AUT Journal of Modeling and Simulation

AUT J. Model. Simul., 58(2) (2026) 215-228
DOI: 10.22060/miscj.2026.26499.5530

Safe Active Learning of Nonlinear Dynamics with Exploration-Aware Recursive Sparse Gaussian Process

Ali Salmanpour*, Hamed Kebriaei

School of Electrical and Computer Engineering, College of Engineering, University of Tehran, Tehran, Iran.

ABSTRACT: This paper develops an exploration-aware predictive control framework for safe online learning of unknown nonlinear dynamics. A recursive sparse Gaussian process (RSGP) with a fixed inducing set estimates an unknown state and input-dependent residual from closed-loop data. The previously established deterministic bounded-noise certificate is used to construct nested residual bounds, so noisy data cannot enlarge the uncertainty bound used for control. We develop a robust tube-based method where the nominal tube center follows only the known dynamics, while the nested residual envelope determines a rigid tube whose radius is updated causally between sampling instants and remains fixed during each predictive horizon. The fixed nominal dynamics, nested uncertainty envelope, causal tube update, and common terminal ingredients establish hard constraint satisfaction under every RSGP update without a post-update feasibility test. The numerical studies illustrate improved model-learning accuracy under active exploration, with no observed state or input constraint violations in the evaluated examples.

Review History:

Received: Aug. 12, 2026
Revised: Sep. 09, 2026
Accepted: Sep. 09, 2026
Available Online: Sep. 15, 2026

Keywords:

Safe Active Learning
Informative Planning
Nonlinear Model Predictive Control
Recursive Sparse Gaussian Processes
Rigid Tube MPC
Constraint Satisfaction

1- Introduction

Learning nonlinear dynamics from data is important when first-principles models are incomplete or affected by significant model mismatch. Online learning enables the model to be refined continuously during operation, but the quality of the learned model depends strongly on the closed-loop trajectory. Regulation-oriented operation may repeatedly visit only a limited region of the state–input space and therefore provide poorly informative data. Active learning addresses this limitation by deliberately selecting measurements or trajectories that improve the learned model efficiently [1, 2].

For physical systems, informative data collection must additionally satisfy safety requirements. Gaussian process (GP)-based safe active-learning methods restrict exploration according to learned safety information [3-5]. For uncertain dynamical systems, safety during learning has also been established through reachability- and Lyapunov-based arguments [6-8], predictive uncertainty propagation [9], contraction-based tracking [10], and sequential exploration–exploitation strategies [11]. These methods provide systematic mechanisms for safe exploration, predominantly using probabilistic or high-confidence descriptions of GP model uncertainty. In contrast, the present work considers bounded disturbances and reproducing kernel Hilbert space (RKHS) regularity and uses a deterministic bounded-noise error certificate. Accordingly, hard state and input constraints are guaranteed for all uncertainties satisfying the assumed bounds rather than on a prescribed confidence event.

Continual online learning introduces an additional difficulty because the uncertainty description changes whenever the model is updated. Castañeda *et al.* [12] address recursive safety using GP-based control barrier functions and pointwise feasibility conditions that determine when additional safe data acquisition is required. Here, the same issue is addressed by constructing a nested deterministic residual envelope. Although a noisy observation may enlarge the newly computed raw certificate, the uncertainty description used for safety cannot increase. Consequently, every bounded-noise observation can be incorporated into the learning update without model-update rejection or an additional post-update feasibility test.

Scalable online GP formulations address the computational growth associated with accumulating data [13, 14]. We adopt the fixed-inducing-point RSGP of Schürch *et al.* [15], which recursively incorporates new measurements while keeping the online model dimension independent of the accumulated data.

*Corresponding author's email: a.salmanpour9@ut.ac.ir

Copyrights for this article are retained by the author(s) with publishing rights granted to Amirkabir University Press. The content of this article is subject to the terms and conditions of the Creative Commons Attribution 4.0 International (CC-BY-NC 4.0) License. For more information, please visit https://www.creativecommons.org/licenses/by-nc/4.0/legalcode.

UNCORRECTED PROOF

Deterministic RKHS error bounds under bounded noise are developed in [16, 17], including a sparse-GP certificate in [18]. We express this certificate using recursively available RSGP quantities and construct a nested residual envelope that remains nonexpanding under successive noisy updates.

Motivated by these observations, this paper develops a safe active-learning framework for partially known nonlinear dynamics. The unknown state- and input-dependent residual is learned online using the RSGP of Schürch *et al.* [15]. The deterministic sparse-GP certificate of [18] is evaluated using recursively available quantities and converted into a nested residual envelope. A certificate-aware exploration criterion combines mutual information with the predicted change in this certificate, so data acquisition accounts for both statistical informativeness and improvement of the safety-relevant model description.

Nonlinear model predictive control (NMPC) is employed as the constrained trajectory-planning mechanism for realizing the safe learning strategy. The nominal trajectory follows only the known dynamics, whereas the nested residual envelope determines a causally updated rigid tube. A baseline problem provides a certified feasible trajectory, and an exploration refinement searches over the same robust feasible set for a more informative trajectory subject to a prescribed performance-degradation budget. This structure preserves hard constraint satisfaction under successive learning updates while allowing every bounded-noise RSGP update to be incorporated.

The main contributions are summarized as follows.

Building on the fixed-inducing-point RSGP of [15], we integrate fixed-dimensional recursive learning with safe active learning of unknown nonlinear residual dynamics, thereby combining scalable online estimation with constrained informative data collection.

Building on the deterministic RKHS error bounds in [16, 17] and the sparse-GP certificate in [18], we construct a nested residual envelope that remains nonexpanding under successive noisy updates, enabling recursive robust constraint handling.

In contrast to the predominantly probabilistic or high-confidence safe-learning approaches in [3-11], the proposed exploration criterion explicitly accounts for the predicted change in the deterministic error certificate in addition to statistical informativeness.

Compared with the update-dependent recursive-safety mechanism in [12], the proposed nested residual envelope is coupled with a causally updated rigid tube to guarantee hard constraint satisfaction under successive bounded-noise RSGP updates without model-update rejection or a post-update feasibility test.

1- 1- Notation

For a positive-definite matrix , define . For a symmetric matrix , denotes its largest eigenvalue. For sets A and B , $A \oplus B$ and $A \ominus B$ denote the Minkowski sum and Pontryagin difference, respectively. The notation $x_{k|t}$ denotes the prediction of x_{t+k} computed at time t . The operator $\mathrm{col}(\cdot)$ denotes column-wise stacking.

2- Problem Formulation

Consider the constrained discrete-time nonlinear system

$$x_{t+1} = f(x_t, u_t) + g(x_t, u_t) + \delta_t, \tag{1}$$

where $x_t \in \mathcal{X} \subset \mathbb{R}^{n_x}$ and $u_t \in \mathcal{U} \subset \mathbb{R}^{n_u}$. The compact sets $\mathcal{X}$ and $\mathcal{U}$ represent hard constraints. The function $f : \mathcal{X} \times \mathcal{U} \to \mathbb{R}^{n_x}$ is a known nominal model, whereas $g : \mathcal{X} \times \mathcal{U} \to \mathbb{R}^{n_x}$ is an unknown residual learned online. The additive term δ_t represents bounded process and residual-observation uncertainty.

Define $\xi_t := \mathrm{col}(x_t, u_t) \in \Xi := \mathcal{X} \times \mathcal{U}$ and use $g(\xi_t)$ as shorthand for $g(x_t, u_t)$. After applying u_t and measuring x_{t+1}, the residual observation is

$$y_t = x_{t+1} - f(x_t, u_t) = g(\xi_t) + \delta_t. \tag{2}$$

Assumption 1 (Regularity and bounded uncertainty). *The functions f and g are continuous on an open neighborhood of Ξ, and f is continuously differentiable. The uncertainty satisfies $\|\delta_t\|_2 \le \bar{\delta}$ for all $t \ge 0$. Each component g_j belongs to the RKHS $\mathcal{H}_q$ associated with the continuous bounded sparse kernel q introduced in Section 3 and satisfies $\|g_j\|_{\mathcal{H}_q} \le \Gamma_j$.*

The objective is to learn g from (2) while actively selecting informative inputs and guaranteeing $x_t \in \mathcal{X}$ and $u_t \in \mathcal{U}$ for all $t \ge 0$. The desired state is denoted by $x_s \in \mathcal{X}$ and need not be an admissible nominal equilibrium.

Because the learned residual changes after each measurement, we separate learning from robust constraint handling. The nominal trajectory follows only the known dynamics, while the residual and disturbance are handled through a certified tube. This architecture is used for both safe regulation and informative input selection.

2- 1- Control Architecture

Let z and v denote the nominal state and input, respectively, and consider an ancillary feedback law $u = \kappa(x, z, v)$ with $\kappa(z, z, v) = v$. The nominal prediction follows the known dynamics, i.e., $z_{k+1|t} = f(z_{k|t}, v_{k|t})$. For a prediction horizon N , define

$$\chi_t := (\mathbf{z}_t, \mathbf{v}_t, \theta_t^s, u_t^s), \qquad z_t := z_{0:N|t}, \qquad v_t := v_{0:N-1|t},$$

where (θ_t^s, u_t^s) denotes the artificial steady-state pair introduced later in the NMPC formulation. The candidate input associated with χ_t is therefore $u_t(\chi_t) = \kappa(x_t, z_{0|t}, v_{0|t})$. The residual is next learned and certified, after which the certificate is used for informative input selection and robust tube construction.

UNCORRECTED PROOF

3- RSGP Learning and Deterministic Certificate

To enable continual learning without increasing the dimension of the online estimator, we employ an RSGP with a fixed inducing set $R = \{\xi_i^r\}_{i=1}^M \subset \Xi$. Let k denote the underlying kernel, with $K_{RR} := [k(\xi_i^r, \xi_j^r)]_{i,j=1}^M$, $K_{\xi R} := \left[k\left(\xi, \xi_1^r\right), \ldots, k\left(\xi, \xi_M^r\right)\right]$, and $K_{R\zeta} := K_{\zeta R}^{\top}$, where K_{RR} is nonsingular. The sparse kernel and corresponding feature map are $q(\xi, \zeta) := K_{\xi R} K_{RR}^{-1} K_{R\zeta}$ and $H(\xi) := K_{\xi R} K_{RR}^{-1}$, respectively.

3- 1- Recursive Sparse GP

An independent RSGP is assigned to each component of g. Its posterior mean is

$$\hat{g}_{j,t}(\xi) = H(\xi)\mu_{j,t}, \qquad j = 1, \ldots, n_x. \tag{3}$$

The recursion is initialized with $\mu_{j,0} = 0$ and $\Sigma_{j,0} = K_{RR}$. For a new sample (ξ_t, y_t),

$$\begin{aligned} S_{j,t+1} &= H(\xi_t)\Sigma_{j,t}H(\xi_t)^{\top} + \sigma_j^2 && \text{(a)}\\ G_{j,t+1} &= \Sigma_{j,t}H(\xi_t)^{\top}S_{j,t+1}^{-1} && \text{(b)}\\ \mu_{j,t+1} &= \mu_{j,t} + G_{j,t+1}\left(y_{j,t} - H(\xi_t)\mu_{j,t}\right) && \text{(c)}\\ \Sigma_{j,t+1} &= \Sigma_{j,t} - G_{j,t+1}S_{j,t+1}G_{j,t+1}^{\top} && \text{(d)} \end{aligned} \tag{4}$$

where $\sigma_j^2 > 0$ is the likelihood-noise parameter. The Gaussian likelihood is used for estimation and information scoring, whereas $\bar{\delta}$ is used for deterministic safety. For the deterministic certificate, initialize $\Omega_0 = 0$ and update the fixed-dimensional statistic

$$\Omega_{t+1} = \Omega_t + H(\xi_t)^{\top}H(\xi_t). \tag{5}$$

Because R is fixed, the RSGP state, Ω_t, and the per-sample recursions have dimensions independent of the number of observations [15].

3- 2- Deterministic Error Certificate and Nested Residual Envelope

The RSGP posterior provides the quantities used for estimation and information scoring, whereas hard constraint certification relies on the deterministic bounded-noise RKHS error bound. Let $D_t := \{(\xi_i, y_i)\}_{i=1}^{n_t}$ denote the data available at time t. When used as a kernel index, D_t denotes its input set. Define $Q_{D_tD_t} := K_{D_tR}K_{RR}^{-1}K_{RD_t}$, $Q_{D_t\xi} := K_{D_tR}K_{RR}^{-1}K_{R\xi}$. Let $H_{D_t} := \mathrm{col}\left(H(\xi_1), \ldots, H(\xi_{n_t})\right)$, so that $\Omega_t = H_{D_t}^{\top}H_{D_t}$, and let $y_{j,D_t} := \mathrm{col}\left(y_{j,1}, \ldots, y_{j,n_t}\right)$. The recursive mean admits the equivalent batch representation

$$\hat{g}_{j,t}(\xi) = Q_{\xi D_t}\left(Q_{D_tD_t} + \sigma_j^2 I\right)^{-1} y_{j,D_t}.$$

Moreover, the matrix inversion lemma gives

$$\begin{aligned} \left(Q_{D_tD_t} + \sigma_j^2 I\right)^{-1} Q_{D_t\xi} &= \frac{1}{\sigma_j^2} H_{D_t}\Sigma_{j,t}H(\xi)^{\top}\\ \left\|\left(Q_{D_tD_t} + \sigma_j^2 I\right)^{-1} Q_{D_t\xi}\right\|_2^2 &= \frac{1}{\sigma_j^4} H(\xi)\Sigma_{j,t}\Omega_t\Sigma_{j,t}H(\xi)^{\top} \end{aligned} \tag{6}$$

Following the bounded-noise RKHS argument in [18], applied to the sparse kernel q, define

$$\begin{aligned} \nu_{j,t}(\xi) &:= \Gamma_j\sqrt{H(\xi)\Sigma_{j,t}H(\xi)^{\top}}\\ &+ \frac{\bar{\delta}\sqrt{n_t}}{\sigma_j^2}\sqrt{H(\xi)\Sigma_{j,t}\Omega_t\Sigma_{j,t}H(\xi)^{\top}} \end{aligned} \tag{7}$$

The RSGP recursion in [15] is algebraically equivalent to the corresponding batch sparse-GP posterior associated with the same fixed inducing set and sparse kernel q. Therefore, the deterministic sparse-GP certificate of [18] applies to the recursive representation used here under the RKHS and bounded-noise conditions of Assumption 1, yielding

$$\begin{aligned} &\left|g_j(\xi) - \hat{g}_{j,t}(\xi)\right| \le \nu_{j,t}(\xi),\\ &\forall \xi \in \Xi, \quad j = 1, \ldots, n_x. \end{aligned} \tag{8}$$

Define the raw pointwise certificate

$$\mathcal{B}_t(\xi) := \left\{d \in \mathbb{R}^{n_x} \mid \left|d_j - \hat{g}_{j,t}(\xi)\right| \le \nu_{j,t}(\xi),\ j = 1, \ldots, n_x\right\}$$

and the nested certificate

$$\begin{aligned} \mathcal{C}_0(\xi) &:= \mathcal{B}_0(\xi),\\ \mathcal{C}_{t+1}(\xi) &:= \mathcal{C}_t(\xi) \cap \mathcal{B}_{t+1}(\xi). \end{aligned} \tag{9}$$

Equation (8) gives $g(\xi) \in \mathcal{C}_{t+1}(\xi) \subseteq \mathcal{C}_t(\xi), \forall \xi \in \Xi$. The pointwise certificate retains the state–input dependence of the residual uncertainty and is useful for assessing candidate measurements. Robust tube construction, in contrast, requires a uniform disturbance bound over the admissible domain. We therefore convert each valid pointwise certificate into a certified global residual bound and retain the smallest bound

UNCORRECTED PROOF

obtained up to the current sampling instant. For the rigid-tube construction, choose a certified raw global bound $\bar{b}_t$ satisfying

$$\bar{b}_t \geq \sup_{\xi\in\Xi}\left(\sum_{j=1}^{n_x}\left(\left|\hat{g}_{j,t}(\xi)\right|+\nu_{j,t}(\xi)\right)^2\right)^{1/2} \tag{10}$$

and update

$$\bar{g}_0 := \bar{b}_0, \qquad \bar{g}_{t+1} := \min\{\bar{g}_t, \bar{b}_{t+1}\}. \tag{11}$$

Here, $\mathcal{B}_t(\xi)$ and $\mathcal{C}_t(\xi)$ are pointwise residual certificates, with $\mathcal{C}_t(\xi)$ obtained by nesting successive raw certificates. In contrast, $\bar{b}_t$ is the raw global scalar bound over Ξ, while $\bar{g}_t$ is its nonincreasing nested counterpart used for the robust tube construction.

The upper bound in (10) is evaluated once per sampling instant using a certified global method on the compact domain Ξ, such as interval branch-and-bound with valid interval extensions of the terms in (10).

Lemma 1 (Validity and monotonicity of the residual envelope). *For all* $t \geq 0$ *and* $\xi \in \Xi$, $\|g(\xi)\|_2 \leq \bar{g}_t$, $\bar{g}_{t+1} \leq \bar{g}_t$.

Proof. Equation (8) implies $\left|g_j(\xi)\right| \leq \left|\hat{g}_{j,i}(\xi)\right| + \nu_{j,i}(\xi)$ for every $i \leq t$. Hence, $\|g(\xi)\|_2 \leq \bar{b}_i$ for every $i \leq t$, and therefore $\|g(\xi)\|_2 \leq \min_{0\leq i\leq t}\bar{b}_i = \bar{g}_t$. Monotonicity follows directly from (11).

Remark 1. *The posterior mean in (3) is the RSGP point estimate and centers each raw certificate. The nominal tube center is propagated only by the known dynamics. Safety uses the nested envelope (11), which remains valid even when noisy data move the posterior mean or enlarge the new raw certificate.*

4- Certificate-Aware Exploration Score

The preceding learning module provides two complementary indicators of the value of a candidate measurement: the RSGP covariance quantifies statistical informativeness, while the deterministic certificate quantifies its predicted effect on the safety-relevant model description. We combine these quantities to construct the exploration criterion. For Gaussian observations, the one-step mutual information between a candidate measurement at ξ and the inducing variables is

$$\mathcal{I}_t(\xi) := \frac{1}{2}\sum_{j=1}^{n_x}\log\left(1+\frac{H(\xi)\Sigma_{j,t}H(\xi)^\top}{\sigma_j^2}\right). \tag{12}$$

The output-independent covariance after sampling at ξ is

$$\Sigma_{j,t+1|t}(\xi) := \Sigma_{j,t} - \frac{\Sigma_{j,t}H(\xi)^\top H(\xi)\Sigma_{j,t}}{H(\xi)\Sigma_{j,t}H(\xi)^\top + \sigma_j^2}.$$

Let $\mathcal{G} := \{\zeta^\ell\}_{\ell=1}^{N_g} \subset \Xi$ be a fixed design set used only to evaluate the exploration score. Let $\nu_{j,t+1|t}(\zeta;\xi)$ be obtained from (7) by replacing

$$\Sigma_{j,t} \leftarrow \Sigma_{j,t+1|t}(\xi),$$
$$\Omega_t \leftarrow \Omega_t + H(\xi)^\top H(\xi),$$
$$n_t \leftarrow n_t + 1$$

Define $\bar{\nu}_t := \max_{\zeta\in\mathcal{G}}\left(\sum_{j=1}^{n_x}\nu_{j,t}^2(\zeta)\right)^{1/2}$ and $\bar{\nu}_{t+1|t}(\xi) := \max_{\zeta\in\mathcal{G}}\left(\sum_{j=1}^{n_x}\nu_{j,t+1|t}^2(\zeta;\xi)\right)^{1/2}$. The next residual observation is collected at $\xi_t^{obs}(\chi_t) := \mathrm{col}\left(x_t, \kappa\left(x_t, z_{0|t}, v_{0|t}\right)\right)$. The certificate-aware exploration score is

$$\mathcal{A}_t(\chi_t) := \mathcal{I}_t\left(\xi_t^{\mathrm{obs}}(\chi_t)\right) + \lambda_\nu\frac{\bar{\nu}_t - \bar{\nu}_{t+1|t}\left(\xi_t^{\mathrm{obs}}(\chi_t)\right)}{\max\{\bar{\nu}_t, \epsilon_\nu\}} \tag{13}$$

where $\lambda_\nu \geq 0$ and $\varepsilon_\nu > 0$. The first term rewards covariance reduction, while the second rewards the predicted reduction and penalizes the predicted growth of the maximum certificate width over $\mathcal{G}$. The score is used only for exploration; safety follows from the certified nested envelope over the full domain Ξ.

5- Computable Rigid Tube Construction

We now translate the deterministic residual envelope into a robust constraint-tightening mechanism. Since the nominal trajectory is propagated only by f, the complete mismatch between the plant and nominal dynamics is $w_t := g(\xi_t) + \delta_t$. Accordingly, define the certified additive uncertainty bound

$$\bar{w}_t := \bar{g}_t + \bar{\delta}. \tag{14}$$

By Lemma 1, $\|w_t\|_2 \leq \bar{w}_t$ and $\bar{w}_{t+1} \leq \bar{w}_t$. The following incremental stabilizability condition converts this disturbance bound into a robust tube around the nominal trajectory.

Assumption 2 (Incremental robust stabilizability). *There exists a continuous function* $V_\delta : \mathbb{R}^{n_x}\times\mathbb{R}^{n_x} \to \mathbb{R}_{\geq 0}$, *a continuous ancillary feedback* κ, *constants* $c_l, c_u > 0$, $\rho \in (0,1)$, *and* $\gamma > 0$ *such that* $\kappa(z, z, v) = v$ and

$$c_l \|x - z\|_2^2 \leq V_\delta(x,z) \leq c_u \|x - z\|_2^2$$
$$V_\delta\begin{pmatrix} f(x,\kappa(x,z,v)) \\ +w, f(z,v)\end{pmatrix} \leq \rho V_\delta(x,z) + \gamma\|w\|_2^2 \tag{15}$$

UNCORRECTED PROOF

for all admissible (x, z, v) and all $\|w\|_2 \le \bar{w}_0$.

Remark 2. *Assumption 2 is closely related to incremental input-to-state stabilizability (δ-ISS) conditions used in robust nonlinear MPC [19, 20]. The incremental Lyapunov function and ancillary feedback are offline controller-design ingredients and need only be certified over the prescribed operating region. For a quadratic metric $V_\delta(x,z) = (x-z)^\top P(x-z)$ and affine ancillary feedback $\kappa(x,z,v) = v + K(x-z)$, the parameters P, K, ρ, and γ can be obtained and verified using polytopic/LPV matrix-inequality constructions based on bounds on the incremental system Jacobian. For more general nonlinear systems, sum-of-squares or control-contraction-metric constructions provide state-dependent alternatives [21, 22]. These computations are performed offline and do not increase the online complexity of the proposed RSGP-NMPC scheme. Such incremental-stabilizability constructions have also been demonstrated for nonlinear quadrotor models [20, 22].*

For a causally determined radius $s_t \ge 0$, define the rigid tube employed in the NMPC problem at time t as

$$\mathcal{Z}_t(z) := \{x \in \mathbb{R}^{n_x} \mid V_\delta(x,z) \le s_t^2\}. \tag{16}$$

The following result provides a valid causal choice of s_t and establishes the corresponding tube propagation and nonexpansion properties.

Proposition 1 (Causal tube radius and propagation). *Suppose Assumption 2 holds and $\bar{w}_{t+1} \le \bar{w}_t$ for all t.* Initialize

$$s_0^2 := \frac{\gamma}{1-\rho}\bar{w}_0^2 \tag{17}$$

and update the radius according to

$$s_{t+1}^2 := \rho s_t^2 + \gamma \bar{w}_t^2 \tag{18}$$

If $x_t \in \mathcal{Z}_t(z_t)$, $u_t = \kappa(x_t, z_t, v_t)$, and $z_{t+1} = f(z_t, v_t)$, then $x_{t+1} \in \mathcal{Z}_{t+1}(z_{t+1})$. Moreover, $s_{t+1} \le s_t$, $\mathcal{Z}_{t+1}(z) \subseteq \mathcal{Z}_t(z)$.

Proof. Using (1), Assumption 2, and equation (14),

$$\begin{aligned} V_\delta(x_{t+1}, z_{t+1}) &\le \rho V_\delta(x_t, z_t) \\ &+ \gamma \bar{w}_t^2 \le \rho s_t^2 + \gamma \bar{w}_t^2 = s_{t+1}^2 \end{aligned} \tag{19}$$

which proves $x_{t+1} \in \mathcal{Z}_{t+1}(z_{t+1})$.

It remains to establish nonexpansion. From equation (17) suppose that $s_t^2 \ge \gamma \bar{w}_t^2 / (1-\rho)$. Then,

$$s_{t+1}^2 = \rho s_t^2 + \gamma \bar{w}_t^2 \ge \frac{\gamma}{1-\rho}\bar{w}_t^2 \ge \frac{\gamma}{1-\rho}\bar{w}_{t+1}^2$$

$$s_{t+1}^2 - s_t^2 = -(1-\rho)s_t^2 + \gamma \bar{w}_t^2 \le 0.$$

where the last inequality follows from $\bar{w}_{t+1} \le \bar{w}_t$. Furthermore,

Induction therefore yields $s_{t+1} \le s_t$. The tube inclusion follows directly from (16).

Hence, s_t is a causally updated controller parameter rather than an optimization variable. The resulting tube can therefore be used to translate the original plant constraints into nominal constraints that are fixed throughout each NMPC optimization.

Define the robustly admissible nominal sets

$$\begin{aligned} &\bar{\mathcal{X}}_t := \{z \in \mathbb{R}^{n_x} \mid \mathcal{Z}_t(z) \subseteq \mathcal{X}\} && \text{(a)} \\ &\bar{\mathcal{U}}_t(z) := \{v \in \mathcal{U} \mid \kappa(x,z,v) \in \mathcal{U},\ \forall x \in \mathcal{Z}_t(z)\} && \text{(b)} \end{aligned} \tag{20}$$

Tube nonexpansion implies $\bar{\mathcal{X}}_t \subseteq \bar{\mathcal{X}}_{t+1}$ and $\bar{u}_t(z) \subseteq \bar{u}_{t+1}(z)$.

Remark 3. *For $\kappa(x,z,v) = v + K(x-z)$ and a translation-invariant tube $\mathcal{Z}_t(z) = z \oplus \mathcal{E}_t$, the robust constraints reduce to $z \in \mathcal{X} \ominus \mathcal{E}_t, v \in \mathcal{U} \ominus K\mathcal{E}_t$. Hence, the nested residual envelope affects the NMPC only through the scalar recursion (18) and the corresponding constraint tightening.*

6- Exploration-Aware Safe NMPC

The preceding construction provides, at each sampling instant, a certified tube and the corresponding tightened state and input sets. These ingredients are now used to formulate the safe NMPC problem. The controller first computes the best regulation trajectory over the certified feasible set and then searches over the same set for a more informative trajectory subject to a prescribed regulation-performance degradation.

Since the desired state x_s need not be an admissible equilibrium of the nominal dynamics, regulation is formulated with respect to an artificial steady state (θ, u^s). Define the fixed admissible set

$$\mathcal{S} := \{(\theta, u^s) \mid \theta = f(\theta, u^s),\ \theta \in \bar{\mathcal{X}}_0,\ u^s \in \bar{\mathcal{U}}_0 \tag{21}$$

Because the tightened sets expand, every element of $\mathcal{S}$ remains admissible for all $t \ge 0$. Define

$$\ell(z, v; \theta, u^s) := \| z - \theta \|_Q^2 + \| v - u^s \|_R^2, \tag{22}$$

UNCORRECTED PROOF

where $Q \succ 0$ and $R \succ 0$. To close the finite prediction horizon while preserving hard constraint satisfaction, terminal ingredients are introduced to keep the terminal nominal state within an admissible invariant region under a local terminal controller. Since the tightened sets expand as the certified tube contracts, these ingredients can be constructed for the initial tube and retained thereafter.

Assumption 3 (Common terminal conditions). *For every $(\theta,u^s)\in\mathcal{S}$, there exists a compact terminal set $\bar{\chi}_f(\theta,u^s)\subseteq\bar{\chi}_0$, a terminal controller v_f, and a continuous nonnegative terminal cost V_f such that, for every $z\in\mathcal{X}_f(\theta,u^s)$, with $z^+ := f(z, v_f(z;\theta,u_s))$,*

$$\begin{aligned} &v_f(z;\theta,u_s) \in \bar{\mathcal{U}}_0(z), && \text{(a)}\\ &z^+ \in \mathcal{X}_f(\theta,u_s), && \text{(b)}\\ &V_f(z^+;\theta) - V_f(z;\theta) \le && \text{(c)}\\ &\quad -\ell(z,v_f(z;\theta,u_s);\theta,u_s). \end{aligned} \tag{23}$$

The terminal ingredients are constructed and verified offline for the initial tube. Since tube nonexpansion implies $\bar{\chi}_0\subseteq\bar{\chi}_t$ and $\bar{u}_0(z)\subseteq\bar{u}_t(z)$, the same terminal ingredients remain valid for all $t\ge 0$.

Remark 4. *For a translation-invariant tube $\mathcal{Z}_0(z)=z\oplus\mathcal{E}_0$, let $\mathcal{X}_f^r(\theta,u^s)\subseteq\mathcal{X}$ be a robust terminal state set. The corresponding admissible nominal-center set is $\mathcal{X}_f^c(\theta,u^s):=\{z\,\square\,\mathcal{Z}_0(z)\subseteq\mathcal{X}_f^r(\theta,u^s)\}=\mathcal{X}_f^r(\theta,u^s)\ominus\mathcal{E}_0$. Hence, $\mathcal{X}_f(\theta,u^s)$ in Assumption 3 may be chosen as any compact positively invariant subset of $\mathcal{X}_f^c(\theta,u^s)$ satisfying the terminal input and decrease conditions. Since $\mathcal{E}_t\subseteq\mathcal{E}_0$, this construction remains valid for all $t\ge 0$.*

For the measured state x_t, define $\mathcal{F}_t(x_t)$ as the set of all χ_t satisfying

$$\begin{aligned} &x_t \in \mathcal{Z}_t(z_{0|t}) && \text{(a)}\\ &z_{k+1|t} = f(z_{k|t},v_{k|t}), \quad k=0,\dots,N-1 && \text{(b)}\\ &z_{k|t}\in\bar{\mathcal{X}}_t, \qquad k=0,\dots,N && \text{(c)}\\ &v_{k|t}\in\bar{\mathcal{U}}_t(z_{k|t}), \qquad k=0,\dots,N-1 && \text{(d)}\\ &z_{N|t}\in\mathcal{X}_f(\theta_t^s,u_t^s) && \text{(e)}\\ &(\theta_t^s,u_t^s)\in\mathcal{S} && \text{(f)} \end{aligned} \tag{24}$$

The initial tube constraint connects the measured state to the nominal prediction, while the tightened state and input constraints ensure admissibility of the corresponding plant trajectory. The terminal constraint provides a feasible continuation at subsequent sampling instants. The tube radius s_t is fixed during the optimization; only the nominal states, nominal inputs, and artificial steady state are optimized.

For $\chi_t\in\mathcal{F}_t(x_t)$, define the finite-horizon regulation cost

$$\begin{aligned} J^{\mathrm{reg}}(\chi_t) :=& \sum_{k=0}^{N-1}\ell\left(z_{k|t},v_{k|t};\theta_t^s,u_t^s\right)\\ &+V_f\left(z_{N|t};\theta_t^s\right)+\lambda_\theta\parallel\theta_t^s-x_s\parallel_2^2 \end{aligned} \tag{25}$$

where $\lambda_\theta>0$. The baseline safe NMPC problem is

$$\chi_t^{b,*}\in\arg\min_{\chi_t\in\mathcal{F}_t(x_t)} J^{\mathrm{reg}}(\chi_t), \qquad V_t(x_t):=J^{\mathrm{reg}}\left(\chi_t^{b,*}\right). \tag{26}$$

The baseline problem therefore determines the minimum regulation cost attainable within the certified feasible set. This optimal value is used as a reference for exploration: informative trajectories are permitted only if their regulation cost remains within the prescribed degradation budget ε_t.

For a prescribed exploration budget $\varepsilon_t\ge 0$, the exploration refinement is defined as

$$\begin{aligned} &\chi_t^{e,*}\in\arg\max_{\chi_t}\mathcal{A}_t(\chi_t) && \text{(a)}\\ &\text{s.t.} \quad \chi_t\in\mathcal{F}_t(x_t) && \text{(b)}\\ &J^{\mathrm{reg}}(\chi_t)\le V_t(x_t)+\varepsilon_t && \text{(c)} \end{aligned} \tag{27}$$

Proposition 2 (Feasibility of the exploration refinement). *If the baseline problem (26) is feasible and $\varepsilon_t\ge 0$, then (27) is feasible.*

Proof. The baseline optimizer satisfies $\chi_t^{b,*}\in\mathcal{F}_t(x_t)$ and $J^{\mathrm{reg}}\left(\chi_t^{b,*}\right)=V_t(x_t)\le V_t(x_t)+\varepsilon_t$.

Hence, $\chi_t^{b,*}$ is a feasible candidate for (27).

Thus, exploration introduces no additional safety mechanism; it only selects a more informative trajectory from the same robust feasible set while limiting the loss of regulation performance.

When the exploration problem is solved, set $\chi_t^{a,*}:=\chi_t^{e,*}$; otherwise, use the baseline optimizer as a numerical fallback and set $\chi_t^{a,*}:=\chi_t^{b,*}$. The applied input is

$$u_t=\kappa\left(x_t,z_{0|t}^{a,*},v_{0|t}^{a,*}\right). \tag{28}$$

After measuring x_{t+1}, the radius is updated by (18) the RSGP is updated by (4), and the residual envelope is updated by (4). Every bounded-noise sample is used; no model-update acceptance or next-step feasibility test is performed.

UNCORRECTED PROOF

The resulting closed-loop implementation is summarized in Algorithm 1.

Algorithm 1: Exploration-aware RSGP-NMPC with a causal rigid tube

1. Initialize $\mu_{j,0}=0$, $\Sigma_{j,0}=K_{RR}$, $\Omega_0=0$, $\bar{g}_0$, $\bar{w}_0$, and s_0 from (17) construct the common terminal ingredients.
2. For $t=0,1,\ldots$ do.
3. Construct $\mathcal{Z}_t$, $\bar{\chi}_t$, and $\bar{u}_t$ from s_t and solve (26).
4. Solve (27); use the baseline optimizer as the numerical fallback.
5. Apply (28) and measure x_{t+1}.
6. Set $s_{t+1}^2=\rho s_t^2+\gamma\bar{w}_t^2$.
7. Form y_t from (2), update the RSGP by (4), and update Ω_{t+1} by (5).
8. Compute $\bar{b}_{t+1}$, set $\bar{g}_{t+1}=\min\{\bar{g}_t,\bar{b}_{t+1}\}$, and set $\bar{w}_{t+1}=\bar{g}_{t+1}+\bar{\delta}$.
9. end for

The following result establishes hard constraint satisfaction of the proposed closed-loop learning scheme under successive RSGP updates.

Proposition 3 (Hard constraint satisfaction). *Suppose Assumptions 1–3 hold, $\varepsilon_t\geq 0$ and the baseline problem is feasible at $t=0$. Then Algorithm 1 is well defined for all $t\geq 0$, and the implemented closed loop satisfies $x_t\in\mathcal{X}, u_t\in\mathcal{U}, \forall t\geq 0$.*

Proof. We proceed by induction. At $t=0$, the baseline problem is feasible by assumption. Suppose it is feasible at time t. By Proposition 2, the exploration problem is feasible, and hence $\chi_t^{a,*}\in\mathcal{F}_t(x_t)$. Therefore, $x_t\in\mathcal{X}$ and $u_t\in\mathcal{U}$ by the definitions of $\bar{\chi}_t$ and $\bar{u}_t$.

Since $z_{1|t}^{a,*}=f\left(z_{0|t}^{a,*},v_{0|t}^{a,*}\right)$, Proposition 1 gives

$$x_{t+1}\in\mathcal{Z}_{t+1}\left(z_{1|t}^{a,*}\right)\subseteq\mathcal{X}.$$

Shifting the implemented nominal sequence by one step, retaining $\left(\theta_t^{s,a,*},u_t^{s,a,*}\right)$, and completing it with the terminal controller yields an admissible candidate at $t+1$. Indeed, $\bar{\chi}_t\subseteq\bar{\chi}_{t+1}$ and $\bar{u}_t(z)\subseteq\bar{u}_{t+1}(z)$ preserve the shifted constraints, while terminal invariance completes the candidate. Hence, the baseline problem remains feasible at $t+1$.

Induction therefore establishes that Algorithm 1 is well defined for all $t\geq 0$ and that $x_t\in\mathcal{X}$ and $u_t\in\mathcal{U}$ for all $t\geq 0$.

7- Numerical Example

The numerical study first evaluates hard constraint satisfaction and learning performance on a constrained nonlinear system, and then examines computational scalability on a higher-dimensional quadrotor benchmark.

7- 1- Safe Active-Learning Performance

The numerical example uses the same two-state nonlinear nominal dynamics as before, with a revised unknown residual. Specifically,

$$x_{1,t+1}=x_{1,t}+T_s x_{2,t}+g_1(x_t,u_t)+\delta_{1,t} \quad \text{(a)}$$

$$x_{2,t+1}=x_{2,t}+T_s\begin{bmatrix}-0.6\sin(x_{1,t})\\-0.2x_{2,t}+u_t\end{bmatrix} \quad \text{(b)} \quad (29)$$

$$+g_2(x_t,u_t)+\delta_{2,t}$$

Define

$$b(x,u)=\exp\begin{bmatrix}-\left(\frac{x_1-0.75}{0.38}\right)^2\\-\left(\frac{x_2-0.15}{0.30}\right)^2-\left(\frac{u-0.65}{0.25}\right)^2\end{bmatrix} \quad (30)$$

and let the unknown residual be

$$g(x,u)=T_s\begin{bmatrix}0.04\tanh(x_1)+0.012x_2u+0.14b(x,u)\\0.06\sin(2x_1)+0.018u^2+0.18b(x,u)\end{bmatrix}$$

The sampling time is $T_s=0.1s$, and the hard constraints are $x_i\in[-7,7]$ and $u\in[-2,2]$. The process disturbance satisfies $\|\delta_t\|_\infty\leq 10^{-5}$; accordingly, $\bar{\delta}=\sqrt{2}10^{-5}$ is used for the Euclidean-norm bound in Assumption 1.

The nominal prediction model is

$$z_{1,k+1|t}=z_{1,k|t}+T_s z_{2,k|t} \quad \text{(a)}$$

$$z_{2,k+1|t}=z_{2,k|t}+T_s\begin{bmatrix}-0.6\sin(z_{1,k|t})\\-0.2z_{2,k|t}+v_{k|t}\end{bmatrix} \quad \text{(b)} \quad (31)$$

The RSGP mean is used for residual estimation and certificate centering, while the nominal prediction remains governed by (31) and robust constraint tightening is determined by the nested deterministic residual envelope.

The prediction horizon is $N=10$, with $Q=\mathrm{diag}(4,6)$ and $R=0.1$. The RSGP uses a squared-exponential kernel with length scales $(0.75,0.45,0.30)$, a fixed $7\times3\times5$ inducing-point grid over $[-1.5,1.5]\times[-0.6,0.6]\times[-1.2,1.2]$, likelihood standard deviations $\sigma_1=\sigma_2=0.2$, and RKHS bounds $(\Gamma_1,\Gamma_2)=(0.02,0.03)$. Six initial samples are generated in $[-0.2,0.2]\times[-0.1,0.1]\times[-0.2,0.2]$ and processed through the RSGP recursion before closed-loop operation.

For the tube construction, we use $V_\delta(x,z)=\|x-z\|_P^2$ and $\kappa(x,z,v)=v+K(x-z)$, with

UNCORRECTED PROOF

$$P = \begin{bmatrix} 6.8325 & 4.2868 \\ 4.2868 & 4.1507 \end{bmatrix}, \quad K = [-2.9182 \quad -2.7677], \tag{32}$$

and $\rho = 0.9055$, $\gamma = 55$. These quantities are obtained offline using the sector-bounded polytopic LMI construction discussed in Remark 2, with the sector bound associated with the $\sin(x_1)$ nonlinearity.

For (10) let $h_t(\xi) := \left(\sum_{j=1}^{n_x} \left(|\hat{g}_{j,t}(\xi)| + \nu_{j,t}(\xi) \right)^2 \right)^{1/2}$. Interval branch-and-bound is applied over $\Xi = [-7, 7]^2 \times [-2, 2]$. For each current box $\mathcal{Q}_i$, an interval bound $\bar{h}_t(\mathcal{Q}_i) \geq \sup_{\xi \in \mathcal{Q}_i} h_t(\xi)$ is computed. With $\bar{h}_t = \max_i \bar{h}_t(\mathcal{Q}_i)$, $\underline{h}_t = \max_i h_t(c_i)$, where c_i is the center of $\mathcal{Q}_i$, subdivision continues until $\bar{h}_t - \underline{h}_t \leq \dot{o}_{BB}$. The certified value $\bar{b}_t := \bar{h}_t$ is then used in (11) For the numerical implementation, mutual information is accumulated over $N_e = 10$ predicted state–input pairs:

$$\mathcal{A}_t^{\text{num}}(\chi_t) := \sum_{k=0}^{N_e-1} \mathcal{I}_t(\xi_{k|t}) + \lambda_v \frac{\bar{v}_t - \bar{v}_{t+1|t}\left(\xi_t^{\text{obs}}(\chi_t)\right)}{\max\{\bar{v}_t, \epsilon_v\}}, \qquad N_e = 10 \tag{33}$$

The certificate-change term is evaluated at the next implemented observation. This numerical extension modifies only the exploration objective and leaves the robust feasible set unchanged. The exploration search uses 21 candidate inputs. The proposed method uses $\lambda_v = 1200$, while the proposed method without the certificate term sets $\lambda_v = 0$; all remaining settings are identical.

The simulation is divided into an initial learning phase and a subsequent common tracking phase to separate the effect of informative data collection from the later regulation task. The initial state is $x_0 = \text{col}(0,0)$. To isolate the effect of active exploration, the exploration budget is nonzero only during the first $30s$:

$$\varepsilon_t = \begin{cases} 5(0.997)^t, & 0 \leq tT_s < 30s, \\ 0, & tT_s \geq 30s. \end{cases}$$

During this phase, the reference is fixed at the origin. Consequently, a purely regulation-oriented controller provides limited excitation, whereas the exploration-aware controllers can actively collect informative data.

After $30s$, all controllers track the same time-varying reference. Let $\tau = tT_s$. For $30 \leq \tau \leq 90s$,

$$r_1(\tau) = 7.5\sin\left(2\pi(0.0125)(\tau - 30)\right)$$
$$r_2(\tau) = 7.5\, 2\pi(0.0125) \times \cos\left(2\pi(0.0125)(\tau - 30)\right)$$

Before $30s$, $r(\tau) = \text{col}(0,0)$. The amplitude of r_1 exceeds the hard constraint $|x_1| \leq 7$, providing a direct test of hard constraint satisfaction, while r_2 is consistent with the kinematic relation $\dot{r}_1 = r_2$.

During the moving-reference phase, only the regulation objective is replaced by the tracking cost

$$J_t^{\text{trk}} := \sum_{k=0}^{N-1} \left(\begin{matrix} \| z_{k|t} - r_{t+k} \|_Q^2 \\ + \| v_{k|t} - u_{t+k}^r \|_R^2 \end{matrix} \right) + \| z_{N|t} - r_{t+N} \|_P^2 \tag{34}$$

where u_t^r is obtained from the nominal dynamics. The RSGP update, tube construction, and robust constraint handling are otherwise unchanged.

Four controllers are compared. The *proposed method* uses both mutual information and the certificate-change term. The *method without the certificate term* sets $\lambda_v = 0$ while retaining active information seeking. The *method without information seeking* performs the same online RSGP updates but applies only the baseline safe controller. Finally, the *robust NMPC without learning* retains its initial RSGP model and uncertainty description throughout the simulation. These comparisons isolate the contributions of certificate-aware exploration, active information seeking, and continual online learning, respectively. All methods use the same constraints, initial data, disturbance realization, horizon, tube parameters, and regulation weights.

The initial root-mean-square error (RMSE) is evaluated before the first online update, whereas RMSE30− is evaluated immediately before the moving-reference phase.

Figures 1and 2 show that the certificate-aware term produces a clear learning benefit without sacrificing hard-constraint satisfaction. At 30− s, the proposed method reduces the validation RMSE by 83.43%, compared with 66.79% for the proposed method without the certificate term. The corresponding RMSE values are 0.000276 and 0.000552. At the end of the 90 s simulation, the proposed method achieves an RMSE of 0.000225, compared with 0.000452 without the certificate term.

The certificate-related quantities exhibit the same trend. Initially, $\bar{g}_0 = 0.036127$ and $s_0 = 0.871890$. At 30 s, the proposed method gives $\bar{g}_{30} = 0.032198$ and $s_{30} = 0.779031$, corresponding to reductions of 10.87% and 10.65%, respectively. Without the certificate term, $\bar{g}_{30} = 0.035405$ and $s_{30} = 0.854507$, corresponding to reductions of 2.00% and 1.99%. Thus, the certificate-aware objective changes the selected informative trajectory in a way that accelerates contraction of both the nested residual envelope and the resulting rigid tube.

To assess statistical significance, the 30 s experiment was repeated over 20 independent paired realizations. The proposed method achieves an RMSE reduction of 81.60%±3.27%, compared with 65.87%±0.86% without the certificate term. The mean paired improvement is 15.74 percentage points with a 95% confidence interval of [14.17,17.30] percentage points. A paired t test gives

UNCORRECTED PROOF

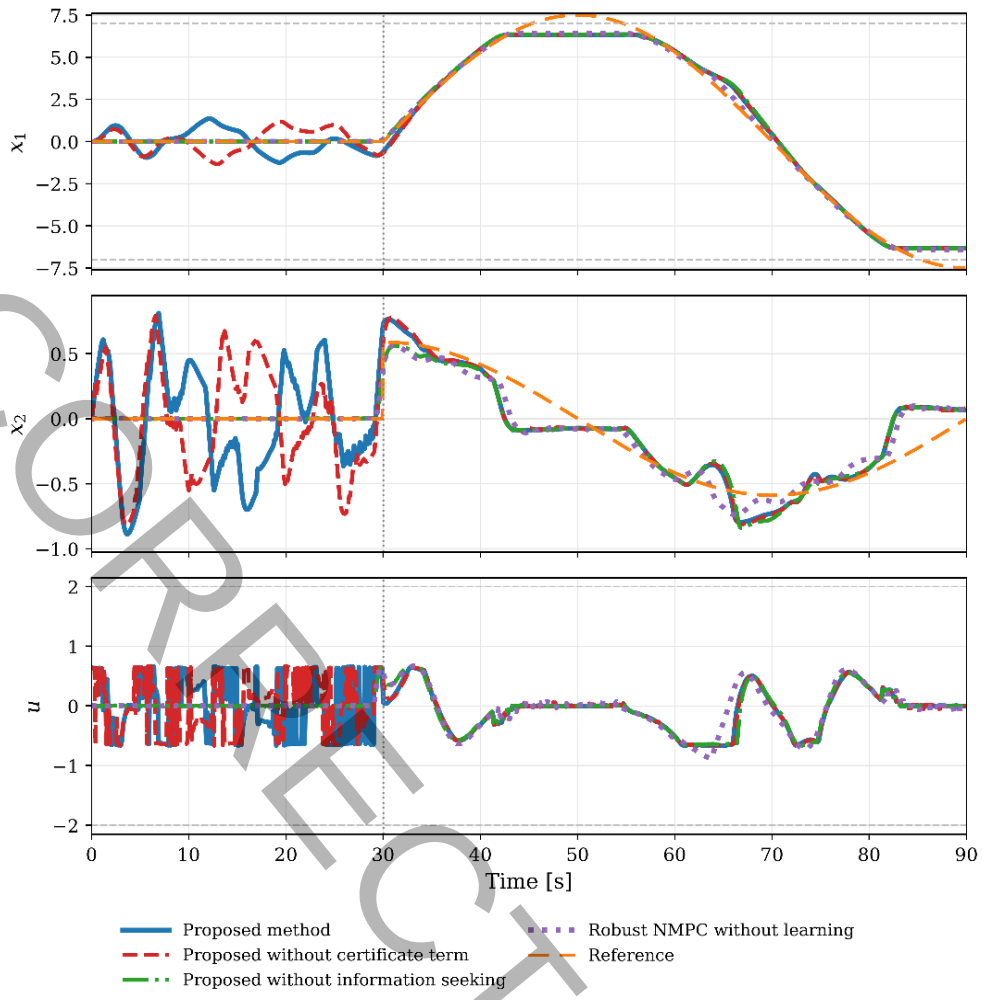


Fig. 1. Closed-loop responses of the four controllers. The panels show x_1, x_2, and the applied input, respectively. The vertical marker at 30 s indicates the end of active exploration. The requested x_1 trajectory exceeds the hard state constraint near its extrema.

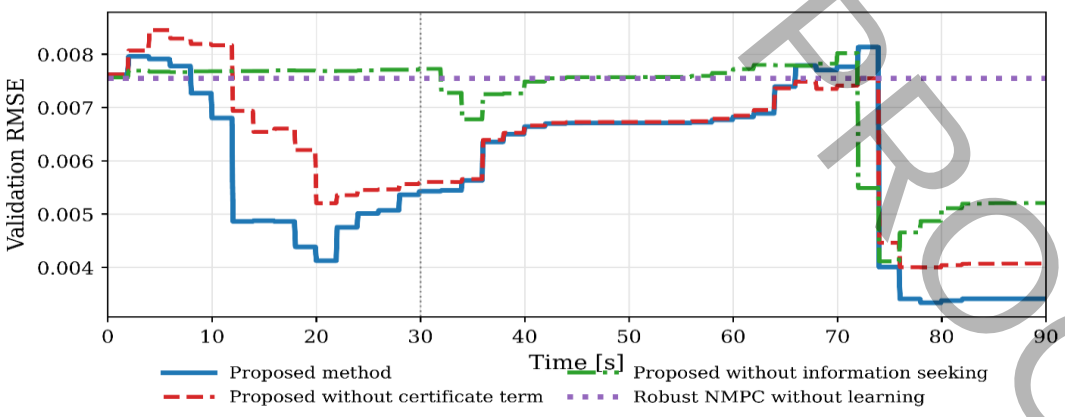


Fig. 2. Evolution of the model RMSE. The vertical marker at 30 s denotes the end of active exploration and the beginning of the moving-reference phase.

UNCORRECTED PROOF

Table 1. Learning and closed-loop comparison. The 30− s values are evaluated immediately before the moving-reference phase.

λ_v	RMSE	$\bar{g}_{30}$ (Δ[%])	s_{30} (Δ[%])
0	0.000552	0.035405 (2.00)	0.854507 (1.99)
600	0.000393	0.033658 (6.83)	0.814166 (6.62)
900	0.000360	0.033873 (6.24)	0.817988 (6.18)
1000	0.000287	0.033161 (8.21)	0.803642 (7.83)
1200	0.000276	0.032198 (10.87)	0.779031 (10.65)

$p = 1.22 \times 10^{-14}$, confirming a statistically significant improvement.

The proposed method, the method without the certificate term, and the method without information seeking achieve tracking RMSEs of 0.275, 0.277, and 0.261, respectively, while robust NMPC without learning achieves 0.261. Thus, the main benefit of active exploration is improved model accuracy, while all controllers satisfy the hard constraints.

7- 2- Effect of the Certificate Weight $\ddot{e}_i$

To isolate the effect of the certificate-aware weight, the same revised nonlinear setup is simulated with all quantities fixed except λ_v. Table 2 reports the validation RMSE, nested residual bound, and tube radius after the 30 s active-learning phase.

The sensitivity study shows that the certificate-aware term affects both learning accuracy and the safety-relevant uncertainty quantities. With $\lambda_v = 0$, the nested residual envelope and tube radius contract by about 2% during active learning. Increasing the weight redirects exploration toward observations predicted to reduce the worst deterministic certificate. In this example, $\lambda_v = 1200$ gives the strongest tested contraction, reducing $\bar{g}_{30}$ by 10.87% and s_{30} by 10.65%, while also producing the lowest 30 s RMSE. The comparison therefore illustrates the intended role of λ_v: it trades statistical informativeness against predicted reduction of the deterministic safety certificate rather than modifying the robust feasible set itself.

7- 3- Computational Scalability

To examine scalability on a higher-dimensional physical system, we consider a Crazyflie-type quadrotor with $x = [p^T v^T, \eta^T, \omega^T]^T \in \mathrm{R}^{12}$ and $u = [T, \tau^T]^T \in \mathrm{R}^4$, where η and ω denote the Euler angles and body angular rates, respectively. The nominal Newton–Euler dynamics are

$$\dot{p} = v, \qquad \dot{v} = R(\eta)e_3 T/m - g e_3$$
$$\dot{\eta} = E(\eta)\omega, \qquad \dot{\omega} = J^{-1}(\tau - \omega \times J\omega) \tag{35}$$

Table 2. Effect of λ_v after 30 s of active learning.

Metric	Proposed	Without certificate	Without information	Without learning
Initial RMSE	0.001663	0.001663	0.001663	0.001663
RMSE30−	0.000276	0.000552	0.001646	0.001663
Reduction at 30−	83.43%	66.79%	1.05%	0.00%
Final RMSE	0.000225	0.000452	0.001408	0.001663
Final reduction	86.49%	72.83%	15.36%	0.00%
Tracking RMSE	0.275	0.277	0.261	0.261

UNCORRECTED PROOF

Table 3. Computational comparison for the quadrotor benchmark.

Method	Time [ms]	$RMSE_0$	$RMSE_f$	Red. [%]
Proposed	14.446	0.06642	0.04618	30.48
Without certificate	14.011	0.06642	0.05565	16.22
Without information	10.618	0.06642	0.06307	5.05
Without learning	8.365	0.06642	0.06642	0.00

where $R(\eta)$ and $E(\eta)$ are defined in [23]. The model is discretized with $T_s = 0.02$ s using $m = 0.029$ kg and $J = \mathrm{diag}(1.43, 1.43, 2.89) \times 10^{-5}$ kg m^2, while the simulated plant uses $m_{true} = 0.031$ kg and $J_{true} = \mathrm{diag}(1.58, 1.50, 3.08) \times 10^{-5}$ kg m^2. In addition to this parametric mismatch, a synthetic residual represents unmodeled aerodynamic and actuator effects:

$$g(x,u) = F_{\mathrm{rb}}(x,u;m_{\mathrm{true}},J_{\mathrm{true}}) - F_{\mathrm{rb}}(x,u;m,J) + T_s\mathrm{col}\left(0_3, d_v, d_\eta, d_\omega\right)$$

where F_{rb} denotes one discretized step of (35) and $d_\eta = 0.02\tanh(\omega)$. Moreover,

$$d_v = \mathrm{col}\begin{pmatrix} -0.30v_x|v_x| + 0.12\theta\bar{T},\ -0.30v_y|v_y| \\ -0.12\phi\bar{T},\ -0.22v_z|v_z| + 0.14\bar{T}|\bar{T}| \end{pmatrix}$$

$$d_\omega = \mathrm{col}\begin{pmatrix} -0.18p|p| + 0.14\bar{\tau}_\phi|\bar{\tau}_\phi|,\ -0.18q|q| \\ +0.14\bar{\tau}_\theta|\bar{\tau}_\theta|,\ -0.12r|r| + 0.10\bar{\tau}_\psi|\bar{\tau}_\psi| \end{pmatrix}$$

with $\bar{T} = \dfrac{T - mg}{mg}$ and $\bar{\tau} = \mathrm{col}\left(\dfrac{\tau_\phi}{0.005}, \dfrac{\tau_\theta}{0.005}, \dfrac{\tau_\psi}{0.0025}\right)$.

These terms are introduced only to provide smooth state- and input-dependent model mismatch and are not intended as an identified aerodynamic model of the Crazyflie. The resulting system provides a physically motivated nonlinear MIMO benchmark and increases the RSGP input dimension to 16.

For the corresponding tube ingredients, we consider $V_{\delta,q}(x, z) = (x - z)^T P_q (x - z)$, $\kappa_q(x, z, v) = v + K_q(x - z)$, where $P_q \in S_{++}^{12}$, $K_q \in R^{4\times 12}$, and the associated ρ_q and γ_q are offline design quantities. These quantities can be constructed offline using an LPV/LMI representation of the incremental Newton–Euler dynamics over the prescribed flight envelope, following Remark 2. These quantities remain fixed during online operation and are not part of the RSGP recursion or exploration optimization.

The RSGP uses $M = 36$ inducing points with $N = 12$, $N_e = 4$, and five exploration candidates. The reported results are averaged over four disturbance realizations, and the RMSE is evaluated on the same fixed forward-flight validation set for all methods. All runtimes are measured using a Python 3.13 implementation on an Intel Xeon Platinum 8370C processor at 2.80 GHz and are reported as the mean online computation time per sampling instant.

The proposed method reduces the final RMSE by 17.0% relative to certificate-free exploration and by 26.8% relative to passive online learning. Its mean online runtime is approximately 1.73 times that of the no-learning controller, indicating that the improved learning accuracy is obtained with moderate computational overhead.

8- Conclusion and Future Work

This paper developed an exploration-aware NMPC framework for safe online learning of unknown nonlinear residual dynamics using recursive sparse Gaussian processes. A deterministic bounded-noise sparse-GP certificate was combined with a nested residual envelope so that successive noisy measurements cannot enlarge the uncertainty bound used for safety. The resulting envelope was incorporated into a causal rigid-tube construction whose nominal center follows only the known dynamics and whose radius remains fixed during each NMPC optimization. A certificate-aware exploration criterion was then used to select informative inputs while preserving the same robust feasible set as the baseline controller. This construction guarantees hard state and input constraint satisfaction under successive bounded-noise RSGP updates without model-update rejection or a post-update feasibility test. The numerical results demonstrated that active information seeking substantially improves model accuracy while maintaining the prescribed hard constraints, with the certificate-aware term providing an additional improvement over mutual-information-only exploration. The present formulation assumes known RKHS and disturbance bounds, a fixed inducing set, and offline construction of the incremental-stabilizability and terminal ingredients. The fixed-dimensional RSGP keeps the online learning complexity independent of the accumulated data, while certified bound evaluation and exploration remain the

UNCORRECTED PROOF

main computational components as the problem dimension increases. Future work will investigate more scalable certificate evaluation, adaptive sparse representations, and experimental validation on physical platforms.

References

[1] M. Buisson-Fenet, F. Solowjow, and S. Trimpe, "Actively learning Gaussian process dynamics," in Proceedings of the 2nd Conference on Learning for Dynamics and Control, ser. Proceedings of Machine Learning Research, vol. 120. PMLR, 2020, pp. 5–15.

[2] A. Capone, G. Noske, J. Umlauft, T. Beckers, A. Lederer, and S. Hirche, "Localized active learning of Gaussian process state space models," in Proceedings of the 2nd Conference on Learning for Dynamics and Control, ser. Proceedings of Machine Learning Research, vol. 120. PMLR, 2020, pp. 490–499.

[3] J. Schreiter, D. Nguyen-Tuong, M. Eberts, B. Bischoff, H. Markert, and M. Toussaint, "Safe exploration for active learning with Gaussian processes," in Machine Learning and Knowledge Discovery in Databases, ser. Lecture Notes in Computer Science, vol. 9286. Springer, 2015, pp. 133–149.

[4] C. Zimmer, M. Meister, and D. Nguyen-Tuong, "Safe active learning for time-series modeling with Gaussian processes," in Advances in Neural Information Processing Systems, vol. 31, 2018, pp. 2730–2739.

[5] C.-Y. Li, B. Rakitsch, and C. Zimmer, "Safe active learning for multi-output Gaussian processes," in Proceedings of the 25th International Conference on Artificial Intelligence and Statistics, ser. Proceedings of Machine Learning Research, vol. 151. PMLR, 2022, pp. 4512–4551.

[6] A. K. Akametalu, S. Kaynama, J. F. Fisac, M. N. Zeilinger, J. H. Gillula, and C. J. Tomlin, "Reachability-based safe learning with Gaussian processes," in 53rd IEEE Conference on Decision and Control. IEEE, 2014, pp. 1424–1431.

[7] F. Berkenkamp, R. Moriconi, A. P. Schoellig, and A. Krause, "Safe learning of regions of attraction for uncertain, nonlinear systems with Gaussian processes," in 2016 IEEE 55th Conference on Decision and Control (CDC). IEEE, 2016, pp. 4661–4666.

[8] F. Berkenkamp, M. Turchetta, A. P. Schoellig, and A. Krause, "Safe model-based reinforcement learning with stability guarantees," in Advances in Neural Information Processing Systems, vol. 30, 2017, pp. 908–918.

[9] T. Koller, F. Berkenkamp, M. Turchetta, and A. Krause, "Learning-based model predictive control for safe exploration," in 2018 IEEE Conference on Decision and Control (CDC). IEEE, 2018, pp. 6059–6066.

[10] D. Sun, M. J. Khojasteh, S. Shekhar, and C. Fan, "Uncertainty-aware safe exploratory planning using Gaussian process and neural control contraction metric," in Proceedings of the 3rd Conference on Learning for Dynamics and Control, ser. Proceedings of Machine Learning Research, vol. 144. PMLR, 2021, pp. 728–741.

[11] T. Lew, A. Sharma, J. Harrison, A. Bylard, and M. Pavone, "Safe active dynamics learning and control: A sequential exploration–exploitation framework," IEEE Transactions on Robotics, vol. 38, no. 5, pp. 2888–2907, 2022.

[12] F. Castañeda, J. J. Choi, W. Jung, B. Zhang, C. J. Tomlin, and K. Sreenath, "Recursively feasible probabilistic safe online learning with control barrier functions," IEEE Open Journal of Control Systems, vol. 4, pp. 531–548, 2025.

[13] J. Umlauft, T. Beckers, A. Capone, A. Lederer, and S. Hirche, "Smart forgetting for safe online learning with Gaussian processes," in Proceedings of the 2nd Conference on Learning for Dynamics and Control, ser. Proceedings of Machine Learning Research, vol. 120. PMLR, 2020, pp. 160–169.

[14] A. Lederer, A. J. Ordóñez Conejo, K. A. Maier, W. Xiao, J. Umlauft, and S. Hirche, "Gaussian process-based real-time learning for safety-critical applications," in Proceedings of the 38th International Conference on Machine Learning, ser. Proceedings of Machine Learning Research, vol. 139. PMLR, 2021, pp. 6055–6064.

[15] M. Schürch, D. Azzimonti, A. Benavoli, and M. Zaffalon, "Recursive estimation for sparse Gaussian process regression," Automatica, vol. 120, p. 109127, 2020.

[16] E. T. Maddalena, P. Scharnhorst, and C. N. Jones, "Deterministic error bounds for kernel-based learning techniques under bounded noise," Automatica, vol. 134, p. 109896, 2021.

[17] P. Scharnhorst, E. T. Maddalena, Y. Jiang, and C. N. Jones, "Robust uncertainty bounds in reproducing kernel Hilbert spaces: A convex optimization approach," IEEE Transactions on Automatic Control, vol. 68, no. 5, pp. 2848–2861, 2023.

[18] J. Wakulicz, K. M. B. Lee, C. Yoo, T. Vidal-Calleja, and R. Fitch, "Informative planning for worst-case error minimisation in sparse Gaussian process regression," in 2022 International Conference on Robotics and Automation (ICRA). IEEE, 2022, pp. 11 066–11 072.

[19] F. A. Bayer, M. Bürger, and F. Allgöwer, "Discrete-time incremental ISS: A framework for robust nmpc," in 2013 European Control Conference (ECC), 2013, pp. 2068–2073.

[20] J. Köhler, M. A. Müller, and F. Allgöwer, "Robust output feedback model predictive control using online estimation bounds," 2021.

[21] I. R. Manchester and J.-J. E. Slotine, "Control contraction metrics: Convex and intrinsic criteria for nonlinear feedback design," IEEE Transactions on Automatic Control, vol. 62, no. 6, pp. 3046–3053, 2017.

[22] A. Sasfi, M. N. Zeilinger, and J. Köhler, "Robust

UNCORRECTED PROOF

adaptive MPC using control contraction metrics," Automatica, vol. 155, p. 111169, 2023.

[23] H. Elkholy and M. K. Habib, "Dynamic modeling and control techniques for a quadrotor," in Handbook of Research on Advancements in Robotics and Mechatronics. IGI Global, 2014, pp. 408–454.

UNCORRECTED PROOF

HOW TO CITE THIS ARTICLE
A. Salmanpour, H. Kebriaei, Safe Active Learning of Nonlinear Dynamics with Exploration-Aware Recursive Sparse Gaussian Process, AUT J. Model. Simul., 58(2) (2026) 189-202.

DOI: 10.22060/miscj.2026.26499.5530

UNCORRECTED PROOF