

Learning-based Observer for Coupled Disturbance

Jindou Jia^{1,2,3,*}, Meng Wang^{1,2,*}, Zihan Yang^{1,2}, Bin Yang^{1,2}, Yuhang Liu^{1,2}, Kexin Guo^{1,2,†}, and Xiang Yu^{1,2}

Abstract—Achieving high-precision control for robotic systems is hindered by the low-fidelity dynamical model and external disturbances. Especially, the intricate coupling between internal uncertainties and external disturbances further exacerbates this challenge. This study introduces an effective and convergent algorithm enabling accurate estimation of the coupled disturbance via combining control and learning philosophies. Concretely, by resorting to *Chebyshev* series expansion, the coupled disturbance is effectively decomposed into an unknown parameter matrix and two known structures dependent on system state and external disturbance respectively. A regularized least squares process is subsequently formalized to learn the parameter matrix using historical time-series data. Furthermore, a polynomial disturbance observer is specifically devised to achieve a high-precision estimation of the coupled disturbance by utilizing the learned structure portion. Extensive simulations and real flight tests valid the effectiveness of the proposed framework. We believe this work can offer a new pathway to integrate learning approaches into control frameworks for addressing longstanding challenges in robotic applications.

I. INTRODUCTION

High-precision control is crucial for robotic systems where states-related model uncertainties and external factors-conditioned disturbances are pervasive. A plethora of estimation schemes have been developed regarding model uncertainties and external disturbances as a lumped disturbance [1]–[3]. Few studies have penetrated the difficult-to-model coupling between internal states and external factors to reduce control conservativeness.

In the control area, focus on the coupled disturbance (depending not only on external disturbance but also internal model uncertainties), many classical methods attempt to estimate or attenuate it with a bounded (or bounded derivative) assumption, such as \mathcal{L}_1 adaptive observer [4], [5], extended state observer (ESO) [6], nonlinear disturbance observer (NDO) [7]–[9], composite adaptive control [1], [10], sliding mode control [11], and \mathcal{H}_∞ control [12]. The bounded assumption on coupled disturbances has explicit limitations from a theoretical perspective. A causality dilemma arises that the state should be preemptively restricted within a bound whose upper limit may be unreasonable, before proving the stability [13]. Moreover, this assumption usually results in a bounded estimation error [7]. A smaller bound of the coupled disturbance (or its derivative) is imposed to

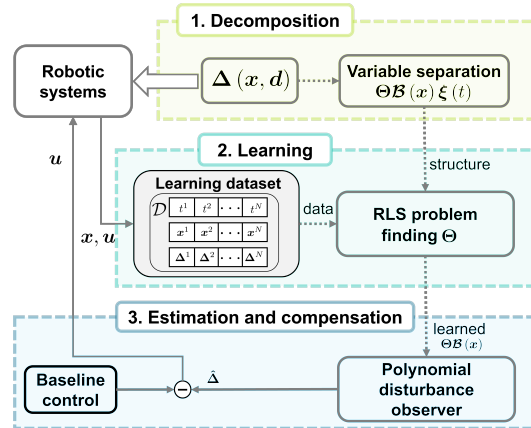


Fig. 1. Overall framework of our proposed learning-based observer. Supplementary Materials: https://anonymous.4open.science/r/Coupled_disturbance-33DF.

enable a small estimation/control error, which may not be satisfied.

Benefiting from the growing computing power, massively available data, and advanced algorithms, data-driven learning approaches appear to be an alternative for handling robotic disturbances. In the data-driven paradigm, most approaches attempt to learn the unknown lumped disturbance with neural networks (NNs) [3], [14], [15]. Focus on the coupled disturbance, a major challenge emerges that the external factor-conditioned disturbances cannot be entirely sampled as learning input (e.g., the real wind speed for quadrotors). Inspired by the time-delay embedding theorem [16], historical states are concentrated as the learning input to reflect unmeasurable factors [17], [18]. With the assistance of meta-learning, several works establish a bi-level optimization to handle coupled wind disturbance [1], [2], [19]. By merging with composite adaptive control [10], the meta-learning scheme can remarkably improve the dynamic performance, but this scheme is limited to specific disturbance and has yet to result in a zero-error estimation theoretically. Moreover, training process for these methods is black-box and labor-intensive.

More recently, the interest in combining control-theoretic approaches with data-driven learning techniques is thriving for achieving stable and high-precision control performance [1], [2], [5], [20]–[25]. For more details refer to Supplementary Materials I-A. Following this routine, we present a convergent estimation framework specifically for the coupled disturbance. Concretely, a learning algorithm is employed to learn the latent invariable parameters offline, while a

*Equal contribution.

†Corresponding author(kxguo@buaa.edu.cn).

¹Beihang University, 100191, Beijing, China. ² Hangzhou Innovation Institute, Beihang University, 310051, Hangzhou, China. ³Nanyang Technological University, 639798, Singapore.

polynomial observer is used to estimate the time-varying part online by utilizing the learned portion. As depicted in Fig. 1, the main contributions are summarized as follows.

Decomposition: A variable decomposition principle (Theorem 1) is established to decompose the coupled disturbance into an unknown parameter matrix, a system-state-related matrix, and an external-disturbance-related vector, with an arbitrarily small residual. The *Chebyshev* polynomial is utilized due to its suitable property. **Learning:** With an analytically derivable assumption on external disturbance, a corollary (Corollary 1) is further developed, which enables the unknown parameter matrix to be learned in a supervised way. Afterward, the learning objective is formalized as a regularized least squares (RLS) problem with a closed-form solution. **Estimation and compensation:** By leveraging the learned knowledge, a polynomial disturbance observer is designed to estimate the final coupled disturbance (Theorem 2), which can improve control performance after compensation.

In the proposed framework, 1) there is no need to manually model complex disturbance, 2) the bounded assumption on the coupled disturbance (or its derivative) in traditional control field [1], [4], [6], [8], [11], [12] can be avoided, and 3) the implemented learning strategy is explainable and lightweight compared with deep neural networks (DNNs) based methods [1]–[3], [14]. Multiple simulations and flight tests (indoor and outdoor) are implemented to verify the proposed scheme.

A. Related Work

As our proposed method falls into the realm of the scheme combining control and data-driven learning, the related advanced research is also reviewed.

1) *Analytical disturbance estimation:* The basic idea of the disturbance estimation approach is to design an ad-hoc observer to estimate the disturbance by utilizing its influence on the system. The estimation method is a two-degree-of-freedom control framework [7], [26], which can achieve tracking and anti-disturbance performances simultaneously. For most disturbance observers, like frequency domain disturbance observer [27], ESO [6], unknown input observer (UIO) [28], generalized proportional integral observer (GPIO) [29], and NDO [8], zero-error estimation can be usually achieved in the event of constant disturbances. For more complicated time-varying disturbances, accurate estimation usually requires *a priori* knowledge of disturbance features. For example, UIO [28] and NDO [30] can accurately estimate the harmonic disturbance if its frequency is known. GPIO [29] and higher-order NDO [31] can achieve an asymptotic estimation of the disturbance represented by a high-order polynomial of time series. More recently, for multi-disturbance with limited priori information, the simultaneous attenuation and compensation approach appears to be a nascent solution [7].

Some researchers attempt to estimate the unknown coupled disturbance with a bounded derivative assumption, such as ESO [6] and NDO [8]. The strict assumption presents an inherent theoretical limitation, as it requires the system

state to be preemptively bounded [13] (Detailed expression is provide in Section II). Moreover, a large derivative bound can result in a large estimation error. In order to avoid the strict premise on system states, a two-stage active disturbance rejection control strategy is designed in [13]. The controller in the first stage guarantees the boundness of the system state by a special auxiliary function, while a linear ESO in the second stage is employed to estimate the total disturbance. However, the existence of the auxiliary function is not discussed.

2) *Combining analytical control and data-driven learning:* Fueled by the advancement of updated learning algorithms, the interest in combining control-theoretic approaches with data-driven learning techniques is thriving for achieving stable and high-precision control performance. In [20], DNNs are utilized to synthesize control certificates such as *Lyapunov* functions and barrier functions, providing formal guarantee of the safety and stability. Moreover, the mass matrix and the potential energy in *Lagrangian* mechanics and *Hamiltonian* mechanics are learned by DNNs [21]. Compared to naive black-box model learning, a more interpretable and plausible model can be obtained from energy conservation principle. Furthermore, an interesting conclusion is found that a higher-order nonlinear system controller by the reinforcement learning (RL) policy behaves like a linear system [32]. The stability of the RL policy can be analyzed by the identified linear closed-loop system with the pole-zero method. [22] combine robust control and echo state networks (ESN) to control nonlinear systems, where ESN is employed to learn the inverse dynamics. Nevertheless, the bound of disturbance and learning output need to be known in advance. [5] develop a \mathcal{L}_1 adaptive observer-argument RL, enabling the quadrotor to track infeasible trajectories in the presence of disturbances.

Even with these advances, for nonlinear systems perturbed by external time-varying disturbances that cannot be accurately sampled, data-driven supervised learning methods would no longer be applicable. Several studies addressing the handling of coupling disturbances have been conducted by formulating a bi-level optimization problem [1], [2]. Within the framework of adaptive control, the nonlinear features depending on the system state are learned via meta-learning offline in [1]. This work breaks through the assumption that the unknown dynamics are linearly parameterizable in the traditional adaptive control method. [2] develop a control-oriented meta-learning framework, which uses the adaptive controller as the base learner to attune learning to the downstream control objective. Both methods attribute the effect of external disturbance in the last layer of the neural networks, which is estimated adaptively online. However, the above scheme ensures zero-error convergence only when the external disturbance is constant. Moreover, laborious offline training is imposed.

In the proposed framework, 1) there is no need to manually model complex disturbance, 2) the bounded assumption on the coupled disturbance (or its derivative) in traditional control field [1], [4], [6], [8], [11], [12] can be avoided,

and 3) the implemented learning strategy is explainable and lightweight compared with deep neural networks (DNNs) based methods [1]–[3], [14]. Multiple simulation and flight tests (indoor and outdoor) are implemented to verify the proposed scheme.

B. Notation

Throughout the paper, \mathbb{R} denotes the real number set; \mathbb{Z}^+ denotes the non-negative integer set; $|x|$ denotes the absolute value of a scalar x ; $\|\mathbf{x}\|$ denotes the 2-norm of a vector \mathbf{x} ; $\|\mathbf{A}\|_F$ denotes the *Frobenius*-norm of a matrix \mathbf{A} ; \mathbf{x}_i denotes the i -th element of a vector \mathbf{x} ; \mathbf{X}_{ij} denotes the i -th row, j -th column element of a matrix \mathbf{X} ; $\mathbf{X}(i, \cdot)$ denotes the i -th row vector of a matrix \mathbf{X} ; \cdot^r denotes a unit right shift operator; $\lambda_m(\cdot)$ represents the minimum eigenvalue of a matrix; \mathbf{I} and $\mathbf{0}$ represent the identity and zero matrices with appropriate sizes, respectively. Moreover, Mean Absolute Error (MAE) is defined as $MAE = \frac{1}{n_d} \sum_{i=1}^{n_d} \|\mathbf{x}_i - \mathbf{x}_{d,i}\|$ to evaluate test results, where n_d denotes the size of collected data, \mathbf{x}_i and $\mathbf{x}_{d,i}$ denote i -th evaluated variable and its desired value, respectively.

II. PROBLEM FORMULATION

Consider a general control affine system of the form

$$\dot{\mathbf{x}} = \mathbf{f}_x(\mathbf{x}) + \mathbf{f}_u(\mathbf{x})\mathbf{u} + \Delta(\mathbf{x}, \mathbf{d}), \quad (1)$$

where $\mathbf{x} \subset \mathcal{X} \in \mathbb{R}^n$ and $\mathbf{u} \subset \mathcal{U} \in \mathbb{R}^o$ denote the state and the control input, respectively; \mathcal{X} and \mathcal{U} are the state and control spaces of dimensionality n and o , respectively; \mathcal{X} includes the unstable set; $\mathbf{f}_x(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ and $\mathbf{f}_u(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times o}$ are nonlinear mappings, which are continuously differentiable; $\Delta(\cdot) : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$ represents the coupled disturbance, which is analytic. Δ depends on known system state \mathbf{x} and unknown external disturbance $\mathbf{d} \subset \mathcal{D} \in \mathbb{R}^m$ with \mathcal{D} being the disturbance space of dimensionality m . Supplementary Materials I-B provides the used notations.

Problem Statement: Consider system (1). The objective is to develop an algorithm to accurately estimate the coupled disturbance $\Delta(\mathbf{x}, \mathbf{d})$ only using control input \mathbf{u} and measurable state \mathbf{x} .

Previous works [1], [2], [5], [20]–[22] estimate the coupled disturbance Δ with a bounded or bounded derivative assumption, i.e., there exists an unknown positive value $\gamma \in \mathbb{R}$ such that

$$\|\Delta(\mathbf{x}, \mathbf{d})\| \leq \gamma \quad \text{or} \quad \|\dot{\Delta}(\mathbf{x}, \mathbf{d})\| \leq \gamma. \quad (2)$$

Three limitations exist in the assumption. **(L-1)** Logical paradox: the bounded assumption on Δ actually requires the system state \mathbf{x} to be preemptively bounded before proving the stability of the system [13]. **(L-2)** Dynamic change: the evolution of \mathbf{x} will change after the estimated disturbance is compensated. There is no guarantee that this assumption based on previous experience will always be satisfied. Consider a scenario where \mathbf{x} constitutes the denominator of Δ . **(L-3)** Estimation conservativeness: the final disturbance estimation error usually depends on γ , which may be large.

Our Solution: The core idea is to 1) decompose the coupled disturbance $\Delta(\mathbf{x}, \mathbf{d})$ into an unknown parameter matrix, a \mathbf{x} -related matrix, and a \mathbf{d} -related vector, with an arbitrarily small residual, 2) offline learn the unknown parameter from historical data, and 3) online estimate the remaining \mathbf{d} -related portion convergently. The whole process is schematized in Fig. 1 and Algorithm 1. By resorting to the proposed algorithm, the limitations **(L-1)**–**(L-3)** are expected to be circumvented.

Algorithm 1 Estimating coupled disturbance via the learning-based observer.

Input:

Collect training dataset $\mathcal{D}_{tra} = \{t_f^n, \mathbf{x}^n, \Delta^n\}$; set parameter δ ;

Solve the RLS problem (6);

return Output parameter matrix Θ ;

Result:

Set observer gain $\Lambda_h(\mathbf{x})$;

while running **do**

Sample current feedback $\{\mathbf{x}, \mathbf{u}\}$;

Solve state observer gain Γ_h for time-varying system (10);

Run polynomial disturbance observer (12);

return Estimated coupled disturbance $\hat{\Delta}$;

end while

III. METHOD

A. Decomposition of the coupled disturbance

Before introducing the variable decomposition theorem for the coupled disturbance $\Delta(\mathbf{x}, \mathbf{d})$, a preliminary lemma from [1, Theorem 3] is reviewed for a scalar $\Delta_i(\mathbf{x}, \mathbf{d})$ firstly.

Lemma 1: [1, Theorem 3] Assume an analytic function $\Delta_i(\mathbf{x}, \mathbf{d}) : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ for all $[\mathbf{x}, \mathbf{d}] \in \mathcal{X} \times \mathcal{D}$. For any small value $\epsilon > 0$, there always exist $p = O\left(\frac{\log(1/\epsilon)}{\sqrt{n+m}}\right) \in \mathbb{Z}^+$, $\phi_i(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{R}^{1 \times s}$ consisting of *Chebyshev* polynomials and unknown constant parameters, $\xi(\mathbf{d}) \in \mathbb{R}^s$ consisting only of *Chebyshev* polynomials such that

$$\sup_{[\mathbf{x}, \mathbf{d}] \in \mathcal{X} \times \mathcal{D}} |\Delta_i(\mathbf{x}, \mathbf{d}) - \phi_i(\mathbf{x})\xi(\mathbf{d})| \leq \epsilon, \quad (3)$$

and $s = (p+1)^m = O(\log(1/\epsilon)^m)$.

Note that Lemma 1 is initially build on $[\bar{\mathbf{x}}, \bar{\mathbf{d}}] \in [-1, 1]^n \times [-1, 1]^m$ in [1, Theorem 3]. By normalization, its result can be generalized to $[\mathbf{x}, \mathbf{d}] \in \mathcal{X} \times \mathcal{D}$. Lemma 1 concludes that the analytic coupled disturbance can be decoupled to a \mathbf{x} -related portion and a \mathbf{d} -related portion with an arbitrarily small residual. The \mathbf{x} -related portion remains unchanged in operation. Intuitively, it will be helpful to estimate the \mathbf{d} -related portion if the knowledge of \mathbf{x} -related portion can be exploited beforehand. In [1], [2], DNNs are adopted to learn the \mathbf{x} -related portion, which needs laborious offline training and lacks interpretability. A more lightweight learning strategy is pursued here. To achieve that, we need to exploit Lemma 1 to drive a more explicit decomposition form.

Theorem 1: $\Delta_i(\mathbf{x}, \mathbf{d})$ is a function satisfying the assumptions in Lemma 1, for all $i \in [1, 2, \dots, n]$. For any small value $\epsilon' > 0$, there always exist $s_1 \in \mathbb{Z}^+$; $s_2 \in \mathbb{Z}^+$; an unknown constant parameter matrix $\Theta \in \mathbb{R}^{n \times s_1}$, two functions $\mathcal{B}(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{R}^{s_1 \times s_2}$ and $\xi(\mathbf{d}) : \mathbb{R}^m \rightarrow \mathbb{R}^{s_2}$ that both consist only of *Chebyshev* polynomials such that

$$\sup_{[\mathbf{x}, \mathbf{d}] \in \mathcal{X} \times \mathcal{D}} \|\Delta(\mathbf{x}, \mathbf{d}) - \Theta \mathcal{B}(\mathbf{x}) \xi(\mathbf{d})\| \leq \epsilon', \quad (4)$$

where $s_1 = (p+1)^{m+n} = O(\log(\sqrt{n}/\epsilon')^{m+n})$ and $s_2 = (p+1)^m = O(\log(\sqrt{n}/\epsilon')^m)$.

Proof: See Supplementary Materials V-A. ■

Remark 1: Theorem 1 extends the result of Lemma 1 to the multidimensional case and obtains a more explicit decomposed structure. It is proven that all unknown constant parameters of the coupled disturbance can be gathered into a matrix, which enables the coupled disturbance to be learned in a more explainable way compared with DNNs-based methods [1]–[3].

In comparison with the existing approaches [1], [2], [5], [20]–[22] requiring bounded or bounded derivative assumption, the only premise of our framework is that $\mathbf{d}(t)$ is analytic. The premise can be readily fulfilled by many disturbances in real-world. Therefore, a theoretical foundation for the reliable implementation of an explicit learning process is achieved. Subsequently, the following corollary is given.

Corollary 1: $\Delta_i(\mathbf{x}, \mathbf{d}(t))$ is a function satisfying the assumptions in Lemma 1, for all $i \in [1, 2, \dots, n]$. Assume $\mathbf{d}(t)$ is analytic with respect to t . For any small value $\epsilon' > 0$, there always exist $s_1 \in \mathbb{Z}^+$; $s_2 \in \mathbb{Z}^+$; an unknown constant parameter matrix $\Theta \in \mathbb{R}^{n \times s_1}$, two functions $\mathcal{B}(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{R}^{s_1 \times s_2}$ and $\xi(t) : \mathbb{R} \rightarrow \mathbb{R}^{s_2}$ that consist only of *Chebyshev* polynomials such that

$$\sup_{[\mathbf{x}, t] \in \mathcal{X} \times \mathcal{D}} \|\Delta(\mathbf{x}, \mathbf{d}(t)) - \Theta \mathcal{B}(\mathbf{x}) \xi(t)\| \leq \epsilon', \quad (5)$$

with $s_1 = (p+1)^{n+1} = O(\log(\sqrt{n}/\epsilon')^{n+1})$, $s_2 = p+1 = O(\log(\sqrt{n}/\epsilon'))$, and $\xi(t) = [T_0(t) \ T_1(t) \ \dots \ T_p(t)]^T$, where T_i represents the i -th order *Chebyshev* polynomial.

Proof: The proof procedure is similar to Theorem 1 by replacing the argument \mathbf{d} with t . ■

Remark 2: Based on Theorem 1, Corollary 1 further resorts the unsamplable external disturbance \mathbf{d} to the samplable feature t , which allows the unknown parameter matrix to be learned in a supervised way. The external disturbance \mathbf{d} is ultimately decomposed into parameters in Θ and t -related structure $\xi(t)$. A practical example is provided in Supplementary Materials V-B to instantiate Corollary 1.

B. Learning the parameter matrix

In this part, an RLS optimization algorithm is established to learn parameter matrix Θ . Construct the training dataset $\mathcal{D}_{tra} = \left\{ \left(t_f^n, \mathbf{x}^n, \Delta^n \right) \mid n = 1, 2, \dots, N \right\}$ with N samples. t_f represents the time in the offline training dataset.

Note that Δ^n can be calculated by using $\Delta^n = \dot{\mathbf{x}}^n - \mathbf{f}_x(\mathbf{x}^n) - \mathbf{f}_u(\mathbf{x}^n) \mathbf{u}^n$, where $\dot{\mathbf{x}}^n$ can be obtained by offline high-order polynomial fitting.

Denote Θ^* as the optimal solution under the dataset \mathcal{D}_{tra} . The learning objective is formalized as

$$\Theta^* = \arg \min_{\Theta} \frac{1}{2} \left[\sum_{n=1}^N \|\Delta^n - \Theta \mathcal{B}_n \xi_n\|^2 + \delta \|\Theta\|_F^2 \right], \quad (6)$$

where $\mathcal{B}_n := \mathcal{B}(\mathbf{x}^n)$, $\xi_n := \xi(t_f^n)$, and δ regularizes Θ . Furthermore, the closed-form solution of (6) is given as:

$$\Theta^* = \sum_{n=1}^N \Delta^n \xi_n^T \mathcal{B}_n^T \cdot \left(\sum_{n=1}^N \mathcal{B}_n \xi_n \xi_n^T \mathcal{B}_n^T + \delta \mathbf{I} \right)^{-1}. \quad (7)$$

Until now, the \mathbf{x} -related portion of $\Delta(\mathbf{x}, \mathbf{d})$ has been separated and the unknown constant parameters Θ can be learned from the historical data.

Denote t_l as the online time. Note that the offline learning phase and online estimating process are in different time domains. In other words, the offset between t_f and t_l is unknown. Therefore Δ cannot be directly obtained by $\Theta \mathcal{B}(\mathbf{x}) \xi(t)$ online. Moreover, update $\xi(t)$ using feedback state can also improve the generalization to other disturbances. The remaining issue is to estimate $\xi(t)$ online from control input \mathbf{u} and measured \mathbf{x} . By resorting to the polynomial disturbance observer to be designed, $\xi(t)$ can be exponentially estimated.

C. Estimation via a polynomial disturbance observer

Before proceeding, $\xi(t)$ in Corollary 1 can be further decomposed due to the structure of *Chebyshev* polynomial (Eq. (20)). It can be rendered that

$$\xi(t) = \mathcal{H} \varsigma(t), \quad (8)$$

where $\mathcal{H} \in \mathbb{R}^{s_2 \times s_2}$,

$$\mathcal{H}(i, :) = \begin{cases} \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix}, & i = 1, \\ \begin{bmatrix} 0 & 1 & \dots & 0 \end{bmatrix}, & i = 2, \\ 2\vec{\mathbf{H}}(i-1, :) - \mathbf{H}(i-2, :), & 2 < i \leq s_2, \end{cases}$$

$\mathbf{H}(i, :)$ denotes the i -th row vector of \mathbf{H} , $\vec{\cdot}$ denotes a unit right shift operator, and $\varsigma(t)$ consists of polynomial basis functions, i.e.,

$$\varsigma(t) = [1 \ t \ \dots \ t^p]^T \in \mathbb{R}^{s_2}. \quad (9)$$

From (5) and (8), the coupled disturbance $\Delta(\mathbf{x}, \mathbf{d})$ is finally represented as

$$\begin{cases} \dot{\zeta}(t) = \mathcal{A} \zeta(t), \\ \Delta = \Theta \mathcal{B}(\mathbf{x}) \mathcal{H} \varsigma(t), \end{cases} \quad (10)$$

where $\mathcal{A} \in \mathbb{R}^{s_2 \times s_2}$,

$$\mathcal{A}(i, j) = \begin{cases} j, & i = j + 1, \\ 0, & i \neq j + 1. \end{cases}$$

Define $\hat{\zeta}$ and $\tilde{\zeta}$ as the estimation of $\zeta(t)$ and the estimation error $\tilde{\zeta} = \zeta(t) - \hat{\zeta}$, respectively. The expected objective of the subsequent disturbance observer is to achieve

$$\dot{\hat{\zeta}} = \Lambda_h(\mathbf{x}) \tilde{\zeta}, \quad (11)$$

where $\Lambda_h(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{R}^{s_2 \times s_2}$ denotes the observer gain matrix which is designed to ensure the error dynamics (11) exponentially stable. To achieve (11), the observer is designed as

$$\begin{cases} \dot{\mathbf{z}}_h = \mathcal{A}\hat{\boldsymbol{\zeta}} - \Gamma_h(\mathbf{f}_x(\mathbf{x}) + \mathbf{f}_u(\mathbf{x})\mathbf{u} + \Theta\mathcal{B}(\mathbf{x})\mathcal{H}\hat{\boldsymbol{\zeta}}), \\ \hat{\boldsymbol{\zeta}} = \mathbf{z}_h + \Gamma_h\mathbf{x}, \\ \hat{\Delta} = \Theta\mathcal{B}(\mathbf{x})\mathcal{H}\hat{\boldsymbol{\zeta}}, \end{cases} \quad (12)$$

where $\mathbf{z}_h \in \mathbb{R}^{s_2}$ is an auxiliary variable and $\Gamma_h \in \mathbb{R}^{s_2 \times n}$ is designed such that $\Lambda_h(\mathbf{x}) = \mathcal{A} - \Gamma_h\Theta\mathcal{B}(\mathbf{x})\mathcal{H}$.

Theorem 2: Consider the nonlinear system (1). Under the designed polynomial disturbance observer (12), the estimation error $\hat{\Delta}$ will converge to zero exponentially if Γ_h can be chosen to make the error dynamics

$$\dot{\boldsymbol{\zeta}} = (\mathcal{A} - \Gamma_h\Theta\mathcal{B}(\mathbf{x})\mathcal{H})\boldsymbol{\zeta} \quad (13)$$

exponentially stable.

Proof: See Supplementary Materials V-C. ■

According to (13), the design of Γ_h for polynomial disturbance observer (12) is equivalent to the design of the state observer gain for the disturbance system (10). The methods of state observer design for linear time-varying systems have been developed in many previous works, such as the least-squares-based observer [33], the extended linear observer [34], and the block-input/block-output model-based observer [35]. Due to the convenience of adjusting the pole, the method proposed in [34] is employed here to decide Γ_h online. Moreover, the computational burden is mainly concentrated on the inverse of an observable matrix, which is suitable for practical systems. A detailed discussion can refer to Supplementary Materials.

Remark 3: The polynomial disturbance observer (12) utilizes the offline learned knowledge $\Theta\mathcal{B}(\mathbf{x})$ and real-time $\{\mathbf{u}, \mathbf{x}\}$ to estimate $\boldsymbol{\zeta}(t)$ online. Two advantages can be achieved. One is that the offset between offline time t_f and online time t_l need not be known. The other is that the generalization to other kinds of external disturbance \mathbf{d} can be improved according to real-time feedback (Section IV-C).

IV. EMPIRICAL STUDY

A. Learning performance on nonlinear functions

A supervised learning strategy is synthesized in Section III-B to extract Θ from disturbance Δ . In this part, the learning performance is exemplified and analyzed by following nonlinear functions

$$\begin{cases} \Delta(\mathbf{x}, \mathbf{d}) = \sin(\mathbf{x})\sin(\mathbf{d}), \quad \mathbf{d} = t, \\ \Delta(\mathbf{x}, \mathbf{d}) = \mathbf{x} - \frac{1}{12}\mathbf{x}^3 - \frac{1}{4}\mathbf{d}, \quad \mathbf{d} = t^2, \\ \Delta(\mathbf{x}, \mathbf{d}) = -\frac{1}{9}\sin(\mathbf{x})\mathbf{d}, \quad \mathbf{d} = t^3. \end{cases} \quad (14)$$

a) Setup: The learning dataset is constructed from $\mathbf{x} \in [-2, 2]$ and $t \in [0, 4]$. 10000 samples are collected and scrambled, where 5000 for training and 5000 for testing. The hyperparameter δ used for training is set as 0.01. Moreover, the influences of measurement noise and the selection of p (Corollary 1) are also analyzed. The state \mathbf{x} in the training dataset is corrupted by noise $\mathcal{N}(\mathbf{0}, \sigma_x^2)$. The learning performance under different σ_x^2 and p is tested.

b) Results: The learning errors of three nonlinear functions (14) under different σ_x^2 and p are presented in Fig. 2A-C. Two phenomena can be observed. On the one hand, the learning performance degrades as the noise variance increases. On the other hand, proper p can achieve decent performance, since small and large p can lead to underfitting and overfitting problems, respectively.

B. Estimation performance on a second-order dynamics

In this part, the estimation performance of the proposed polynomial disturbance observer is demonstrated through a toy example for intuitive understanding. Considering a second-order *Newton* system perturbed by a coupled disturbance

$$\begin{cases} \dot{\eta} = v, \dot{v} = a, \\ ma = u + \Delta(v, d), \end{cases} \quad (15)$$

with position $\eta \in \mathbb{R}$, velocity $v \in \mathbb{R}$, acceleration $a \in \mathbb{R}$, mass $m \in \mathbb{R}$, control input $u \in \mathbb{R}$, and coupled disturbance $\Delta(v, d) \in \mathbb{R}$. Note that the measured v is corrupted by noise $\mathcal{N}(0, \sigma_v^2)$. The coupled disturbance used in the simulation is modeled as

$$\Delta(v, d(t)) = -v^2 + 50 - 10t - 0.5t^2. \quad (16)$$

The truth value of Θ of (16) can be derived, i.e., $\Theta = [49.75, 0, -0.5, -10, 0, 0, 0.25, 0, 0]$.

The objective is to design control input u so as to ensure that η tracks the desired state η_d . The baseline controller adopts the proportional-derivative control, and the estimated disturbance $\hat{\Delta}$ by the proposed polynomial disturbance observer is compensated via feedforward, i.e.,

$$u = K_\eta e_\eta + K_v e_v - \hat{\Delta}, \quad (17)$$

with gains $K_\eta \in \mathbb{R}$ and $K_v \in \mathbb{R}$, and tracking errors $e_\eta = \eta_d - \eta$ and $e_v = v_d - v$.

a) Setup: The learning dataset is constructed from $v \in [-10, 10]$ and $t \in [0, 100]$. 10000 samples are collected. The hyperparameter δ used for training is set as 0.01. p in Theorem 1 is chosen as 2. The desired tracking trajectory is set as $\eta_d = \sin(\frac{1}{2}t)$. The variance σ_v^2 of imposed noise is set as 0.1. The baseline controller gains K_η and K_v are tuned as 10 and 25, respectively. The NDO [8], ESO [6], [36], \mathcal{L}_1 adaptive [5], [37], EVOLVER [38], and the baseline controller (without compensation) are taken as comparisons. For the sake of fairness, the observer gains (in charge of the convergence speed) of the compared NDO, \mathcal{L}_1 adaptive, EVOLVER, and the proposed one are set to be the same. All eigenvalues of $\Lambda_h(\mathbf{x})$ are set as 0.4.

b) Results: Denote $\hat{\Theta}$ as the learning result of Θ . By employing the proposed learning strategy designed in Section III-B, the learning error $\|\Theta - \hat{\Theta}\|$ is 1.1473×10^{-3} , which demonstrates the effectiveness of the proposed learning strategy in Section III-B. As shown in Fig. 3A, the tracking performance of learning-based methods (EVOLVER and proposed one (12)) outperform feedback-based methods (Baseline, NDO [8], \mathcal{L}_1 adaptive [5], [37], ESO [6], [36]).

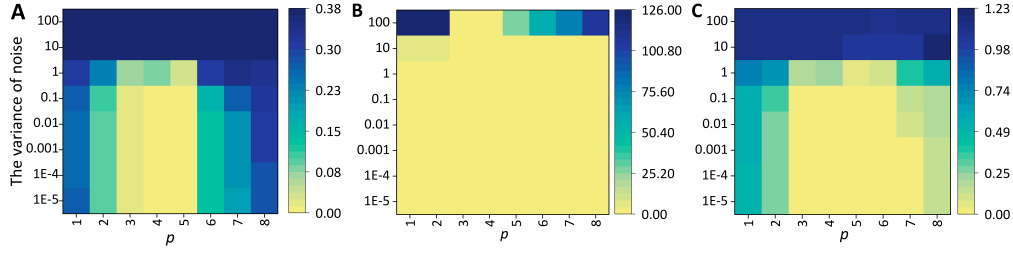


Fig. 2. Learning errors (MAE) of three nonlinear functions (14) under different noise variances σ_x^2 and parameters p on the test dataset.

Focusing on feedback-based methods, there are always estimations lags from Fig. 3B. Compared with EVOLVER, a *Koopman*-based online learning framework, the proposed one can avoid the dataset construction process and allows for a more in-depth learning of the coupled structure on disturbance (16), leading to the smallest tracking and estimation error (Figs. 3A-C).

C. Control performance on a flying drone

The proposed learning-based observer is further verified in real flight tests, which consider a drone flying under multiple disturbances. The conventional cascaded baseline control framework is employed here, where the coupled disturbances are further estimated by the proposed polynomial disturbance observer (12) to enhance the control accuracy.

a) Setup: Three scenarios are implemented indoor, including hovering under wind disturbance (Fig. 4A), circle flight under model uncertainty (Fig. 4B), and circle flight under both model uncertainty and wind disturbance (Fig. 4C). Θ is offline trained under single mass uncertainty with 2500 samples. p in Theorem 1 is chosen as 2. The hyperparameter δ used for training is set as 0.01. Since the mass uncertainty mainly depends on unmeasurable acceleration, inspired by the time-delay embedding theory [16], 9 historical velocities are utilized and reduced to 4 through an autoencoder, which is finally regarded as x of $\Delta(x, d)$. The mass uncertainty in tests is around 1 kg, constituting more than 35% of the entire system. The wind is considered to verify the generalization ability. The maximum wing speed can reach at 5 m/s. The radius and period of the commended circle are set as 1 m and 7 s, respectively. The observer (12) is running on an onboard Intel NUC with 100 HZ.

In the outdoor test, the drone is commanded to follow a circular trajectory in the presence of both mass uncertainty (around 1.2 kg) and natural wind, as illustrated in Fig. 6E. The radius and period of the commanded circle are set to be 1.2 m and 7 s, respectively. The observer gain Γ_h in Eq. (12) is decided online [34] so that the eigenvalues of $\Lambda_h(x)$ are all -0.6 .

b) Results: The control performance of all 3 indoor tests is summarized in Fig. 5. The control errors (MAE) are improved 44.3%, 20.0%, and 35.0%, respectively. The results show that the proposed one can successfully enhance the tracking performance with the compensation of estimated disturbance, even in the face of unseen wind disturbance. Fig. 5G further provides the control and estimation performance

along the X -axis in Test #3, under both mass and wind disturbances. It can be seen the changing trends of tracking error and estimated disturbance are roughly consistent. Moreover, the computation burden is mainly around the solution of gain Γ_h . The computation cost is provided in Fig. 6G.

We further conduct an outdoor flight test under natural wind. As depicted in Fig. 6F, the tracking delay and overshoot issues of the baseline method are mitigated by the proposed algorithm, particularly along the X -axis. Fig. 6D further presents the distributions of the tracking errors, which effectively demonstrates the robustness of the proposed approach in the outdoor scenario.

V. CONCLUSION

In this article, we propose a learning-based disturbance observer for robotic systems perpetuated by a coupled disturbance. The considered coupled disturbance is difficult to model explicitly. Firstly, a variable decomposition principle is presented by leveraging the *Chebyshev* series expansion. An RLS-based learning strategy is subsequently developed to learn the separated unknown parameter matrix using historical data, which maintains a low computational complexity. Finally, the polynomial observer is developed by utilizing the learned structure, which can theoretically achieve a precise estimation of the coupled disturbance. The learning and estimation performance of the proposed method is verified by several simulational and experimental examples.

Although the proposed framework showcases compelling theoretical and practical promise, the convergence of the presented uncertainty estimation algorithm is based on the analytic assumption on d . Therefore, effectively handling the nonanalytic disturbances with a theoretical guarantee, such as contact-induced, step-like scenarios, remains an open challenge that warrants further investigation.

REFERENCES

- [1] M. O’Connell, G. Shi, X. Shi, K. Azizzadenesheli, A. Anandkumar, Y. Yue, and S.-J. Chung, “Neural-fly enables rapid learning for agile flight in strong winds,” *Science Robotics*, vol. 7, no. 66, p. eabm6597, 2022.
- [2] S. M. Richards, N. Azizan, J.-J. Slotine, and M. Pavone, “Control-oriented meta-learning,” *The International Journal of Robotics Research*, vol. 42, no. 10, pp. 777–797, 2023.
- [3] M. Bisheban and T. Lee, “Geometric adaptive control with neural networks for a quadrotor in wind fields,” *IEEE Transactions on Control Systems Technology*, vol. 29, no. 4, pp. 1533–1548, 2021.
- [4] N. Hovakimyan and C. Cao, *\mathcal{L}_1 adaptive control theory: Guaranteed robustness with fast adaptation*. SIAM, 2010.

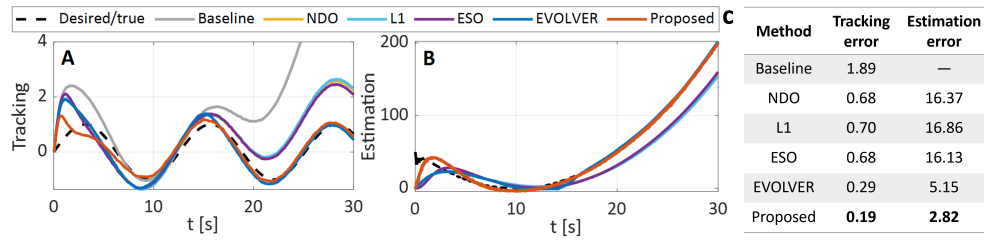


Fig. 3. The tracking and estimation performance of the second-order dynamics example. **A**, The trajectory tracking performance. **B**, The disturbance estimation performance. **C**, MAEs of all compared methods: NDO [8], ESO [6], [36], \mathcal{L}_1 adaptive [5], [37], EVOLVER [38], the baseline controller (without compensation) and the proposed one (12).

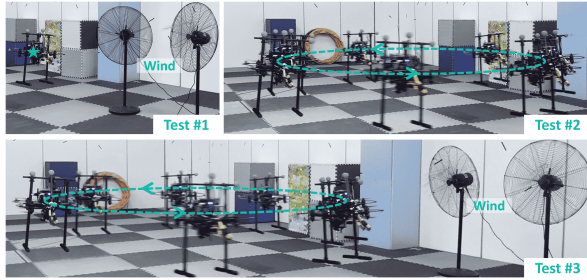


Fig. 4. Experimental scenarios in the drone example. The location state of the drone is provided by the NOKOV motion capture system. **A**, Hovering under wind disturbance. **B**, Circle flight under model uncertainty. **C**, Circle flight under model uncertainty and wind disturbance.

[5] K. Huang, R. Rana, A. Spitzer, G. Shi, and B. Boots, "DATT: Deep adaptive trajectory tracking for quadrotor control," in *7th Annual Conference on Robot Learning*, 2023.

[6] J. Han, "From PID to active disturbance rejection control," *IEEE Transactions on Industrial Electronics*, vol. 56, no. 3, pp. 900–906, 2009.

[7] L. Guo and W. Chen, "Disturbance attenuation and rejection for systems with nonlinearity via DOBC approach," *International Journal of Robust and Nonlinear Control*, vol. 15, no. 3, pp. 109–125, 2005.

[8] W. Chen and L. Guo, "Analysis of disturbance observer based control for nonlinear systems under disturbances with bounded variation," in *Proceedings of International Conference on Control*, 2004, pp. 1–5.

[9] M. Wang, S. Lyu, Q. Liu, Z. Yang, K. Guo, and X. Yu, "Precise end-effector control for an aerial manipulator under composite disturbances: Theory and experiments," *IEEE Transactions on Automation Science and Engineering*, vol. 22, pp. 4006–4021, 2025.

[10] J.-J. E. Slotine and W. Li, "Composite adaptive control of robot manipulators," *Automatica*, vol. 25, no. 4, pp. 509–519, 1989.

[11] J.-M. Yang and J.-H. Kim, "Sliding mode control for trajectory tracking of nonholonomic wheeled mobile robots," *IEEE Transactions on Robotics and Automation*, vol. 15, no. 3, pp. 578–587, 1999.

[12] I. Khalil, J. Doyle, and K. Glover, *Robust and optimal control*. Prentice hall New York, 1996, vol. 2.

[13] H. Wang, Z. Zuo, W. Xue, Y. Wang, and H. Yang, "Switching longitudinal and lateral semi-decoupled active disturbance rejection control for unmanned ground vehicles," *IEEE Transactions on Industrial Electronics*, vol. 71, no. 3, pp. 3034–3043, 2024.

[14] A. Saviolo and G. Loianno, "Learning quadrotor dynamics for precise, safe, and agile flight control," *Annual Reviews in Control*, vol. 55, pp. 45–60, 2023.

[15] M. Wang, Z. Chen, K. Guo, X. Yu, Y. Zhang *et al.*, "Millimeter-level pick and peg-in-hole task achieved by aerial manipulator," *IEEE Transactions on Robotics*, vol. 40, pp. 1242–1260, 2024.

[16] F. Takens, "Detecting strange attractors in turbulence," in *Dynamical Systems and Turbulence*. Springer, 2006, pp. 366–381.

[17] D. A. Haggerty, M. J. Banks, E. Kamenar, A. B. Cao, P. C. Curtis, I. Mezić, and E. W. Hawkes, "Control of soft robots with inertial dynamics," *Science robotics*, vol. 8, no. 81, p. eadd6864, 2023.

[18] S. Lale, P. I. Renn, K. Azizzadenesheli, B. Hassibi, M. Gharib, and

A. Anandkumar, "FALCON: Fourier adaptive learning and control for disturbance rejection under extreme turbulence," *npj Robotics*, vol. 2, no. 1, p. 6, 2024.

[19] F. Xie, G. Shi, M. O'Connell, Y. Yue, and S.-J. Chung, "Hierarchical meta-learning-based adaptive controller," in *Proceedings of IEEE International Conference on Robotics and Automation*, 2024.

[20] C. Dawson, S. Gao, and C. Fan, "Safe control with learned certificates: A survey of neural lyapunov, barrier, and contraction methods for robotics and control," *IEEE Transactions on Robotics*, vol. 39, no. 3, pp. 1749–1767, 2023.

[21] M. Lutter and J. Peters, "Combining physics and deep learning to learn continuous-time dynamics models," *The International Journal of Robotics Research*, vol. 42, no. 3, pp. 83–107, 2023.

[22] A. Banderchuk, D. Coutinho, and E. Camponogara, "Combining robust control and machine learning for uncertain nonlinear systems subject to persistent disturbances," *arXiv preprint arXiv:2303.11890*, 2023.

[23] J. Jia, K. Guo, Y. Wang, S. Zhou, J. Zhang, Y. Liu, X. Yu, Y. Shi, and L. Guo, "FORESEER: Recognize and utilize uncertainties by integrating data-based learning and symbolic feedback," *The International Journal of Robotics Research*, p. 02783649251364000, 2025.

[24] J. Jia, Z. Yang, M. Wang, K. Guo, J. Yang, X. Yu, and L. Guo, "Feedback favors the generalization of neural ODEs," in *Proceedings of International Conference on Learning Representations*, 2025.

[25] J. Jia, G. Li, X. Chen, T. An, Y. Hu, J. Li, X. Guo, and J. Yang, "Action-to-action flow matching," *arXiv preprint arXiv:2602.07322*, 2026.

[26] X. Li, Q. Wang, X. Li, K. Tan, and L. Xie, "Feedforward control with disturbance prediction for linear discrete-time systems," *IEEE Transactions on Control Systems Technology*, vol. 27, no. 6, pp. 2340–2350, 2019.

[27] K. Ohishi, M. Nakao, K. Ohnishi, and K. Miyachi, "Microprocessor-controlled DC motor for load-insensitive position servo system," *IEEE Transactions on Industrial Electronics*, vol. 34, no. 1, pp. 44–49, 1987.

[28] C. Johnson, "Further study of the linear regulator with disturbances—The case of vector disturbances satisfying a linear differential equation," *IEEE Transactions on Automatic Control*, vol. 15, no. 2, pp. 222–228, 1970.

[29] H. Sira-Ramirez and M. A. Oliver-Salazar, "On the robust control of buck-converter DC-motor combinations," *IEEE Transactions on Power Electronics*, vol. 28, no. 8, pp. 3912–3922, 2013.

[30] W. Chen, "Disturbance observer based control for nonlinear systems," *IEEE/ASME Transactions on Mechatronics*, vol. 9, no. 4, pp. 706–710, 2004.

[31] K.-S. Kim, K.-H. Rew, and S. Kim, "Disturbance observer for estimating higher order disturbances in time series expansion," *IEEE Transactions on Automatic Control*, vol. 55, no. 8, pp. 1905–1911, 2010.

[32] Z. Li, J. Zeng, A. Thirugnanam, and K. Sreenath, "Bridging model-based safety and model-free reinforcement learning through system identification of low dimensional linear models," in *Proceedings of Robotics: Science and Systems*, 2022.

[33] M.-S. Chen and J.-Y. Yen, "Application of the least squares algorithm to the observer design for linear time-varying systems," *IEEE Transactions on Automatic Control*, vol. 44, no. 9, pp. 1742–1745, 1999.

[34] K. Busawon and M. Saif, "A state observer for nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 44, no. 11, pp. 2098–2103, 1999.

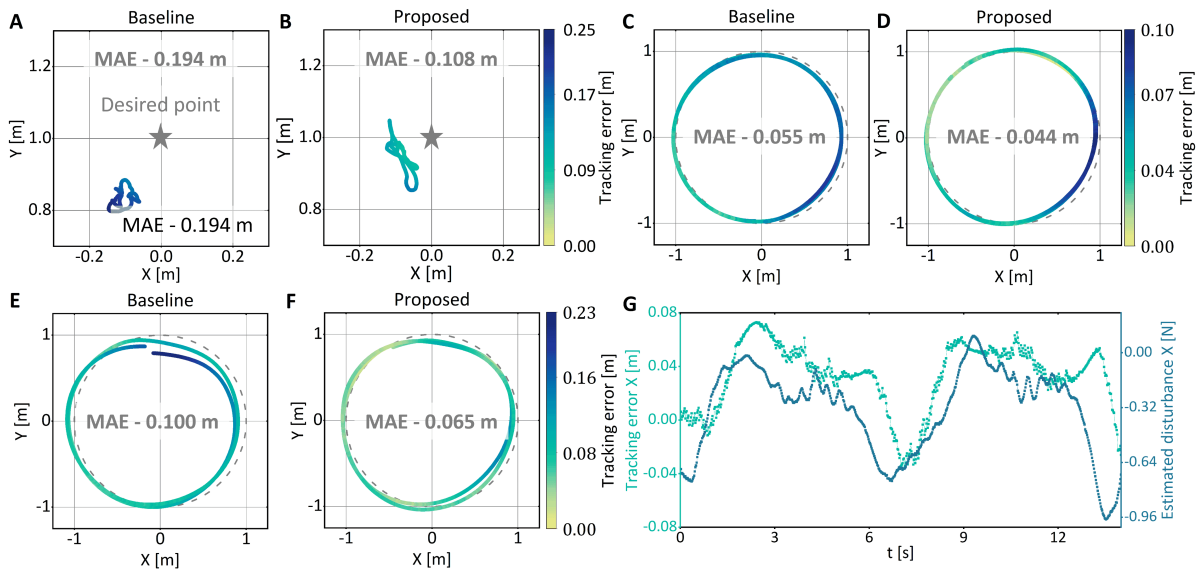


Fig. 5. The tracking and estimation performance in the drone example. A-B, Test #1. C-D, Test #2. E-F, Test #3. G, Tracking error and estimated disturbance along the X -axis in Test #1.

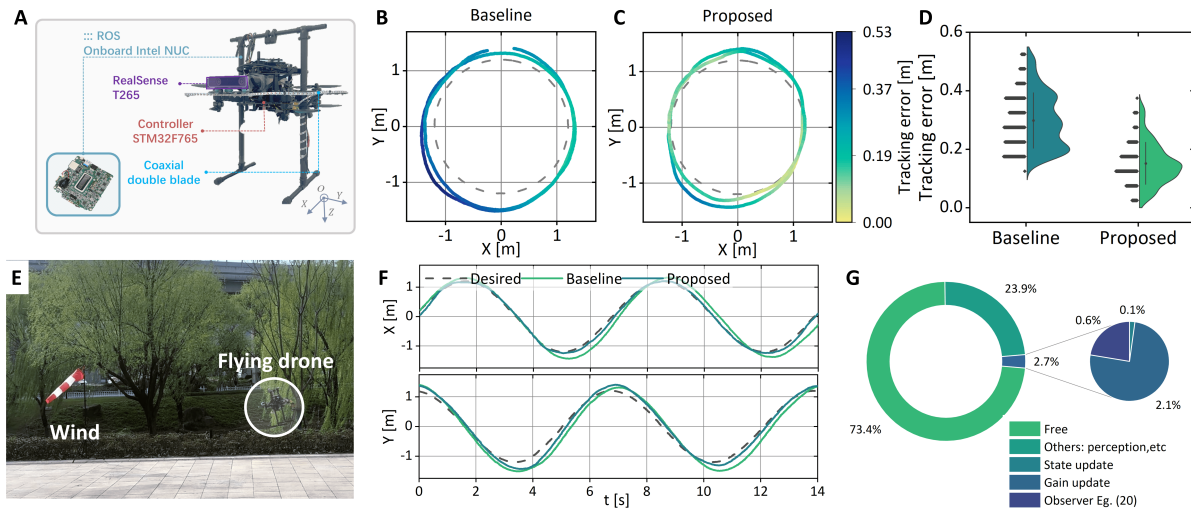


Fig. 6. Flight results in the outdoor scenario. A, The employed flight platform. B, 2D view of the trajectory tracking by the baseline method. C, 2D view of the trajectory tracking by the proposed method. D, Violin plot of the tracking errors. E, The outdoor experimental scenario. F, Detailed tracking performance along X and Y directions of the compared two methods. G, The computation cost of the onboard Inter NUC.

[35] B. Kamen, "Block-form observers for linear time-varying discrete-time systems," in *Proceedings of IEEE Conference on Decision and Control*, 1993, pp. 355–356.

[36] N. Gu, D. Wang, Z. Peng, J. Wang, and Q.-L. Han, "Disturbance observers and extended state observers for marine vehicles: A survey," *Control Engineering Practice*, vol. 123, p. 105158, 2022.

[37] D. Hanover, P. Foehn, S. Sun, E. Kaufmann, and D. Scaramuzza, "Performance, precision, and payloads: Adaptive nonlinear MPC for quadrotors," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 690–697, 2021.

[38] J. Jia, W. Zhang, K. Guo, J. Wang, X. Yu, Y. Shi, and L. Guo, "EVOLVER: Online learning and prediction of disturbances for robot control," *IEEE Transactions on Robotics*, vol. 40, pp. 382–402, 2024.

[39] J.-M. Kai, G. Allibert, M.-D. Hua, and T. Hamel, "Nonlinear feedback control of quadrotors exploiting first-order drag effects," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 8189–8195, 2017.

[40] M. Faessler, A. Franchi, and D. Scaramuzza, "Differential flatness of quadrotor dynamics subject to rotor drag for accurate tracking of high-speed trajectories," *IEEE Robotics and Automation Letters*, vol. 3, no. 2, pp. 620–626, 2017.

[41] I. Abraham and T. D. Murphey, "Active learning of dynamics for data-driven control using Koopman operators," *IEEE Transactions on Robotics*, vol. 35, no. 5, pp. 1071–1083, 2019.