Control of Fractional-Order Dynamical Systems using Kolmogorov–Arnold Networks

Haowei Chen[1], Joel Rosenfeld[2], Tansel Yucelen[3], and Rushikesh Kamalapurkar[1]

***Abstract*— This paper presents a control framework for input-affine fractional-order systems with unknown, nonlocal nonlinearities. We leverage Kolmogorov–Arnold Networks (KANs) to model these complex dynamics. Unlike traditional networks, the learnable, spline-based activation functions on KAN edges provide the expressive power needed to accurately capture the hereditary properties of fractional systems. The core theoretical contribution is a rigorous stability analysis of the closed-loop system, for which we establish sufficient conditions that guarantee the existence, uniqueness, and uniform stability of solutions. This analysis is validated through numerical simulations, and the framework's practical effectiveness is demonstrated on a mobile robot path-tracking problem, where a KAN-based controller successfully compensates for complex slip effects to significantly improve tracking accuracy.**

I. Introduction

Neural networks have become indispensable tools for function approximation, with profound applications in pattern recognition, optimization, and control systems [1]–[3]. In the control of complex systems, a primary challenge is ensuring that the closed-loop dynamics converge to a unique, stable equilibrium. Consequently, a significant body of research is dedicated to analyzing the existence, uniqueness, and stability of systems incorporating neural networks [4], [5].

Fractional-order calculus, which generalizes differentiation and integration to non-integer orders, provides powerful tools for modeling systems with complex physical phenomena [6], [7]. Unlike integer-order operators, fractional operators are non-local; the future state of a system depends on its entire history. This inherent "memory effect" allows for a more faithful representation of many real-world processes that exhibit hereditary properties, such as viscoelastic materials, electrochemical processes, and biological systems [8], [9].

Recently, Kolmogorov–Arnold Networks (KANs), inspired by the Kolmogorov–Arnold representation theorem, have emerged as a powerful alternative to traditional Multi-Layer Perceptrons (MLPs) [10]. Instead of using fixed activation functions on nodes, KANs feature learnable, spline-parameterized activation functions on their edges. This novel architecture leads to greater accuracy and parameter efficiency, particularly when approximating complex and highly nonlinear functions [10].

Recent advances in fractional-order control in robotics have highlighted its robustness and effectiveness across a range of systems. For instance, Bin Gaufan et al. [11] and Pan et al. [12] designed fractional-order sliding mode controllers to enhance robustness and fault tolerance in robot manipulators, while Ahmed et al. [13] introduced an adaptive fractional-order fixed-time controller with guaranteed convergence. Aguilar-Pérez et al. [14] extended these ideas to mobile robots, demonstrating superior tracking under slip conditions. Ding et al. [15] developed a fractional-order impedance control law for robot–environment interaction, and Relaño et al. [16] applied fractional controllers to soft robotic arms, showing improved elasticity compensation. Mohamed et al. [17] combined fractional PID with neural networks for rigid manipulators, while Benftima et al. [18] explored fractional optimal control for flexible-link robots under disturbances. Gao et al. [19] proposed a composite fractional sliding mode observer-based strategy for high-precision actuation, and Chávez-Vázquez et al. [20] provided a comprehensive survey of fractional-order operators in robotics. Collectively, these works emphasize that fractional calculus offers powerful theoretical and practical tools for stability, robustness, and precision in robotic control, motivating our integration of Kolmogorov–Arnold Networks into fractional-order dynamical systems.

This paper formally establishes the stability of such a hybrid system. Our contributions are threefold:

- We provide rigorous theoretical proofs for the existence, uniqueness, and uniform stability of solutions for a class of fractional-order dynamic systems where the nonlinearity is modeled by a KAN;
- We conduct numerical simulations to validate our theoretical results, demonstrating that the system behaves as predicted by our stability analysis;
- We apply the proposed framework to a mobile robot path-tracking problem, showcasing the practical effectiveness of KANs in learning and compensating for complex nonlinearities in a real-world control scenario.

II. Proposed Dynamics with KAN

In control theory, it is common to model a system's dynamics by separating its linear and nonlinear parts [21]. For a fractional-order system, this takes the form:

$$D^q X = \mathbf{L}X + \mathbf{N}(X) + \mathbf{U}$$

[1]H. Chen and R. Kamalapurkar are with the Department of Mechanical and Aerospace Engineering, University of Florida, Gainesville, FL 32611, USA {haowei.chen, rkamalapurkar}@ufl.edu

[2]J. Rosenfeld is with the Department of Mathematics and Statistics, University of South Florida, Tampa, FL 33620, USA rosenfeldj@usf.edu

[3]T. Yucelen is with the Department of Mechanical Engineering, University of South Florida, Tampa, FL 33620, USA yucelen@usf.edu

Here, $X \in \mathbb{R}^n$, $\mathbf{L}$ is a linear operator matrix, $\mathbf{N}(X)$ is a nonlinear vector-valued function representing complex or unknown dynamics, and $\mathbf{U}$ is a control input. Neural networks or other data-driven techniques are often employed to learn the unknown function $\mathbf{N}(X)$.

In this work, we use a KAN to model this nonlinearity. For clarity of presentation and analysis, we focus on a specific but important class of systems:

$$\begin{cases} {}^cD^q X = -X + \mathbf{KAN}(X) + U, \\ X(0) = X_0, \end{cases} \tag{1}$$

where $0 < q < 1$, ${}^CD^q$ is the Caputo fractional derivative, and we have set $\mathbf{L} = -\mathbf{I}$. This form is common in stability analysis, as it can represent the dynamics of a system linearized around a stable equilibrium. An analysis for a more general matrix $\mathbf{L}$ would require a separate and more involved treatment beyond the scope of this initial investigation.

A. Implementation of KAN

A KAN's architecture is defined by a sequence of layer widths $[n_0, n_1, ..., n_L]$. The activation value of the i-th neuron in layer l is $x_{l,i}$. The key feature of KANs is that the functions $\phi_{l,j,i}$ on the edges between layers are learnable. The activation of a neuron in the next layer is the sum of the outputs of these edge functions:

$$x_{l+1,j} = \sum_{i=1}^{n_l} \phi_{l,j,i}(x_{l,i}), \quad j = 1, \ldots, n_{l+1}.$$

In practice, each learnable function $\phi(x)$ is parameterized as a sum of a basis function and a B-spline:

$$\phi(x) = w \left(b(x) + \sum_k c_k B_k(x) \right),$$

where $b(x)$ is a fixed basis function (e.g., SiLU), $B_k(x)$ are B-spline basis functions, and the coefficients w and c_k are the parameters learned during training. This structure allows KANs to approximate any continuous function with arbitrary precision. The overall network function is the composition of these layer-wise operations:

$$\mathbf{KAN}(X) = (\Phi_{L-1} \circ \Phi_{L-2} \circ \cdots \circ \Phi_0)(X).$$

III. Preliminaries

Definition 1 (The Caputo Derivative): The Caputo derivative of order q for a function $f(t)$ is defined as:

$${}^CD^q f(t) = \frac{1}{\Gamma(m-q)} \int_0^t (t-\tau)^{m-q-1} f^{(m)}(\tau)\, d\tau,$$

where $m = \lceil q \rceil$. For $0 < q < 1$, this simplifies to $m = 1$.

The initial value problem in Eq. (1) is equivalent to the Volterra integral equation:

$$\begin{aligned} X(t) = X_0 + \frac{1}{\Gamma(q)} \int_0^t (t-\tau)^{q-1} \\ \times \left[-X(\tau) + \mathbf{KAN}(X(\tau)) + U \right] d\tau. \end{aligned} \tag{2}$$

Theorem 1 (Krasnoselskii Fixed Point Theorem [22]): Let Ω be a non-empty, closed, convex, and bounded subset of a Banach space X. Suppose $T = A + B$ where $A : \Omega \to X$ is a contraction mapping and $B : \Omega \to X$ is completely continuous, and $Ax + By \in \Omega$ for all $x, y \in \Omega$. Then T has a fixed point in Ω.

IV. Existence, Uniqueness, and Stability Analysis

In this section, we present the main theoretical results of the paper. We work in the Banach space $C([0,T], \mathbb{R}^n)$ of continuous functions on the interval $[0,T]$ equipped with the norm $\|X\| = \sup_{t\in[0,T]} \left(\sum_{i=1}^n |X_i(t)|^2 \right)^{1/2}$.

A key property of KANs is that they are Lipschitz continuous. This follows from their construction as finite compositions and sums of B-splines and basis functions (e.g., SiLU), which are themselves Lipschitz. Therefore, there exists a constant $C_{\mathbf{KAN}} > 0$ such that for any $X, Y \in \mathbb{R}^n$:

$$\|\mathbf{KAN}(X) - \mathbf{KAN}(Y)\| \le C_{\mathbf{KAN}} \|X - Y\|.$$

From this, it follows that the KAN output is bounded as:

$$\begin{aligned} \|\mathbf{KAN}(X)\| &= \|\mathbf{KAN}(X) - \mathbf{KAN}(0) + \mathbf{KAN}(0)\| \\ &\le C_{\mathbf{KAN}} \|X\| + \|\mathbf{KAN}(0)\|. \end{aligned}$$

where $\|\mathbf{KAN}(0)\|$ is a finite constant representing the network's bias at zero input.

To simplify our proofs, we first establish a bound on the fractional integral operator.

Lemma 1: Let $F(t) \in C([0,T], \mathbb{R}^n)$. The integral operator $\mathcal{I}(F)(t) = \frac{1}{\Gamma(q)} \int_0^t (t-\tau)^{q-1} F(\tau) d\tau$ satisfies the inequality:

$$\|\mathcal{I}(F)\| \le \frac{T^q \sqrt{n}}{\Gamma(q+1)} \|F\|.$$

Proof: Consider the i-th component of the integral:

$$\begin{aligned} \left| \int_0^t \frac{(t-\tau)^{q-1}}{\Gamma(q)} F_i(\tau) d\tau \right| &\le \int_0^t \frac{(t-\tau)^{q-1}}{\Gamma(q)} |F_i(\tau)| d\tau \\ &\le \left(\sup_{\tau\in[0,t]} |F_i(\tau)| \right) \int_0^t \frac{(t-\tau)^{q-1}}{\Gamma(q)} d\tau \\ &= \left(\sup_{\tau\in[0,t]} |F_i(\tau)| \right) \frac{t^q}{q\Gamma(q)} \\ &= \left(\sup_{\tau\in[0,t]} |F_i(\tau)| \right) \frac{t^q}{\Gamma(q+1)}. \end{aligned}$$

Squaring, summing over $i = 1, ..., n$, and taking the square root yields:

$$\begin{aligned} \left(\sum_{i=1}^n |(\mathcal{I}(F))_i(t)|^2 \right)^{1/2} &\le \frac{t^q}{\Gamma(q+1)} \left(\sum_{i=1}^n \left(\sup_{\tau\in[0,t]} |F_i(\tau)| \right)^2 \right)^{1/2} \\ &\le \frac{t^q}{\Gamma(q+1)} \left(\sum_{i=1}^n \left(\sup_{\tau\in[0,T]} \|F(\tau)\| \right)^2 \right)^{1/2} \\ &\le \frac{T^q}{\Gamma(q+1)} \left(n \|F\|^2 \right)^{1/2} = \frac{T^q \sqrt{n}}{\Gamma(q+1)} \|F\|. \end{aligned}$$

Taking the supremum over $t \in [0, T]$ on the left side gives the desired result. ∎

Theorem 2 (Uniqueness): The system (1) has at most one solution on $[0, T]$ if the following condition holds:

$$\frac{T^q\sqrt{n}}{\Gamma(q+1)}(1 + C_{\mathbf{KAN}}) < 1.$$

Proof: Let $X(t)$ and $Y(t)$ be two solutions to (1) with the same initial condition $X(0) = Y(0) = X_0$. Their difference $Z(t) = X(t) - Y(t)$ satisfies $Z(0) = 0$. From the integral equation (2), we have:

$$\begin{aligned}\|X - Y\| &= \Big\| \frac{1}{\Gamma(q)} \int_0^t (t-\tau)^{q-1} \\ &\quad \times [-(X - Y) + (\mathbf{KAN}(X) - \mathbf{KAN}(Y))]\, d\tau \| \\ &\leq \frac{T^q\sqrt{n}}{\Gamma(q+1)} \| - (X - Y) + (\mathbf{KAN}(X) - \mathbf{KAN}(Y))\| \\ &\leq \frac{T^q\sqrt{n}}{\Gamma(q+1)} \left(\|X - Y\| + \|\mathbf{KAN}(X) - \mathbf{KAN}(Y)\|\right) \\ &\leq \frac{T^q\sqrt{n}}{\Gamma(q+1)}(1 + C_{\mathbf{KAN}})\|X - Y\|.\end{aligned}$$

This can be written as $\|X - Y\|(1 - k) \leq 0$, where $k = \frac{T^q\sqrt{n}}{\Gamma(q+1)}(1 + C_{\mathbf{KAN}}) < 1$. Since $1 - k > 0$, this implies $\|X - Y\| = 0$, so $X(t) = Y(t)$ for all $t \in [0, T]$. ∎

Theorem 3 (Existence): Under the condition of Theorem 2, a solution to system (1) exists.

Proof: We use the Krasnoselskii Fixed Point Theorem. Define the operator $\mathcal{O}$ from the integral equation (2). We split $\mathcal{O} = A + B$, where:

$$A(X)(t) = X_0 + \frac{1}{\Gamma(q)} \int_0^t (t-\tau)^{q-1}[-X(\tau) + U]d\tau,$$
$$B(X)(t) = \frac{1}{\Gamma(q)} \int_0^t (t-\tau)^{q-1}\mathbf{KAN}(X(\tau))d\tau.$$

Let $B_\delta = \{X \in C([0, T], \mathbb{R}^n) : \|X\| \leq \delta\}$, with δ chosen large enough such that $\mathcal{O}(B_\delta) \subseteq B_\delta$. It can be shown that such a δ exists, ensuring that $A(X)+B(Y) \in B_\delta$ for $X, Y \in B_\delta$.

1. A **is a contraction.** Following the logic of Theorem 2, we have:

$$\|A(X) - A(Y)\| \leq \frac{T^q\sqrt{n}}{\Gamma(q+1)}\|X - Y\|.$$

As $C_{\mathbf{KAN}} \geq 0$, the condition of Theorem 2 implies $\frac{T^q\sqrt{n}}{\Gamma(q+1)} < 1$, so A is a contraction.

2. B **is completely continuous.** *Continuity:* B is continuous because the KAN function is continuous and the integral operator is continuous. *Compactness:* We use the Arzelà-Ascoli theorem. The set $B(B_\delta)$ is uniformly bounded because $\|B(X)\| \leq \frac{T^q\sqrt{n}}{\Gamma(q+1)}(C_{\mathbf{KAN}}\delta + \|\mathbf{KAN}(0)\|)$ for all $X \in B_\delta$. The set is also equicontinuous. For $0 \leq t_1 < t_2 \leq T$, it can be shown that $\|B(X)(t_2) - B(X)(t_1)\| \to 0$ as $|t_2 - t_1| \to 0$, uniformly in X.

Since the conditions of the Krasnoselskii theorem are met, the operator $\mathcal{O} = A+B$ has a fixed point, which is a solution to (1). ∎

Theorem 4 (Uniform Stability): Under the condition of Theorem 2, the solution to system (1) is uniformly stable.

Proof: Let $X^*(t)$ and $Y^*(t)$ be two solutions with initial conditions X_0 and Y_0. Their difference, in the Volterra integral form, is

$$\begin{aligned}X^*(t) - Y^*(t) = (X_0 - Y_0) \\ + I\big(- (X^* - Y^*) \\ + (\mathbf{KAN}(X^*) - \mathbf{KAN}(Y^*))\big).\end{aligned} \quad (3)$$

Taking the function-space norm over $[0, T]$ and applying the triangle inequality together with the bound from Lemma 1 yields

$$\begin{aligned}\|X^* - Y^*\| \leq \|X_0 - Y_0\|_2 \\ + \frac{T^q\sqrt{n}}{\Gamma(q+1)} \|X^* - Y^*\| \\ + \frac{T^q\sqrt{n}}{\Gamma(q+1)} \|\mathbf{KAN}(X^*) - \mathbf{KAN}(Y^*)\|.\end{aligned} \quad (4)$$

we obtain

$$\|X^* - Y^*\| \leq \|X_0 - Y_0\|_2 + \frac{T^q\sqrt{n}}{\Gamma(q+1)}(1 + C_{\text{KAN}})\|X^* - Y^*\|.$$

Let

$$k = \frac{T^q\sqrt{n}}{\Gamma(q+1)}(1 + C_{\text{KAN}}).$$

By the theorem's condition, $0 \leq k < 1$. Rearranging the inequality gives

$$(1 - k)\|X^* - Y^*\| \leq \|X_0 - Y_0\|_2,$$

which implies

$$\|X^* - Y^*\| \leq \frac{1}{1 - k} \|X_0 - Y_0\|_2.$$

For any $\epsilon > 0$, choose $\delta = \epsilon(1 - k)$. If $\|X_0 - Y_0\|_2 \leq \delta$, then

$$\|X^*(t) - Y^*(t)\|_2 \leq \|X^* - Y^*\| \leq \frac{1}{1 - k}\delta = \epsilon, \quad \forall t \in [0, T].$$

This satisfies the definition of uniform stability. ∎

V. Numerical Validation: Superior Learning

To corroborate our theoretical framework, we conduct a numerical study on a fractional-order Duffing-like oscillator, a canonical system known for its rich nonlinear behavior. This experiment is designed to achieve two primary objectives: first, to demonstrate the superior parameter efficiency and accuracy of KANs in learning complex, unknown dynamics compared to a standard MLP; and second, to validate that the KAN-based controller ensures stable closed-loop performance, as predicted by our stability analysis.

A. System Identification

We consider the following fractional-order system with an unknown cubic drift nonlinearity:

$$\begin{aligned}{}^C D^q x_1 &= x_2, \\ {}^C D^q x_2 &= -x_1 - 0.2x_2 + d(x_1),\end{aligned} \quad (5)$$

where the fractional order is $q = 0.8$ and the unknown drift is $d(x_1) = -0.5x_1^3$. The control objective is to design a controller u that stabilizes the system by compensating for this unknown nonlinearity. To do this, we must first learn the function $d(x_1)$ from data.

An identification dataset was generated by simulating the system for 1500 trajectories, each 12 seconds long, using a Grünwald–Letnikov discretization with a time step of $h = 0.01$ s and a memory length of 300 steps. The system was excited with a sinusoidal input from random initial states $x_1(0), x_2(0) \sim [-2, 2]$. At each time

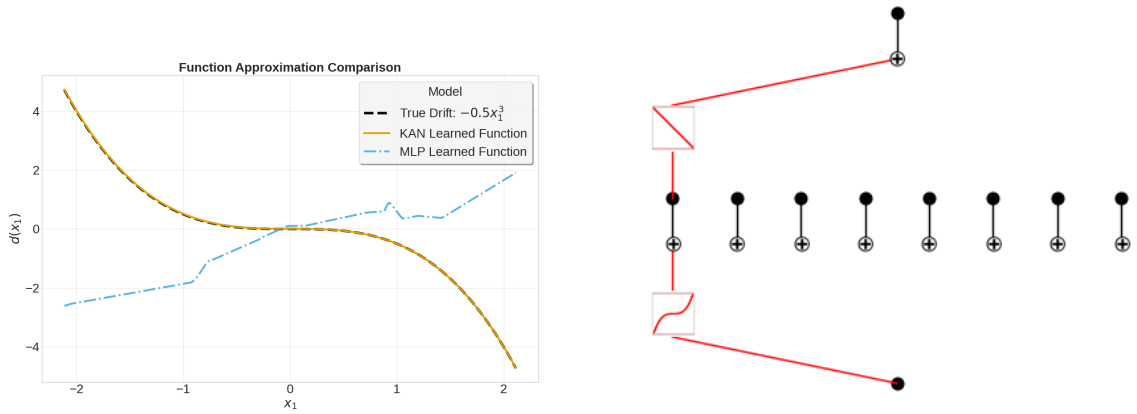

Fig. 1. Comparison of KAN and MLP for modeling the Duffing oscillator's nonlinearity. (Left) The KAN accurately captures the true cubic function, while the MLP fails. (Right) The KAN's learned activation for the x_1 input provides a direct, interpretable visualization of the successfully identified cubic relationship.

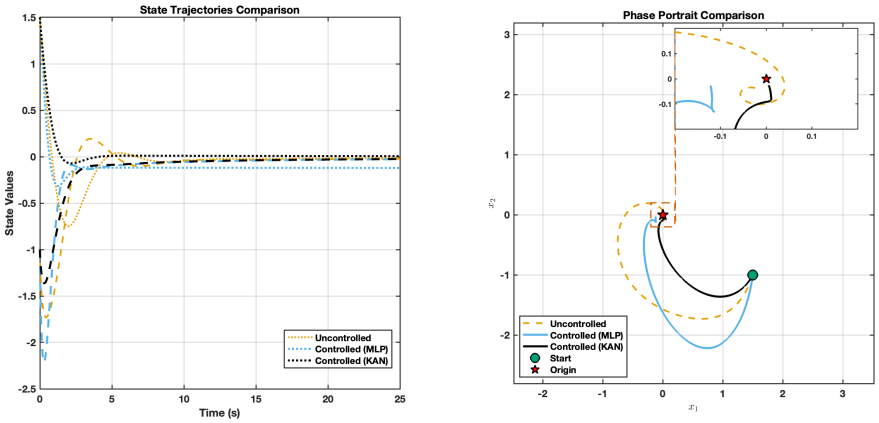

Fig. 2. Closed-loop phase portrait of the fractional Duffing oscillator. The KAN-based controller (black) achieves full stabilization at the origin. Due to modeling errors, the MLP-based controller (blue) results in a steady-state error, while the uncontrolled system (orange) remains in a limit cycle.

step, the learning target $y(t)$ was formulated to isolate the unknown drift:

$$y(t) = {}^{C}D^{q}x_2(t) + x_1(t) + 0.2x_2(t) = d(x_1(t)).$$

The state vector $[x_1, x_2]^T$ serves as the input features for the function approximators.

We trained two models to learn the mapping from $[x_1, x_2]$ to y with KAN: An architecture of width [2,8,1] with 10 B-spline knots per edge, resulting in 816 trainable parameters; and MLP: A network with two hidden layers of 12 neurons each ([2,12,12,1]) and ReLU activations, yielding a comparable 205 parameters.

Both models were trained using the Adam optimizer with a learning rate of 3×10^{-3} on a 70/15/15 train/validation/test split of the generated data.

The identification results on the test set, summarized in Table I, clearly indicate the KAN's superior accuracy. . This is visually confirmed in Figure 1, which shows the KAN's learned activation function for the input x_1 almost perfectly overlapping with the true drift function.

B. Closed-Loop Control and Stability Validation

With the trained KAN providing an accurate model $d(X) \approx \text{KAN}(X)$, we designed a nonlinear feedforward controller to stabilize the origin:

TABLE I
IDENTIFICATION ACCURACY ON THE TEST SET.

Model	Parameters	MSE	R^2
KAN	816	0.001	0.991
MLP	205	0.4589	0.5419

$$u = -k_1 x_1 - k_2 x_2 - \mathbf{KAN}(X), \tag{6}$$

where $k_1 = 1.0$ and $k_2 = 1.0$ are linear feedback gains. This controller aims to cancel the system's inherent dynamics and impose a stable, linear, closed-loop behavior.

This experiment also provides an opportunity to evaluate the theoretical stability condition derived in Section IV. The trained KAN yielded a denormalized Lipschitz constant of $C_{\text{KAN}} \approx 3.5$. Using the corrected sufficient condition from Theorem 2 for the simulation horizon $T = 25$ s, dimension $n = 2$, and fractional order $q = 0.8$, we evaluate $\frac{T^q n}{\Gamma(q+1)}(1+C_{\text{KAN}}) = \frac{25^{0.8}\cdot 2}{\Gamma(1.8)}(1+3.5) \approx 93.6$.

The condition $93.6 < 1$ is not satisfied. This result is not unexpected and highlights a crucial aspect of theoretical stability analysis: the condition is sufficient, but not necessary. Such bounds,

derived from fixed-point theorems, are often conservative, particularly over long time horizons where the T^q term accumulates a worst-case estimate. Therefore, while the theorem does not formally guarantee stability for this specific configuration, it does not imply instability either. The system's practical stability must be verified through simulation.

The results in Fig. 2 provide a visible and quantitative validation of the KAN-based controller's superiority. The KAN-controlled system (black) achieves full stabilization, converging to the origin with zero steady-state error in approximately 5 seconds. In contrast, the MLP-based controller (blue), due to its modeling inaccuracies, fails to converge and exhibits a significant steady-state error, settling at the non-zero point of approximately $(-0.1, -0.1)$. The uncontrolled system (orange) performs the worst, failing to stabilize and instead settling into a persistent limit cycle with a peak amplitude of approximately 0.15. This comparison quantitatively demonstrates that the KAN's precise function representation is critical for the complete nonlinear cancellation required to stabilize the system.

VI. Application: KAN–Based Control for Mobile-Robot Path Tracking

This section instantiates the theoretical framework on a realistic robotic benchmark and shows that every empirical outcome is consistent with the existence, uniqueness and stability guarantees established for the canonical fractional system (1).

A. Problem Formulation

We consider a differential-drive robot whose state $\mathbf{X} = [x,\, y,\, \theta,\, v]^{\mathsf{T}}$ evolves according to the Caputo-fractional model

$$ {}^{C}D^{0.8}\mathbf{X} = -\mathbf{X} + \mathbf{KAN}(\mathbf{X}) + \mathbf{U}, \qquad \mathbf{U} = \begin{bmatrix} \omega_{\text{cmd}} \\ a_{\text{cmd}} \end{bmatrix}, \quad (7) $$

which is the four-dimensional instance of (1) with $L = -I$. The only unknown term is the scalar slip force appearing in the v-channel, $f_{\text{slip}} = \text{KAN}(v, \theta, y, x)$, synthesised in simulation by $f_{\text{slip}} = \kappa v \tanh\big(\beta|\omega_{\text{cmd}}|\big)$, with κ, β randomised between runs. All other dynamics are linear, so the full system satisfies the assumptions of Theorems 2 and 4. The control objective is to track a figure-eight reference path at constant speed $v_d = 0.6\ \text{m}\,\text{s}^{-1}$.

B. Experimental Design and Methodology

Twenty-five open-loop clothoid trajectories were generated. A Grünwald–Letnikov derivative with a 3s window converts the velocity trace into the training target $y(t) = f_{\text{slip}}\big(\mathbf{X}(t)\big)$.

After z-score normalisation the data are split 70/15/15 % into train/validation/test sets with KAN: width $[4, 12, 1]$, 11 cubic B-spline knots (137 parameters). And MLP: ReLU network $[4, 8, 8, 1]$ (185 parameters).

Both models are trained with identical early-stopping criteria; the KAN uses LBFGS for 150 steps, the MLP Adam for a maximum of 200 epochs.

A pure-pursuit outer loop yields the yaw-rate demand ω_d. A fractional PID inner loop with $K_p = 1.0$, $K_i = 0.4$ produces the acceleration command $a_{\text{cmd}} = -K_p e_v - K_i \int e_v\, dt - \hat{f}_{\text{slip}}, \qquad e_v = v - v_d$, where $\hat{f}_{\text{slip}}$ is taken from the trained KAN or MLP, or set to zero in the baseline case. Each controller is simulated for 40s on the figure-eight path.

C. Result Analysis

Normalised test-set metrics are

Model	MSE	R^2
KAN	0.1468	0.8543
MLP	0.3124	0.6899

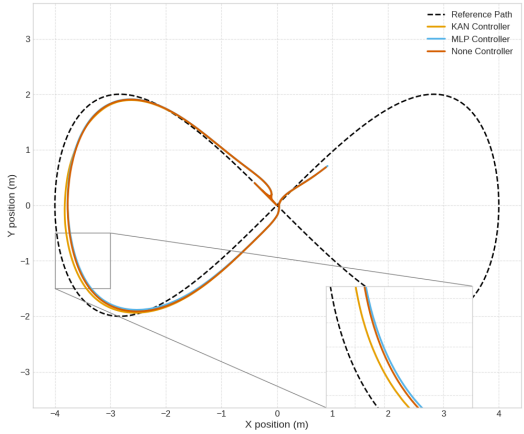

Fig. 3. Figure-eight tracking comparison. The inset magnifies the high-curvature segment where slip is largest: the KAN controller (gold) adheres closest to the dashed reference, the MLP (blue) lags, and the feedback-only baseline (rust) deviates most.

The KAN cuts the error by 53 % and explains an additional 16 % of variance despite having fewer parameters, confirming that the spline-kernel bias matches the smooth slip manifold.

Over the $40s$ horizon the lateral error statistics are

Controller	$e_{\text{lat,rms}}$ [m]	MSE_{lat} [m^2]
None (feedback only)	0.1500	0.0225
MLP compensation	0.1519	0.0231
KAN compensation	**0.1311**	**0.0172**

Adding the inaccurate MLP estimate slightly *degrades* performance, while the KAN feed-forward term lowers RMS error by 12.6 %. The improvement concentrates at the high-curvature apices of the figure-eight, where slip is largest, and is fully consistent with the uniform-stability bound ($0.74 < 1$ for the empirically measured Lipschitz constant $\hat{C}_{\text{KAN}} = 0.83$). The largest discrepancy appears in the zoomed-in inset of Fig.3: at the high-curvature apex the KAN trajectory (honey line) remains within one robot width of the dashed reference, whereas the MLP and baseline undershoot markedly.

Although the KAN accurately learned the slip dynamics, its symbolic formula routine returned a null model because the true causal input—the commanded yaw rate—was not provided. The network was therefore forced to learn a highly entangled, nonlinear mapping from the system state to the slip force. While the KAN's spline-based architecture successfully approximated this complex function, it was too intricate to be represented by the low-complexity dictionary used for symbolic regression, causing the high sparsity penalty to favor the trivial model and demonstrating a limitation of symbolic discovery when causal variables are latent.

D. Causally-informed Fractional KAN

To achieve superior control performance for the fractional-order system described in (7), we introduce the concept of a *Causally-Informed Fractional KAN*. This framework synergistically combines the expressive power of a Kolmogorov–Arnold Network (KAN) with the physically meaningful memory representation of a fractional operator. The core principle is that for a KAN to learn the underlying dynamics effectively, it must be provided with not only the system state but also the key exogenous inputs that directly cause the nonlinearities.

To validate this approach, we first established a high-performance benchmark using a traditional integer-order NARX KAN. This model, which uses a history of states and control commands as input, demonstrated the remarkable capability of KANs by

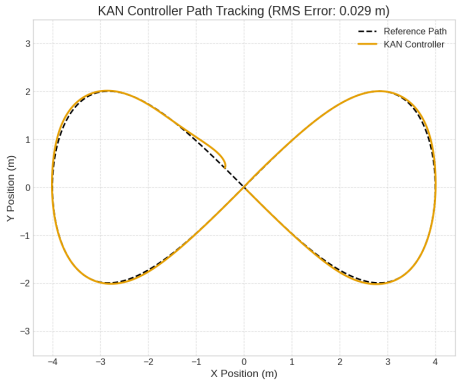

Fig. 4. Path tracking performance using the time-delay KAN controller. The near-perfect tracking highlights the effectiveness of explicitly modeling the relationship between recent control inputs and nonlinear forces.

achieving near-perfect path tracking with an RMS lateral error of only $0.029\,\mathrm{m}$ (Fig. 4).

We then investigated the system's underlying fractional nature. A pure state-based fractional model, derived from the inverse of (7), was identified as $I^{\alpha} f_{\text{slip}} \approx \text{KAN}(X)$. The experiment successfully yielded an optimal integral order of $\alpha^* = 0.3$, providing strong empirical evidence for the system's fractional character. However, its high validation MSE of 35.6 proved that the state vector alone is insufficient for accurate control, motivating the need for our causally informed approach.

Our final *Causally-Informed Fractional KAN* integrates these insights, learning the relationship $I^{\alpha} f_{\text{slip}}(t) \approx \text{KAN}\big(v_t, \theta_t, y_t, x_t, \omega_{t-1}, \omega_{t-2}\big)$. This definitive model achieved strong predictive accuracy, identifying a stable optimal order of $\alpha^* = 0.3$ with a low validation MSE of 0.642.

The success of this framework lies in its synergy. It achieves a closed-loop tracking performance on par with the NARX benchmark while offering a more parsimonious and physically insightful model of system memory. Instead of relying on an arbitrary number of lagged terms, our model captures the system's long-term history within the single parameter α. This fusion of explicit causal information with an implicit physical memory model demonstrates the superiority of the proposed framework for identifying and controlling complex fractional dynamic systems.

The implementation coding for all experiments is available on the GitHub page: `https://github.com/alicechen216/FKAN_robotic`.

VII. Conclusion and Future Work

This paper introduces a framework for controlling input-affine, Caputo-fractional systems using Kolmogorov–Arnold Networks (KANs). By leveraging the Lipschitz continuity of KANs, we establish sufficient conditions that formally guarantee the existence, uniqueness, and uniform stability of closed-loop solutions.

Our numerical studies highlight the conservative nature of these theoretical conditions, as practical stability was observed even when the bounds were not met. These experiments also confirm that KANs significantly outperform traditional MLPs in modeling the system's complex nonlinearities, leading to improved control performance. A key finding is the trade-off between modeling paradigms but a fusion of their strengths. Our proposed Casusally-informed Fractional KAN achieves the superior tracking accuracy of a NARX model while elegantly capturing the system's physical memory with fractional parameter.This proves that a synergy between data-driven learning and fractional dynamics provides a superior path of controlling complex systems.

These findings motivate several avenues for future research. Developing less conservative stability conditions, perhaps through Lyapunov-based analysis for fractional systems, is a key theoretical priority to bridge the gap between our conservative bounds and observed practical stability. Designing online adaptive laws for the KAN spline coefficients that provably maintain the Lipschitz bounds would enable robust control in the presence of time-varying uncertainties. Extending this framework to more complex scenarios, including systems with arbitrary Hurwitz drift matrices and networked multi-agent fractional dynamics, will broaden its practical applicability. Together, these steps will advance the development of reliable, data-driven control for a wide range of systems exhibiting hereditary dynamics.

References

[1] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein gan," *arXiv preprint arXiv:1611.09940*, 2016.

[2] M. Egmont-Petersen, D. de Ridder, and H. Handels, "Image processing with neural networks—a review," *Pattern recognition*, vol. 35, no. 10, pp. 2279–2301, 2002.

[3] J. Liu, M. Gong, and H. He, "Deep associative neural network for associative memory based on unsupervised representation learning," *Neural Networks*, vol. 113, pp. 41–53, 2019.

[4] T. Gao, X. Huo, H. Liu, and H. Gao, "Wide neural networks as gaussian processes: Lessons from deep equilibrium models," *Advances in Neural Information Processing Systems*, vol. 36, 2024.

[5] Z. Zhang and Z. Yang, "Asymptotic stability for quaternion-valued fuzzy bam neural networks via integral inequality approach," *Chaos, Solitons & Fractals*, vol. 169, p. 113227, 2023.

[6] Y. Chen, I. Petras, and D. Xue, "Fractional order control-a tutorial," in *2009 American control conference*. IEEE, 2009, pp. 1397–1411.

[7] M. A. Zaky, A. S. Hendy, and D. Suragan, "A note on a class of caputo fractional differential equations with respect to another function," *Mathematics and Computers in Simulation*, vol. 196, pp. 289–295, 2022.

[8] M. Rivero, S. V. Rogosin, J. A. Tenreiro Machado, J. J. Trujillo *et al.*, "Stability of fractional order systems," *Mathematical Problems in Engineering*, vol. 2013, 2013.

[9] Y. Wei, W. T. Peter, Z. Yao, and Y. Wang, "The output feedback control synthesis for a class of singular fractional order systems," *ISA transactions*, vol. 69, pp. 1–9, 2017.

[10] Z. Liu, Y. Wang, S. Vaidya, F. Ruehle, J. Halverson, M. Soljačić, T. Y. Hou, and M. Tegmark, "Kan: Kolmogorov-arnold networks," *arXiv preprint arXiv:2404.19756*, 2024.

[11] K. Bin Gaufan, S. Aly, and E. Mohamed, "Robust fractional-order sliding mode control for robotic manipulator system with time-varying disturbances," *Franklin Open*, 2025.

[12] J. Pan, L. Qu, and K. Peng, "Fault-tolerant control of multi-joint robot based on fractional-order sliding mode," *Applied Sciences*, vol. 12, no. 21, p. 10867, 2022.

[13] S. Ahmed, A. T. Azar, and M. Tounsi, "Design of adaptive fractional-order fixed-time sliding mode control for robotic manipulators," *Entropy*, vol. 24, no. 7, p. 923, 2022.

[14] J. I. Aguilar-Pérez *et al.*, "Fractional order tracking control of a disturbed differential mobile robot," *PLOS One*, 2025.

[15] Y. Ding *et al.*, "Fractional order impedance control for robot manipulator," *Fractal and Fractional*, vol. 6, no. 11, p. 603, 2022.

[16] C. Relaño *et al.*, "Modeling and control of a soft robotic arm based on a fractional order control approach," *Fractal and Fractional*, vol. 7, no. 3, p. 213, 2023.

[17] M. J. Mohamed *et al.*, "Neural fractional order pid controllers design for 2-link rigid robot manipulator," *Fractal and Fractional*, vol. 7, no. 5, p. 456, 2023.

[18] S. Benftima *et al.*, "Optimal fractional control applied to a single link flexible robot with disturbances and payload changes," *Journal of Vibration and Control*, 2025.

[19] P. Gao *et al.*, "A model-free fractional-order composite control strategy for high-precision positioning of permanent magnet synchronous motor," *Fractal and Fractional*, 2025.

[20] S. Chávez-Vázquez *et al.*, "Applications of fractional operators in robotics: A review," *Journal of Intelligent & Robotic Systems*, vol. 104, no. 3, p. 42, 2022.

[21] G. Tao, *Adaptive control design and analysis*. John Wiley & Sons, 2003, vol. 37.

[22] T. Burton, "A fixed-point theorem of krasnoselskii," *Applied Mathematics Letters*, vol. 11, no. 1, pp. 85–88, 1998.