption complements the previous one, as it considers the actuator dynamics to evolve significantly faster than the system states. Such an assumption is reasonable for many practical actuation systems, whose dynamics can often be approximated by linear first- or second-order models [42,43].

Assumption 4. Complete and accurate knowledge of the system state is assumed. Sensors providing state derivatives are also considered to be available for this study. Some acceleration measurements may be obtained directly, while others may be estimated by differentiating other measurements and applying appropriate filtering.

Under this assumption, the incremental dynamics can be reformulated as follows:

$$\begin{aligned} \hat{\dot{\mathbf{x}}}_i &= \dot{\mathbf{x}}_{i_0} + \mathbf{B}_{x_i} \Delta \hat{\mathbf{x}}_{i+1} + \mathcal{O}(\Delta \hat{\mathbf{x}}_i^2), \quad i = 1, \dots, h-1, \\ \hat{\dot{\mathbf{x}}}_h &= \dot{\mathbf{x}}_{h_0} + \mathbf{B}_u \Delta \mathbf{u} + \mathcal{O}(\Delta \hat{\mathbf{x}}_h^2). \end{aligned} \quad (6)$$

From a practical standpoint, Assumption 4 is quite stringent, as sensor faults or inaccuracies can lead to significant risks in real-world applications. To address this, various mitigation strategies may be employed, including sensor fusion, hardware redundancy, filtering, and state estimation techniques. In this work, we propose an additional strategy based on GP adaptation, which enables the

system to compensate for sensor imperfections in a data-driven manner.

3. Incremental backstepping control

In this section, we follow a conventional IBKS procedure [14,44,45] to design a control law for the linearised system (6). The error for system (6) is defined as follows:

$$\mathbf{z}_i = \mathbf{x}_{i\text{ref}} - \hat{\mathbf{x}}_i, \quad i = 1, \dots, h. \quad (7)$$

The error dynamics of the system (6) are governed by the following equations:

$$\dot{\mathbf{z}}_i = \dot{\mathbf{x}}_{i\text{ref}} - \dot{\hat{\mathbf{x}}}_i, \quad i = 1, \dots, h. \quad (8)$$

For $\mathbf{x}_1$ -subsystem, the Lyapunov function is constructed as follows

$$V_1 = \frac{1}{2} \mathbf{z}_1^T \mathbf{z}_1, \quad (9)$$

such that

$$\dot{V}_1 = \mathbf{z}_1^T \dot{\mathbf{z}}_1 = -\mathbf{z}_1^T \mathbf{W}_1 \mathbf{z}_1 \quad (10)$$

and $\mathbf{W}_1$ is a diagonal matrix with positive main diagonal coefficients. Considering the error dynamics in Eqs. (6) and (8), Eq. (10) can be re-written as

$$-\mathbf{z}_1^T \mathbf{W}_1 \mathbf{z}_1 \cong \mathbf{z}_1^T \left(\dot{\mathbf{x}}_{1\text{ref}} - \dot{\mathbf{x}}_{10} - \mathbf{B}_{x_1} \Delta \hat{\mathbf{x}}_2 \right). \quad (11)$$

Desired virtual control for $\hat{\mathbf{x}}_1$ -subsystem $\nu_1 = \Delta \hat{\mathbf{x}}_2$ can be derived to guarantee the stability of the $\hat{\mathbf{x}}_1$ -sub-system:

$$\nu_1 = \mathbf{B}_{x_1}^{-1} \left(\dot{\mathbf{x}}_{1\text{ref}} - \dot{\mathbf{x}}_{10} + \mathbf{W}_1 \mathbf{z}_1 \right). \quad (12)$$

In a similar way, step-by-step, we can introduce the Lyapunov function for the whole system (6), which guarantees the stability of the extended system dynamics described in Eq. (8):

$$V = V_1 + \sum_{i=2}^h \frac{1}{2} \mathbf{z}_i^T \mathbf{z}_i = \frac{1}{2} \mathbf{z}^T \mathbf{z}, \quad (13)$$

$$\dot{V} = \mathbf{z}^T \dot{\mathbf{z}} = -\mathbf{z}^T \mathbf{W} \mathbf{z}$$

where $\mathbf{z} = [\mathbf{z}_1^T \mathbf{z}_2^T \dots \mathbf{z}_h^T]^T$, $\mathbf{W}$ is a diagonal matrix with positive main diagonal coefficients.

Then, the corresponding virtual control can be derived as follows:

$$\nu_i = \mathbf{B}_{x_i}^{-1} \left(\dot{\mathbf{x}}_{i\text{ref}} - \dot{\mathbf{x}}_{i0} + \mathbf{W}_i \mathbf{z}_i \right), \quad i = 1, \dots, h-1$$

$$\Delta \mathbf{u} = \nu_{h+1} = \mathbf{B}_u^{-1} \left(\dot{\mathbf{x}}_{h\text{ref}} - \dot{\mathbf{x}}_{h0} + \mathbf{W}_h \mathbf{z}_h \right) \quad (14)$$

IBKS does not perform a coordinate transformation or globally cancel system nonlinearities, as is done in Isidori's feedback linearisation. However, it implements a control law that creates linear behaviour locally, centred at the previously measured state and input. IBKS does not linearise the plant model itself, but rather linearises the effect of the control action within a local neighbourhood.

The IBKS control derived using the conventional incremental approach in Eq. (14) does not account for the error introduced by approximating the original system in Eq. (1) with the incremental system in Eq. (6). The next section provides a theoretical analysis of the bounds on the control error resulting from this approximation.

4. Stability analysis of incremental backstepping

The incremental system dynamics in Eq. (6) differ from the original system dynamics in Eq. (1), which may cause instability of the control law designed in Eq. (14). This section analyses the effect of this difference on system stability.

Theorem 1. *Let us consider the system represented by Eq. (1) with incremental control law given by Eq. (14) $|\mathbf{x}_{i+10}| \leq \mathbf{x}_{i+1\text{max}}$, $i = 1, \dots, h$ and $|\mathbf{u}| \leq u_{\text{max}}$ are bounded due to the boundedness of the initial conditions. The system's tracking error is globally uniformly ultimately bounded, and the ultimate error bound is given by*

$$\lim_{t \rightarrow \infty} |\mathbf{z}| \leq \max \left\{ \tau \frac{\sqrt{\sum_{i=1}^{h-1} (\Gamma_{F_i} + \Gamma_{G_i} \delta_i \tau)^2 + (\Gamma_{F_h} + \Gamma_{G_h} u_{\text{max}})^2}}{\min_i (\mathbf{W}_i)}, \tau^2 \sqrt{\sum_{i=1}^{h-1} (\Gamma_{F_i} + \Gamma_{G_i} \delta_i \tau)^2 + (\Gamma_{F_h} + \Gamma_{G_h} u_{\text{max}})^2} \right\}.$$

The tracking error of the original system in Eq. (1) under the incremental control laws in Eqs. (12) and (14) can be defined as:

$$\dot{\mathbf{e}}_i = \dot{\mathbf{x}}_{i_{ref}} - \dot{\mathbf{x}}_i = \dot{\mathbf{z}}_i - \dot{\Delta}_i, \quad i = 1, \dots, h, \quad (15)$$

where $\Delta_i = \hat{\mathbf{x}}_i - \mathbf{x}_i$ is the difference between the incremental dynamics in Eq. (6) and the original dynamics in Eq. (1).

Consider the difference dynamics arising from the incremental control in Eq. (14):

$$\dot{\Delta} = \begin{bmatrix} \hat{\mathbf{x}}_1 - \mathbf{x}_1 \\ \dots \\ \hat{\mathbf{x}}_h - \mathbf{x}_h \end{bmatrix} = \begin{bmatrix} \dot{\mathbf{x}}_{1_0} + \mathbf{B}_{x_1} \Delta \hat{\mathbf{x}}_2 - \mathbf{F}_1(\mathbf{w}_1) + \mathbf{G}_1(\mathbf{w}_1) \hat{\mathbf{x}}_2 \\ \dots \\ \dot{\mathbf{x}}_{h_0} + \mathbf{B}_u \Delta \hat{\mathbf{u}} - \mathbf{F}_h(\mathbf{w}_h) + \mathbf{G}_u(\mathbf{w}_h) \mathbf{u} \end{bmatrix} \quad (16)$$

Considering that $\dot{\mathbf{x}}_{i_0}$ represents the real system dynamics in Eq. (1) at the measurement instant, and using the definitions of $\mathbf{B}_{x_i}$, and $\mathbf{B}_u$, the difference dynamics can be expressed as:

$$\dot{\Delta} = \begin{bmatrix} \mathbf{F}_1(\mathbf{w}_{1_0}) - \mathbf{F}_1(\mathbf{w}_1) + [\mathbf{G}_1(\mathbf{w}_{1_0}) - \mathbf{G}_1(\mathbf{w}_1)](\mathbf{x}_{2_0} + \Delta \hat{\mathbf{x}}_2) \\ \dots \\ \mathbf{F}_h(\mathbf{w}_{1_0}) - \mathbf{F}_h(\mathbf{w}_1) + [\mathbf{G}_u(\mathbf{w}_{h_0}) - \mathbf{G}_u(\mathbf{w}_h)](\mathbf{u}_0 + \Delta \mathbf{u}) \end{bmatrix} = \Delta \mathbf{F} + \Delta \mathbf{G}^T \mathbf{U}, \quad (17)$$

where

$$\Delta \mathbf{F} = \begin{bmatrix} \mathbf{F}_1(\mathbf{w}_{1_0}) - \mathbf{F}_1(\mathbf{w}_1) \\ \dots \\ \mathbf{F}_h(\mathbf{w}_{1_0}) - \mathbf{F}_h(\mathbf{w}_1) \end{bmatrix}, \quad \Delta \mathbf{G} = \begin{bmatrix} \mathbf{G}_1(\mathbf{w}_{1_0}) - \mathbf{G}_1(\mathbf{w}_1) \\ \dots \\ \mathbf{G}_h(\mathbf{w}_{1_0}) - \mathbf{G}_h(\mathbf{w}_1) \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} \mathbf{x}_{2_0} + \Delta \hat{\mathbf{x}}_2 \\ \dots \\ \mathbf{x}_{h-1_0} + \Delta \hat{\mathbf{x}}_{h-1} \\ \mathbf{u}_0 + \Delta \mathbf{u} \end{bmatrix}. \quad (18)$$

In this case the tracking error of the original system (15) is given by the following equation:

$$\dot{\mathbf{e}} = \dot{\mathbf{z}} + \Delta \mathbf{F} + \Delta \mathbf{G}^T \mathbf{w}. \quad (19)$$

The above relationship clearly illustrates that the tracking error of the original system in Eq. (1) under the incremental control in Eq. (14) arises not only from the IBKS tracking error $\mathbf{z}$ but also from the dynamics error $\dot{\Delta}$, which is caused by the incremental linearisation errors $\Delta \mathbf{F}$ and $\Delta \mathbf{G}$.

The error in this case is as follows

$$\mathbf{e}(\tau) = \mathbf{z} + \int_0^\tau (\Delta \mathbf{F} + \Delta \mathbf{G}^T \mathbf{w}) dt. \quad (20)$$

A candidate Lyapunov function can be composed as follows:

$$V = \frac{1}{2} \mathbf{e}^T \mathbf{e}. \quad (21)$$

The derivative is given by

$$\dot{V} = \mathbf{e}^T \dot{\mathbf{e}} = \left(\mathbf{z} + \int_0^\tau (\Delta \mathbf{F} + \Delta \mathbf{G}^T \mathbf{w}) dt \right)^T (\dot{\mathbf{z}} + \Delta \mathbf{F} + \Delta \mathbf{G}^T \mathbf{w}) \quad (22)$$

Considering that $\dot{\mathbf{z}} = -\mathbf{W}\mathbf{z}$, we have

$$\dot{V} = \left(\mathbf{z} + \int_0^\tau (\Delta \mathbf{F} + \Delta \mathbf{G}^T \mathbf{w}) dt \right)^T (-\mathbf{W}\mathbf{z} + \Delta \mathbf{F} + \Delta \mathbf{G}^T \mathbf{w}). \quad (23)$$

The derivative $\dot{V}$ is negative when

$$|\mathbf{z}| \geq \max \left\{ \frac{|\Delta \mathbf{F} + \Delta \mathbf{G}^T \mathbf{w}|}{\min_i \mathbf{W}}, \left| \int_0^\tau (\Delta \mathbf{F} + \Delta \mathbf{G}^T \mathbf{w}) dt \right| \right\}. \quad (24)$$

Due to Lipschitz conditions in Eq. (2), the following inequalities hold: $\Delta F_i \leq \Gamma_{F_i} \tau$, $i = 1 \dots h$, $\Delta G_i \leq \Gamma_{G_i} \tau$, $i = 1 \dots h$, $\Delta \hat{\mathbf{x}}_i = \delta_i \tau$, $i = 1 \dots h$. Additionally, the following relationship holds:

$$\left| \int_0^\tau (\Delta \mathbf{F} + \Delta \mathbf{G}^T \mathbf{w}) dt \right| \leq \int_0^\tau |\Delta \mathbf{F} + \Delta \mathbf{G}^T \mathbf{w}| dt. \quad (25)$$

Therefore,

$$|\mathbf{z}| \geq \max \left\{ \tau \frac{\sqrt{\sum_{i=1}^{h-1} (\Gamma_{F_i} + \Gamma_{G_i} \delta_i \tau)^2 + (\Gamma_{F_h} + \Gamma_{G_h} u_{\max})^2}}{\min(\mathbf{W}_i)}, \tau^2 \sqrt{\sum_{i=1}^{h-1} (\Gamma_{F_i} + \Gamma_{G_i} \delta_i \tau)^2 + (\Gamma_{F_h} + \Gamma_{G_h} u_{\max})^2} \right\}. \quad (26)$$

This condition holds for sufficiently large $|\mathbf{z}|$ and, therefore, the tracking error $|\mathbf{z}|$ is globally uniformly ultimately bounded. ■

These results clearly demonstrate that the IBKS can stabilise the original system in Eq. (1). At the same time, due to the linearisation of the dynamics with respect to the previous time step, this control strategy still exhibits a tracking error that is globally uniformly ultimately bounded. The tracking error is determined by the ratio of the system dynamics growth rates Γ_{F_i} , Γ_{G_i} , Γ_{F_h} , Γ_{G_h} and the sampling time τ . [Theorem 1](#) compellingly demonstrates that as the system's growth rate decreases and the sampling rate increases, the tracking error is reduced. One possible way to reduce the error to some extent is to increase the control gains $\mathbf{W}$.

The original system dynamics in Eq. (1) are formulated as a continuous nonlinear strict-feedback system that is affine in the control input. Within the IBKS approach, at each step, a control increment is synthesised to ensure that the tracking error dynamics possess a strict Lyapunov function whose derivative is negative definite in a neighbourhood of the origin. Consequently, the closed-loop system is locally asymptotically stable, and the incremental control law behaves as a locally linear feedback around the previously measured state. Therefore, it acts as a discrete controller, constructing control laws that yield a locally stabilising effect for the original continuous system in Eq. (1). The system's ultimate error bound is reduced with an increase in the sampling frequency, allowing the discrete controller to approximate linear behaviour more closely.

[Theorem 1](#) shows the limitations of IBKS, which could be even worse in the case of sensor noise or failure, as it requires precise measurements or estimations of the state vector derivatives and actuator positions obtained at high rates. In this case, the expression for the Lyapunov function derivative should be modified to include the uncertainty due to measurement noise:

$$\dot{V} = \left(\mathbf{z} + \int_0^\tau (\Delta \mathbf{F} + \Delta \mathbf{G}^T \mathbf{w}) dt + \Delta_m \right)^T (-\mathbf{W}\mathbf{z} + \Delta \mathbf{F} + \Delta \mathbf{G}^T \mathbf{w} + \dot{\Delta}_m) \quad (27)$$

Thus, the stability and performance of the system can be significantly compromised by inadequate measurements or faulty conditions. In the next section, we introduce a sparse online Gaussian Process adaptation to overcome these obstacles.

5. Sparse on-line Gaussian processes

In IBKS, the control design proceeds recursively through a sequence of blocks or layers. At each stage, a virtual control signal is constructed for block i , which subsequently becomes part of the dynamics in block $i+1$. Consequently, the derivative of the virtual control signal at block i appears explicitly in the dynamics of block $i+1$. Additionally, any uncompensated perturbations or modelling errors in block i propagate forward, effectively acting as disturbances in block $i+1$. This coupling between layers is an inherent feature of the backstepping structure, whether classical or incremental. Therefore, in IBKS, these residual perturbations directly influence the subsequent block. Although IBKS operates incrementally—constructing the control law around a nominal input or state—the propagation of errors or unmodelled dynamics remains a challenge.

In this section, we propose using GP to address the uncertainties caused by approximation errors from incremental linearization, as well as other potential negative effects such as measurement noise or failures at each block level i .

Although a multivariate GP formulation exists [46] and can be employed to approximate each uncertainty in (27) with a single GP, in our approach, we employ multiple smaller GPs, with each GP assigned to approximate the individual uncertainties Δ at a specific block level of the incremental controller. This design choice enables more targeted modelling of local uncertainties and avoids the computational and modelling complexity associated with a single multivariate GP for the entire system.

Usage of multiple scalar GPs increases the number of GP models to maintain, but each individual GP deals with a lower-dimensional input space and smaller covariance matrices, keeping computational costs manageable. In contrast, a single multivariate GP would require handling a significantly larger covariance matrix, leading to cubic scaling in computational burden with the number of data points and increased difficulty in hyperparameter tuning, which would make online adaptation impractical for real-time control.

For brevity, we will designate the total uncertainty due to incremental linearization and measurement noise as Δ_{x_i} at each block i . In our design, GPs learn the unknown dynamics Δ_{x_i} for each block i .

Assumption 5. Due to [Theorem 1](#) and a high sampling rate, Δ_{x_i} are assumed to be bounded. These assumptions are common for incremental control.

GP employ a Bayesian framework for adaptive control, treating identification as a statistical problem [31]. GP is non-parametric, meaning that the "parameters" to be identified are functions f_x of an input variable $x \in \mathbb{R}^d$. These functions are characterized by their statistical properties, namely the mean $\langle f \rangle$ and the covariance, also known as the kernel $K_0(x, x') = \text{Cov}(\zeta, \zeta')$ [47]. The prior assumption is that f follows a Gaussian process. According to the Central Limit Theorem, any sufficiently large set of random samples f_i is assumed to have a normal distribution. Within this Bayesian framework, the posterior distribution of the process f_x is computed

based on prior knowledge and likelihood given a set of input-output observations (x_n, ς_n) ($n = 1, \dots, N$).

A representation of the posterior means $\langle f_x \rangle_t = \langle f(x_t) \rangle_t$ and the posterior covariance $K_t(x, x')$, where t denotes the number of data points, can be expressed as finite linear combinations of kernels $K_0(x, x_i)$ evaluated at the training inputs x_i , as shown in [24]. By sequentially projecting the posterior process onto the manifold of Gaussian processes, approximate recursions for the effective parameters of these representations can be derived. To manage the size of these representations and prevent excessive growth, an elegant algorithm for extracting a smaller subset of input data is introduced [31]. Such a subset allows an on-line sparse representation of the posterior process, which is used to predict the GP model.

The posterior expectations within the Bayesian approach are conventionally expressed by high-dimensional integrals. Obviously, this is not applicable for on-line identification. However, it was shown in [31] that the posterior mean and the posterior covariance of the process arbitrary inputs can be expressed as a combination of a finite set of parameters which depend on the training data only. To make Bayesian interference trackable on-line, the posterior is projected to the closest Gaussian process by a single sequential sweep through the examples. When a new feature vector is projected to a subspace spanned by previous feature vectors it results in a projection and the orthogonal residual.

The posterior GP approximation, including its means and covariance, can be estimated using the initial kernel $K_0(x, x')$ and the likelihoods.

$$\begin{aligned} \langle f_x \rangle_t &= \alpha_t^T \mathbf{k}_x \\ K_t(x, x') &= K_0(x, x') + \mathbf{k}_x^T \mathbf{C}_t \mathbf{k}_{x'} \end{aligned} \quad (28)$$

where $\mathbf{k}_x = [K_0(x_1, x), \dots, K_0(x_t, x)]^T$ are the kernel functions, $\alpha_t = [\alpha_t(1), \dots, \alpha_t(t)]^T$ is the coefficient vector, $\mathbf{C}_t = \{C_t(ij)\}_{i,j=1,\dots,t}$ is the coefficient matrix. It should be noted that coefficients $\alpha_t(i)$ and $C_t(ij)$ do not depend on x and x' [31]. For regression problems, Radial Basis Functions (RBF) are a widely used choice for kernel functions (see, for example, [31,35,48,49])

$$K(x, x') = \exp\left(-\frac{\|x - x'\|^2}{2\sigma_x^2}\right). \quad (29)$$

5.1. Online learning

The recursive update of the GP parameters in Eq. (28) can be performed via the following equations:

$$\begin{aligned} \alpha_{t+1} &= T_{t+1}(\alpha_t) + q^{(t+1)} \mathbf{s}_{t+1}, \\ \mathbf{C}_t &= U_{t+1}(\mathbf{C}_t) + r^{(t+1)} \mathbf{s}_{t+1} \mathbf{s}_{t+1}^T, \\ \mathbf{s}_{t+1} &= T_{t+1}(\mathbf{C}_t \mathbf{k}_{t+1}) + \mathbf{e}_{t+1}, \end{aligned} \quad (30)$$

where $\mathbf{k}_{t+1} = \mathbf{k}_{x_{t+1}}$ and $\mathbf{e}_{t+1}$ is the $(t+1)$ -th unit vector and $\mathbf{s}_{t+1}$ is introduced for clarity. The operators T_{t+1} and U_{t+1} expand a t -dimensional vector and matrix to a $t+1$ -dimensional vector and matrix, respectively, by appending zeros to the end of the vector and adding a zero row and column to the matrix.

For the RBF kernel functions the $q^{(t+1)}$ and $r^{(t+1)}$ are defined as

$$\begin{aligned} q^{(t+1)} &= (\varsigma - \alpha_t^T \mathbf{k}_x) / \sigma_x^2, \\ r^{(t+1)} &= -1 / \sigma_x^2, \end{aligned} \quad (31)$$

where $\sigma_x^2 = \sigma_0^2 + \mathbf{k}_x^T \mathbf{C}_t \mathbf{k}_x + k_x^*$, $k_x^* = K_0(x, x)$. It is easy to see that the dimension of the vector α and the size of matrix $\mathbf{C}$ increases with each data point added since $\mathbf{e}_{t+1}$ is the $(t+1)$ -th unit vector.

The updates in the form of Eqs. (30) have a drawback in that the number of parameters grows quadratically with the number of training examples. An effective method for managing the number of parameters was introduced in [31], which involves incorporating sparsity within the GP framework. This approach allows for updating GP parameters without increasing the number of parameters α and $\mathbf{C}$ when, according to a specified criterion, the approximation error remains within acceptable limits.

If the new input x_{t+1} satisfies the condition

$$\mathbf{K}_0(x, x_{t+1}) = \sum_{i=1}^t \hat{e}_{t+1}(i) K_0(x, x_i) \quad (32)$$

for all x , then the update can be performed exactly. In this case, the updated process, as described by Eq. (28), can be represented using only the initial t inputs, but with "renormalized" parameters $\hat{\alpha}$ and $\hat{\mathbf{C}}$. The update in Eq. (30) can be carried out without increasing the size of the parameters α and $\mathbf{C}$ and $\mathbf{s}_{t+1}$ as follows

$$\begin{aligned} \hat{\alpha}_{t+1} &= \alpha_t + q^{(t+1)} \hat{\mathbf{s}}_{t+1}, \\ \hat{\mathbf{C}}_t &= \mathbf{C}_t + r^{(t+1)} \hat{\mathbf{s}}_{t+1} \hat{\mathbf{s}}_{t+1}^T, \\ \hat{\mathbf{s}}_{t+1} &= \mathbf{C}_t \mathbf{k}_{t+1} + \hat{\mathbf{e}}_{t+1}, \end{aligned} \quad (33)$$

where $\hat{\alpha}_{t+1}$, $\hat{\mathbf{C}}_t$ and $\hat{\mathbf{S}}_{t+1}$ are t -th unit vectors.

Clearly, for most kernels and inputs $\mathbf{x}_{t+1}$, the relationship in (32) does not hold for all inputs $\mathbf{x}$. Nevertheless, updates in the form of (33) can be used for approximations if $\hat{\mathbf{e}}_{t+1}$ is determined by minimizing the distance:

$$\|K_0(\cdot, \mathbf{x}_{t+1}) - \sum_{i=1}^t \hat{\mathbf{e}}_{t+1}(i) K_0(\cdot, \mathbf{x}_i)\|^2, \quad (34)$$

where $\|\cdot\|$ is a norm in a space of functions of inputs $\mathbf{x}$. If the norm is defined via the inner product of the Reproducing Kernel Hilbert Space (RKHS) generated by the kernel K_0 , then minimising (34), one can obtain the following expression:

$$\hat{\mathbf{e}}_{t+1} = \mathbf{K}_t^{-1} \mathbf{k}_{t+1}, \quad (35)$$

where $\mathbf{K}_t = \{K_0(\mathbf{x}_i, \mathbf{x}_j)\}_{i,j=1,t}$ is the Gram matrix. In this case, the equation

$$\hat{K}_0(\mathbf{x}, \mathbf{x}_{t+1}) = \sum_{i=1}^t \hat{\mathbf{e}}_{t+1}(i) K_0(\cdot, \mathbf{x}_i), \quad (36)$$

gives the orthogonal projection of the function $K_0(\mathbf{x}, \mathbf{x}_{t+1})$ on the linear span of the functions $K_0(\mathbf{x}, \mathbf{x}_i)$.

The update rule (33) is performed when a squared norm of the residual vector from the projection in the RKHS

$$\gamma_{t+1} = \mathbf{k}_{t+1}^* - \mathbf{k}_{t+1}^T \mathbf{K}_t^{-1} \mathbf{k}_{t+1}, \quad (37)$$

does not exceed some tolerance level $\epsilon_{tol} > 0$. Here, $\mathbf{k}_{t+1}^* = K_0(\mathbf{x}_{t+1}, \mathbf{x}_{t+1})$. The Eq. (37) has a geometrical interpretation: it represents the square norm of the “residual vector” from the projection in the RKHS. Alternatively, it measures the “novelty” of the current input. If γ_{t+1} exceeds a threshold value, the current input contains additional information compared to the existing set of inputs, known as the “basis vector set” or *BV* set, and should be added to this set. Sequentially, some inputs are excluded while others are included in the basis vector set. However, due to the projection described in (32) the inputs excluded from the *BV* set still contribute to the final GP configuration used for prediction and measuring posterior uncertainties. These inputs are not stored and do not increase the size of the parameter set [31].

5.2. Removing a basis vector

The recursive update of the GP parameters (33) is carried out as long as the *BV* set remains within the budget, which is the maximum allowed number of elements. Therefore, a pruning procedure is necessary. When a new input is identified as novel, this procedure should remove one of the basis vectors and replace it with the new input vector. Two strategies can be used for selecting which vector to remove from the *BV* set. The first strategy suggests replacing the oldest basis vector with the new one [48]. The second strategy [31] recommends replacing the basis vector with the smallest error. The first strategy may be preferable for rapidly changing processes. However, we will follow the latter approach here, as it enhances the richness of the *BV* set and preserves long-term memory.

The removal procedure assumes that the respective *BV* was added and the previous update step (33) was implemented. In this scenario, α_{t+1} has $t+1$ elements, and $\mathbf{C}_{t+1}$ and $\mathbf{Q}_{t+1}$ are the $(t+1) \times (t+1)$ matrices. If we assume that the last added element should be deleted, the decomposition of the α_{t+1} , $\mathbf{C}_{t+1}$ and $\mathbf{Q}_{t+1}$ could be represented as follows:

$$\alpha_{t+1} = \begin{bmatrix} \alpha_t^l \\ \alpha^r \end{bmatrix}, \quad \mathbf{C}_{t+1} = \begin{bmatrix} \mathbf{C}_t^l & \mathbf{c}^r \\ \mathbf{c}^{rT} & c^r \end{bmatrix}, \quad \mathbf{Q}_{t+1} = \begin{bmatrix} \mathbf{Q}_t^l & \mathbf{q}^r \\ \mathbf{q}^{rT} & q^r \end{bmatrix}, \quad (38)$$

where $\mathbf{C}_t^l$ and $\mathbf{Q}_t^l$ are $t \times t$ sub-matrices extracted from the $(t+1) \times (t+1)$ matrices $\mathbf{C}_{t+1}$ and $\mathbf{Q}_{t+1}$. For simplicity, this representation illustrates the case where the last element is removed. However, similar partitioning can be applied to a more general case. The updating equations for removing an element are as follows:

$$\begin{aligned} \hat{\alpha} &= \alpha_t^l - \alpha^r \frac{q^r}{q^r}, \\ \hat{\mathbf{C}} &= \mathbf{C}_t^l + \mathbf{c}^r \frac{\mathbf{q}^r \mathbf{q}^{rT}}{q^r} - \frac{1}{q^r} [\mathbf{q}^r \mathbf{c}^{rT} + \mathbf{c}^r \mathbf{q}^{rT}], \\ \hat{\mathbf{Q}} &= \mathbf{Q}_t^l - \frac{\mathbf{q}^r \mathbf{q}^{rT}}{q^r}, \end{aligned} \quad (39)$$

where $\hat{\alpha}$, $\hat{\mathbf{C}}$ and $\hat{\mathbf{Q}}$ are the parameters after the deletion of the last basis vector and α_t^l , $\mathbf{C}_t^l$, $\mathbf{Q}_t^l$, α^r , $\mathbf{c}^r$, $\mathbf{q}^r$, c^r and q^r are taken from GP parameters before deletion.

To determine which element of the basis vector set should be deleted, a score measure is calculated for each element i :

$$\varepsilon_i = \frac{|\alpha_{t+1}(i)|}{Q_{t+1}(i, i)}. \quad (40)$$

Subsequently, the basis vector with the lowest score (40) is deleted. This method ensures the removal of a basis vector from the BV set with minimal loss of information.

6. Convergence analysis

In this section, the convergence of the GP adaptive IBKS closed-loop system is analysed.

Let us first evaluate the estimation precision achievable within the proposed framework when measurements are available in advance and the GP model is trained offline. This assumption reflects a best-case scenario, in which data rich enough to train an accurate model can be collected beforehand. Subsequently, we will analyse the effect of collecting data online and evaluate the influence of online training on the overall estimation precision.

Each uncertainty is characterised by a GP model, specified with its mean and covariance:

$$\hat{\Delta}_{\mathbf{x}_i} \sim \mathcal{GP}(\langle \Delta_{\mathbf{x}_i} \rangle, \mathbf{K}_{\Delta_{\mathbf{x}_i}}), \quad i = 1 \dots h.$$

After N measurements for each uncertainty $\Delta_{\mathbf{x}_i}$ the maximum information gain is introduced with the following equation [32]

$$\gamma_i^N = \max_{\mathbf{w}_i^{(1)}, \dots, \mathbf{w}_i^{(N)} \in \mathcal{W}_i} \frac{1}{2} \log |\mathbf{I} + \sigma_{i,j}^{-2} \tilde{\mathbf{K}}_{i,j}|, \quad (41)$$

where $[\tilde{\mathbf{K}}]_{ab} = K(\tilde{\mathbf{w}}_i^{(a)}, \tilde{\mathbf{w}}_i^{(b)})$ and $|\cdot|$ is the determinant operator. γ_i^N can be interpreted as a measure of the reduction in uncertainty achievable when measurements are taken in the optimal manner [32]. On a compact set $\mathcal{W}_i$, γ_i^N has a sublinear dependence on N for many kernels and can be efficiently approximated up to a small constant using the approach outlined in [50]. The following theorem from [51] provides estimates of the bound for the model error when using a GP trained with noisy measurements.

Theorem 2. [51]: Let $f: \mathcal{W}_i \rightarrow \mathbb{R}$ be a nonlinear function, $B_f \in \mathbb{R}$ be a bound for the corresponding RKHS norm w.r.t. K_i , i.e., $\|f\|_{K_i} \leq B_f$, and let $\delta \in (0, 1)$. For all $N \in \mathbb{N}$, define $\beta_N = B_f + 4\sigma \sqrt{\gamma_i^N + 1 + \ln\left(\frac{1}{\delta}\right)}$, where γ_i^N is determined by (41). Then, for all $N \geq 1$ and $\mathbf{w}_i \in \mathcal{W}_i$, the following holds with probability of at least $1 - \delta$

$$|\Delta(\mathbf{w}_i) - \langle \Delta(\mathbf{w}_i) \rangle| \leq \beta_N \sigma_{N-1}(\mathbf{w}_i), \quad (42)$$

where N is a number of system measurements. If B_f is not available a priori, a guess-and-doubling strategy can be employed to obtain an estimate, σ can be estimated by sampling the same data point multiple times [32,50]. Since γ_i^N grows sub-linearly with N on a compact set, the term β_N grows slowly with N on a compact set. The covariance term $\sigma_{N-1}(\mathbf{w}_i)$ typically is very small next to training data points [46]. Hence, the variance term $\sigma_{N-1}(\mathbf{w}_i)$ can be decreased if the region of interest is sampled densely enough. This is achieved on a compact set [32,50] by selecting training data points $\mathbf{w}_i$ corresponding to highest model uncertainty, i.e.,

$$\mathbf{w}_i^{(N+1)} = \arg \max_{\mathbf{w}_i \in \mathcal{W}_i} \sigma_N(\tilde{\mathbf{w}}_i), \quad (43)$$

In Bayesian learning, all information about the parameters that we wish to infer is encoded in probability distributions [31]. In the GP framework, the parameters are functions and the GP priors specify a Gaussian distribution over a function space. The posterior process is entirely specified by all its finite dimensional marginals and predictive quantities like the posterior mean and the posterior covariance of the process at arbitrary inputs can be expressed as a combination of a finite set of parameters, which depend on the training data only.

However, within the Sparse Online Gaussian Process framework the posterior distribution is no longer Gaussian, and a variational technique is used in order to project it to the closest Gaussian process. To take into account projection error, we proved Lemma 1, which gives an estimate of this effect and helps to establish a boundness of the estimation error. Following approach proposed in [35], we can state the next Lemma, showing that the error $\varepsilon_f = f(\mathbf{w}_i) - \langle f(\mathbf{w}_i) \rangle$ induced by the projection is bounded.

Lemma 1. Let the unknown function f be representable by a GP, $\langle f_x \rangle_t$ be an on-line estimated GP representation of the function via a sparse subset BV of the data selected by deleting either the oldest basis vector or the basis vector with minimal score (40), then error $\varepsilon_f = f(\mathbf{w}_i) - \langle f(\mathbf{w}_i) \rangle$ is bounded with the probability of at least $1 - \delta$.

Proof. According to [31] and the nonparametric Representer theorem [52] the sample averaged posterior mean of the GP has the following error due to the sparse approximation:

$$|\varepsilon_{t+1}| = \frac{|f(\mathbf{w}_i) - \langle f(\mathbf{w}_i) \rangle|}{\sigma_x^2} (\mathbf{k}_{t+1}^* - \mathbf{k}_{t+1}^T \mathbf{K}_t^{-1} \mathbf{k}_{t+1}), \quad (44)$$

Since $\gamma_{t+1} = \mathbf{k}_{t+1}^* - \mathbf{k}_{t+1}^T \mathbf{K}_t^{-1} \mathbf{k}_{t+1} < \epsilon_{\text{tol}}$, then using [Theorem 2](#) it could be derived that $|f(\mathbf{w}_i) - \langle f(\mathbf{w}_i) \rangle| \leq \beta_N \sigma_{N-1}(\mathbf{w}_i)$ due to (42). Thus, it can be concluded that $|\epsilon_{t+1}| < \epsilon_{\text{tol}} \beta_N \sigma_{N-1}(\mathbf{w}_i) / \sigma_x^2$ with a probability of at least $1 - \delta$. ■

The boundedness of the tracking error can be established now.

Theorem 3. Consider the tracking error dynamics given by (15) and the control law given with (12) and (14). The system tracking error $\mathbf{e}$ is globally uniformly ultimately bounded. With probability, of at least $1 - \delta$, the system ultimate error is given with

$$\lim_{t \rightarrow \infty} |\mathbf{z}| \leq \sqrt{\sum_i \left(\frac{\epsilon_{\text{tol}} \beta_{N_i} \sigma_{N-1i}}{\sigma_{x_i}^2 \min \mathbf{W}_i} \right)^2}. \quad (45)$$

For the brevity, let us designate the estimation error as $\tilde{\Delta} = [\tilde{\Delta}_{x_1}, \tilde{\Delta}_{x_2}, \dots, \tilde{\Delta}_{x_h}]^T$, where $\tilde{\Delta}_{x_i} = \Delta_{x_i} - \langle \Delta_{x_i} \rangle$

In this case the Lyapunov function derivative can be written

$$\dot{V} = -\frac{1}{2} \mathbf{z}^T \mathbf{W} \mathbf{z} - \mathbf{z}^T \tilde{\Delta}. \quad (46)$$

The derivative is negative if

$$|\tilde{\Delta}| < \frac{1}{2} \min_i \mathbf{W}_i |\mathbf{z}|. \quad (47)$$

Given Assumption 4 and the boundedness of the kernel functions, $|\tilde{\Delta}|$ is also bounded. Consequently, (47) holds for sufficiently large $|\mathbf{z}|$, indicating that $|\mathbf{z}|$ is globally uniformly ultimately bounded. Considering [Theorem 2](#), we obtain the desired result. ■

It should be noted that σ_{N-1i} and σ_{x_i} are different; the former arises from approximation after N measurements, while the latter results from the online approximation. The proof demonstrates that the ultimate error bound of the controlled system with GP adaptation can be made arbitrarily small with high probability by increasing the control gains W_i , reducing the tolerance level ϵ_{tol_i} , which minimizes the "residual vector" from the projection in the RKHS, or selecting data points to reduce the terms $\beta_{N_i} \sigma_{N-1i}$. The latter can be achieved by collecting new measurements at points of high uncertainty. This approach is particularly useful when the control input is constrained and high gains W_i are not feasible. Additionally, the estimation error bound benefits from the baseline control law, which enforces compactness in the portion of the state space explored by the state trajectory. This compactness ensures sub-linear dependence of γ_i^N on N , allowing the ultimate bound of $|\mathbf{z}|$ to be efficiently reduced with new training data. Furthermore, as shown in [53], the linear independence of BV ensures that persistency of excitation in the state space is reflected in the RKHS. Since the proposed algorithm enforces this independence subject to the tolerance ϵ_{tol_i} , persistency of excitation is maintained, ensuring $\mathbf{K}_t$ remains invertible.

As previously mentioned, IBKS is fundamentally designed as a discrete-time control method, and its GP adaptation follows naturally from this structure.

Comparison of Indirect and Direct GP-adaptation to IBKS

One of the advantages of the incremental approach is that it reduces requirements for model knowledge because of usage the simplified model of system dynamics (6). However, some knowledge about the model cascade coupling matrices $\mathbf{B}_{\xi_1}, \mathbf{B}_{\xi_{i+1}}, i = 2, \dots, h-1, \mathbf{B}_u$ are still required. Indirect GP-adaptive IBKS for flight control was studied in [20], where the model uncertainties were assumed in the model cascade coupling matrices (control effectiveness matrix $\mathbf{B}_u$ in particular). Let us compare the GP-indirect adaptation approach proposed in [20] with the GP-direct adaptation elaborated in this paper. In addition to uncertainties due to linearization and measurement errors $\Delta \mathbf{x}, \Delta \xi_i$, let us distinguish uncertainties related to imperfect knowledge of $\mathbf{B}_{\xi_1}, \mathbf{B}_{\xi_{i+1}}, i = 2, \dots, h-1, \mathbf{B}_u$ such that

$$\begin{aligned} \mathbf{B}_{\xi_1} &= \hat{\mathbf{B}}_{\xi_1} + \Delta \mathbf{B}_{\xi_1}, \\ \mathbf{B}_{\xi_{i+1}} &= \hat{\mathbf{B}}_{\xi_{i+1}} + \Delta \mathbf{B}_{\xi_{i+1}}, \quad i = 2, \dots, h-1, \\ \mathbf{B}_u &= \hat{\mathbf{B}}_u + \Delta \mathbf{B}_u, \end{aligned}$$

where $\hat{\mathbf{B}}_{\xi_1}, \hat{\mathbf{B}}_{\xi_{i+1}}, \hat{\mathbf{B}}_u$ are parameters used in the model, $\Delta \mathbf{B}_{\xi_1}, \Delta \mathbf{B}_{\xi_{i+1}}, \Delta \mathbf{B}_u$ are the model uncertainties. Thus, the system dynamics (6) becomes:

$$\begin{aligned} \dot{\mathbf{x}} &\cong \dot{\mathbf{x}}_0 + \hat{\mathbf{B}}_{\xi_1} \Delta \xi_1 + \hat{\Delta}_{\mathbf{x}} + \Delta \mathbf{x}, \\ \dot{\xi}_i &\cong \dot{\xi}_{i0} + \hat{\mathbf{B}}_{\xi_{i+1}} \Delta \xi_{i+1} + \hat{\Delta}_{\xi_i} + \Delta \xi_i, \quad i = 2, \dots, h-1, \\ \dot{\xi}_h &\cong \dot{\xi}_{h0} + \hat{\mathbf{B}}_u \Delta \mathbf{u} + \hat{\Delta}_{\xi_h} + \Delta \xi_h, \end{aligned} \quad (48)$$

where $\hat{\Delta}_{\mathbf{x}} = \Delta \mathbf{B}_{\xi_1} \Delta \xi_1, \hat{\Delta}_{\xi_i} = \Delta \mathbf{B}_{\xi_{i+1}} \Delta \xi_{i+1}, \hat{\Delta}_{\xi_h} = \Delta \mathbf{B}_u \Delta \mathbf{u}$ are uncertainties due to imperfect knowledge of the model cascade coupling matrices $\mathbf{B}_{\xi_1}, \mathbf{B}_{\xi_{i+1}}, i = 2, \dots, h-1, \mathbf{B}_u$. From (48) one can see that the indirect GP adaptation tackles only the uncertainties related to these matrices, namely, $\hat{\Delta}_{\mathbf{x}}, \hat{\Delta}_{\xi_i}, \hat{\Delta}_{\xi_h}$. The proposed GP adaptation can handle the total uncertainties $\Delta \mathbf{x}_{\text{total}} = \hat{\Delta}_{\mathbf{x}} + \Delta \mathbf{x}, \Delta \xi_{i\text{total}} = \hat{\Delta}_{\xi_i} +$

$\Delta_{\xi_i}, \Delta_{\xi_{h_{total}}} = \hat{\Delta}_{\xi_h} + \Delta_{\xi_h}$, which could be not only due to imperfect knowledge of the model cascade coupling matrices, but also because of linearization of the system (1), assumptions (5), possible failed estimation/measurement of state vector derivatives and/or faulty actuation. In such a way, the proposed approach provides improved adaptation capabilities.

It should be noted that both indirect and direct GP adaptations can handle uncertainties only when the distribution parameters can be inferred from the data, that is, when the system is excited.

7. Longitudinal flight control

7.1. Flight dynamics model

In this section, we evaluate the performance of the proposed approach using a tracking control scenario for the longitudinal missile model from [54]. This second-order nonlinear model captures the longitudinal dynamics of a vehicle traveling at an altitude of approximately 6000 m, with aerodynamic coefficients expressed as third-order polynomials in terms of the angle of attack α and Mach number M . The nonlinear equations of motion in the pitch plane are given by:

$$\dot{x}_1 = x_2 + f_1(x_1) + g_1 u, \quad \dot{x}_2 = f_2(x_1) + g_2 u, \quad (49)$$

where $x_1 = \alpha$, $x_2 = q$, $g_1 = C_1 b_z$ and $g_2 = C_2 b_m$ and

$$f_1(x_1) = C_1 [\varphi_{z1}(x_1) + \varphi_{z2}(x_1)M], \quad C_1 = \frac{\bar{q}S}{mV},$$

$$f_2(x_1) = C_2 [\varphi_{m1}(x_1) + \varphi_{m2}(x_1)M], \quad C_2 = \frac{\bar{q}Sd}{I_{yy}}.$$

Here

$$\begin{aligned} \varphi_{z1}(x_1) &= -288.7x_1^3 + 50.32x_1|x_1| - 23.89x_1 \\ \varphi_{z2}(x_1) &= -13.53x_1|x_1| + 4.185x_1, \\ \varphi_{m1}(x_1) &= 303.1x_1^3 - 246.3x_1|x_1| - 37.56x_1, \\ \varphi_{m2}(x_1) &= 71.51x_1|x_1| + 10.01x_1. \end{aligned}$$

The model is valid for $-10^\circ \leq \alpha \leq 10^\circ$ and $1.8 \leq M \leq 2.6$. Details of the aerodynamic coefficients can be found in [54].

7.2. IBKS longitudinal controller

An IBKS longitudinal controller that tracks a smooth command reference y_r with the pitch rate x_2 is developed as a baseline controller. Here we are following a procedure from [42]. In its design, it was assumed that the aerodynamic force and moment functions are accurately known, and the Mach number M is treated as a measurable parameter. Moreover, for this second-order system, which is not in a lower triangular form due to $g_1 u$ and $f_2(x_1)$, pitch rate control using IBKS is feasible because of the timescale separation principle [6]. Additionally, during the control design, the influence of fin deflection on the $\dot{x}_1$ -equation Eq. (49) is ignored to simplify the system dynamics into a lower-triangular form. This assumption is valid for most types of aircraft and aerodynamically controlled missiles and is a common practice in flight control system design. In the simulation, the effect of the fin deflection is preserved.

The incremental dynamics of is as follows:

$$\dot{x}_1 \cong \dot{x}_{10} + x_2 \dot{x}_2 \cong \dot{x}_{20} + \bar{g}u.$$

Here $\bar{g}_2$ is the estimate of g_2 . It should be noted here that the accurate knowledge of f_2 is replaced with measurements (or estimates), such that $\dot{f}_2 \cong \dot{x}_{20}$, reducing the control law's dependence on the model.

The rate tracking error is defined as:

$$\begin{aligned} z_1 &= x_{1_{ref}} - x_1, \\ z_2 &= x_{2_{ref}} - x_2. \end{aligned}$$

The system's error dynamics are given by:

$$\dot{z}_1 = \dot{x}_{1_{ref}} - \dot{x}_1, \quad \dot{z}_2 = \dot{x}_{2_{ref}} - \dot{x}_2.$$

For the z_1 -system the following Lyapunov function is constructed

$$V_1 = \frac{1}{2}z_1^2,$$

whose derivative is:

$$\dot{V}_1 = \mathbf{z}_1 \dot{\mathbf{z}}_1 = \mathbf{z}_1 \left(\dot{\mathbf{x}}_{1ref} - \dot{\mathbf{x}}_{10} + \mathbf{x}_{20} + \mathbf{z}_2 - \mathbf{x}_{2ref} \right).$$

The following control law stabilizing the $\mathbf{z}_1$ -dynamics can be chosen:

$$\mathbf{x}_{2ref} = \mathbf{C}_1 \mathbf{z}_1 + \mathbf{x}_{20} + \dot{\mathbf{x}}_{1ref} - \dot{\mathbf{x}}_{10}.$$

In this case, the derivative of the Lyapunov function becomes:

$$\dot{V}_1 = -C_1 \mathbf{z}_1^2 + \mathbf{z}_1 \mathbf{z}_2. \quad (50)$$

The term $\mathbf{z}_1 \mathbf{z}_2$ will be cancelled out in the next stage. To stabilize the inner loop, we can penalize the Lyapunov function with $\mathbf{z}_2$ -dynamics:

$$V_2 = V_1 + \frac{1}{2} \mathbf{z}_2^2. \quad (51)$$

The derivative is the following

$$\dot{V}_2 = \dot{V}_1 + \mathbf{z}_2 \dot{\mathbf{z}}_2 = \dot{V}_1 + \mathbf{z}_2 \left(\dot{\mathbf{x}}_{2ref} + \mathbf{z}_1 - \dot{\mathbf{x}}_{20} - \bar{\mathbf{g}} \mathbf{u} \right). \quad (52)$$

To make the derivative $\dot{V}_2$ in (52) negative, we need to ensure that the second term is negative everywhere except $\mathbf{z}_2 = 0$. Therefore, the control law can be constructed as follows:

$$\mathbf{u} \cong \mathbf{u}_0 + \bar{\mathbf{g}}^{-1} \left(\dot{\mathbf{x}}_{2ref} + \mathbf{z}_1 - \dot{\mathbf{x}}_{20} + C_2 \mathbf{z}_2 \right)$$

7.3. GP-based adaptive augmentation

As previously noted, an incremental-type controller relies on the quality of the state derivative measurements (or estimations). The performance of such controllers can degrade, potentially leading to system instability in the presence of unknown nonlinearities, such as transport delays or unknown actuator dynamics [20]. To enhance the robustness of the controller against uncertainties, we introduced a GP-based adaptation and modified the controller as follows:

$$\mathbf{u} \cong \mathbf{u}_0 + \bar{\mathbf{g}}^{-1} \left(\dot{\mathbf{x}}_{2ref} + \mathbf{z}_1 - \dot{\mathbf{x}}_{20} + \langle \Delta_x \rangle \right),$$

where $\langle \Delta_x \rangle$ is a GP estimation of uncertainties. The maximum number d of the BV set was to 5, $\epsilon_{tol} = 1e-5$, $\sigma_0^2 = 0.9$. To ensure appropriate overlap between neighbouring kernels, the RBF width was specified as follows:

$$\sigma = \frac{10}{2(d-1)}.$$

8. Simulation results

In this section, we evaluate the proposed controller's performance in the presence of unknown actuator dynamics, modelled as $F(s) = (s + 0.25)^{-1}$ and sensor noise characterized by zero-mean Gaussian white noise with a standard deviation of $1e^{-6}$ rad and rad/s for the q and $\dot{q}$ correspondingly. The proposed IBKS with GP adaptation is compared with the baseline IBKS controller with gains $C_1 = 10$, $C_2 = 10$. Values of d , ϵ_{tol} , σ_0^2 , C_1 and C_2 were selected based on conducted parametric analysis and provide the optimized system performance. The results are presented in Fig. 1. Normalized Root Mean Square Tracking Error (NRMSTE), calculated for the algorithms as

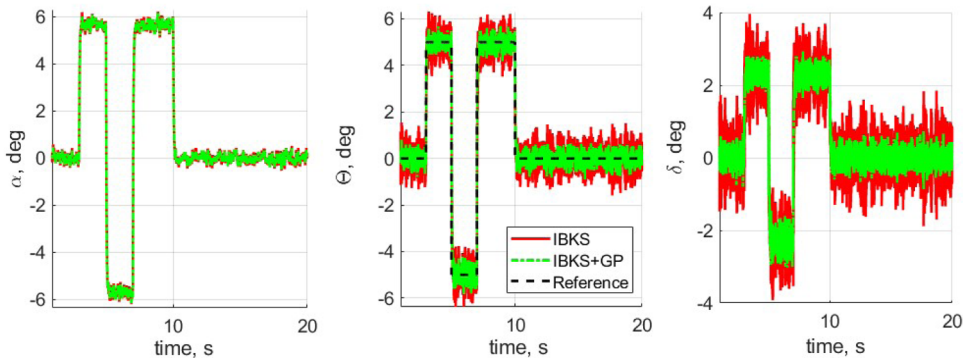


Fig. 1. Actuator dynamics and noise.

$$err = \frac{\sqrt{\frac{1}{N-1} \sum_{j=1}^N (x_j^{ref} - x_j^{meas})^2}}{\Delta x},$$

where x_j^{ref} , x_j^{meas} , $j = 1 \dots N$ are the reference and measured state, Δx is the state range, is given in Table 1.

From the figure and the table, it is evident that the baseline IBKS controller exhibits high-amplitude oscillations, which significantly degrade system performance. However, the proposed GP-based adaptation compensates for the existing uncertainties and effectively reduces the undesired oscillations. Unlike the method described in [32], the GP in this approach identifies unknown functions online. Furthermore, thanks to its design, this method can handle a wide range of uncertainties, not just those related to actuation as addressed in previous studies [20].

It should be noted that Sparse Online Gaussian Adaptation do not require persistent excitation due to their pruning procedure and the way the BV set is updated—specifically, only when a significant novel datum is detected relative to the existing BV set. This approach enhances the richness of the BV set and preserves long-term memory. This effect can be observed in Fig. 1, where the system is no longer actively excited after $t = 10$ s, yet its performance does not degrade over time.

The initial adaptation of the GP algorithm is illustrated in Fig. 2. It can be seen that the adaptation process is quite rapid, requiring less than one second to learn the system dynamics while maintaining bounded behaviour throughout the transition period. During training, the residual error exhibits large but bounded values, which decrease quickly once the GP has learned the dynamics. Due to the presence of adaptation error, the total NRMSTE for the IBKS+GP combination is higher than for IBKS alone. However, the IBKS+GP demonstrates lower tracking error once the learning is completed. Given that a relatively large portion of the simulation time is allocated to training, this result is not unexpected. In real scenarios, where adaptation occupies only a small fraction of the overall operating time, the total NRMSTE would be lower. Although GP augmentation with $d=5$ increases the computational burden to the pure IBKS, the increase is not substantial.

The sparseness introduced in the GP prevents the uncontrollable growth of the system's parameters, thereby keeping the adaptation manageable and limiting the computational burden. To demonstrate the effect of increasing the size of the BV set, we compared the system's behaviour with maximum BV set sizes of $d=5$ and $d=50$, as presented in Table 1. The results indicate that increasing the number of basis vectors does not always lead to improved precision. However, it does cause a significant increase in the time required to calculate the covariance matrices, thereby reducing the algorithm's real-world applicability. These results confirm the observations in [31], which show that due to the orthogonal projection mechanism, the approximation error is concentrated on the most recent data point, enabling the effectiveness of the sparseness approach.

To isolate the effects of noise and unknown dynamics, and to demonstrate the impact of GP adaptation in the case of noiseless measurements, we conducted a separate simulation analysis, presented in Fig. 3. In this simulation, the unknown actuator dynamics is introduced in the model, without the measurement noise being considered. GP-IBKS shows fast adaptation at the beginning with a bounded error; after the GP learned the dynamics, it follows the reference signal quite well. The pure IBKS controller, also shown in the same figure for comparison, exhibits oscillations caused by unknown actuator dynamics throughout the simulations.

9. Conclusions

The Incremental Backstepping (IBKS) method is a promising control approach based on the idea that nonlinear system dynamics can be linearised relative to the system states at the previous time step and represented incrementally. This allows partial substitution of model knowledge with measurements or estimations of the state vector derivatives.

In this study, we have theoretically proved that the IBKS controller designed for the incremental dynamics can stabilise the original system, resulting in a globally uniformly ultimately bounded tracking error. The error estimate can be reduced by increasing the sampling frequency relative to the system's growth between consecutive measurements and by increasing the control gains. However, there is a natural limitation to reducing the tracking error.

To address linearisation errors and errors arising from potential sensor noise and system failures, we introduced a Gaussian Process (GP) adaptation to the IBKS. This approach combines the advantages of sensor-based incremental methods and data-driven techniques, reducing the negative impact of model and measurement uncertainties on controller performance. GPs provide a flexible, nonparametric data-driven modelling framework that balances data fitting and regularisation in noisy conditions. Additionally, GPs support long-term learning. The resulting tracking error is globally uniformly ultimately bounded, and its ultimate bound can be reduced by increasing the control gains, minimising the residual vector from the projection in the Reproducing Kernel Hilbert Space (RKHS), or incorporating additional training data.

Our simulation results have demonstrated enhanced tracking performance and stability of the IBKS with GP-based adaptation. The adaptive loop not only captures uncertainties but also filters noise signals. Simulations showed that with multiple sensor and actuator

Table 1
Normalized Root Mean Square Tracking Error.

Controller	Total process	Training	Excl. initial training	Elapsed time, s
IBKS	0.0647	0.0647	0.0647	6.43
IBKS + GP ($d = 5$)	0.1262	0.3692 ($t < 2$ s)	0.0507 ($t > 2$ s)	11.89
IBKS + GP ($d = 50$)	0.3126	0.4007 ($t < 3$ s)	0.1681 ($t > 3$ s)	93.34

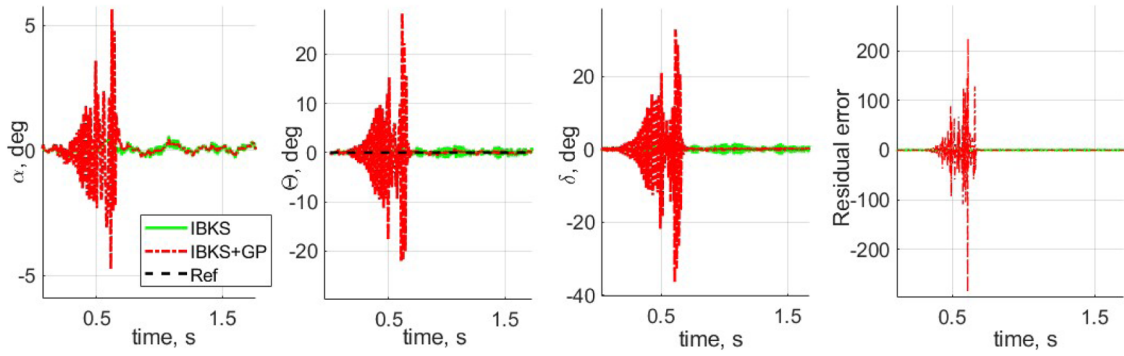


Fig. 2. Initial adaptation process.

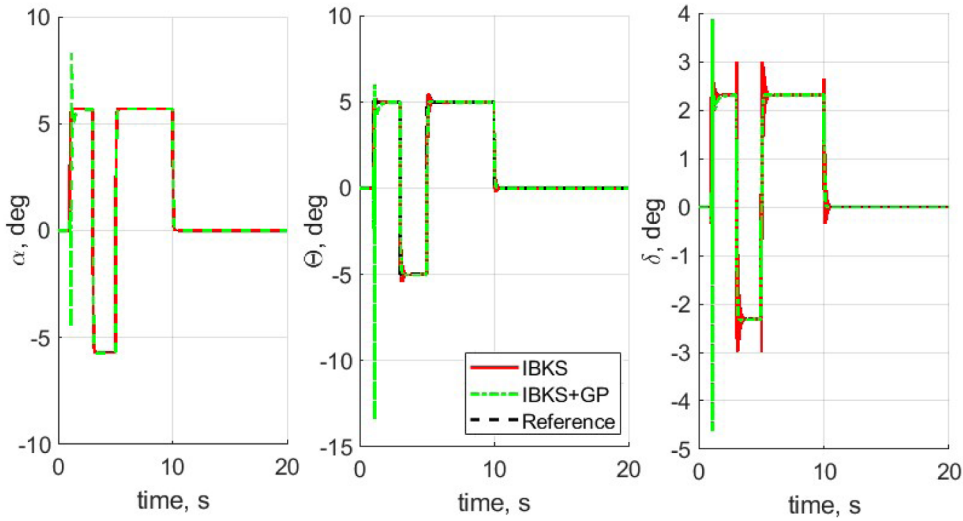


Fig. 3. Actuator failure.

failures, the system dynamics might lose its input-affine property, jeopardising the stability of the closed-loop system when using pure IBKS. The proposed GP-adaptive augmentation compensates for the unmodelled dynamics, ensuring the required stability and performance characteristics.

The proposed IBKS with GP adaptation framework offers a practical solution for real-time control of complex, safety-critical systems, such as unmanned aerial vehicles or precision-guided munitions, where measurement imperfections and unmodelled dynamics frequently arise. The computational requirements of the GP adaptation remain manageable with the proposed sparsity strategy, supporting real-time implementation on modern embedded processors.

Nevertheless, the performance of the proposed approach may still depend on the availability and quality of state derivative measurements at high sampling rates, and potential delays in measurement or actuation could affect stability margins.

Future research will focus on extending the proposed methodology to systems where only partial state information is available, integrating observers or estimators to relax the assumption of complete state measurements. Additionally, investigating hardware-in-the-loop experiments will help validate the method in realistic flight control environments.

Overall, this work bridges the gap between model-dependent control strategies and purely data-driven adaptation, providing a unified framework that enhances the robustness of incremental backstepping controllers through direct online Gaussian Process adaptation.

CRedit authorship contribution statement

Dmitry Ignatyev: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization. **António Tsourdos:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Dmitry Ignatyev reports financial support was provided by European Union.

Acknowledgements

This research is partially funded by the European Union in the scope of INCEPTION project, which has received funding from the EU's Horizon2020 Research and Innovation Programme under grant agreement No. 723515.

References

- [1] J.H. Blakelock, *Automatic Control of Aircraft and Missiles*, 2nd ed., Wiley, New York, 1991.
- [2] J.-J.E. Slotine and W. Li, *Applied nonlinear control*, vol. 62, no. 7. 1991. doi: 10.1366/000370208784909643.
- [3] M. Krstic, P.V. Kokotovic, I. Kanellakopoulos, *Nonlinear and Adaptive Control Design*, 1st, John Wiley & Sons, Inc., USA., 1995.
- [4] G. Tao, L. Wen, Higher-order tracking properties of adaptive backstepping control systems, *Automatica* 153 (Jul. 2023) 111019, <https://doi.org/10.1016/J.AUTOMATICA.2023.111019>.
- [5] A. Tsourdos, D. Ignatyev, S. Mondal, M. Satm, *Uncrewed Aerial Systems : Guidance and Control*, Elsevier Ltd., 2024, <https://doi.org/10.1016/B978-0-443-14081-5.00015-5>.
- [6] H.K. Khalil, *Nonlinear Systems* (1996) [Online]. Available: http://iie.fing.edu.uy/~monzon/Nonlinear_Systems_2nd-Hassan_Khalil.pdf.
- [7] A.G. Loukianov, V. Utkin, Methods of reducing equations of dynamic systems to regular form, *Autom. Remote Control* 42 (4) (1981) 413–420.
- [8] S.V. Drakunov, D.B. Izosimov, A.G. Luk'yanov, V.A. Utkin, V.I. Utkin, The block control principle. I, *Avtomatika i Telemekhanika* (5) (1990) 38–47.
- [9] V.I. Utkin, S.V. Drakunov, D.E. Izosimov, A.G. Lukyanov, V.A. Utkin, A hierarchical principle of the control system decomposition based on motion separation, *IFAC Proc. Volumes* 17 (2) (1984) 1553–1558, [https://doi.org/10.1016/S1474-6670\(17\)61197-2](https://doi.org/10.1016/S1474-6670(17)61197-2).
- [10] L.G. Sun, C.C. de Visser, Q.P. Chu, W. Falkena, Hybrid sensor-based backstepping control approach with its application to fault-tolerant flight control, *J. Guidance, Control, Dyn.* 37 (1) (2013) 59–71, <https://doi.org/10.2514/1.61890>.
- [11] S. Sieberling, Q.P. Chu, J.A. Mulder, Robust flight control using incremental nonlinear dynamic inversion and angular acceleration prediction, *J. Guidance, Control, Dyn.* 33 (6) (Nov. 2010) 1732–1742, <https://doi.org/10.2514/1.49978>.
- [12] E. Tal, S. Karaman, Accurate tracking of aggressive quadrotor trajectories using incremental nonlinear dynamic inversion and differential flatness, *IEEE Trans. Control Syst. Technol.* 29 (3) (2021) 1203–1218, <https://doi.org/10.1109/TCST.2020.3001117>.
- [13] G.P. Falconi, V.A. Marvakov, F. Holzapfel, Fault tolerant control for a hexarotor system using Incremental Backstepping, in: 2016 IEEE Conference on Control Applications, CCA 2016, 2016, pp. 237–242, <https://doi.org/10.1109/CCA.2016.7587842>, no. 1.
- [14] D.I. Ignatyev, H.S. Shin, A. Tsourdos, Two-layer adaptive augmentation for incremental backstepping flight control of transport aircraft in uncertain conditions, *Aerosp Sci. Technol.* 105 (2020) 106051, <https://doi.org/10.1016/j.ast.2020.106051>.
- [15] B.J. Jeon, M.G. Seo, H.S. Shin, A. Tsourdos, Understandings of classical and incremental backstepping controllers with model uncertainties, *IEEE Trans. Aerosp Electron. Syst.* 56 (4) (2020) 2628–2641, <https://doi.org/10.1109/TAES.2019.2952631>.
- [16] R. Steffensen, A. Steinert, E.J.J. Smeur, Nonlinear dynamic inversion with actuator dynamics: an incremental control perspective, *J. Guidance, Control Dyn.* 46 (4) (2023) 709–717, <https://doi.org/10.2514/1.g007079>.
- [17] Y. Liu, C. Duan, L. Liu, and L. Cao, 'Discrete-time incremental backstepping control with extended kalman filter for UAVs', 2023. doi: 10.3390/electronics12143079.
- [18] M.P. Deisenroth, R.D. Turner, M.F. Huber, U.D. Hanebeck, C.E. Rasmussen, Robust filtering and smoothing with Gaussian processes, *IEEE Trans. Automat Control* 57 (7) (2012) 1865–1871, <https://doi.org/10.1109/TAC.2011.2179426>.
- [19] J. Suraj Nandiganahalli, C. Kwony, I. Hwangz, Delay-tolerant adaptive robust control for unmanned aerial system, in: AIAA Information Systems-AIAA Infotech at Aerospace, 2018, 2018, pp. 1–12, <https://doi.org/10.2514/6.2018-0078>, no. January.
- [20] D.I. Ignatyev, H.S. Shin, A. Tsourdos, Sparse online Gaussian process adaptation for incremental backstepping flight control, *Aerosp Sci. Technol.* 136 (2023) 108157, <https://doi.org/10.1016/j.ast.2023.108157>.
- [21] D. Ignatyev, H.-S. Shin, A. Tsourdos, Gaussian process adaptive incremental backstepping flight control, in: AIAA 2022-2032, 2022, <https://doi.org/10.2514/6.2022-2032>.
- [22] M.R. Kartal, D. Ignatyev, A. Zolotas, Non-linear control of a quadrotor with actuator delay, in: AIAA SciTech Forum and Exposition, 2024, 2024, <https://doi.org/10.2514/6.2024-0569> no. January.
- [23] R.A. Cordeiro, A.S. Marton, J.R. Azinheira, J.R.H. Carvalho, A. Moutinho, Increased robustness to delay in incremental controllers using input scaling gain, *IEEE Trans. Aerosp Electron. Syst.* 58 (2) (2022) 1199–1210, <https://doi.org/10.1109/TAES.2021.3123215>.
- [24] J. Chang, R. De Breuker, X. Wang, Adaptive nonlinear incremental flight control for systems with unknown control effectiveness, *IEEE Trans Aerosp. Electron. Syst.* 59 (1) (2023) 228–240, <https://doi.org/10.1109/TAES.2022.3187057>.
- [25] P. van Gils, E.-J. Van Kampen, C.C. de Visser, Q.P. Chu, Adaptive incremental backstepping flight control for a high-performance aircraft with uncertainties, in: AIAA 2016-1380, 2016, <https://doi.org/10.2514/6.2016-1380>.
- [26] T.L. Grigorie, R.M. Botez, A.V. Popov, Fuzzy logic control of a smart actuation system in a morphing wing. *Fuzzy Controllers- Recent Advances in Theory and Applications*, 2012, pp. 1–22, <https://doi.org/10.5772/48778>.
- [27] S. Shimomura, S. Sekimoto, A. Oyama, K. Fujii, and H. Nishida, 'Experimental study on application of distributed deep reinforcement learning to closed-loop flow separation control over an airfoil', no. January, pp. 1–11, 2020, doi: 10.2514/6.2020-0579.
- [28] A. Calise, Y. Shin, and M. Johnson, 'A comparison study of classical and neural network based adaptive control of wing rock', no. August, pp. 1–17, 2004, doi: 10.2514/6.2004-5320.
- [29] D.I. Ignatyev, H.S. Shin, A. Tsourdos, Bayesian calibration for multiple source regression model, *Neurocomputing* 318 (2018) 55–64, <https://doi.org/10.1016/j.neucom.2018.08.027>.
- [30] K.S. Narendra, Neural networks for control: theory and practice, *Proc. IEEE* 84 (10) (1996) 1385–1406, <https://doi.org/10.1109/5.537106>.
- [31] L. Csato, M. Opper, Sparse on-line gaussian processes, *Neural Comput.* 14 (3) (Mar. 2002) 641–668, <https://doi.org/10.1162/089976602317250933>.
- [32] A. Capone, S. Hircche, Backstepping for partially unknown nonlinear systems using gaussian processes, *IEEE Control Syst. Lett.* 3 (2) (2019) 416–421, <https://doi.org/10.1109/LCSYS.2018.2890467>.
- [33] P.A. Ioannou, P.V. Kokotovic, Instability analysis and improvement of robustness of adaptive control, *Automatica* 20 (5) (1984) 583–594, [https://doi.org/10.1016/0005-1098\(84\)90009-8](https://doi.org/10.1016/0005-1098(84)90009-8).
- [34] B.W. Karl, Johan Åström, *Adaptive Control*, 2nd ed., Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1994.
- [35] G. Chowdhary, H.A. Kingravi, J.P. How, P.A. Vela, Bayesian nonparametric adaptive control using gaussian processes, *IEEE Trans Neural Netw. Learn Syst.* 26 (3) (2015) 537–550, <https://doi.org/10.1109/TNNLS.2014.2319052>.
- [36] D. Ignatyev, A. Tsourdos, Incremental nonlinear dynamic inversion with sparse online gaussian processes adaptation for partially unknown systems, in: 2022 30th Mediterranean Conference on Control and Automation, MED 2022, 2022, pp. 233–238, <https://doi.org/10.1109/MED54222.2022.9837175>.

- [37] M.R. Kartal, D.I. Ignatyev, Sparse online Gaussian process adaptive control of unmanned aerial vehicle with slung payload, *Drones* 8 (11) (2024) 687, <https://doi.org/10.3390/drones8110687>.
- [38] D. Li, D. Ignatyev, A. Tsourdos, Z. Wang, Estimation of non-symmetric and unbounded region of attraction using shifted shape function and R-composition, *ISA Trans* 136 (2023) 308–322, <https://doi.org/10.1016/j.isatra.2022.11.015>, no. xxxx.
- [39] B. Biswas, D. Ignatyev, A. Zolotas, Construction of control lyapunov function with region of attraction using union theorem in sum-of-squares optimization, in: 2023 31st Mediterranean Conference on Control and Automation, MED 2023, 2023, pp. 771–776, <https://doi.org/10.1109/MED59994.2023.10185818>.
- [40] B. Biswas, D. Ignatyev, A. Zolotas, A. Tsourdos, Design of controller with enlarged region of attraction using union theorem in sum-of-squares optimization for the F-8 aircraft, in: *AIAA SciTech Forum and Exposition*, 2024, pp. 1–10, <https://doi.org/10.2514/6.2024-1018>, no. January2024.
- [41] J. Lee, J. Seo, J. Choi, Output feedback control design using Extended high-gain observers and dynamic inversion with projection for a small scaled helicopter, *Automatica* 133 (Nov. 2021) 109883, <https://doi.org/10.1016/J.AUTOMATICA.2021.109883>.
- [42] P. Acquatella, E. Van Kampen, P.Q. Chu, Incremental backstepping for robust nonlinear flight control, in: *Proceedings of the EuroGNC2013, 2nd CEAS Specialist Conference on Guidance, Navigation & Control*, Delft University of Technology, Delft, The Netherlands, 2013, [https://doi.org/10.1016/S0140-6736\(01\)92136-1](https://doi.org/10.1016/S0140-6736(01)92136-1), April 10–12, 2013.
- [43] R. Kristiansen, D. Hagen, Modelling of actuator dynamics for spacecraft attitude control, *J. Guidance, Control, Dyn.* 32 (3) (2009) 3–6, <https://doi.org/10.2514/1.42574>.
- [44] R.A. Cordeiro, R. Azinheira, A. Moutinho, Cascaded incremental backstepping controller for the attitude tracking of fixed-wing aircraft, in: *5th CEAS Conference on Guidance, Navigation and Control*, 2019.
- [45] S. Zhang, C. Huang, K. Ji, H. Zhang, Prescribed performance incremental adaptive optimal fault-tolerant control for nonlinear systems with actuator faults, *ISA Trans* 120 (2022) 99–109, <https://doi.org/10.1016/j.isatra.2021.03.011>.
- [46] C.E. Rasmussen, C.K.I. Williams, *Gaussian Processes for Machine Learning*, The MIT Press, 2005, <https://doi.org/10.7551/mitpress/3206.001.0001>, Nov. 23.
- [47] V.N. Vapnik, *The Nature of Statistical Learning Theory*, Second, Springer, 1999, <https://doi.org/10.1007/978-1-4757-3264-1>.
- [48] R.C. Grande, G. Chowdhary, J.P. How, Experimental validation of bayesian nonparametric adaptive control using gaussian processes, *J. Aerospace Infor. Syst.* 11 (9) (2014) 565–578, <https://doi.org/10.2514/1.i010190>.
- [49] D.I. Ignatyev, M.E. Sidoryuk, K.A. Kolinko, A.N. Khrabrov, Dynamic rig for validation of control algorithms at high angles of attack, *J. Aircr* 54 (5) (2017) 1760–1771, <https://doi.org/10.2514/1.C034167>.
- [50] N. Srinivas, A. Krause, S.M. Kakade, M.W. Seeger, Information-theoretic regret bounds for gaussian process optimization in the bandit setting, *IEEE Trans Inf Theory* 58 (5) (2012) 3250–3265, <https://doi.org/10.1109/TIT.2011.2182033>.
- [51] S.R. Chowdhury, A. Gopalan, On kernelized multi-armed bandits, in: *Proceedings of the 34th International Conference on Machine Learning (PMLR)* 70, 2017, pp. 844–853.
- [52] B. Schölkopf, R. Herbrich, A.J. Smola, A generalized representer theorem, *Lecture Notes Comput. Sci.* 2111 (2001) 416–426, https://doi.org/10.1007/3-540-44581-1_27.
- [53] H.A. Kingravi, G. Chowdhary, P.A. Vela, E.N. Johnson, Reproducing kernel hilbert space approach for the online update of radial bases in neuro-adaptive control, *IEEE Trans. Neural Netw. Learn Syst.* 23 (7) (2012) 1130–1141, <https://doi.org/10.1109/TNNLS.2012.2198889>.
- [54] S.H. Kim, Y.S. Kim, C. Song, A robust adaptive nonlinear control approach to missile autopilot design, *Control Eng. Pract.* 12 (2) (2004) 149–154, [https://doi.org/10.1016/S0967-0661\(03\)00016-9](https://doi.org/10.1016/S0967-0661(03)00016-9).



DSPACE

<https://dspace.org/>

Addressing incremental backstepping control limitations with direct online Gaussian process adaptation

Ignatyev, Dmitry; Tsourdos, Antonios

2025-09-01

Elsevier

<https://doi.org/10.1016/j.jfranklin.2025.107948>

Ignatyev D, Tsourdos A. (2025) Addressing incremental backstepping control limitations with direct online Gaussian process adaptation. Journal of the Franklin Institute, Volume 362, Issue 14, September 2025, Article number 107948