KYBERNETIKA — VOLUME *62* (2026), NUMBER 2, PAGES 305–331

CHANCE-CONSTRAINED REACHABILITY ANALYSIS FOR DATA-DRIVEN PREDICTIVE CONTROL OF UNKNOWN NONLINEAR SYSTEMS

Teketel Ketema, Surafel Luleseged Tilahun, Simon D. Zawka and Abebe Geletu

This study presents a novel data-driven predictive control approach for unknown nonlinear systems under bounded process and measurement noises. A chance-constrained reachability analysis control framework is proposed to provide probabilistic safety and robustness guarantees by characterizing the likely evolution of system behavior under risk-aware control. A zonotopic deep Koopman reachability analysis is used to design a data-driven controller without acquiring prior knowledge of the statistical properties of the noise. Unlike previous set-based approaches that enforce hard constraints under worst-case scenarios, the proposed method balances robustness and performance more effectively, reducing conservatism while still ensuring safety with bounded risk. It also guarantees recursive feasibility using a first-step constraint technique. A simulation study is conducted on a stirred-tank reactor system and a cart–damper–spring system to demonstrate the effectiveness of the proposed approach, with numerical results supporting the theoretical claims and highlighting its practical applicability.

Keywords: reachability analysis, chance constraint, first-step constraint, predictive control, recursive feasibility

Classification: 93B03, 93C10, 93E20, 93E35

1. INTRODUCTION

Amid the rising complexity of dynamic systems and expanding data accessibility, predictive control has played a pivotal role in advancing optimization-driven approaches to system management and control. Predictive control in this realm encompasses two primary paradigms: model predictive control, which relies on an explicit mathematical model of the system, and data-driven predictive control, which operates without requiring prior knowledge of the system dynamics [18, 31]. Model predictive control approach offers superior performance by explicitly handling system constraints and optimizing future behavior, yet its reliance on an accurate system model limits its applicability to complex, uncertain, or unknown systems [11]. Data-driven predictive control paradigm mitigates this limitation by learning the predictive controller directly from past input-output data [1].

DOI: 10.14736/kyb-2026-2-0305

Data-driven predictive control can be categorized into two types: indirect (or parametric) and direct (or nonparametric). The indirect approach involves a system identification phase, which is usually costly and time-consuming for complex systems [4]. In contrast, the direct approach bypasses the explicit model identification step, and it learns the predictive controller directly from the data [15]. Nevertheless, both approaches follow a model predictive control-like receding-horizon principle: at each control step, the controller optimizes predictions over a finite future horizon, only the first control action from the optimized sequence is applied to the system, and the horizon recedes forward at the next time step using the updated measurements [3]. In fact, the latter can be viewed as a relaxation of the former, as its regularization implicitly captures the system identification step [10].

One of the forefront challenges in real-world control processes is the fact that they lie in uncertain and dynamic environments [23]. Even more concerning is that the growing complexity and nonlinearity of modern industrial systems are compelling most control investigations to operate under unknown dynamics and probabilistic safety [18]. In such cases, strict constraint satisfaction becomes too conservative; and needs to be managed within acceptable risk limits or a predetermined level of constraint violation [21]. Chance constraint is an appropriate measure in such scenarios to increase the region of attraction without changing the prediction horizon, and address random inequality constraints in terms of probability measures. A chance-constrained data-driven predictive control is then a control strategy that uses historical data to furnish structural information about the system to be controlled, under probabilistic constraints.

The concept of chance-constraint was introduced in [6] for finance planning optimization. Since its inception, it has remained dynamic, evolving into a pivotal area of research in optimal control problems under uncertainty. It gained traction in model predictive control applications with its introduction by [25] and [17]. In data-driven predictive control problems of unknown nonlinear systems, the consideration of chance-constrained optimization enhances the controller to be robust, reliable, and capable of handling intrinsic real-world uncertainties. It addresses uncertainties and variations in many data-driven control applications, such as energy management [2] and chemical process [13], where traditional control approaches are insufficient.

The notable difficulty in chance constraint-based control is to adjust the confidence level of the chance constraint with no prior knowledge of the uncertain variables. To address this, various data-driven approaches were proposed in [8, 9, 28, 29, 30]. These approaches used historical data to determine a nonparametric probability density function of the random variable with no prior statistical assumptions. Then, they adjusted the chance constraint and reformulated it to its equivalent algebraic form based on the estimated function. While these methods effectively handle chance constraint implementation, they are all done under the assumption of known system dynamics.

Existing predictive control approaches of chance-constrained unknown nonlinear systems in the literature are designed based on model predictive control approaches, except [5] and [26]. This shows that the model-free approach in this specific topic is still in its early stages. This mainly happens because the controller should simultaneously explore new actions to learn the system and stay within the system constraint with a high probability. This makes the system's safety delicate and yields difficulty in predicting and

quantifying future uncertainties with an unknown system model. However, reliance on an accurate model raises many gaps in real-world applications, which in turn indicates the necessity of model-free predictive control approaches, despite the computational and safety hurdles they involve [7].

Chance-constrained optimal control problem of stochastic nonlinear systems with unknown dynamics under Gaussian noise was studied by [5]. This work helped bridge the gaps between model-free predictive control and safety-critical systems with safety considerations, but the Gaussian noise assumption is not usually the case in real-world scenarios. A chance-constrained data-driven control method of unknown nonlinear systems with no prior assumptions of both system dynamics and uncertainty was proposed in [26]. They combined kernel methods and stochastic optimization to reformulate the original nonlinear problem as a tractable linear program using only a dataset of observed trajectories. While this method is promising, its open-loop stochastic policy lacks feedback and adaptivity, which potentially affects the robustness of the controller.

In this paper, we introduce a chance-constrained reachability analysis technique for a data-driven predictive control of unknown nonlinear systems in the presence of process and measurement noises. Our chance-constrained approach differs from previous related works [5] and [26] in two key respects: first, it does not assume any prior statistical knowledge of the noises; second, it is distributionally robust yet cautiously adaptive with a closed-loop policy. While our feasibility guarantee approach builds on the recursive feasibility methodology introduced in [12], it differs by relying solely on input-output data rather than an explicit system model.

Our proposed approach is a conservativeness-performance trade-off approach that has practical relevance in numerous real-world settings. For instance, in a chemical process context, a conservative controller must be designed to prevent thermal runaway or incomplete reactions under uncertainty; however, this safety margin often reduces efficiency and optimal yield. Chance constraint in such cases offers a principled way to balance safety and performance by allowing a quantifiable risk level, making it possible to constrain the process output probabilistically without being overly conservative. On the other hand, the unavailability or impracticality of full-state feedback is a common reality in chemical process industries. Under these circumstances, a data-driven output-based predictive control design (of the kind proposed in this study) provides a practical framework for directly meeting safety and product quality standards.

1.1. Problem statement

Consider a discrete-time nonlinear system,

$$x(k+1) = f(x(k), u(k)) + w(k), w(k) \in \mathcal{Z}_w \tag{1a}$$

$$y(k) = \mathcal{H}x(k) + v(k), v(k) \in \mathcal{Z}_v \tag{1b}$$

with input-output constraints,

$$u(k) \in \mathcal{U}_k \subseteq \mathcal{Z}_u, \tag{2a}$$

$$y(k) \in \mathcal{Y}_k \subset \mathbb{R}^o, \tag{2b}$$

where $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ is an unknown nonlinear function; $x(k) \in \mathbb{R}^n$, $u(k) \in \mathbb{R}^m$, and $y(k) \in \mathbb{R}^o$ are respectively the state, input, and output variables with time step $k \in \mathbb{Z}_{\geq 0}$; $\mathcal{H} \in \mathbb{R}^o \times \mathbb{R}^n$ is an output matrix of full row rank; $w(k)$ and $v(k)$ are process and measurement noises with zonotope bounds $\mathcal{Z}_w$ and $\mathcal{Z}_v$; $\mathcal{Z}_u$ is a control input zonotope; $\mathcal{U}_k$ and $\mathcal{Y}_k$ are time-varying input and output constraints, respectively.

For a given initial state $\mathcal{X}_0$, we define the set of all possible states and outputs that system (1) can reach after P time steps as (3a) and (3b), respectively.

$$\mathcal{R}_P^x = \left\{ x(P) \in \mathbb{R}^n \;\middle|\; \begin{array}{l} x(k+1) = f(x(k), u(k)) + w(k), \\ x(0) \in \mathcal{X}_0,\ w(k) \in \mathcal{Z}_w, \\ u(k) \in \mathcal{Z}_u, k = 0, \ldots, P-1 \end{array} \right\} \tag{3a}$$

$$\mathcal{R}_P^y = \mathcal{H}\mathcal{R}_P^x + \mathcal{Z}_v \tag{3b}$$

subject to the chance constraint (4) with a violation probability Ξ.

$$\mathbb{P}(\mathcal{R}_P^y \subseteq \mathcal{Y}_P) \geq 1 - \Xi. \tag{4}$$

We aim to over-approximate (3) with the unknown f directly from the available noisy input-output data. The over-approximated sets are then used for an implicit design of a data-driven predictive controller, instead of relying on the usual first principles modeling which is usually error-prone, and in certain cases, not achievable at all. In this regard, the chance constraint in (4) is imposed on the system's reachable output sets to manage the trade-off between safety and precision.

Our main contributions are the following:

1. We propose a zonotopic and deep Koopman-based chance-constrained reachability analysis technique for the data-driven predictive control of (1) to strike the balance between the two conflicting fundamental ideas found in prior deterministic approaches, specifically:
   - Reachable set over-approximation has to be conservative enough to enclose all potential trajectories and provide a guarantee that every essential system behavior is considered.
   - Conservativeness always comes at the expense of efficiency as control inputs may become overly restricted to satisfy overly inflated reachable sets.
2. We introduce a data-driven first-step constraint method combined with a Q-learning algorithm to compute distributionally robust, recursively feasible, and risk-aware safe control policy for noisy nonlinear systems with unknown system dynamics.

2. NOTATION AND PRELIMINARY CONCEPTS

Notation:

$A^{\dagger}$	right Moore-Penrose pseudoinverse of matrix A
X^{-}	trajectory of X starting from the initial time step
X^{+}	trajectory of X starting one step later
$x_{k\|t}$	$(t+k)$-step ahead prediction of x
0_n	n-dimensional all-zeros column vector
1_n	n-dimensional all-ones column vector
I_n	$n \times n$ identity matrix
$I_{x\|y}$	interval of x constructed from known y
$Y_{.,j}$	column j of Y
$y_{\Psi,k\|t}$	$(t+k)$-step ahead prediction of y_{Ψ}
$\mathcal{Y}_{\psi,k}$	constraint of y_{Ψ} at time step k
$\mathcal{Y}_{FS,t}$	first-step constraint of y_{Ψ} at time t
$\|\cdot\|_R^2$	weighted squared norm with weighting matrix R

A Long Short-Term Memory (LSTM) neural network is a sequential neural network that is mainly designed to model long-term dependencies within a data sequence, which cannot be done with the classical recurrent neural networks [19]. The general structure of LSTM neural networks is represented in Figure 1, where the LSTM parameterizations introduced in this section are intended solely for the LSTM descriptions presented here. The cell and hidden states in Figure 1 store long-term and short-term memories, respectively. The forget gate F assigns weights in $[0, 1]$ to keep or discard knowledge from these states. The input gate I uses sigmoid (σ) and tanh functions that respectively determine how much to update and what to update to the cell memory. The output gate O takes the product of the sigmoid's output, and the tanh-squashed updated cell state to yield a new hidden state.

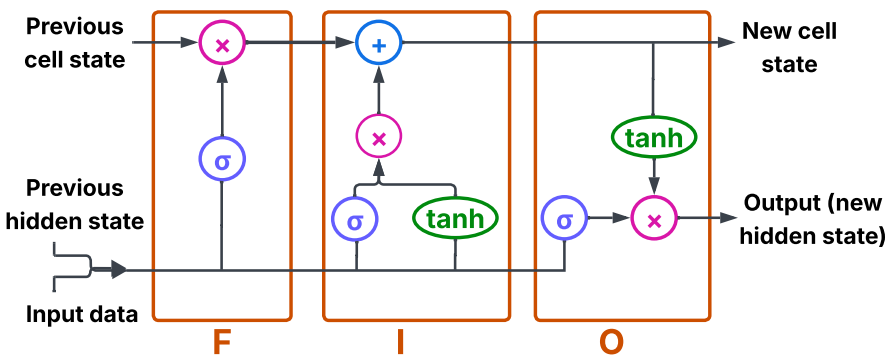

Fig. 1. Structure of LSTM.

Suppose c and h represent the cell and hidden states, respectively; w and w' denote

the weight matrices and b the associated bias vector. The structure illustrated in Figure 1 is mathematically described by (5).

$$F_t = \sigma(w_F x_t + w_F' h_{t-1} + b_F) \tag{5a}$$
$$I_t = \sigma(w_I x_t + w_I' h_{t-1} + b_I) \tag{5b}$$
$$O_t = \sigma(w_O x_t + w_O' h_{t-1} + b_O) \tag{5c}$$
$$c_t = F_t \dot{c}_{t-1} + I_t \dot{\tanh}(w_c x_t + w_c' h_{t-1} + b_c) \tag{5d}$$
$$h_t = O_t \dot{\tanh}(c_t). \tag{5e}$$

Definition 2.1. (Zonotope) A zonotope $\mathcal{Z}_y = \langle c_{\mathcal{Z}}, G_{\mathcal{Z}} \rangle$ of an output $y \in \mathbb{R}^o$ about a center $c_{\mathcal{Z}} \in \mathbb{R}^o$ with generator vector $G_{\mathcal{Z}} = [g^1_{\mathcal{Z}}, \ldots, g^p_{\mathcal{Z}}] \in \mathbb{R}^{o \times p}, p \geq o$ is defined as

$$\mathcal{Z}_y = \left\{ y \in \mathbb{R}^o \middle| y = c_{\mathcal{Z}} + \sum_{i=1}^{p \geq o} \beta^{(i)} g^i_{\mathcal{Z}}, -1 \leq \beta^{(i)} \leq 1 \right\}.$$

Definition 2.2. (Zonotope operations) (Farjadnia et al. [11]) The Minkowski sum and Cartesian product of zonotopes $\mathcal{Z}_1$ and $\mathcal{Z}_2$ are defined as:

$$\begin{aligned} \mathcal{Z}_1 + \mathcal{Z}_2 &= \left\{ z_1 + z_2 \middle| z_1 \in \mathcal{Z}_1, z_2 \in \mathcal{Z}_2 \right\} \\ &= \left\langle c_{\mathcal{Z}_1} + c_{\mathcal{Z}_2}, [G_{\mathcal{Z}_1}, G_{\mathcal{Z}_2}] \right\rangle, \\ \mathcal{Z}_1 \times \mathcal{Z}_2 &= \left\{ \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \middle| z_1 \in \mathcal{Z}_1, z_2 \in \mathcal{Z}_2 \right\}, \\ &= \left\langle \begin{bmatrix} c_{\mathcal{Z}_1} \\ c_{\mathcal{Z}_2} \end{bmatrix}, \begin{bmatrix} G_{\mathcal{Z}_1} & 0 \\ 0 & G_{\mathcal{Z}_2} \end{bmatrix} \right\rangle. \end{aligned}$$

Definition 2.3. (Koopman operator) Let $\mathcal{T}$ be a topological space and the snapshots of data

$$X = [x_1, \ldots, x_M], Y = [y_1, \ldots, y_M]$$

be the trajectory of a discrete-time dynamical system $x^+ = F(x^-)$; $F : \mathcal{T} \to \mathcal{T}$ with $y_k = F(x_k), k = 1, \ldots, M$. Given a vector space of observables $\mathcal{F}$ such that $\Psi : \mathcal{T} \to \mathbb{C}$ and $\Psi \circ F \in \mathcal{F}$ for every $\Psi \in \mathcal{F}$, the Koopman operator $K : \mathcal{F} \to \mathcal{F}$ is defined as:

$$\Psi(x^+) = \mathcal{K}\Psi(x^-) = \Psi \circ F(x^-),$$

where $\circ$ denotes the pointwise function composition.

Definition 2.4. (Extended Dynamic Mode Decomposition (EDMD)) (Korda and Mezić [16]) Let $\Psi_i \in \mathcal{F}, i = 1, 2, \ldots, N$ be sequences of linearly independent observable functions that span a finite-dimensional subspace $\mathbb{F}_N$ of $\mathcal{F}$. The EDMD is a finite-dimensional approximation $\mathcal{K}_{N,M} : \mathbb{F}_N \to \mathbb{F}_N$ that solves the least-squares problem

$$\min_{A \in \mathbb{C}^{N \times N}} \|A\Psi(X) - \Psi(Y)\|_F^2 = \min_{A \in \mathbb{C}^{N \times N}} \sum_{i=1}^{M} \|A\Psi(x_i) - \Psi(y_i)\|_2^2,$$

where $|| \cdot ||_F^2$ represents the squared Frobenius norm, and $\Psi(x) = [\psi_1(x) \ldots \psi_N(x)]^\top$.

3. CHANCE-CONSTRAINED DATA-DRIVEN PREDICTIVE CONTROL

3.1. Data-driven reachable set computation

Let $\mathcal{D}_{[t-T,t]} = (U_{[t-T,t]}, Y_{[t-T,t]})$ be a system trajectory input-output data of length T defined by (6) with a persistently exciting $U_{[t-T,t]}$. In this case, t represents the current time, and T is the depth of memory from which we learn. At $t = 0$, this memory is obtained from offline experiments.

$$\begin{aligned} U_{[t-T,t]} &= [u(t-T), u(t-T+1), \ldots, u(t)] \\ Y_{[t-T,t]} &= [y(t-T), y(t-T+1), \ldots, y(t)]. \end{aligned} \tag{6}$$

For the known input-output measurements (6), we define a zonotope $\mathcal{Z}_\alpha = \langle 0_n, \alpha I_n \rangle$ that inscribes the state space X for a known $\alpha > 0$, see [11]. In particular, for a given output measurement $y(k)$, the zonotope

$$\mathcal{Z}_{x|y} = \mathcal{H}^\dagger(y(k) - \mathcal{Z}_v) + (I_n - \mathcal{H}^\dagger\mathcal{H})\mathcal{Z}_\alpha \tag{7}$$

contains the corresponding possible state $x(k)$, where $\mathcal{H}^\dagger = \mathcal{H}^\top(\mathcal{H}\mathcal{H}^\top)^{-1}$, $c_{x|y} = \mathcal{H}^\dagger(y(k) - c_{\mathcal{Z}_v})$, and $G_{x|y} = \mathcal{H}^\dagger G_{\mathcal{Z}_v}\alpha(I_n - \mathcal{H}^\dagger\mathcal{H})$. The time step k here is the relative counter within the prediction horizon and it represents the prediction index at time t.

The sequences of process and measurement noises corresponding to the input-output trajectory in (6) are denoted by $W_{[t-T,t]}$ and $V_{[t-T,t]}$, respectively; and their stacked matrix zonotopes are denoted by $\mathcal{M}_w = \langle C_{\mathcal{M}_w}, G_{\mathcal{M}_w} \rangle$ and $\mathcal{M}_v = \langle C_{\mathcal{M}_v}, G_{\mathcal{M}_v} \rangle$. We represent the state zonotope estimation (7) as a sequence of intervals using the interval approximation given in (8). This constructs a state trajectory $X_{[t-T,t]} \in \mathbb{R}^n$ from the available input-output trajectory data $\mathcal{D}_{[t-T,t]}$.

$$I_{x|y} = \left[c_{x|y} - \sum_{i=1}^{\gamma_\mathcal{Z}} |g^i_{x|y}|, \quad c_\mathcal{Z} + \sum_{i=1}^{\gamma_\mathcal{Z}} |g^i_{x|y}| \right]. \tag{8}$$

We aim to lift $X_{[t-T,t]}$ into a higher-dimensional space $n' >> n$ by employing a data-driven deep Koopman approach, specifically an LSTM-based Koopman technique to represent the original nonlinear dynamics within the space of observable functions Ψ, learned directly from $X_{[t-T,t]}$.

$$\Psi(x(k)) = \Psi(I_{x|y}(k)) = \begin{bmatrix} \psi_1(I_{x|y}(k)) \\ \psi_2(I_{x|y}(k)) \\ \vdots \\ \psi_{n'}(I_{x|y}(k)) \end{bmatrix}. \tag{9}$$

Equation (9) is learned automatically using the LSTM neural network. For process noise $w_\Psi \in \mathcal{Z}_{w_\Psi}$ in the lifted space, and zonotope $\mathcal{Z}_\eta = \langle 0_{n'}, \eta I_{n'} \rangle$ that inscribes the observable state Ψ for some $\eta > 0$, we define

$$\mathcal{Z}_{\Psi|y_\Psi}(k) = \mathcal{H}^\dagger_\Psi(y_\Psi(k) - \mathcal{Z}_v) + (I_{n'} - \mathcal{H}^\dagger_\Psi\mathcal{H}_\Psi)\mathcal{Z}_\eta, \tag{10}$$

where $\mathcal{H}_\Psi$ represents the output matrix that maps Ψ to y_Ψ. That is, for each output $y_\Psi(k)$, the zonotope in (10) contains all the possible observable states $\Psi(k)$ consistent with $y_\Psi(k)$ at each step k. These sequences of observable functions are intended to represent the complex nonlinearities and hidden behaviors of the original system.

Algorithm 1 Data-driven reachable set computation

Input: $\mathcal{D}_{[t-T,t]}, \mathcal{R}^{y_\Psi}_{t|t}, \mathcal{Z}_u, \mathcal{Z}_w, \mathcal{Z}_v, \alpha$, and LSTM inputs.

Output: Output reachable sets $\mathcal{R}^{y_\Psi}_{k+1|t}, k = 0, \ldots, P-1$.

1: $\mathcal{D}_{[t-T,t]} \xrightarrow{(8)} \mathcal{Z}_{X|Y} \overset{\substack{LSTM \\ EDMD}}{\longrightarrow} \begin{bmatrix} \Psi \\ \mathcal{K} \\ \mathcal{H}_\Psi \end{bmatrix}$

2: $\sigma_- = \min_j \left(Y^+_{.,j} - \mathcal{H}_\Psi \mathcal{K} \begin{bmatrix} \mathcal{H}^\dagger_\Psi (Y^-_{.,j} - (C_{Mv})_{.,j}) \\ U^-_{.,j} \end{bmatrix} \right)$

3: $\sigma_+ = \max_j \left(Y^+_{.,j} - \mathcal{H}_\Psi \mathcal{K} \begin{bmatrix} \mathcal{H}^\dagger_\Psi (Y^-_{.,j} - (C_{Mv})_{.,j}) \\ U^-_{.,j} \end{bmatrix} \right)$

4: $\mathcal{Z}_\sigma = zonotope(\sigma_-, \sigma_+)$

5: **for** $k = 0 : P-1$ **do**

6: $\quad \mathcal{R}^\Psi_{k|t} = \mathcal{H}^\dagger_\Psi (\mathcal{R}^{y_\Psi}_{k|t} - \mathcal{Z}_v) + (I_M - \mathcal{H}^\dagger_\Psi \mathcal{H}_\Psi) \mathcal{Z}_\eta$

7: $\quad \mathcal{R}^{y_\Psi}_{k+1|t} = \mathcal{H}_\Psi \mathcal{K} \big(\mathcal{R}^\Psi_{k|t} \times \mathcal{Z}_u \big) + \mathcal{H}_\Psi \mathcal{Z}_{w_\Psi} + \mathcal{Z}_v + \mathcal{Z}_\sigma$

8: **end for**

Algorithm 1 lifts the original state space into an n'-dimensional observable space, and over-approximates system reachable sets based on the resulting lifted representation. The steps are as follows: In Line 1, the LSTM network learns the Koopman observable functions Ψ, and the EDMD nested in the LSTM loop computes the Koopman operator $\mathcal{K} = \begin{bmatrix} A & B \end{bmatrix}$ and output matrix $\mathcal{H}_\Psi$. Lines 2–4 compute the model mismatch. Lines 5–8 over-approximate output reachable sets over the horizon P based on the knowledge from the reachable observable states. The EDMD solves

$$\mathcal{K} = (\Psi^+ - W^-_\Psi) \begin{bmatrix} \Psi^- \\ U^- \end{bmatrix}^\dagger \text{ and } \mathcal{H}_\Psi = (Y^+ - V^+)(\Psi^+)^\dagger, \tag{11}$$

where W_Ψ represents the process noise matrix associated with w_Ψ.

3.2. Data-driven chance constraint

Since Algorithm 1 over-approximates the set of all possible output reachable sets of the system at each time step, inferring the chance constraint in terms of the output reachable sets instead of the respective predicted outputs enables safe and robust control by considering the system's worst-case scenarios. Based on the lifted system approximation in Algorithm 1, we have

$$\Psi_{k+1|t} = \mathcal{H}^\dagger_\Psi (y_{\Psi,k+1|t} - \mathcal{Z}_v) + (I_M - \mathcal{H}^\dagger_\Psi \mathcal{H}_\Psi) \mathcal{Z}_{\delta,k}, \tag{12a}$$

$$y_{\Psi,k+1|t} = \mathcal{H}_\Psi \mathcal{K} \begin{bmatrix} \mathcal{H}^\dagger_\Psi (y_{\Psi,k|t} - v(k)) \\ u_{k|t} \end{bmatrix} + \mathcal{H}_\Psi w_\Psi(k) + v(k+1) + \mathcal{Z}_\sigma, \tag{12b}$$

subject to the constraints

$$u_{k|t} \in \mathcal{U}_{t+k} \in \mathcal{Z}_u, \tag{13a}$$

$$y_{\Psi,k+1|t} \in \mathcal{Y}_{\Psi,t+k+1} \subset \mathcal{R}^o. \tag{13b}$$

The interval of output reachable set $int(\mathcal{R}^{y_\Psi}_{k+1|t})$ and time-varying output constraint $int(\mathcal{Y}_{t+k+1})$ after $k+1$ time steps are represented in (14). We aim to establish a system criterion that ensures the over-approximated reachable set at each time step satisfies the output constraint with a specified confidence level $1-\Xi$. In this regard, we define the random variable $\hat{\mathcal{R}}_k = ||I_c^{(k)} - J_c^{(k)}||_\infty + ||I_r^{(k)}||_\infty$ and its bound $\|J_r^{(k)}\|_\infty$, where the subscripts $\_c$ and $\_r$ denote the center and the radius of the corresponding interval, respectively.

$$int(\mathcal{R}^{y_\Psi}_{k+1|t}) = \begin{bmatrix} I^l_{k+1|t} & I^u_{k+1|t} \end{bmatrix}, int(\mathcal{Y}_{t+k+1}) = \begin{bmatrix} J^l_{t+k+1} & J^u_{t+k+1} \end{bmatrix}. \tag{14}$$

If $\mathbb{P}$ denotes the set of all plausible probability distributions of the random variable $\hat{\mathcal{R}}$, we define the system chance constraint in (15). This constraint gently manages reachable set predictions in the way that a user-defined risk level Ξ allows our set-based prediction to include conditionally feasible outputs that appear around the boundaries of the system constraints.

$$\underset{\mathbb{P}\in\mathcal{D}}{\text{Inf}}\, \mathbb{P}(\hat{\mathcal{R}}_k \leq \mathcal{V}_k) \geq 1-\Xi, \text{ where } \mathcal{V}_k = \|J_r^{(k)}\|_\infty. \tag{15}$$

Direct utilization of (15) is not practical for online control scenarios, as it requires much computational effort and numerical integration to calculate the exact probability. As a result, we reformulate it using a data-driven nonparametric approximation technique. More specifically, we employed an adaptive kernel density estimation technique with confidence set $\mathcal{D}$,

$$\mathcal{D} = \left\{ \mathbb{P} \in \mathcal{M}^+ : \mathcal{D}_{\chi^2}(h||\hat{h}_{\hat{\mathcal{R}}}) \leq d, h = \frac{d\mathbb{P}}{d\hat{\mathcal{R}}} \right\}, \tag{16}$$

where h and $\hat{h}$ respectively are the true and estimated probability distributions of $\hat{\mathcal{R}}$, $\mathcal{M}^+$ represents the set of all valid probability distributions over the output space $\mathbb{R}^o$, d is the confidence set size, and $\mathcal{D}_{\chi^2}(h||\hat{h}_{\hat{\mathcal{R}}})$ is a divergence given by:

$$\mathcal{D}_{\chi^2}(h||\hat{h}_{\hat{\mathcal{R}}}) = \int_{\mathbb{R}} \frac{(h-\hat{h}_{\hat{\mathcal{R}}})^2}{\hat{h}_{\hat{\mathcal{R}}}} \mathrm{d}\hat{\mathcal{R}}. \tag{17}$$

Both h and $\hat{h}$ are assumed to have the same confidence set $\mathcal{D}$ to establish a framework where the χ^2-divergence between the true and estimated distribution is less than the confidence set size d. For z-data samples $\hat{\mathcal{R}} = [\hat{r}_1, \hat{r}_2, \hat{r}_3, \ldots, \hat{r}_z]$ and bandwidth $b \in \mathbb{R}$,

$$\hat{h}_{\hat{\mathcal{R}}}(\hat{r}) = \frac{1}{zb}\sum_{i=1}^{z} K\left(\frac{\hat{r}-\hat{r}_i}{b}\right), K(\hat{r}) = \frac{1}{\sqrt{2\pi}} exp(-0.5\hat{r}^2). \tag{18}$$

To improve the robustness of our approximation, (19) adjusts Ξ to a d-driven reduced risk level Ξ' which ensures the satisfaction of (15) within $\mathcal{D}$. Equation (19) generally works for χ^2-divergence of risk level $\Xi < 1/2$; its detailed closed-form derivation can be found in [14].

$$\Xi' = \max(\Xi - \frac{\sqrt{d^2 + 4d(\Xi - \Xi^2)} - (1 - 2\Xi)d}{2d + 2}, 0). \tag{19}$$

Suppose $p = \hat{\mathbb{H}}_{\hat{\mathcal{R}}}(\mathcal{V}_k)$ for $\hat{\mathbb{H}}_{\hat{\mathcal{R}}}(\mathcal{V}_k) = \hat{\mathbb{P}}(\hat{\mathcal{R}}_k \leq \mathcal{V}_k)$, where $\hat{\mathbb{H}}_{\hat{\mathcal{R}}}$ and $\hat{\mathbb{P}}$ represent the estimated cumulative distribution function and its corresponding probability measure, respectively. The point that the empirical cumulative distribution function reaches (or exceeds) the probability level p is determined using the $p-$quantile $\mathcal{Q}_p(\hat{\mathcal{R}})$ (20a). Consequently, the equivalent algebraic equation of (15) can be reformulated as (20b). For further details on quantile constraint reformulation, see [22].

$$\mathcal{Q}_p(\hat{\mathcal{R}}) = \hat{\mathbb{H}}^{-1}_{\hat{\mathcal{R}}}(p) = \text{Inf}\{\hat{r} | \hat{\mathbb{H}}_{\hat{\mathcal{R}}}(\hat{r}) \geq p\} = \mathcal{V}_k \tag{20a}$$

$$\mathcal{V}_k \geq \hat{\mathbb{H}}^{-1}_{\hat{\mathcal{R}}}(1 - \Xi'). \tag{20b}$$

We quantified the confidence set size d using the uncertainty in the estimated probability distribution through a bootstrapping resampling technique. Let $\mathbb{P}(t_{\hat{\mathcal{R}}} \leq q_{\Xi/2}) = \Xi/2$ and $\mathbb{P}(t_{\hat{\mathcal{R}}} \leq q_{1-(\Xi/2)}) = 1 - (\Xi/2)$ respectively be the lower and upper quantiles of the t-statistic $t_{\hat{\mathcal{R}}}$ defined by:

$$t_{\hat{\mathcal{R}}}(\hat{r}) = \frac{\hat{h}_{\hat{\mathcal{R}}}(\hat{r}) - \mathbb{E}[\hat{h}_{\hat{\mathcal{R}}}(\hat{r})]}{s_{\hat{\mathcal{R}}}(\hat{r})}, \tag{21}$$

where

$$s_{\hat{\mathcal{R}}}(\hat{r}) = \sqrt{\frac{1}{zb^2} \sum_{i=1}^{z} K\left(\frac{\hat{r} - \hat{r}_i}{b}\right)^2 - \frac{\hat{h}_{\hat{\mathcal{R}}}(\hat{r})^2}{z}}, \tag{22}$$

$$\hat{h}_{\hat{\mathcal{R}}}(\hat{r}) - s_{\hat{\mathcal{R}}} q_{1-(\Xi/2)} \leq \mathbb{E}[\hat{h}_{\hat{\mathcal{R}}}(\hat{r})] \leq \hat{h}_{\hat{\mathcal{R}}}(\hat{r}) - s_{\hat{\mathcal{R}}} q_{\Xi/2}. \tag{23}$$

We intend to determine d from the confidence interval (23) such that $t_{\hat{\mathcal{R}}}$ is unbiased with $q_{\Xi/2} \leq 0$ and $q_{1-(\Xi/2)} \geq 0$. By taking different bootstrapped datasets $\hat{\mathcal{R}}^*$ that are constructed by performing a random sampling from the original samples with replacement, (18) and (22) are computed for each sample. More specifically, bootstrapped estimate $\hat{h}^*_{\hat{\mathcal{R}}}$ and bootstrapped standard deviation $s^*_{\hat{R}}$ are calculated by replacing the data points $\hat{r}_i$ with the bootstrapped data points $\hat{r}^*_i$. Based on this, the unbiased t-statistic $t^*_{\hat{\mathcal{R}}}$ is computed as:

$$t^*_{\hat{\mathcal{R}}}(\hat{r}) = \frac{\hat{h}^*_{\hat{\mathcal{R}}}(\hat{r}) - \hat{h}_{\hat{\mathcal{R}}}(\hat{r})}{s^*_{\hat{\mathcal{R}}}(\hat{r})}. \tag{24}$$

Consequently, using the quantiles $\mathbb{P}(t^*_{\hat{\mathcal{R}}} \leq q^*_{\Xi/2}) = \Xi/2$ and $\mathbb{P}(t^*_{\hat{\mathcal{R}}} \leq q^*_{1-(\Xi/2)}) = 1-(\Xi/2)$ associated with (24), the adjusted confidence set size d is calculated as:

$$d = \mathcal{Q}_{1-\Xi}((\mathbb{D}_U - \mathbb{D}_L)^2), \tag{25}$$

where

$$\mathbb{D}_U = \hat{h}_{\hat{\mathcal{R}}}(\hat{r}) - s_{\hat{\mathcal{R}}} q^*_{\Xi/2},$$
$$\mathbb{D}_L = \hat{h}_{\hat{\mathcal{R}}}(\hat{r}) - s_{\hat{\mathcal{R}}} q^*_{1-(\Xi/2)}.$$

3.3. Control design

This section introduces a chance-constrained data-driven predictive control technique for unknown nonlinear systems under process and measurement noises. At each time step t, Algorithm 1 is executed using past input-output data. The results are then used as an implicit system representation to construct a data-driven receding horizon control problem (26), which computes the control input trajectory subject to system constraints over the horizon P.

$$\underset{y_\Psi, u}{\text{Min}} \quad \mathcal{J} = \sum_{k=0}^{P-1} ||y_{\Psi,k+1|t} - y_{r,k+1|t}||^2_R + ||u_{k|t} - u_{r,k|t}||^2_S \tag{26a}$$

$$\text{subject to} \quad \mathcal{V}_{t+k+1} \geq \hat{\mathbb{H}}^{-1}_{\hat{\mathcal{R}}_{k+1|t}}(1 - \Xi') \tag{26b}$$

$$y_{\Psi,k+1|t} \in \mathcal{R}^{y_\Psi}_{k+1|t} \tag{26c}$$

$$u_{k|t} \in \mathcal{U}_{t+k} \tag{26d}$$

$$y_{\Psi,1|t} \in \mathcal{Y}_{FS,t}. \tag{26e}$$

The stage cost function in (26a) penalizes the deviations of the predicted outputs $y_{\Psi,k+1|t}$ and inputs $u_{k|t}$ from their time-varying references $y_{r,k+1|t}$ and $u_{r,k|t}$ with positive definite weight matrices R and S, respectively. For a deviation ζ, $\|\zeta\|^2_R = \zeta^\top R\zeta$ and $\|\zeta\|^2_S = \zeta^\top S\zeta$. The input-output references are set to be adaptive with the help of a Q-learning algorithm. They are updated iteratively to adjust the desired system behavior; while remaining consistent with the input-output constraints. This creates a learning-based controller that empowers the control system with crucial intelligence and adaptability layers to maintain performance and safety under noisy environments.

Equation (26b) is the reformulated deterministic actionable constraint of the probabilistic constraint (15). It assures the satisfaction of the output constraint within a desired confidence level under noise and system uncertainties. The constraint in (26c) makes the reachable sets conservative, yet likely to be shrunk by (26b), not always to plan the worst-case scenarios. The control input constraint in (26d) restricts the actions the control inputs must obey. Constraint (26e) is the first-step constraint, which serves a critical role in ensuring the recursive feasibility of the proposed data-driven predictive control.

Definition 3.1. A subset $\mathcal{Y}_R \subseteq \mathcal{Y}_\Psi$ is considered a robust control invariant set if there exists a control input u such that (27) holds for all $y_\Psi \in \mathcal{Y}_R, v \in \mathcal{Z}_v$ and $w_\Psi \in \mathcal{Z}_{w_\Psi}$.

$$\mathcal{H}_\Psi \mathcal{K} \begin{bmatrix} \mathcal{H}^\dagger_\Psi (y_\Psi - v) \\ u \end{bmatrix} + \mathcal{H}_\Psi w_\Psi + v + \mathcal{Z}_\sigma \in \mathcal{Y}_R. \tag{27}$$

Assumption 1. Given the time-varying output constraint $\mathcal{Y}_{\Psi,k}$, there exists a robust output-invariant proximity $\tilde{\mathcal{Y}}_\Psi$ and initial auxiliary output $\tilde{y}_{\Psi,0}$ such that all outputs in $\tilde{\mathcal{Y}}_\Psi$ around the solution of the system $\tilde{y}_\Psi(k+1) = \Gamma \tilde{y}_\Psi(k)$ are in $\mathcal{Y}_{\Psi,k}$ for $\Gamma = \mathcal{H}_\Psi A \mathcal{H}_\Psi^\dagger$. That is, it holds that the Minkowski sum $\tilde{y}_\Psi(k) + \tilde{\mathcal{Y}}_\Psi \subseteq \mathcal{Y}_{\Psi,k}$.

The condition for Assumption 1 to hold in practice is that the control objective and the available data should allow the construction of a safe trajectory and an invariant set around it, such that their Minkowski sum is contained by the time-variant constraint. In this case, the robust control-invariant set $\tilde{\mathcal{Y}}_\Psi$ is computed offline using iterative backward reachability with $\mathcal{H}_\Psi$, and $\mathcal{Y}_{\Psi,k} = \mathcal{Y}_k + \phi, \phi \in (\mathcal{Z}_\sigma + \mathcal{Z}_v)$.

We define the time-varying first-step constraint $\mathcal{Y}_{FS,t}$ in (28), which allows flexible trajectories by avoiding a strict terminal-output constraint that requires the entire prediction horizon to converge to a fixed terminal set.

$$\mathcal{Y}_{FS,t} = \left\{ y_{\Psi,1|t} \;\middle|\; \begin{array}{l} \forall v(t) \in \mathcal{Z}_v \text{ and } w_\Psi(t) \in \mathcal{Z}_{w_\Psi}, \exists \mathcal{U}_t \subseteq \mathcal{Z}_u : \\ \left(y_{\Psi,P|t} + \Sigma_P\right) \in \tilde{y}_\Psi(t+P) + \tilde{\mathcal{Y}}_\Psi \end{array} \right\}, \tag{28}$$

where Σ_P calculates the cumulative impact of noises and model mismatches over the entire prediction horizon P, and it is defined by:

$$\Sigma_P = \sum_{i=1}^{P} \sum_{j=0}^{i-1} \Gamma^j \left(\mathcal{H}_\Psi w_\Psi(t+i-1-j) + v(t+i-j) + \mathcal{Z}_\sigma \right) - \sum_{i=1}^{P} \Gamma^i v(t). \tag{29}$$

Algorithm 2 solves (26a) subject to the constraints (26b) – (26e) over a moving prediction window. It finds the optimal control input directly from noisy data, bypassing the model identification step. It provides a distributionally robust chance constraint satisfaction and recursive feasibility guarantees under unknown statistical properties of the noises, as shown in Theorem 3.2. At each time step, Algorithm 2 starts by computing the reachable sets of the system using Algorithm 1. Within the prediction horizon P, it evaluates the deviation between the current input-output trajectories and their adaptive reference counterparts, which are iteratively updated using the Q-learning policy. The Q-learning phase explores or exploits input-output adjustments to improve control performance over time. The adaptive references then guide the receding horizon, initialized with a candidate solution (31). The algorithm then updates internal states and observed trajectories based on the current optimal control action, enabling continuous learning and robust prediction in the presence of noise.

The proposed Q-learning module's exploration strategy follows an ε-greedy policy. It selects a random action a from the action space with probability ε; otherwise, it finds the best a that maximizes the current Q-value. It uses the negative of (26a) as its reward function to penalize deviations from the desired trajectory. Its learning process continues through the Q-update (30) with a learning rate α_q and discount factor γ_q, allowing a stable but gradual policy improvement by decaying ε according to $\varepsilon_{\text{decay}}$. That is, its exploration rate decays progressively to reduce exploration over time while respecting $\varepsilon_{\min}$, encouraging more exploitation of the learned Q-values as the agent gains confidence in its policy. The Q-update is given by (30), where $r_{k|t}$ represents the $t+k$-reward.

$$Q(s_{k+1|t}, a_{k+1|t}) \leftarrow Q(s_{k|t}, a_{k|t}) + \alpha_q \left(r_{k|t} + \gamma \max_a Q(s_{k+1|t}, a) - Q(s_{k|t}, a_{k|t}) \right). \tag{30}$$

Algorithm 2 Chance-constrained nonlinear zonotopic predictive control

Input: Algorithm 1 inputs, input-output constraints, time horizon P, Q-learning inputs
Initialize: $Q(s,a), u_{\text{prev}}, y_{\Psi,\text{prev}}$
while $t \in \mathbb{Z}_{\geq 0}$ **do**
  Execute Algorithm 1
  for $k = 0 : P-1$ **do**
    $e_{u,k|t} = u_{k|t} - u_{r,k|t}, e_{y,k+1|t} = y_{\Psi,k+1|t} - y_{r,k+1|t}$
    $s_{k|t} = \left[u_{r,k|t}, e_{u,k|t}, y_{r,k+1|t}, e_{y,k+1|t}\right]$
    if $rand < \varepsilon$ **then**
      Choose a random action $a_{k|t} = (\Delta u_r, \Delta y_r) \in a$
    else
      $a_{k|t} = \arg\max_a Q(s_{k|t}, a)$
    end if
    $y_{r,k+1|t} = y_{\Psi,k+1|t} + \Delta y_r, \quad u_{r,k|t} = u_{k|t} + \Delta u_r$
    $y_{\Psi,\text{cand}} \leftarrow y_{\Psi,\text{prev}}$
    Solve (26) initialized with:
    $y_\Psi \leftarrow y_{\Psi,\text{cand}}$
  end for
  Apply $u^*_{t|t}$ to the system
  $y_{\Psi,\text{prev}} \leftarrow y_\Psi$
  $t = t+1$
  Update the trajectory $(u_{[t-T,t]}, y_{[t-T,t]})$
end while

The control optimization at time $t+1$ is set to begin from the extended version of the previous solution at t as shown in (31). It determines a control sequence starting from $y_{\Psi,0|t+1} = y^*_{\Psi,1|t} + \Sigma_1$ and encodes the entire forward reachability tube.

$$Y_{\Psi,t+1} = \begin{bmatrix} y_{\Psi,0|t+1} \\ \vdots \\ y_{\Psi,P-1|t+1} \\ y_{\Psi,P|t+1} \end{bmatrix} = \begin{bmatrix} y^*_{\Psi,1|t} + \Sigma_1 \\ \vdots \\ y^*_{\Psi,P|t} + \Sigma_P \\ y_{\Psi,P|t+1} \end{bmatrix}. \tag{31}$$

Theorem 3.2. (Recursive feasibility) Suppose that Assumption 1 holds, and

$$Y^*_{\Psi,t} = [y^*_{\Psi,0|t}\ y^*_{\Psi,1|t}\ \cdots\ y^*_{\Psi,P|t}]^T$$

is the solution of (26) at time t, and the subsequent candidate solution is defined by (31). Then, (26) is recursively feasible and distributionally robust.

Proof. Let $\{y_{\Psi,k+1|t}, u_{k|t}\}_{k=0}^{P-1}$ be the feasible solution of (26) at time t. We want to prove that such solution sequences exist at $t+1$. Based on (31), we have

$$y_{\Psi,0|t+1} = y_{\Psi,1|t} + \Sigma_1.$$

With this initial output, (28) guarantees the existence of a trajectory such that $y_{\Psi,P|t} + \Sigma_P \in \tilde{y}_\Psi(t+P) + \tilde{\mathcal{Y}}_\Psi$. Since $y_{\Psi,P|t} + \Sigma_P = y_{\Psi,P-1|t+1}$, it follows that $y_{\Psi,P-1|t+1} \in \tilde{y}_\Psi(t+P) + \tilde{\mathcal{Y}}_\Psi$.

For a predicted output $y_{\Psi,k|t} \in \tilde{y}_\Psi(t+k) + \tilde{\mathcal{Y}}_\Psi$, we have $y_{\Psi,k|t} = \tilde{y}_\Psi(t+k) + y_\Psi$ for some $y_\Psi \in \tilde{\mathcal{Y}}_\Psi$. Employing this in (12b) yields,

$$\begin{aligned} y_{\Psi,k+1|t} &= \Gamma[\tilde{y}_\Psi(t+k) + y_\Psi] - \Gamma v(k) + \mathcal{H}_\Psi B u_{k|t} + \mathcal{H}_\Psi w_\Psi(k) \\ &\quad + v(k+1) + \mathcal{Z}_\sigma \\ &= \tilde{y}_\Psi(t+k+1) + \Gamma y_\Psi - \Gamma v(k) + \mathcal{H}_\Psi B u_{k|t} + \mathcal{H}_\Psi w_\Psi(k) \\ &\quad + v(k+1) + \mathcal{Z}_\sigma \\ &\in \tilde{y}_\Psi(t+k+1) + \tilde{\mathcal{Y}}_\Psi, \text{using Definition 3.1.} \end{aligned}$$

Hence, for every time step t and each k in the prediction horizon, (32) is valid.

$$y_{\Psi,k|t} \in \tilde{y}_\Psi(t+k) + \tilde{\mathcal{Y}}_\Psi \implies y_{\Psi,k+1|t} \in \tilde{y}_\Psi(t+k+1) + \tilde{\mathcal{Y}}_\Psi. \tag{32}$$

Based on (32), $y_{\Psi,P|t+1} \in \tilde{y}_\Psi(t+P+1) + \tilde{\mathcal{Y}}_\Psi$ follows from the previously established step $y_{\Psi,P-1|t+1} \in \tilde{y}_\Psi(t+P) + \tilde{\mathcal{Y}}_\Psi$. Therefore,

$$y_{\Psi,1|t} \in \mathcal{Y}_{FS,t} \implies y_{\Psi,1|t+1} \in \mathcal{Y}_{FS,t+1}. \tag{33}$$

The first-step constraints $\mathcal{Y}_{FS,t}$ and $\mathcal{Y}_{FS,t+1}$ implicitly enforce the existence of control input sequences $U_t = \{u_{0|t}, \ldots, u_{P-1|t}\}$ and $U_{t+1} = \{u_{0|t+1}, \ldots, u_{P-1|t+1}\}$, respectively. Hence, (26d) can be directly ensured from (33). On the other hand, each reachable set $\mathcal{R}^{y_\Psi}_{k+1|t+1}$ is computed based on $\mathcal{R}^{y_\Psi}_{k|t+1}$ as shown in Algorithm 1. Given the fact that $y_{\Psi,k|t+1}$ is constructed by a disturbance shift of $y_{\Psi,k+1|t}$ using (31); and $\mathcal{R}^{y_\Psi}_{k+1|t+1}$ over-approximates all disturbance-driven variations starting from $y_{\Psi,1|t+1}$ that lie in the terminal tube;

$$y_{\Psi,k+1|t} \in \mathcal{R}^{y_\Psi}_{k+1|t} \implies y_{\Psi,k+1|t+1} \in \mathcal{R}^{y_\Psi}_{k+1|t+1}.$$

Since the Minkowski sum $\mathcal{H}_\Psi \mathcal{Z}_{w_\Psi} + \mathcal{Z}_v + \mathcal{Z}_\sigma$ at each time step $t+1$ has surely a non-zero volume, the volume of $\mathcal{R}^{y_\Psi}_{k+1|t+1}$ is always lower bounded by a function of the previous volume at t. Thus, if

$$\mathcal{V}_{t+k+1} \geq \hat{\mathbb{H}}^{-1}_{\hat{\mathcal{R}}_{k+1|t}}(1-\Xi') \equiv \mathbb{P}\left(\mathcal{R}^{y_\Psi}_{k+1|t} \subseteq \mathcal{Y}_{\Psi,t+k+1}\right) \geq (1-\Xi')$$

holds at time step t, then

$$\mathcal{V}_{t+k+2} \geq \hat{\mathbb{H}}^{-1}_{\hat{\mathcal{R}}_{k+1|t+1}}(1-\Xi') \equiv \mathbb{P}\left(\mathcal{R}^{y_\Psi}_{k+1|t+1} \subseteq \mathcal{Y}_{\Psi,t+k+2}\right) \geq (1-\Xi')$$

follows. That is, the volume of the reachable set at $t+1$ remains large enough to satisfy (26b), provided it was satisfied at t.

This establishes that constraints (26b) – (26e) remain satisfied at subsequent time steps, provided that (26) admits a feasible solution at the current time step. It can thus be concluded that the distributional robust constraint satisfaction of $\mathcal{Y}_{\Psi,k}$ at each time step is guaranteed subject to the feasibility of (26). □

4. NUMERICAL EXAMPLES

4.1. Experimental design

We consider a stirred-tank reactor system and a cart–damper–spring system to demonstrate the practical implementation of the proposed data-driven predictive control approach. It is important to note that the corresponding system dynamics, (36) and (37), are utilized solely for data generation purposes to establish ground truth, not for model-based analysis. For each system, we generated $1,000$ distinct trajectories with a sequence length of 10, total sample $T = 10,000$. Each trajectory is normalized using the z-score normalization to ensure numerical stability. We utilized the Mosek solver along with the toolboxes: COntinuous Reachability Analyzer (CORA), Multi-Parametric Toolbox (MPT), and Yet Another Linear Matrix Inequality Parser (YALMIP) in MATLAB 2023a software.

The LSTM architecture consists of a sequence input layer, 2 stacked LSTM layers $(n1, n2)$, a regression layer, a dropout layer with rate dr to prevent overfitting, and the EDMD module embedded within the LSTM. A systematic grid search is performed across multiple networks and training configurations. Since our benchmark systems are both 2-dimensional, the LSTM is trained with conservative sizes $n1 \in \{10, 20\}$ and $n2 \in \{20, 30\}$ to avoid unnecessary expressive networks. Initial learning rate $lr \in \{10^{-3}, 10^{-2}\}$, $dr \in \{0.2, 0.3\}$, and mini-batch size $mb \in \{50, 100\}$ are applied. Each configuration defines a unique LSTM model, and a total of 32 different configurations are trained exhaustively.

The proposed LSTM approach follows a Koopman-supervised learning approach, instead of the usual train–validation paradigm. Because if the lifting efficiency is evaluated by the internal LSTM loss, then the learned observable functions cannot guarantee the expected linear evolution in the lifted space. As a result, the LSTM internally learns the observable functions, and the network model selection is carried out based on the Koopman consistency metric (34). This yields an accurate linear representation of the original nonlinear system dynamics in the learned lifted space. The Adam optimizer with fixed 50-epoch runs is utilized, where the learning rate lr is set to decay geometrically by a factor of 0.95 every 10 epochs to enhance stable convergence. An $L2$ regularization term of 0.01 is applied to discourage overfitting.

Based on [15], an 8-dimensional lifted state space yields an efficient Koopman-based implicit linear representation of (36). Since the nonlinearity of (37) is no greater than that of (36), we lifted both nonlinear systems into an 8-dimensional space. It is worth noting that our current approach is scalable to any lifting size, with the only compromise being an expected computational expense. The Mean Absolute Error (MAE) that measures the Koopman consistency of each candidate network is defined in (34) for the LSTM-learned Ψ and Koopman-predicted $\hat{\Psi}$.

$$\mathrm{MAE} = \frac{1}{8T} \sum_{j=1}^{T} \sum_{i=1}^{8} |\Psi_{i,j} - \hat{\Psi}_{i,j}|. \tag{34}$$

We used Shannon's entropy difference (35a), mutual information (35b), and Kullback–Leibler divergence (35c) as post-training metrics. The first metric measures the uncertainty and variability between the original and lifted states. A large positive H suggests

an information loss, and a negative H indicates an additional variability in the lifted space. The second metric quantifies the shared information between the original and lifted states, where a large I means the lifted system retains more information about the original system. The third metric compares the probability distributions associated with the original and lifted states. A large D suggests that the lifted state fails to reproduce output statistics.

We set the thresholds: $|H| \leq 0.3, I \geq 0.7$ and $D \leq 1.0$. If the lifting fails to pass these thresholds, then the learned observables do not fully encapsulate the underlying state dynamics. In that case, the LSTM training and Koopman approximation process will be repeated with parameter tuning until the resulting Ψ satisfies the thresholds. Further mathematical details on (35a) – (35c) can be found in [24].

$$H = -\sum P_X(x)\log_2(P_X(x)) + \sum P_\Psi(\psi)\log_2(P_\Psi(\psi)) \tag{35a}$$

$$I = -\sum P_\Psi(\psi)\log_2(P_\Psi(\psi)) + \sum P_X(x)P_{\Psi|x}(\psi)\log_2(P_{\Psi|x}(\psi)) \tag{35b}$$

$$D = \sum P_Y(y)\ \log\left(\frac{P_Y(y)}{P_{Y_\Psi}(y_\Psi)}\right). \tag{35c}$$

After the lifted observable state representation is obtained, the projection matrix $\mathcal{H}_\Psi$ is computed using (11). Then, Algorithm 1 is applied to over-approximate all the possible future system outputs over the prediction horizon P. Algorithm 2 alters this robust over-approximation into a distributionally robust scenario by imposing the chance constraint (26b). Finally, Algorithm 2 solves (26) for both (36) and (37) according to their specified input-output constraints and parametrization setups described in their respective forthcoming sections.

4.2. Stirred-tank reactor system

The nonlinear discrete-time state-space model of the considered stirred-tank reactor system can be found in [11] and [20] but with different implementations. For the present case, we used [11]'s representation shown in (36):

$$\begin{cases} x_1(k+1) = & (1/\lambda_1)(1.0075 - \lambda_2 exp(\frac{\beta}{x_2(k)})\tau)x_1(k)) + u_1(k)\tau, \\ x_2(k+1) = & (1/\lambda_3)(5.25 + 0.9775x_2(k) + \rho x_1(k)exp(\frac{\beta}{x_2(k)}) \\ & -0.0945x_1(k) - 0.216x_2(k)) + u_2(k)\tau, \end{cases} \tag{36}$$

where $\lambda_1 = 1.0075, \lambda_2 = 7.2 \times 10^{10}, \lambda_3 = 1.0225, \tau = 0.015, \beta = -8750$, and $\rho = 1.5 \times 10^{13}$.

We used: $\mathcal{H} = [1, 0.001; -0.01, 1], \alpha = 22, \mathcal{X}_0 = \langle [-2; -20.5], diag(0.01, 0.2)\rangle, \mathcal{Z}_w = \langle 0_n, 2 \times 10^{-4} I_n\rangle$, and $\mathcal{Z}_v = \langle 0_o, 10^{-3} I_o\rangle$. The input and output constraints are set to be $\mathcal{Z}_u = \langle [0; 0], diag(5, 3)\rangle$ and $\mathcal{Y} = \text{int}\,([-3; -22], [0.25; 2.7])$, respectively. For the Q-learning parameters: $a = \text{linspace}(-1, 1, 10), \alpha_q = 0.05, \gamma = 0.95, \varepsilon = 1.0, \varepsilon_{\text{decay}} = 0.995$, and $\varepsilon_{\min} = 0.01$ are utilized.

Figure 2 illustrates how effectively the deep Koopman approach lifts the original nonlinear dynamics into a higher-dimensional observable space where linear evolution of

the system can be approximated. The close correspondence between the actual observable directly learned by the LSTM and Koopman-predicted observable across time steps shows that the learned Koopman-invariant subspace accurately captures the essential system behavior. This verifies that the lifted system can be used for efficient analysis of the original nonlinear system in a more tractable linear framework.

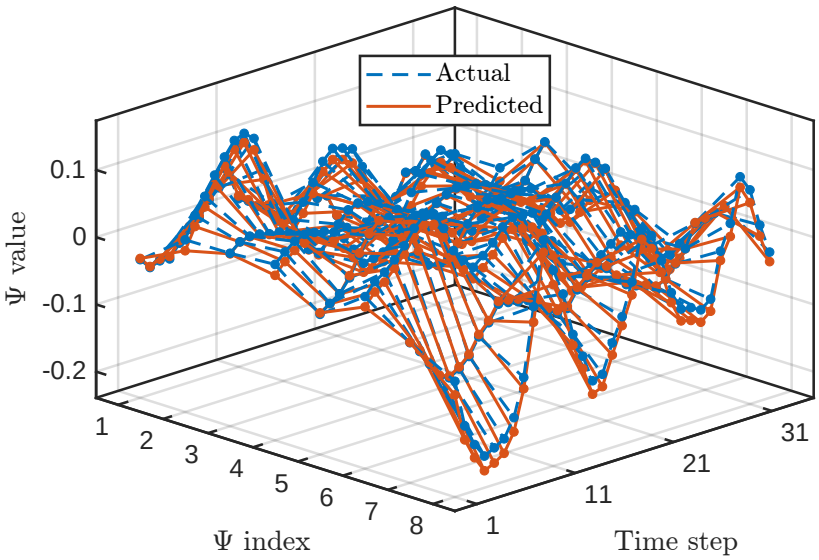

Fig. 2. Evolution of Ψ in the Koopman space.

The effect of the bandwidth selection on the performance of our estimation is illustrated in Figure 3. The smooth curves represent the nonparametric kernel density estimate, and the vertical dashed lines denote the risk-adjusted quantile computed based on (20b). Figure 3 shows a higher maximum density for a low bandwidth b_1, and heavy smoothing for a large bandwidth b_2. The tightness and looseness gap between the risk bound from the bulk of the corresponding estimate shows that bandwidth selection has a direct impact on whether candidate solutions satisfy the probabilistic constraint. Since the risk bounds for all three cases lie in a low-density tail, we can say that the quantile is robust to small smoothing changes.

The confidence size d monotonically decreases for an increasing Ξ in Figure 4 as the estimate should tolerate more uncertainty for a looser risk policy. Similarly, the risk bound $\hat{\mathbb{H}}^{-1}_{\hat{\mathcal{R}}}(1-\Xi')$ gets smaller as risk loosens. Figure 5 illustrates the dynamic evolution of output reachable sets over time with different risk levels. As one might think in a well-designed reachability analysis, Algorithm 1 yields greater predicted output ranges and constraint violations for $\Xi = 0.2$ relative to $\Xi = 0.1$. In this case, high Ξ should be fine-tuned based on the desired margin of safety of the over-approximation. In Figure 6, the predicted output gently follows the reference trajectory across the entire time horizon. The minimal deviation between the two curves suggests that the Q-learning

phase of Algorithm 2 has effectively learned to generate references that guide the system dynamics toward the desired behavior. The root-mean-square error of this deviation is 0.2797, which is much smaller than the standard deviation of the actual output 0.9271. This shows the strong predictive accuracy of Algorithm 2 with small prediction errors compared to the actual system fluctuations.

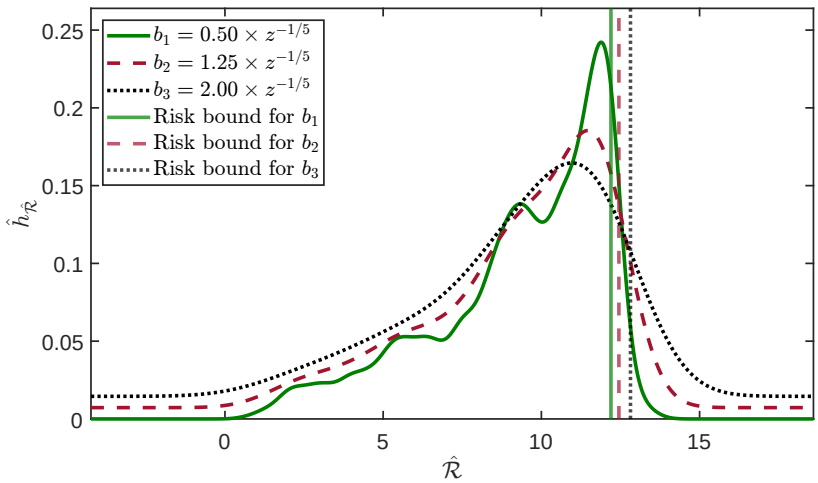

Fig. 3. Sensitivity of $\hat{h}_{\hat{\mathcal{R}}}$ to b with $\Xi = 0.1$.

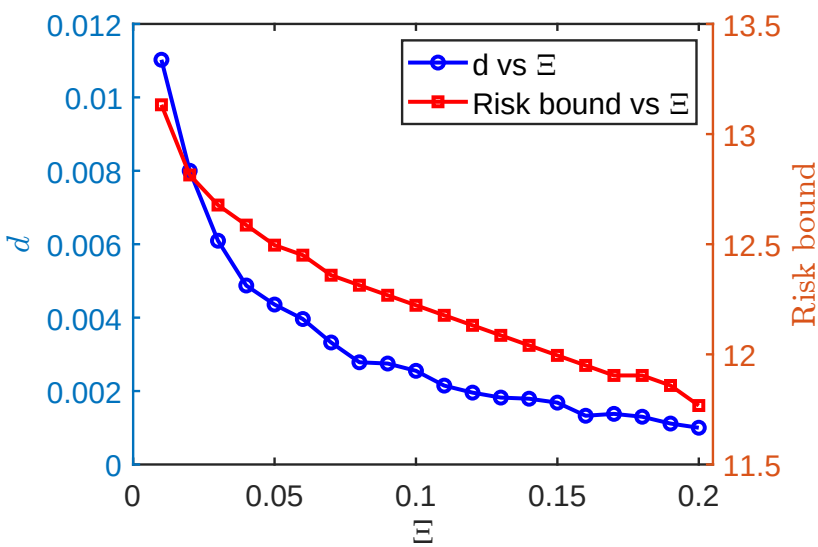

Fig. 4. Sensitivity of confidence size and risk bound to Ξ.

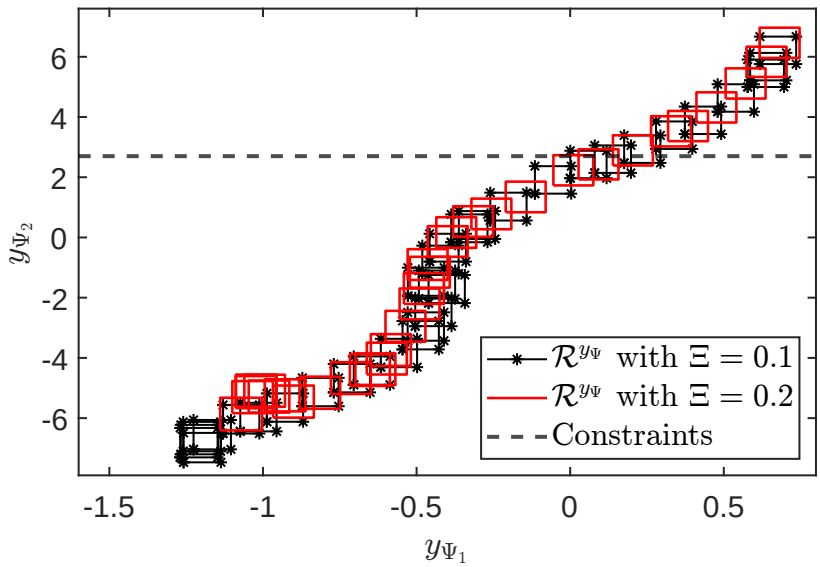

Fig. 5. Output reachable sets.

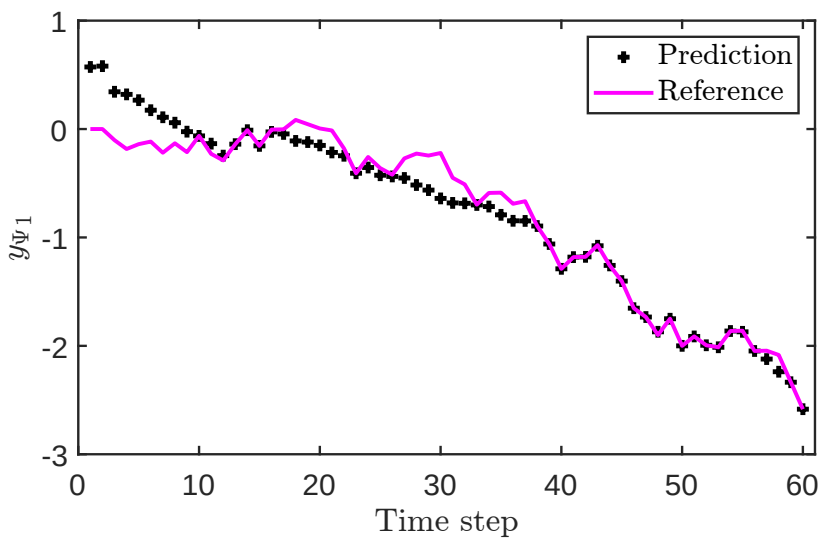

Fig. 6. Evolution of output with adaptive reference for $\Xi = 0.1$

In Figure 7, the optimal control input tracks the references effectively across both risk levels 0.1 and 0.2. However, the optimal control input associated with the higher risk level exhibits a delayed convergence compared to the lower risk case. This reflects the controller's sensitivity to noises and uncertainties, with more conservative behavior evident at the lower risk level. Overall, it illustrates how the controller adapts its actions based on the specified risk, maintaining effective performance while balancing robustness and flexibility.

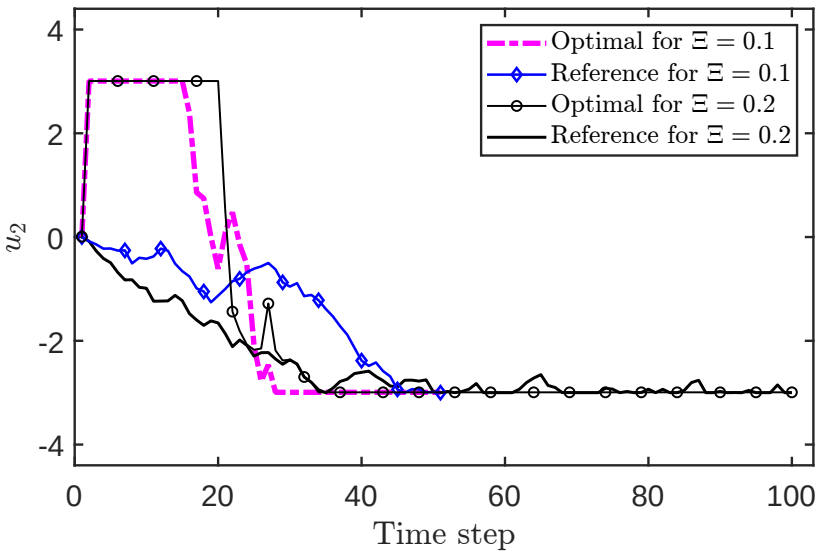

Fig. 7. Optimal control input.

4.3. Cart–damper–spring system

To further evaluate the efficacy and generalizability of the proposed data-driven predictive control approach, we consider the cart–damper–spring system (37) given in [27]. The system dynamics is given by:

$$\begin{cases} x_1(k+1) = & x_1(k) + k_s \cdot x_2(k) + w_1(k), \\ x_2(k+1) = & x_2(k) + k_s \cdot \Big(-\frac{\Upsilon}{m_c} \exp^{-x_1(k)} x_1(k) \\ & -\frac{c_d}{m_c} x_2(k) + \frac{u(k)}{m_c} + \frac{w_2(k)}{m_c} \Big). \end{cases} \tag{37}$$

In our simulation, we used $k_s = 0.1, \Upsilon = 0.9, m_c = 1.25, c_d = 0.42, \mathcal{Z}_u = \langle 0, 0.4 \rangle, \mathcal{Y} = \text{int}\left([-1;-2],[1;2]\right), \alpha = 3, \mathcal{X}_0 = \langle [0.3;-1], \text{diag}(0.02;0.1) \rangle, R = \text{diag}(0.06, 0.06), S = 0.001, \mathcal{Z}_w = \langle 0_2, \text{diag}(10^{-4}, 10^{-3}) \rangle$, and $\mathcal{Z}_v = \langle 0_2, \text{diag}(10^{-3}, 10^{-3}) \rangle$. For the Q-learning parameters: we used $a = \text{linspace}(-0.5, 0.5, 10), \alpha_q = 0.05, \gamma = 0.95, \varepsilon = 1.0, \varepsilon_{\text{decay}} = 0.995$, and $\varepsilon_{\min} = 0.01$. Note that the implementation of the process noise w in our illustrative example differs from [27]. Furthermore, we set the output matrix

$\mathcal{H} = [1, 0.05; -0.02, 1]$.

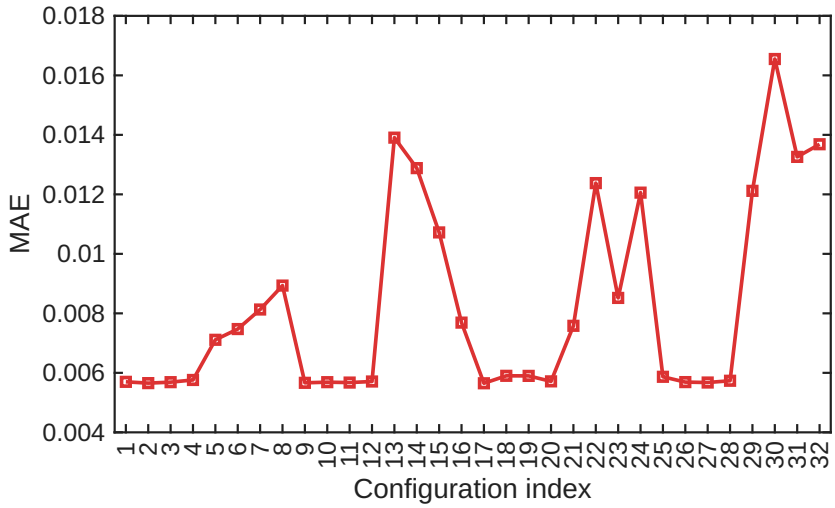

Fig. 8. Koopman consistency error across 32 LSTM configurations.

Figure 8 shows the Koopman consistency error of the employed deep Koopman approach across the 32 configured LSTM networks. The non-uniform distribution of MAE across the configurations highlights Koopman's sensitivity to a hyperparameter selection. The learned observable functions corresponding to the configuration with the lowest Koopman consistency error represents the system in the lifted space.

Figure 9 depicts the lifting accuracy as evidenced by the coherence between the actual LSTM observables and Koopman-predicted observables. The close correspondence and identical patterns between the two trajectories show that the evolution of the observable state is consistent under the approximated Koopman dynamics. Figure 10 illustrates the fundamental trade-off between the performance and robustness of the proposed approach. That is, the controller relinquishes a robust worst-case scenario and follows a probabilistic risk-aware framework by excluding statistically low-probability outputs at the distribution tails. The actual system outputs are gently over-approximated by the computed output reachable sets, yet they breach the boundaries of the reachable sets. This underscores our aim of designing a distributionally robust controller that tolerates infrequent constraint violations to achieve faster convergence with optimized performance by encompassing only the $(1 - \Xi)$-quantile of the possible output reachable sets.

The performance of the proposed Q-learning algorithm in adaptively determining the input and output references within the predictive control loop is illustrated in Figures 11 and 12. In both figures, the controller achieves an accurate real-time tracking between predictions and references. The root-mean-square error in Figure 11 is 0.1061, which is smaller than the standard deviation of the actual output 0.1101. This indicates that the controller predicts the output patterns almost as well as the actual variation allows. In

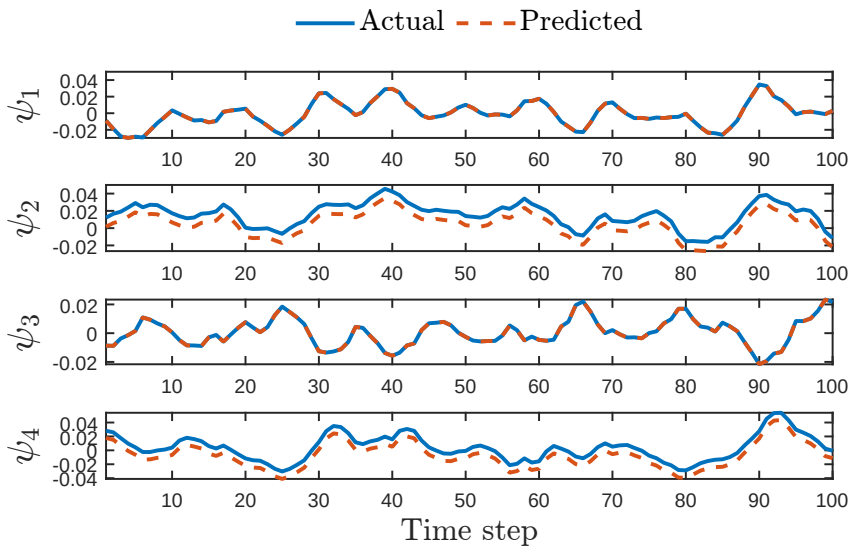

Fig. 9. Koopman prediction accuracy in the observable space.

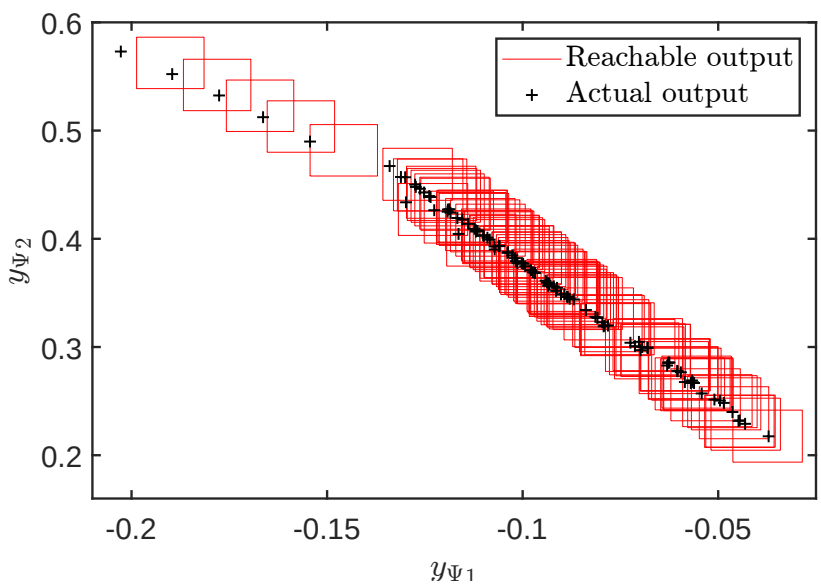

Fig. 10. Ouput reachable set for $\Xi = 0.05$.

Figure 12, the optimal input trajectory and its associated reference for $\Xi = 0.2$ is less tight compared with the resulting trajectories obtained for $\Xi = 0.05$. This shows that a higher controller's tolerance for constraint violation results in a flexible reference behavior. Nonetheless, the reference tracking performance in Figures 11 and 12 is smoother and tighter than the cases in Figures 6 and 7. This is because the stirred-tank reactor system introduces stronger nonlinearities relative to the cart–damper–spring system.

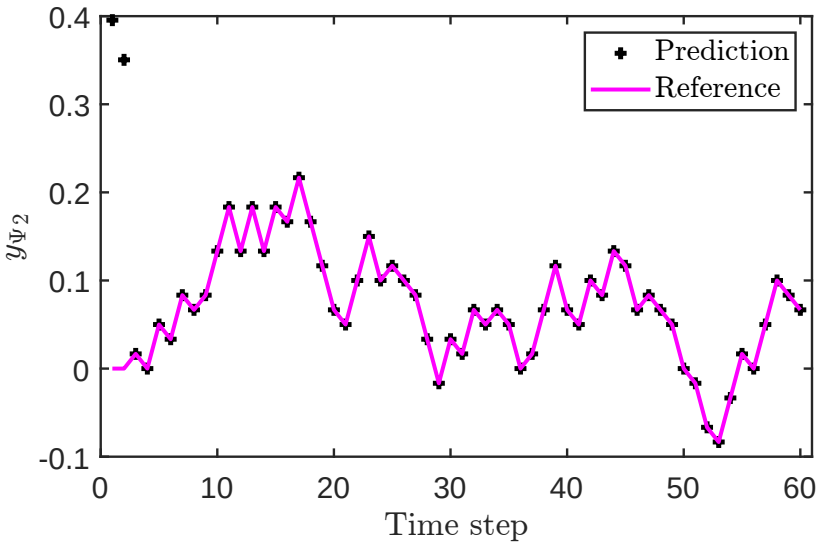

Fig. 11. Evolution of output with adaptive reference for $\Xi = 0.2$.

5. CONCLUSION

This study has presented a novel data-driven predictive control framework for noisy nonlinear systems without assuming any prior knowledge of the system dynamics or noise statistics. A predictive control approach grounded in LSTM-based Koopman operator theory and reinforcement learning is proposed by relying solely on input-output trajectory data. The controller is designed to use a chance-constrained reachability analysis technique in the original system output space while retaining the benefits of a higher-dimensional lifted system representation. A first-step constraint approach is introduced to ensure the recursive feasibility of the control problem, where the controller gradually adjusts its adaptive targets using a Q-learning algorithm, all while adhering to system constraints.

The proposed approach involves two phases: offline Koopman-based reachable set computation and online chance-constrained control optimization. The synergistic integration of these two phases handles the combinatorial generator growth by providing a global system representation directly from data, and reduces computational expense by

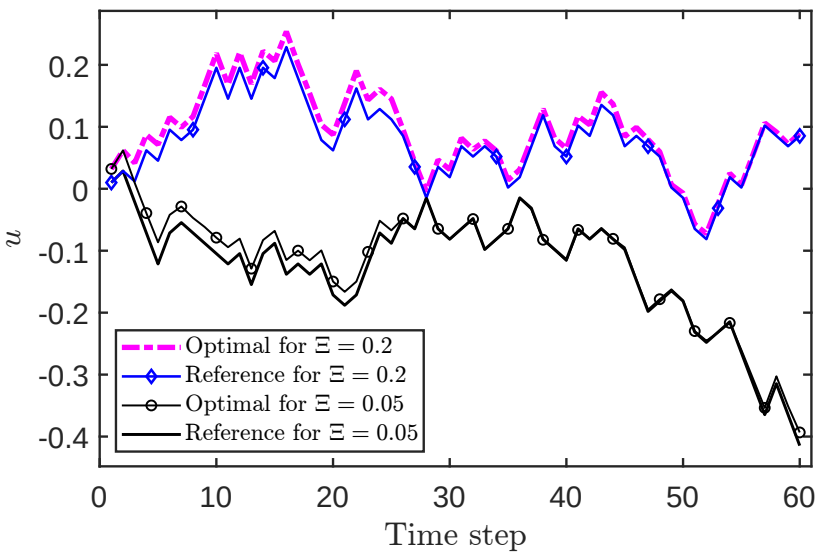

Fig. 12. Optimal control input.

avoiding excessive system conservatism. This enhances the computational performance and scalability of the approach even for high-dimensional systems over long prediction horizons, though the computational burden still grows with the dimension of the lifted space. Numerical examples are presented on the stirred-tank reactor system and the cart–damper–spring system to validate the practical viability of the approach. Both numerical examples show that the proposed integration of LSTM, Koopman theory, chance-constrained reachability analysis, and reinforcement learning can be practically implemented and yield an intelligent, adaptive, and distributionally robust data-driven predictive control of unknown noisy nonlinear systems. Future work may focus on extending the current approach to continuous-time unknown nonlinear systems.

(Received July 2, 2025)

REFERENCES

[1] A. S. Anand, S. Sawant, D. Reinhardt, and S. Gros: Data-driven predictive control and MPC: Do we achieve optimality? IFAC-PapersOnLine *58* (2024), 15, 73–78. DOI:10.1016/j.ifacol.2024.08.507

[2] L. Babić, M. Lauricella, G. Ceusters, and M. Biskoping: Data-driven non-parametric chance-constrained model predictive control for microgrids energy management using small data batches. Front. Control Engrg. *4* (2023), 1237759. DOI:10.3389/fcteg.2023.1237759

[3] J. Berberich and F. Allgöwer: An overview of systems-theoretic guarantees in data-driven model predictive control. Ann. Rev. Control Robotics Autonom. Systems *8* (2024), 77–100. DOI:10.1146/annurev-control-030323-024328

[4] V. Breschi, A. Chiuso, M. Fabris, and S. Formentin: On the impact of regularization in data-driven predictive control. In: 62nd IEEE Conference on Decision and Control (CDC), 2023, pp. 3061–3066. DOI:10.1109/CDC49753.2023.10383820

[5] O. Celik, H. Abdulsamad, and J. Peters: Chance-constrained trajectory optimization for non-linear systems with unknown stochastic dynamics. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2019, pp. 6828–6833. DOI:10.1109/IROS40897.2019.8967794

[6] A. Charnes and W. W. Cooper: Chance-constrained programming. Manage. Sci. *6* (1959), 1, 73–79. DOI:10.1287/mnsc.6.1.73

[7] A. Chiuso, M. Fabris, V. Breschi, and S. Formentin: Harnessing uncertainty for a separation principle in direct data-driven predictive control. Automatica *173* (2025), 112070. DOI:10.1016/j.automatica.2024.112070

[8] O. Ciftci, M. Mehrtash,and A. Kargarian: Data-driven nonparametric chance-constrained optimization for microgrid energy management. IEEE Trans. Ind. Inform. *16* (2019), 4, 2447–2457. DOI:10.1109/TII.2019.2932078

[9] O. Ciftci, M. Mehrtash, F. Safdarian, and A. Kargarian: Chance-constrained microgrid energy management with flexibility constraints provided by battery storage. In: 2019 IEEE Texas Power and Energy Conference (TPEC), 2019, pp. 1–6. DOI:10.1109/TPEC.2019.8662200

[10] F. Dörfler, J. Coulson, and I. Markovsky: Bridging direct and indirect data-driven control formulations via regularizations and relaxations. IEEE Trans. Automat. Control *68* (2022), 2, 883–897. DOI:10.1109/TAC.2022.3148374

[11] M. Farjadnia, A. Alanwar, M. O. B. Niazi, M. Molinari, and K. H. Johansson: Robust data-driven predictive control of unknown nonlinear systems using reachability analysis. European J. Control *74* (2023), 100878. DOI:10.1016/j.ejcon.2023.100878

[12] M. Fink, D. Wollherr, and M. Leibold: Stochastic model predictive control with minimal constraint violation probability for time-variant chance constraints. IEEE Control Syst. Lett. (2024), 1385–1390. DOI:10.1109/LCSYS.2024.3408608

[13] M. He: Data-driven approximated optimal control for chemical processes with state and input constraints. Complexity *2019* (2019), 1, 1396913. DOI:10.1155/2019/1396913

[14] R. Jiang and Y. Guan: Data-driven chance constrained stochastic program. Math. Program. *158* (2016), 1, 291–327. DOI:10.1007/s10107-015-0929-7

[15] T. Ketema, S. L. Tilahun, S. D. Zawka, and A. Geletu: Deep Koopman-based reachability analysis for data-driven predictive control of unknown nonlinear systems. IFAC J. Systems Control *34* (2025), 100339. DOI:10.1016/j.ifacsc.2025.100339

[16] M. Korda and I. Mezić, Igor: On convergence of extended dynamic mode decomposition to the Koopman operator. J. Nonlinear Sci. *28* (2018), 687–710. DOI:10.1007/s00332-017-9423-0

[17] P. Li, M. Wendt, and G. Wozny: A probabilistically constrained model predictive controller. Automatica *38* (2002), 7, 1171–1176. DOI:10.1016/S0005-1098(02)00002-X

[18] X. Li, M. Yan, X. Zhang, M. Han, A. W.-K. Law, Adrian, and X. Yin: Efficient data-driven predictive control of nonlinear systems: A review and perspectives. Digital Chem. Engrg. (2025), 100219. DOI:10.1016/j.dche.2025.100219

[19] Y. Liu, R. Young, Robert, and B. Jafarpour: Long–short-term memory encoder–decoder with regularized hidden dynamics for fault detection in industrial processes. J. Process Control *124* (2023), 166–178. DOI:10.1016/j.jprocont.2023.01.015

[20] L. Magni, G. De Nicolao, M. Magnani, and R. Scattolini: A stabilizing model-based predictive control algorithm for nonlinear systems. Automatica *37* (2001), 9, 1351–1362. DOI:10.1016/S0005-1098(01)00083-8

[21] Sh. Mirkhani and Y. Saboohi: Stochastic modeling of the energy supply system with uncertain fuel price – A case of emerging technologies for distributed power generation. Appl. Energy *93* (2012), 668–674. DOI:10.1016/j.apenergy.2011.12.099

[22] A. Peña-Ordieres, J. R. Luedtke, and A., Wächter: Solving chance-constrained problems via a smooth sample-based nonlinear approximation. SIAM J. Optim. *30* (2020), 3, 2221–2250. DOI:10.1137/19M1261985

[23] Z. Peng, Z. Zhang, R. Luo, Y. Kuang, J. Hu, H. Cheng, and B. K. Ghosh: Event-triggered optimal control of completely unknown nonlinear systems via identifier-critic learning. Kybernetika *59* (2023), 3, 365–391. DOI:10.14736/kyb-2023-3-0365

[24] P. K. Shenoy: Mutual information and Kullback–Leibler divergence in the Dempster-Shafer theory. In: International Conference on Belief Functions, 2024, pp. 225–233. DOI:10.1007/978-3-031-67977-324

[25] A. T. Schwarm and M. Nikolaou: Chance-constrained model predictive control. AIChE J. *45* (1999), 8, 1743–1752. DOI:10.1002/aic.690450811

[26] A. Thorpe, T. Lew, M. Oishi,and M. Pavone: Data-driven chance constrained control using kernel distribution embeddings. In: Learning for Dynamics and Control Conference, 2022, pp. 790–802. URL: https://proceedings.mlr.press/v168/thorpe22a.html

[27] H. Xie, L. Dai, Y. Luo, and Y. Xia: Robust MPC for disturbed nonlinear discrete-time systems via a composite self-triggered scheme. Automatica *127* (2021), 109499. DOI:10.1016/j.automatica.2021.109499

[28] Ch. Wu, A. Mohammadi, M. Mehrtash,and A. Kargarian: Non-parametric joint chance constraints for economic dispatch problem with solar generation. In: 2019 IEEE Texas Power and Energy Conference (TPEC), 2019, pp. 1–6. DOI:10.1109/TPEC.2019.8662145

[29] Ch. Wu, A. Kargarian, and H. M. Jeon: Data-driven nonparametric joint chance constraints for economic dispatch with renewable generation. IEEE Trans. Ind. Appl. *57* (2021), 6, 6537–6546. DOI:10.1109/TIA.2021.3105364

[30] Ch. Wu and A. Kargarian: Computationally efficient data-driven joint chance constraints for power systems scheduling. IEEE Trans. Power Syst. *38* (2022), 3, 2858–2867. DOI:10.1109/TPWRS.2022.3195127

[31] F.- J. Zhao, Y.-F. Gao, X.-F. Wang, H.-Y. Gu, and X.-M. Sun: Robust model predictive control for nonlinear systems with incremental control input constraints. IEEE Trans. Automat. Scie. Engrg. *22* (2024), 9983–9993. DOI:10.1109/TASE.2024.3515172

Teketel Ketema, Corresponding author. Department of Mathematics, Arba Minch University, Arba Minch, 21, Ethiopia and Department of Mathematics, Mekdela Amba University, Tulu Awuliya, 32. Ethiopia.
e-mail: teketelketema1987@gmail.com

Surafel Luleseged Tilahun, HPC and Big Data Analytics Center of Excellence, Addis Ababa Science and Technology University, Addis Ababa, 16417, Ethiopia and Department of Mathematics, Debark University, Debark, 90. Ethiopia.
e-mail: surafel42@gmail.com

Simon D. Zawka, Department of Mathematics, Arba Minch University, Arba Minch, 21. Ethiopia.
e-mail: simondereke@gmail.com

Abebe Geletu, German Research Chair, African Institute for Mathematical Sciences (AIMS), Kigali, 7150. Rwanda.
e-mail: abebe.geletu@aims.ac.rw