Robust Tube-based Model Predictive Control with Koopman Operators

Xinglong Zhang [a], Wei Pan [b], Riccardo Scattolini [c], Shuyou Yu [d], Xin Xu [a]

[a]*College of Intelligence Science and Technology, National University of Defense Technology, Changsha 410073, China.*

[b]*Department of Cognitive Robotics, Delft University of Technology, the Netherlands.*

[c]*Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milan 20133, Italy.*

[d]*Department of Control Science and Engineering, Jilin University at NanLing, Changchun 130025, China.*

Abstract

Koopman operators are of infinite dimension and capture the characteristics of nonlinear dynamics in a lifted global linear manner. The finite data-driven approximation of Koopman operators results in a class of linear predictors, useful for formulating linear model predictive control (MPC) of nonlinear dynamical systems with reduced computational complexity. However, the robustness of the closed-loop Koopman MPC under modeling approximation errors and possible exogenous disturbances is still a crucial issue to be resolved. Aiming at the above problem, this paper presents a robust tube-based MPC solution with Koopman operators, i.e., r-KMPC, for nonlinear discrete-time dynamical systems with additive disturbances. The proposed controller is composed of a nominal MPC using a lifted Koopman model and an off-line nonlinear feedback policy. The proposed approach does not assume the convergence of the approximated Koopman operator, which allows using a Koopman model with a limited order for controller design. Fundamental properties, e.g., stabilizability, observability, of the Koopman model are derived under standard assumptions with which, the closed-loop robustness and nominal point-wise convergence are proven. Simulated examples are illustrated to verify the effectiveness of the proposed approach.

Key words: Model predictive control, Koopman operators, nonlinear systems, robustness, convergence.

1 Introduction

Model predictive control (MPC), is employed as an effective control tool for control of numerous applications, such as robotics and industrial plants, see [1, 2]. In MPC, a prediction model is typically required, with which many MPC algorithms can be formulated, e.g., stabilizing MPC for nominal models in [3] and the references therein, and robust MPC such as min-max MPC in [4] or tube-based MPC in [5, 6] for systems under model uncertainties and disturbances. Focusing on building well-performed prediction models, designing MPC using data-driven models has been noted as a promising direction. Many works on this aspect have been developed, see [7–12] and the references therein. Among them, a unitary learning-based predictive controller for linear systems was addressed in [9], where set membership is adopted to estimate a multi-step linear prediction model to be used for designing a robust MPC. Similarly, resorting to the set membership identification, adaptive MPC algorithms for uncertain time-varying systems were proposed in [13, 14] for reducing the conservativity caused by robust MPC. Relying on the main idea from iterative learning control, a data-driven learning MPC for repetitive tasks was studied in [15] with terminal constraints updated iteratively. In these approaches, a linear robust (or stabilizing) MPC problem is to be solved online since the considered model is linear.

In the control of nonlinear systems, the derivation of a nonlinear prediction model could be a nontrivial task. In [16], a nonlinear MPC algorithm using a machine learning-based model estimation was proposed with stability property guarantees.

However, nonlinear MPC results in nonlinear, even nonconvex, optimization problems, which can be computationally intensive for systems with high nonlinearity. To reduce the computational load, a supervised machine learning algorithm was used to approximate the nonlinear MPC law in [17]. The robustness was guaranteed under bounded control approximation errors with verified statistic empirical risk.

In another line, Koopman operators have been noted to be effective to represent the internal dynamics of nonlinear systems [18, 19]. Specifically, the Koopman operator is typically with infinite dimension, capturing the nonlinear dynamical characteristics through a linear dynamic evolution on a lifted observable function of states. In [20], a finite-dimensional truncation of the Koopman operator was used to form a linear predictor of nonlinear dynamics for designing a linear MPC. This approach paves the way to linear MPC formulations of nonlinear systems with a linear predictor that represents a wide operation range. Among the most related works, the extension to an MPC algorithm using the integrated Koopman operator was addressed in [10], and MPC methods using a deep learning-based Koopman model were developed in [21, 22]. The application of Koopman MPC to flow control was addressed in [18]. In [23], a Lyapunov-based Koopman MPC was developed for the feedback stabilization of nonlinear systems.

Note that, the presence of modeling errors is almost inevitable in lifted Koopman models, due to the data-driven finite-dimensional approximation of Koopman operators and exogenous disturbances, see [20, 24]. An interesting offset-free Koopman MPC extension of [23] was presented in [25] for handling model-mismatch with an estimator. However, the satisfaction of hard state constraints and closed-loop robustness under both non-negligible approximating errors and unknown additive disturbances are crucial concerns, which were not addressed in the previous Koopman-based MPC [10, 20, 21, 23, 25]. This motivated our research work.

In this paper, we present a robust MPC solution with Koopman operators in the framework of tube-based MPC [5, 6]. The contributions are twofold. The first contribution is a linear robust Koopman-based MPC design methodology for nonlinear systems with unknown dynamics and additive disturbances. As opposed to the classic Tube-based MPC [5, 6], our approach allows designing robust MPC from measured data and no explicit model information is required; also, our approach results in a nonlinear MPC law by a linear robust MPC design. The second contribution is the analysis of the closed-loop theoretical guarantees for Koopman-based MPC under modeling errors and additive disturbances. This is achieved via imposing standard prior conditions on the lifting observable functions, allowing for using a truncated Koopman model with a limited system order in the controller design.

The rest of the paper is organized as follows. Section 2 introduces the considered control problem and preliminary solutions. In Section 2 the main idea of the proposed r-KMPC and the associated theoretical results are obtained. Section 4 shows the simulation results obtained by applying the proposed approach to nontrivial simulated systems, while some conclusions are drawn in Section 5. The ingredients for estimating the uncertainty terms are given in the Appendix.

Notation: We denote $\mathbb{N}$ as the set of positive natural numbers and $\mathbb{N}_1^l$ as the numbers $1, \cdots, l$. Given the variable r, we use $\boldsymbol{r}_{k:k+N}$ to denote the sequence $r_k \ldots r_{k+N}$ and $\boldsymbol{r}_k$ to denote $\boldsymbol{r}_{k:k+N}$ after its first appearance, where k is the discrete time index and N is a positive integer. For a vector $x \in \mathbb{R}^n$, we use $\|x\|_Q^2$ to stand for $x^\top Q x$, $\|x\|$ to denote its Euclidean norm; while for a matrix $x \in \mathbb{R}^{n \times m}$, we denote $\|x\|_F$ as the Frobenius norm. Given two sets $\mathcal{Z}$ and $\mathcal{V}$, their Minkowski sum is represented by $\mathcal{Z} \oplus \mathcal{V} = \{z + v | z \in \mathcal{Z}, v \in \mathcal{V}\}$. For a given set of variables $z_i \in \mathbb{R}^{q_i}$, $i = 1, 2, \ldots, M$, we define the vector whose vector-components are z_i in the following compact form: $(z_1, z_2, \cdots, z_\text{M}) = [\, z_1^\top \; z_2^\top \; \cdots \; z_\text{M}^\top \,]^\top \in \mathbb{R}^q$, where $q = \sum_{i=1}^M q_i$.

2 Control problem and preliminaries

2.1 Control problem

Consider a class of nonlinear discrete-time systems with additive disturbances described by

$$x^+ = f(x, u) + w_o, \tag{1}$$

where $x \in \mathcal{X} \subset \mathbb{R}^n$, $u \in \mathcal{U} \subset \mathbb{R}^m$ are the state and control variables, x^+ is the successor state at the next discrete-time instant, $w_o \in \mathcal{W}_o$ is an additive bounded noise which can be unknown and not measurable, $\mathcal{W}_o$ is a compact set containing the origin, $\mathcal{X}$ and $\mathcal{U}$ are convex sets containing the origin in their interiors, f is the state transition function, which can be partially or completely unknown. It is assumed that $f(0, 0) = 0$, $f(x, u)$ is C^∞ on $\mathcal{X} \times \mathcal{U}$, and $\|f(x, u)\| < +\infty$ for all $x \times u \in \mathcal{X} \times \mathcal{U}$. The state x is measurable.

Starting from any initial condition $x_0 \in \mathcal{X}$, the control objective is to minimize a quadratic cost of type $J = \sum_{k=0}^{+\infty} \|x_k\|_Q^2 + \|u_k\|_R^2$, where $Q = Q^\top \in \mathbb{R}^{n \times n}$ and $R = R^\top \in \mathbb{R}^{m \times m}$, $Q, R \succ 0$.

Definition 1 (Local stabilizability [26]) *Model $x_{k+1} = f(x_k, u_k)$ is stabilizable on the domain $\mathcal{X} \times \mathcal{U}$ if, for any $x_0 \in \mathcal{X}$, there exists a feedback function $u(x_k) \in \mathcal{U}$, $u(0) = 0$, such that state x_k of the corresponding closed-loop system asymptotically converges to the origin.*

Definition 2 (Generalized gradient [27]) *The generalized gradient of a Lipschitzian function $h(x) : \mathbb{R}^n \to \mathbb{R}^q$ at x, denoted as $\nabla h(x)$, is the convex hull of all matrices of the form $M = \lim_{i \to \infty} \nabla h(x + \delta x_i)$, where $\delta x_i \to 0$ as $i \to \infty$.*

Definition 3 (Maximal rank [28]) *The generalized gradient $\nabla h(x)$ is of maximal rank, if for every M $\mathrm{rank}(M) = \min\{n, q\}$.*

2.2 Preliminary Koopman MPC

We first review the Koopman operator theory for autonomous dynamical systems and its extension to dynamical systems with controls.

Let us introduce the so-called observable of x defined by a scalar-valued function $\phi(x) : \mathcal{X} \to \mathbb{C}$ and denote $\mathcal{F}$ as a given space of observables. For an autonomous model $x^+ = f(x, 0)$, i.e. model (1) with $u, w_o = 0$, the Koopman operator $\mathcal{K} : \mathcal{F} \to \mathcal{F}$ is defined by [18, 19]

$$\mathcal{K}\phi(x) = \phi(x) \circ f(x, 0), \tag{2}$$

for every observable $\phi(x) \in \mathcal{F}$ ($\mathcal{F}$ is invariant under the action of the Koopman operator), $\circ$ is the composition operator, i.e., $\mathcal{K}\phi(x) = \phi(f(x, 0))$. For any discrete-time instant $k \in \mathbb{N}$, it holds that

$$\phi(x_k) = \mathcal{K}\phi(x_{k-1}) = \cdots = \mathcal{K}^k \phi(x_0),$$

which captures the dynamical characteristics of the original nonlinear dynamics. For a detailed introduction of the Koopman operator please refer to [19, 29, 30].

The Koopman operator for $x^+ = f(x, 0)$ can be generalized to systems with controls (i.e. model (1)) in several ways, see e.g. [20, 31, 32]. In our study, we adopt the practical and rigorous scheme in [20], which relies upon an extended state space $\mathcal{X} \times \ell(\mathcal{U} \times \mathcal{W}_o)$, where $\ell(\mathcal{U} \times \mathcal{W}_o)$ is the space of all the sequences composed of the control and disturbance, i.e. $\boldsymbol{u}_w := \{u_w(i)\}_{i=0}^{+\infty}$ with $u_w(i) = (u(i), w_o(i)) \in \mathcal{U} \times \mathcal{W}_o$. Letting $f_W(x, \boldsymbol{u}_w) := f(x, u) + w_o$ and $\boldsymbol{u}_w(i)$ be the i-th element of $\boldsymbol{u}_w$, one can write the dynamics of the extended state $\boldsymbol{\chi} = (x, \boldsymbol{u}_w)$ as

$$\boldsymbol{\chi}^+ = F(\boldsymbol{\chi}) := (f_W(x, \boldsymbol{u}_w(0)), \Gamma \boldsymbol{u}_w), \tag{3}$$

where Γ is a left shift operator such that $\boldsymbol{u}_w(i + 1) = \Gamma \boldsymbol{u}_w(i)$. In this way, the Koopman operator $\mathcal{K} : \mathcal{F}_e \to \mathcal{F}_e$ associated with (3) is given by [20]

$$\mathcal{K}\phi(\boldsymbol{\chi}) = \phi(\boldsymbol{\chi}) \circ F(\boldsymbol{\chi}) \tag{4}$$

where $\phi(\boldsymbol{\chi}) : \mathcal{X} \times \ell(\mathcal{U} \times \mathcal{W}_o) \to \mathbb{C}$ belongs to the extended observable space $\mathcal{F}_e$, which contains observable functions on arguments x and $\boldsymbol{u}_w$.

A finite-dimensional numerical approximation of $\mathcal{K}$ in (4) is of interest for controller design, which can be computed by resorting to the extended dynamic mode decomposition (EDMD) method in a data-driven manner, see [20]. Let a finite-dimensional approximation of $\mathcal{K}$ be $\mathcal{K}_{N_\phi} \in \mathbb{R}^{N_\phi \times N_\phi}$, associated with an observable vector $\Phi(\boldsymbol{\chi}) = (\phi_1(\boldsymbol{\chi}), \cdots, \phi_{N_\phi}(\boldsymbol{\chi}))$. The goal in EDMD is to compute $\mathcal{K}_{N_\phi}$ via minimizing $\|\Phi(\boldsymbol{\chi}^+) - \mathcal{K}_{N_\phi}\Phi(\boldsymbol{\chi})\|^2$. Note however that $\boldsymbol{\chi} = (x, \boldsymbol{u}_w)$ is of infinite-dimension, which can be problematic from the computational viewpoint. Hence, we choose a computable observable function as

$$\Phi(x, \boldsymbol{u}_w) = (\Psi(x), \boldsymbol{u}_w(0)), \tag{5}$$

where $N_\phi = n_\psi + m + n$, $n_\psi > n$,

$$\Psi(x) := (\psi_1(x), \cdots, \psi_{n_\psi}(x)),$$

and ψ_i, $i \in \mathbb{N}_1^{n_\psi}$, can be chosen as some basis functions or neural networks [21]. To compute $\mathcal{K}_{N_\phi}$ with EDMD, let us assume to have collected M input-state datasets $\{(u_i, \hat{w}_{o,i}, x_i, x_i^+)\}_{i=1}^M$ satisfying $x_i^+ = f(x_i, u_i) + w_{o,i}$, where $\hat{w}_{o,i}$ is the estimation of $w_{o,i}$ which can be computed by resorting to a nonlinear estimation technique or a Koopman operator-based estimator [33]. The following condition is assumed to hold [19, 29, 30].

Assumption 1 *The data points $\{(u_i, x_i)\}_{i=1}^M$ are drawn independently according to a non-negative probability distribution μ.*

This condition can also be replaced with the assumption that $\{(u_i, x_i)\}_{i=1}^M$ are ergodic in $\mathcal{U} \times \mathcal{X}$ with respect to μ, which can be generated by integrating a stochastic dynamics, see [19, 29].

Let $[\mathcal{K}_{N_\phi}]_{1:n_\psi} = [A\,B\,D] \in \mathbb{R}^{n_\psi \times n_\phi}$, where $A \in \mathbb{R}^{n_\psi \times n_\psi}$, $B \in \mathbb{R}^{n_\psi \times m}$, and $D \in \mathbb{R}^{n_\psi \times n}$. Since we are only interested in predicting $\Psi(x^+)$ with $\Psi(x)$ and the estimation of $\boldsymbol{u}_w(0)$ (due to $\hat{w}_o$), the following regularized least squares problem can be

stated:

$$\min_{[\mathcal{K}_{N_\phi}]_{1:n_\psi}} \sum_{i=1}^{M} \|[\mathcal{K}_{N_\phi}]_{1:n_\psi}(\Psi(x_i), \hat{u}_{w,i}) - \Psi(x_i^+)\|^2 + \alpha\|[\mathcal{K}_{N_\phi}]_{1:n_\psi}\|_F^2 := V_\mathcal{K} \tag{6}$$

where $\hat{u}_{w,i} = (u_i, \hat{w}_{o,i})$, $\alpha > 0$ is a tuning parameter on the Frobenius norm regularization of $[\mathcal{K}_{N_\phi}]_{1:n_\psi}$. To recover x using $\Psi(x)$, a linear matrix $C \in \mathbb{R}^{n \times n_\psi}$ is optimized according to the following problem [20, 24, 34]:

$$\min_C \sum_{i=1}^{M} \|C\Psi(x_i) - x_i\|^2 + \beta\|C\|_F^2, \tag{7}$$

where $\beta > 0$ is a tuning parameter. By solving (6) and (7), a linear Koopman predictor of (1) can be obtained, i.e.,

$$\begin{cases} \hat{s}^+ = A\hat{s} + B\hat{u} \\ \hat{x} = C\hat{s}. \end{cases} \tag{8}$$

Note that, in (8), a new (abstract) variable $\hat{s}$ serves as the state due to the lifted observable construction, while the predicted value of the original state, i.e., $\hat{x}$, becomes the output variable through the mapping matrix C from the observable space.
With (8), a linear Koopman MPC (KMPC) problem similar to [20] can be stated as follows:

$$\min_{\boldsymbol{u}_{k:k+N-1}} \sum_{i=0}^{N-1}(\|\hat{x}_{k+i}\|_Q^2 + \|u_{k+i}\|_R^2) + \|\hat{x}_{k+N}\|_{Q_N}^2 \tag{9}$$

subject to model (8) with $\hat{s}_k = \Psi(x_k)$, state constraint $C\hat{s}_{k+i} \in \mathcal{X}$, $\forall i \in \mathbb{N}_1^N$, and control constraint $u_{k+i} = \hat{u}_{k+i} \in \mathcal{U}$, $\forall i \in \mathbb{N}_0^{N-1}$, where $Q_N = Q_N^\top \in \mathbb{R}^{n \times n}$ and $Q_N \succ 0$.
Because of the characteristics of Koopman operators, a merit of KMPC lies in the linear property of the built model (8), leading to a linear MPC problem instead of a nonlinear one. However, the derivation of (8) with (5) by (6) and (7) could bring modeling errors, see [20, 24, 34], whose property and possible (negative) influences on the closed-loop control performance is not yet analyzed. As a consequence, the closed-loop property under modeling errors and additive disturbances remains still a crucial issue, which was not addressed in the prescribed Koopman-based MPC [10, 20, 21, 23, 25]. Peculiarly, with KMPC in [20], the constraint satisfaction $x \in \mathcal{X}$ might not be fulfilled by $C\hat{s} \in \mathcal{X}$ and the closed-loop robustness of the Koopman MPC might not be verified under modeling errors and disturbances. Aiming at this problem, in the following section we propose a robust tube-based MPC using Koopman operators with theoretical guarantees.

3 Robust Koopman MPC

In this section, the proposed robust MPC solution using Koopman operators, i.e., r-KMPC, is presented. First, a Koopman model with approximation errors is derived and its stabilizability and observability properties are proven under standard assumptions. Then, the proposed r-KMPC algorithm using the Koopman model is presented. Finally, the closed-loop theoretical properties of r-KMPC are proven.

3.1 Koopman model for robust MPC

As described in [20], it is not guaranteed that $\mathcal{K}_{N_\phi}$ converges to $\mathcal{K}$ as $N_\phi, M \to +\infty$ even if w_o is measurable and $\{w_{o,i}\}_{i=1}^{M}$ are ergodic samples, because the adopted observable function (5) does not form an orthonormal basis of $\mathcal{F}_e$. Hence, the presence of modeling errors of (8) with (5) is inevitable also due to the existence of estimation errors of w_o and to the practical design with $\alpha, \beta \neq 0$ in (6) and (7). Nonetheless, as pointed out in [35], the model structure like (8) is of interest from the control viewpoint because it could still be effective for approximating (1) in a large state space region; also it permits a linear MPC implementation for nonlinear systems, leading to a computational load reduction compared with the approaches using complex bilinear models and switched models [10, 36]. The consideration of using bilinear or switched models is left as further investigation.
To derive a robust Koopman MPC solution with (8), the fundamental boundedness property of the overall uncertainty caused by

multiple sources is required. To this end, letting $s = \Psi(x)$, one can first write an equivalent Koopman model of (1) considering the effects of model uncertainties, that is

$$\begin{cases} s^+ = As + Bu + \bar{w}(s, u, w_o, \hat{w}_o) \\ x = Cs + v(s), \end{cases} \tag{10}$$

where $\bar{w} = D\hat{w}_o + w(s, u, w_o, \hat{w}_o) \in \bar{\mathcal{W}}$, $w(s, u, w_o, \hat{w}_o) \in \mathcal{W}$ and $v(s) \in \mathcal{V}$ are the modeling errors, where $\mathcal{W}$ and $\mathcal{V}$ are convex sets containing the origin; $\bar{\mathcal{W}} = D\hat{\mathcal{W}}_o \oplus \mathcal{W}$, $\hat{\mathcal{W}}_o$ is a computable convex set where $\hat{w}_o$ lies in. It is assumed that $\hat{\mathcal{W}}_o$ is bounded. A discussion on the boundedness property of sets $\bar{\mathcal{W}}$ (i.e., $\mathcal{W}$) and $\mathcal{V}$ is deferred to Proposition 2. To this end, we first introduce the following assumption about $\Psi(x)$.

Assumption 2

(1) The lifted function $\Psi(x)$ is Lipschitz continuous.
(2) $\{\psi_i(x)\}_{i=1}^{n_\psi}$ are linearly independent.

The verification of the above assumption is easy since many adopted basis functions (BF) such as Gaussian kernel functions, polyharmonic splines, and thinplate splines are in fact C^∞ and linearly independent.

Lemma 1 (Existence of inverse maps [28]) *Letting $\mathcal{S}_\Psi$ be the set such that $\mathcal{S}_\Psi = \{s \in \mathbb{R}^{n_\psi} | s = \Psi(x), x \in \mathcal{X}\}$, if $\nabla\Psi(x)$ is of maximal rank, i.e., $\mathrm{rank}(\nabla\Psi(x)) = n$, there exists a Lipschitzian function $\Psi^{-1} : \mathcal{S}_\Psi \to \mathcal{X}$ such that $\Psi^{-1}(\Psi(x)) = x$ for every $x \in \mathcal{X}$.*

In the following proposition, taking a special type of Gaussian kernels as an example of basis functions, we show indeed that multiple solutions of Ψ^{-1} can be found.

Proposition 1 (Multiple choices of Ψ^{-1}) *Letting $\psi_i(x) = e^{-\|x - c_i\|^2}$, $c_i \in \mathcal{X}$, $i = \mathbb{N}_1^{n_\psi}$, a group of any $n + 1$ basis functions of $\Psi(x)$ can surely define a choice of Ψ^{-1} if and only if Assumption 2.(2) holds.*

Proof. Since $\psi_i(x)$, $i \in \mathbb{N}_1^{n_\psi}$ are Gaussian kernels, the resulting Ψ is continuous differentiable. In this case the generalized gradient $\nabla\Psi(x)$ coincides with the exact gradient $\partial\Psi(x)$. For a generic $n \in \mathbb{N}$, select a group of any $n + 1$ basis functions such that $\|x - c_i\|^2 = -\log\psi_i(x)$, $\forall i \in \mathbb{N}_1^{n+1}$. Letting $\bar{\psi}(x) = (\psi_1(x), \cdots, \psi_{n+1}(x))$, the corresponding gradient $\partial\bar{\psi}(x)$ is

$$\begin{aligned} \partial\bar{\psi}(x) &= -2 \left[\psi_1(x) \cdot (x - c_1) \ \cdots \psi_{n+1}(x) \cdot (x - c_{n+1}) \right]^\top \\ &= -2\mathrm{diag}\{\bar{\psi}(x)\} \left[(x - c_1) \ \cdots (x - c_{n+1}) \right]^\top, \end{aligned} \tag{11}$$

where $\mathrm{diag}\{\bar{\psi}(x)\} = \mathrm{diag}\{\psi_1(x), \cdots, \psi_{n+1}(x)\}$. As $\mathrm{diag}\{\bar{\psi}(x)\}$ is full rank in view of the property of $\psi_i(x)$, for any $x \in \mathcal{X}$, one has

$$\begin{aligned} \mathrm{rank}(\partial\bar{\psi}(x)) &= \mathrm{rank}(\left[(x - c_1) \ \cdots (x - c_{n+1}) \right]^\top) \\ &= n, \end{aligned} \tag{12}$$

if and only if $c_1, \cdots, c_{n+1}$ are linearly independent, where the worst testing scenario is $x = c_i$, $i \in \mathbb{N}_1^{n+1}$. $\square$

Remark 1 *Proposition 1 implies that, with prescribed Gaussian kernels, one can find $\sum_{i=n}^{n_\psi} \frac{n_\psi!}{(n_\psi - i)!i!}$ combinations of basis functions (choices of Ψ^{-1}). In this peculiar case, determining x using multiple $\psi_i's$, can be stated as a feasibility problem with multiple quadratic equality constraints, while its dual problem does not fulfill the Slater condition that ensures strong duality property, see [37].*

We recall that (cf. [20]), a straightforward and well-performed choice of $\Psi(x)$ can be of type $\Psi(x) = (x, \bar{\Psi}(x))$, i.e., with the original state x being included. With this choice, one can promptly find a candidate $\Psi^{-1}(s) = [I_n \ 0]s$ such that $\Psi^{-1}(\Psi(x)) = x$.

Now, it is possible to state the boundedness property of sets $\bar{\mathcal{W}}$ (i.e., $\mathcal{W}$) and $\mathcal{V}$ in the following proposition.

Proposition 2 *If $\Psi(x)$ is such that set $\mathcal{S}_\Psi$ is bounded, then $\bar{\mathcal{W}}$ and $\mathcal{V}$ are bounded.*

Proof. To first prove $\bar{\mathcal{W}}$ is bounded, in view of (1) and (10) and recalling that $s = \Psi(x)$, one has that the modeling uncertainty is $w(s, u, w_o, \hat{w}_o) = \Psi(f(\Psi^{-1}(s), u) + w_o) - As - Bu - D\hat{w}_o$. Letting $\delta w = w_o - \hat{w}_o$, for all $s \in \mathcal{S}_\Psi$, $u \in \mathcal{U}$, $\hat{w}_o \in \hat{\mathcal{W}}_o$, and $w_o \in \mathcal{W}_o$, it is convenient to write the following inequality:

$$\|w(s, u, w_o, \hat{w}_o)\| \le L_s\|s\| + L_u\|u\| + L_{\delta w}\|\delta w_o\| + L_{\hat{w}}\|\hat{w}_o\|$$
$$< +\infty, \tag{13}$$

where L_s, L_u, $L_{\delta w}$, and $L_{\hat{w}}$ are bounded Lipschitz constants and the last inequality holds since $\mathcal{X}, \mathcal{U}, \hat{\mathcal{W}}_o$, and $\mathcal{W}_o$ are bounded. Hence, $\mathcal{W}$ is bounded, leading to $\bar{\mathcal{W}}$ being bounded. As for $\mathcal{V}$, it follows from (10) that $\mathcal{V} = \mathcal{X} \ominus C\mathcal{S}_\Psi$ being bounded. $\square$
Indeed, the boundedness (instead of convergence) of $\bar{w}$ and v from Proposition 2 is sufficient for theoretical guarantees of the proposed r-KMPC, which allows us to reduce the dimension of the approximated Koopman model in (10). It can also be observed through (13) that, the range of $\bar{w}$ (i.e., w) and v hinges upon the Lipschitz constants L_s, L_u, and $L_{\hat{w}}$. To further reduce the size of $\mathcal{W}$ and $\mathcal{V}$, it is convenient to minimize L_s, L_u, and $L_{\hat{w}}$ or their estimations in the modeling phase by modifying the optimization problem (6) as

$$\min_{[\mathcal{K}_{N_\phi}]_{1:n_\psi}, L_s, L_u, L_{\hat{w}}} V_\mathcal{K} + \alpha_s L_s + \alpha_u L_u + \alpha_w L_{\hat{w}} \tag{14}$$

subject to $L_s, L_u, L_{\hat{w}} \ge 0$ and

$$\|Y_i - Y_j\| \le L_s\|s_i - s_j\| + L_u\|u_i - u_j\| + L_{\hat{w}}\|\hat{w}_{o,i} - \hat{w}_{o,j}\|,$$

$i, j \in \mathbb{N}_1^M$, where $Y_i = [\mathcal{K}_{N_\phi}]_{1:n_\psi}(\Psi(x_i), \hat{u}_{w,i}) - \Psi(x_i^+)$, $\alpha_s, \alpha_u, \alpha_w > 0$ are tuning parameters.

In the case that $w_o = 0$, one also has $\hat{w}_o = 0$. Letting $s_r = \Psi(0)$ and $\bar{A} = \begin{bmatrix} I - A^\top & C^\top \end{bmatrix}^\top$, the triple $(u, s, x) = (0, s_r, 0)$ is an equilibrium of model (10) if and only if

$$(w(s_r, 0, 0, 0), v(s_r)) = \bar{A}s_r, \tag{15}$$

which results in a prior knowledge of w and v, since $s_r = \Psi(0)$ is available. In principle, it is convenient to define an MPC prediction model using (10) with $\bar{w}$ replaced by $w(s_r, 0, 0, 0)$. However, for the case that w_o is time-varying, the computation of $w(s_r, 0, 0, 0)$ can be nontrivial because $w_o, \hat{w}_o = 0$ might not be accessible or even exist. As a consequence, (10) with $\bar{w} = w(s_r, 0, 0, 0)$ is no longer suitable for prediction in the MPC algorithm. To solve this issue, we first introduce the following proposition.

Proposition 3 (Equilibrium) $(w(s_r, 0, 0, 0), v(s_r)) = 0$ *if and only if*

$$s_r \subseteq \text{Ker}\,\bar{A} \quad or \quad s_r = 0; \tag{16}$$

moreover, if the pair (A, C) is observable, condition (16) is reduced to

$$s_r = 0. \tag{17}$$

Proof. Condition (16) follows directly from (15) by linear algebras.

Remark 2 *Condition (16) implies that two choices can be chosen for guaranteeing $w(s_r, 0, 0, 0), v(s_r) = 0$. A choice is to allow $s_r \ne 0$, i.e., to enforce $s_r \subseteq \text{Ker}\,\bar{A}$, which can be imposed by $s_r = As_r$ and $Cs_r = 0$. The derivation of the above condition requires to include in the optimization problem (6) the constraint $As_r = s_r$ and in (7) the constraint $Cs_r = 0$. The resulting problems can be solved since s_r, i.e., $\Psi(0)$ can be chosen a priori. Note however that the resulting model is not observable if there exist eigenvalues of A being 1, in view of the Popov-Belevitch-Hautus (PBH) test.*
Also, provided that a prior condition that (A, C) is observable, (16) is equivalent to

$$Cs_r = CAs_r = \cdots = CA^{n_\psi - 1}s_r = 0 \tag{18}$$

since $s_r = As_r = A^2s_r, \cdots$ from $s_r = As_r$, leading to the unique solution $s_r = 0$.
Hence, a simple and elegant way is to choose $s_r = 0$ for verifying (15) (see (17)), which allows to avoid the reformulation of

problems (6) *(or* (14)*), and* (7). □

We now give the following basic assumption.

Assumption 3 *The lifted function is constructed such that $\Psi(0) = 0$.*

Remark 3 *Assumption 3 can be easily verified since, for $\Psi(x)$ being selected as a vector of basis functions with $s_r \neq 0$, a coordinate transformation can be used to define a new function $\Psi(x) = \Psi'(x) - \Psi'(0)$, leading to $\Psi(0) = 0$, consequently to verification of* (15). *Take the prescribed Gaussian kernels as an example. Let $\psi_i'(x) = e^{-\|x-c_i\|^2}$, $i \in \mathbb{N}_1^{n_\psi}$, then it holds that $\psi_i(x) = \psi_i'(x) - \psi_i'(0) = e^{-\|x-c_i\|^2} - e^{-\|0-c_i\|^2}$ with $\psi_i(0) = 0$. Hence the resulting $\Psi(x)$ satisfies $\Psi(0) = 0$. With this choice, Proposition 1 can still be verified because $\partial\bar{\psi}(x)$ becomes $\partial\bar{\psi}(x) = -2\mathrm{diag}\{\bar{\psi}'(x)\}[(x-c_1) \cdots (x-c_{n+1})]^\top$, where $\bar{\psi}'(x) = (\psi_1'(x), \cdots, \psi_{n+1}'(x))$.*

Proposition 4 (Local stabilizability) *Under Assumption 3, model* (10) *with $w_o = 0$ is stabilizable on the set $\mathcal{S}_\Psi \times \mathcal{U}$, if and only if $f(x, u)$ is stabilizable on the domain $\mathcal{X} \times \mathcal{U}$.*

Proof. (i) Sufficient condition: as $f(x, u)$ is stabilizable on $\mathcal{X} \times \mathcal{U}$, there exists a feedback law $u(x) \in \mathcal{U}$ such that $x = 0$ is asymptotically stable, under model $x^+ = f(x, u)$. In view of the fact that (10) with $w_o = 0$ is equivalent to $x^+ = f(x, u)$, it holds that $s = \Psi(x)$ converges to the origin asymptotically under $u(x)$, since x converges to the origin asymptotically and $\Psi(0) = 0$. Hence, model (10) is stabilizable on $\mathcal{S}_\Psi \times \mathcal{U}$.
(ii) Necessary condition: likewise, starting with the stabilizability of model (10) on $\mathcal{S}_\Psi \times \mathcal{U}$, one has s converging to the origin asymptotically under $u(x) \in \mathcal{U}$. The stabilizability of $f(x, u)$ on $\mathcal{X} \times \mathcal{U}$ follows since $x = Cs + v(s) \to 0 + v(0) = 0$ asymptotically. □

Remark 4 *Proposition 4 implies that, in the unperturbed case, i.e., $w_o = 0$, there exists a control law such that the point-wise convergence of the closed-loop evolution of s and x under* (10) *can be guaranteed, which is a fundamental hint for the result stated in Theorem 3 (deferred in Section 3.3).*

In a class of MPC problems such as (9) where the penalization on the output is used in the cost function, the observability property could be a prior condition in deriving the closed-loop theoretical property. In the following, we also analyze the observability property of (10).

Definition 4 (Local observability [38]) *System* (10) *under $w_o = 0$ is locally observable on $\mathcal{S}_\Psi$ if, for each $s_{a,0}, s_{b,0} \in \mathcal{S}_\Psi$, the corresponding future states $x_{a,j} \in \mathcal{X}$, $x_{b,j} \in \mathcal{X}$, under the same control sequence, $u_0, \cdots, u_j \in \mathcal{U}^{j+1}$, $j \geq 1$, are such that $x_{a,j} = x_{b,j}$ implies $s_{a,0} = s_{b,0}$.*

Proposition 5 (Local observability) *System* (10) *is locally observable on $\mathcal{S}_\Psi$ if*

$$\frac{\partial\{f(x_j, u_j) \circ \cdots \circ f(x_0, u_0)\}}{\partial x_0} = n, \tag{19a}$$

for each $x_0 \in \mathcal{X}$, $u_i \in \mathcal{U}$, $\forall i \in \mathbb{N}_0^j$ satisfying $f(x_i, u_i) \in \mathcal{X}$, $\forall i \in \mathbb{N}_0^j$, and

$$\Psi(x_a) \neq \Psi(x_b), \forall\, x_a \neq x_b,\ x_a, x_b \in \mathcal{X}. \tag{19b}$$

Proof. In view of the argument in [38], if condition (19a) is satisfied, it promptly holds that $x_{a,j} = x_{b,j}$ implies $x_{a,0} = x_{b,0}$. Also, $x_{a,0} = x_{b,0}$ implies $\Psi(x_{a,0}) = \Psi(x_{b,0})$ since (19b) holds. Finally, $x_{a,j} = x_{b,j}$ implies $s_{a,0} = s_{b,0}$ due to $s_{a,0} = \Psi(x_{a,0})$ and $s_{b,0} = \Psi(x_{b,0})$. Hence, in view of Definition 4, (10) under $w_o = 0$ is locally observable on $\mathcal{S}_\Psi$. □
In view of Proposition 3 and Assumption 3, (8) becomes the nominal model of (10) and is used as the predictor in the online MPC deferred in Section 3.2. It is also observed that the pair $(\hat{u}, \hat{s}, \hat{x}) = (0, 0, 0)$ is an equilibrium of model (8).

Remark 5 *One can conclude from Proposition 4 and 5 that (A, B) is stabilizable and (A, C) is observable, if the infimum of Lipschitz constants L_s, L_u, i.e., $\underline{L}_s, \underline{L}_u \to 0$ (which leads to $w, v \to 0$), since model* (10) *under $w_o = 0$ is equivalent to* (8) *in*

this peculiar case. If $\underline{L}_s$, $\underline{L}_u$ are nonzeros, Assumption 4 might not be directly derived using Proposition 4 and 5. For instance, consider a special Koopman model of type:

$$\begin{cases} s_{k+1} = \mathrm{diag}\{1.01, 1\}s_k + [0\ 1]^\top u_k + w_k \\ x_k = [0\ 1]s_k + v_k, \end{cases}$$

where $w = -0.02s$, $v = [0.01\ 0]s$; which is stabilizable and observable. However, its nominal model is neither stabilizable nor observable.

Hence, we also require the following assumption about (8).

Assumption 4 *The pair (A, B) is stabilizable and the pair (A, C) is observable.*

Note that, once (6) and (7) are solved, one can verify the above assumption via calculating the stabilizability and observability matrices using the computed A, B, and C.

3.2 Robust Koopman MPC design

Inline with classic tube-based MPC, the proposed control action to Koopman model (10) relies on the feedback term of the state error correction:
$$u = \hat{u} + K(s - \hat{s}), \tag{20}$$
where K is a gain matrix such that $F = A + BK$ is Schur stable, $\hat{u}$ and $\hat{s}$ are decision variables computed with a standard MPC (deferred in (22)) with respect to (8).

Remark 6 *Different from classic tube-based MPC for linear systems [5], $\hat{u}$ and $K(\Psi(x) - \hat{s})$ in (20) are both nonlinear control laws on the original state x since $\hat{u}$ depends on $\hat{s}$ and $\hat{s}$ is related to $s = \Psi(x)$, see (25c).*

From (10) and (20), the error $e_s := s - \hat{s}$ evolves according to the following unforced system (10):

$$\begin{cases} e_s^+ = Fe_s + \bar{w} \\ e_x = Ce_s + v, \end{cases} \tag{21}$$

where $e_x = x - \hat{x}$. Let $\mathcal{Z}_s$ be a robust positively invariant set of e_s such that $\mathcal{Z}_s \subseteq F\mathcal{Z}_s \oplus \bar{\mathcal{W}}$, then it holds that $e_x \in C\mathcal{Z}_s \oplus \mathcal{V} := \mathcal{Z}_x$.

Now we are ready to state the nominal MPC problem to compute $\hat{u}$ and $\hat{s}$ in (20). At any time instant k, the following online quadratic optimization problem is to be solved:
$$\min_{\hat{s}_k, \hat{\boldsymbol{u}}_k} V(\hat{s}_k, \hat{\boldsymbol{u}}_k) \tag{22}$$

where

$$V(\hat{s}_k, \hat{\boldsymbol{u}}_k) = \sum_{i=0}^{N-1}(\|\hat{x}_{k+i}\|_Q^2 + \|\hat{u}_{k+i}\|_R^2) + V_f(\hat{s}_{k+N}), \tag{23}$$

and $\hat{s}_k$, $\hat{\boldsymbol{u}}_k$ are the decision variables, $V_f(\hat{s})$ is the terminal cost with respect to the nominal lifted state $\hat{s}$ and it is chosen as $V_f(\hat{s}) = \hat{s}^\top P\hat{s}$, where the symmetric positive-definite matrix P is the solution to the Lyapunov equation

$$F^\top PF - P = -(\bar{Q} + K^\top RK) \tag{24}$$

and $\bar{Q} = C^\top QC$.

The optimization problem (22) with (23) is performed subject to the following constraints:

1) The nominal Koopman model (8).
2) Tighter state and control constraints:

$$\hat{s}_{k+i} \in \mathcal{S}, \ i = 0, \ldots, N-1 \tag{25a}$$
$$\hat{u}_{k+i} \in \hat{\mathcal{U}}, \ i = 0, \ldots, N-1, \tag{25b}$$

8

where $\mathcal{S} = \{\hat{s}|C\hat{s} \in \mathcal{X} \ominus \mathcal{Z}_x\}, \hat{\mathcal{U}} = \mathcal{U} \ominus K\mathcal{Z}_s$.

3) The initial and terminal state constraints

$$s_k - \hat{s}_k \in \mathcal{Z}_s \tag{25c}$$
$$\hat{s}_{k+N} \in \mathcal{S}_f, \tag{25d}$$

where $\mathcal{S}_f$ is a positive invariant set of (8) under state and control constraints such that $(A + BK)\mathcal{S}_f \subseteq \mathcal{S}_f$.

Assume that, at any time instant k the optimal solution of $\hat{s}_k$ and $\hat{\boldsymbol{u}}_k^*$ can be found and is denoted by $\hat{s}_{k|k}$ and $\hat{\boldsymbol{u}}_{k|k}^*$, then the final control applied to system (1) is given as

$$u_k = \hat{u}_{k|k}^* + K(s_k - \hat{s}_{k|k}). \tag{26}$$

In summary, the pseudocode of the proposed r-KMPC is described in Algorithm 1.

Algorithm 1 Pseudocode of r-KMPC

Off-line designs:

   1) Select $\Psi(x)$ such that Assumption 2 is verified.
   2) Compute A, B, C, and D with (6) (or (14)) and (7), and check that Assumption 4 is verified.
   3) Calculate $\mathcal{W}$ and $\mathcal{V}$ according to Algorithm 2 described in Appendix A.
   4) Compute P and K with (24) and calculate the robust positively invariant set $\mathcal{Z}_s$, $\mathcal{Z}_x$ with K.
   5) Compute $\mathcal{S}, \hat{\mathcal{U}}$ and check that Assumption 5 is verified (see Section 3.3); calculate the terminal set $\mathcal{S}_f$ with K.

. . . .

On-line procedures:

At each discrete-time step $k = 1, 2, \cdots$

   1) Measure x_k and set the lifted state $s_k = \Psi(x_k)$.
   2) Solve (22) with (23) and obtain $\hat{u}_{k|k}^*$, $\hat{s}_{k|k}$.
   3) Set u_k with (26) and apply it to the nonlinear system (1).

3.3   Closed-loop property analysis

Assumption 5 *The computed nominal control and lifted state constraints are non-empty and contain the origin in their interiors, i.e., $\{\mathbf{0}\} \subset \hat{\mathcal{U}}$, $\{\mathbf{0}\} \subset \mathcal{S}$.*

Under Assumptions 1 and 2, to satisfy the above condition, one can choose a proper observable function $\Psi(x)$ and a proper number of samples M used in (6) and (7).

Theorem 1 (Recursive feasibility) *Under Assumptions 1-5, if the optimization problem* (22) *is feasible at the initial time $k = 0$, then it is recursively feasible at all times $k \in \mathbb{N}_1^\infty$, i.e., the recursive feasibility of r-KMPC is guaranteed.*

Proof. Assume at any time instant k, an optimal decision sequence of (22) with (23), i.e., $\hat{s}_{k|k}, \hat{\boldsymbol{u}}_{k|k}^*$, have been found associated with the optimal cost, denoted as V_k^*, such that the state and control constraints $\hat{s}_{k+i} \in \mathcal{S}$, $\hat{u}_{k+i} \in \hat{\mathcal{U}}$, $\forall i \in \mathbb{N}_0^{N-1}$, and the terminal constraint $\hat{s}_{k+N|k} \in \mathcal{S}_f$ is fulfilled. Hence, it holds that the real state and control constraints are verified, i.e., $x_{t|k} \in \mathcal{X}$ and $u_{t|k} \in \mathcal{U}$ for all $t \geq k$. At the next time instant $k + 1$, choose $\hat{s}_{k+1|k+1} = \hat{s}_{k+1|k}, \hat{\boldsymbol{u}}_{k+1}^s = \hat{\boldsymbol{u}}_{k+1:k+N-1|k}^*, K\hat{s}_{k+N|k}$ as a candidate sub-optimal solution, under which it follows that $Cs_{k+1} + v_{k+1} \in \mathcal{X}$, $\hat{s}_{k+1|k} \in \mathcal{S}$, leading to $s_{k+1} - \hat{s}_{k+1|k} \in \mathcal{Z}_s$, $\hat{s}_{k+j|k+1} \in \mathcal{S}$, $\forall j \in \mathbb{N}_1^N$, through inheritance, and $\hat{s}_{k+N+1|k+1} \in \mathcal{S}_f$ in view of the definition of $\mathcal{S}_f$. Hence, (22) is feasible at time $k + 1$ under $\hat{s}_{k+1|k}, \hat{\boldsymbol{u}}_{k+1}^s$, associated with a sub-optimal cost V_{k+1}^s. The recursive feasibility of r-KMPC holds. $\square$
Under Theorem 1, one can state the following theoretical results.

Theorem 2 (Closed-loop robustness) *Under Assumptions 1-5, it holds that:*

(a) the lifted nominal system (8) *using $\hat{s}_k, \hat{u}_k$ computed with* (22) *converges to the origin asymptotically, i.e., $\hat{s}_k$, $\hat{x}_k$, and $\hat{u}_k \to 0$ as $k \to +\infty$;*

9

(b) the state s and the control u of the closed-loop system (10) *with* (20) *are such that*

$$s_k \to \mathcal{Z}_s \text{ and } u_k \to K\mathcal{Z}_s \text{ as } k \to +\infty;$$

and

$$x_k \to \mathcal{Z}_x \text{ as } k \to +\infty.$$

Proof. To prove the closed-loop robustness property, first note that the optimal cost V_k^* associated with the optimal solution $\hat{s}_{k|k}^*, \hat{u}_{k|k}^*$ (cf. Theorem 1), satisfies $V_k^* \geq \|\hat{x}_k\|_Q^2$ for all $\hat{x}_k \oplus \mathcal{Z}_x \in \mathcal{X}$. Also, iterating (24) leads to $V_k^* \leq \hat{s}_k^\top P \hat{s}_k$ for $\hat{s}_k \in \mathcal{S}_f$. Considering the fact that $V_{k+1}^* \leq V_{k+1}^s$, where V_k^s is the sub-optimal cost associated with $\hat{s}_{k+1|k}, \hat{u}_{k+1}^s$ (cf. Theorem 1), and in view of (24), one has the monotonic property:

$$
\begin{aligned}
V_{k+1}^* - V_k^* &\leq -\|\hat{x}_k\|_Q^2 - \|\hat{u}_k\|_R^2 + \|\hat{s}_{k+N}\|_{F^\top PF - P + \bar{Q} + K^\top RK}^2 \\
&= -\|\hat{x}_k\|_Q^2 - \|\hat{u}_k\|_R^2.
\end{aligned}
\tag{27}
$$

Then, from (27), $V_{k+1}^* - V_k^* \to 0$ as $k \to +\infty$. Consider also

$$V_k^* - V_{k+1}^* \geq \|\hat{x}_k\|_Q^2 + \|\hat{u}_k\|_R^2,$$

hence $\hat{u}_k \to 0$, $\hat{x}_k \to 0$, as $k \to +\infty$, in view of the positive-definiteness of R and Q. Also, $\hat{s}_k \to 0$ as $k \to +\infty$ due to $\Psi(0) = 0$ in view of Assumption 3.
Moreover, recall that $x \in \hat{x} \oplus \mathcal{Z}_x$, $u \in \hat{u} \oplus K\mathcal{Z}_s$, and $s \in \hat{s} \oplus \mathcal{Z}_s$, it holds that,

$$x_k \to \mathcal{Z}_x \text{ and } u_k \to K\mathcal{Z}_s \text{ as } k \to +\infty,$$

and

$$s_k \to \mathcal{Z}_s \text{ as } k \to +\infty.$$

$\square$

Remark 7 *Note that, in the MPC problem* (22) *with* (23)*, penalizing the lifted state, i.e., $\|\hat{s}\|_{\tilde{Q}}^2$ can be used instead, where $\tilde{Q} \in \mathbb{R}^{n_\psi \times n_\psi}$ is a positive-definite tuning matrix. Herewith the observability of (A, C) is not necessarily required in* (24) *since a solution surely exists for any $\tilde{Q}$ being positive-definite. In this case, in the proof argument of Theorem 2, $\|\hat{x}\|_Q^2$ in* (27) *is replaced by $\|\hat{s}\|_{\tilde{Q}}^2$, which leads directly to the asymptotic stability of $\hat{s}$ and $\hat{x}$ since $\hat{x} = C\hat{s}$.*

Moreover, under the result in Theorem 2, point-wise asymptotic convergence of r-KMPC can be verified, under $w_o = 0$.

Theorem 3 (Point-wise convergence) *Under Assumptions 1-5, if $w_o = 0$ and*

$$E = (L_s I + L_u K^\top K)\| \sum_{k=0}^{+\infty} F^k \| \text{ is Schur stable,} \tag{28}$$

then the closed-loop systems (10) *with* (20) *and* (1) *with* (20) *converge to the origin asymptotically, i.e.,*

$$x_k \to 0, \ u_k \to 0, \text{ and } s_k \to 0 \text{ as } k \to +\infty.$$

Proof. In view of Theorem 2.(a), assume that $\hat{u}$, $\hat{x}$, and $\hat{s}$ have converged asymptotically to the origin, i.e., $\hat{u} = 0$, $\hat{x} = 0$, and $\hat{s} = 0$. Also, u, x, and s have converged asymptotically to robust tubes, i.e., $x \in \mathcal{Z}_x$, $u \in K\mathcal{Z}_s$, and $s \in \mathcal{Z}_s$. From (21), one can write

$$
\begin{cases}
s_{k+1} = Fs_k + w_k \\
x_k = Cs_k + v_k,
\end{cases}
\tag{29}
$$

where, in view of (13), the uncertainty w satisfies:

$$\begin{aligned}
\|w\| &= \|w(s, u, 0, 0)\| \\
&\leq L_s \|s\| + L_u \|u\| \\
&= L_s \|s\| + L_u \|\hat{u} + K(s - \hat{s})\| \\
&\leq \|s\|_{L_s I + L_u K^\top K}
\end{aligned} \tag{30}$$

and the last inequality is due to (20) and $\hat{u}, \hat{s} = 0$. Inline with [39], we write the evolution $s_{k+1} = F s_k + w_k$ from (29) with two redundant interconnected systems, i.e.,

$$\begin{cases} s_1^+ = F s_1 + w_2 \\ s_2^+ = F s_2 + w_1, \end{cases} \tag{31}$$

where the initial condition $s_{1,0} = s_{2,0} = s_0$ and $w_{1,0} = w_{2,0} = w_0$. In view of the small gain theorem in [39], the interconnected system converges to the origin asymptotically (31) if

$$\bar{E} = \begin{bmatrix} 0 & E \\ E & 0 \end{bmatrix} \tag{32}$$

is Schur stable, which is fulfilled by (28). Hence, the state s converges to the origin asymptotically. Moreover, $x_k \to 0$ as $k \to +\infty$ since $x_k = C s_k + v(s_k)$, $s_k \to 0$ as $k \to +\infty$, and $v(0) = 0$. Also, in view of (20), $u_k \to 0$ as $k \to +\infty$. □

4 Simulation results

4.1 Regulation of a Van der Pol oscillator

Consider a Van der Pol oscillator [20], whose continuous-time model is

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ 2x_2 - 10x_1^2 x_2 - 0.8x_1 - u \end{bmatrix} + w_o, \tag{33}$$

where x_1 and x_2 represent the position and speed respectively, while u is the force input, $\|w_o\|_\infty \leq 0.4$. Peculiarly, a sinusoidal noise $w_o = 0.4\sin(10\pi t)$ is initially adopted. The consideration on other types of disturbances are deferred in Table 1. Denoting $x = (x_1, x_2)$, the state and control are limited as $-(2.5m, 2.5m/s) \leq x \leq (2.5m, 2.5m/s)$, $-10m^2/s \leq u \leq 10m^2/s$.

In order to implement r-KMPC, we first have discretized model (33) to obtain model (1) with a suitable sampling period $T = 0.01s$. We have obtained the data-driven model using Koopman operators. The datasets of $(u, \hat{w}_o, x, x^+)$ in (6) with $M = 8 \cdot 10^5$ have been collected with a uniformly distributed random (UDR) control policy. A thinplate BF has been selected to construct the lifted observable state variables, i.e., $\Psi(x) = (x, \psi_1(x), \psi_2(x)) - (\mathbf{0}, \psi_1(0), \psi_2(0))$, where $\psi_i(x) = \|x - c_i\|^2 \log(\|x - c_i\|)$, $i = 1, 2$, $n_\psi = 4$, $c_1 = (0.381, -0.341)$, $c_2 = (0.267, -0.889)$ are the kernel centers generated with UDR numbers. The parameters of the linear predictor have been computed according to (6) and (7).

The penalty matrices $\tilde{Q}$ and R have been selected as $\tilde{Q} = \text{diag}\{1, 1, 0.1, 0.1\}$, $R = 0.1$. The terminal penalty matrix can be calculated with (24). The sets $\mathcal{W}$, $\mathcal{V}$ have been empirically obtained according to Algorithm 2, see the Appendix. The gain matrix, associated with the computation of $\mathcal{Z}_s$, has been computed as $K = [6.5\ 4.3\ -0.4\ 0.1]$, resulting in the closed-loop poles of the linear system $\{0.91,\ 0.93,\ 0.98 + 0.003i,\ 0.98 - 0.003i\}$. The robust invariant set has been obtained with (21) and the terminal constraint $\mathcal{S}_f$ has been computed, according to the algorithms described in [3]. The prediction horizon has been chosen as $N = 10$.

The proposed r-KMPC has been implemented with an initial condition $x_0 = (1.5, -1.5) \in \mathcal{X}$. The simulation tests with r-KMPC have been performed within Yalmip toolbox [40] installed in Matlab 2019a environment. The simulation results are reported in Figures 1-3. It can be seen from Figure 1 that the nominal lifted state $\hat{s}$ and control $\hat{u}$ converge to the origin; and the real state and control converge to the neighbor of the origin, under sinusoidal noises. Note that, the real state trajectory x and the real control u remain in tubes centered at $\hat{x}$ and $\hat{u}$ respectively, i.e., $x \in \hat{x} \oplus \mathcal{Z}_s$ and $u \in \hat{u} \oplus K\mathcal{Z}_x$, which verifies the robustness of the closed-loop control system, see Figure 2. Also, as shown in Figure 3, under no external perturbation, the state and control of r-KMPC converge asymptotically to the origin.

Numerical comparison with KMPC [20]: We have also compared the proposed r-KPMC with the Koopman MPC in [20] in
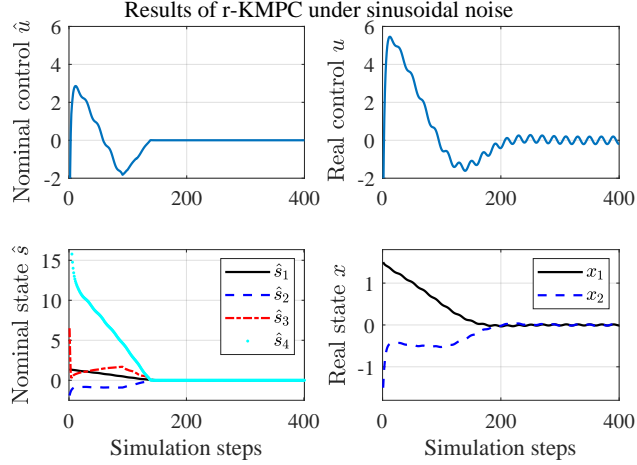
Fig. 1. The state and control of the controlled Van der Pol oscillator with r-KMPC under sinusoidal noise.
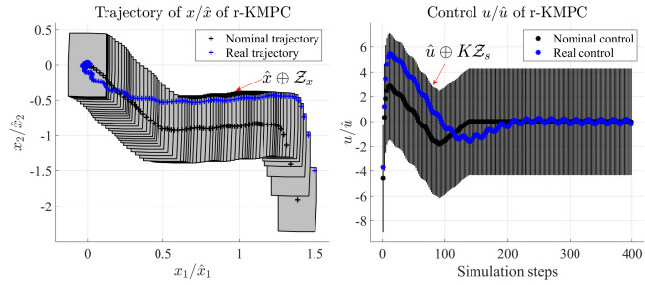

Fig. 2. The state and control tubes of the controlled Van der Pol oscillator under sinusoidal noise. In the left (right) panel, the real trajectory, marked with blue + (∗), lies in the grey tube, i.e., $\mathcal{Z}_x$ ($K\mathcal{Z}_s$), centered at the nominal one, marked with black + (∗).
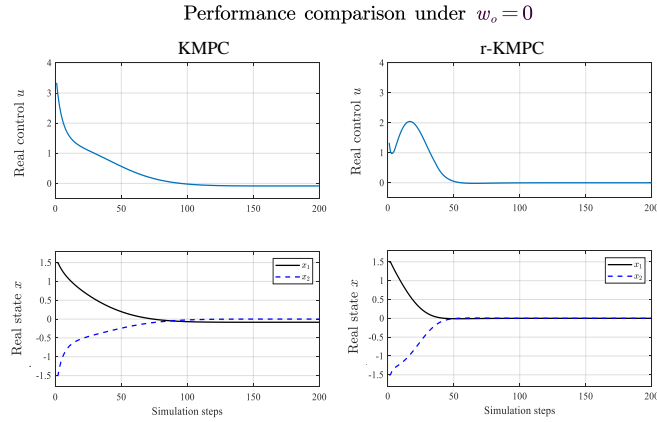

Fig. 3. Performance comparison of the controlled Van der Pol oscillator with r-KMPC and KMPC under $w_o = 0$. The state and control of r-KMPC converge asymptotically to the origin.

terms of control performance. First, we have chosen the lifted function in KMPC as $\Psi(x)$ without the resetting procedure, i.e., $\Psi(0)$ is not identically zero. The resulting one-step prediction collective square error with 50000 different initial state conditions is 56.4, comparable with 55.9 obtained by using the resetting procedure. As shown in Figure 3, the resulting controller using KMPC does not yet converge with 400 steps, while the state and control of our approach converge to the origin. For a numerical comparison, the cumulative costs of both approaches, computed as $J = \sum_{k=1}^{400} \|x_k\|_Q^2 + \|u_k\|_R^2$, are collected in Table 1, which show that the proposed r-KMPC results in a lower cost consumption, compared with KMPC [20] even using a greater number of observable functions.

Table 1
Cumulative cost comparison: "S" and "SW" stand for "sinusoidal" and "Step-wise" respectively.

Algorithm		r-KMPC	KMPC [20]		
		$n_\psi = 4$	$n_\psi = 4$	$n_\psi = 12$	$n_\psi = 22$
Cost J	Thinplate kernel				
	Nominal	258	431	401	360
	S-noise	270	449	418	374
	UDR-noise	248	413	403	371
	SW-noise	262	379	387	359
	Polynomial kernel				
	Nominal	247	427	402	–
	S-noise	257	446	419	–
	UDR-noise	241	416	403	–
	SW-noise	253	416	410	–

4.2 Angle regulation of an inverted pendulum

Consider the problem of angle regulation of an inverted pendulum, whose continuous-time model is

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ 4g\sin(x_1) - 3u\cos(x_1) \end{bmatrix} + w_o, \tag{34}$$

where $x = (x_1, x_2)$ are the state variables that represent the angle and its rate respectively, while u is the control, $w_o = 2\sin(10\pi t)$. The constraints to be enforced are $-(1rad, 2rad/s) \le x \le (1rad, 2rad/s)$, $-20m^2/s \le u \le 20m^2/s$.
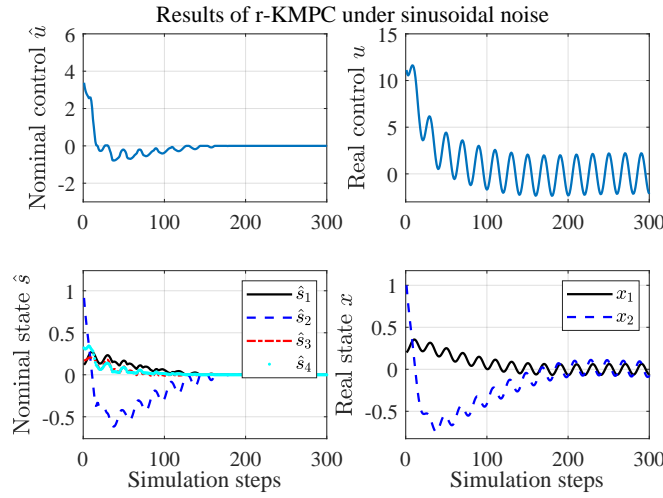

Fig. 4. The state and control of the controlled inverted pendulum with r-KMPC under sinusoidal noise.
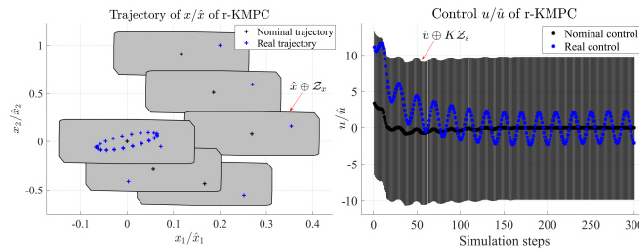

Fig. 5. The state and control tubes of the controlled inverted pendulum under sinusoidal noise. In the left (right) panel, the real trajectory, marked with blue + (∗), lies in the grey tube, i.e., $\mathcal{Z}_x$ ($K\mathcal{Z}_s$), centered at the nominal one, marked with black + (∗).

In line with Section 4.1, model (34) is discretized with $T = 0.005s$. The data samples have been obtained with $M = 5 \cdot 10^4$ under a UDR control policy. A Gaussian kernel has been selected to construct $\Psi(x)$, i.e., $\Psi(x) = (x, \psi_1(x), \psi_2(x), \psi_3(x)) -$
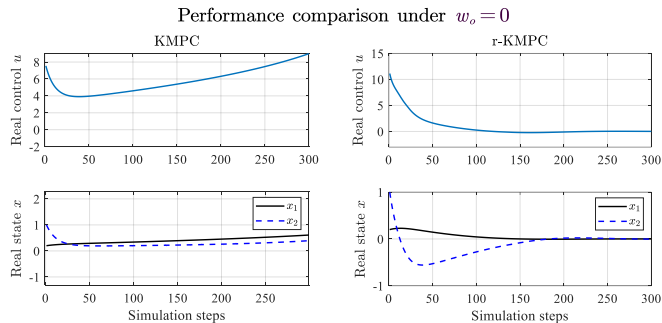
Fig. 6. Performance comparison of of the controlled inverted pendulum with r-KMPC and KMPC under $w_o = 0$. The state and control of r-KMPC converge asymptotically to the origin.

$(\mathbf{0}, \psi_1(0), \psi_2(0), \psi_3(0))$, where $\psi_i(x) = e^{-\|x - c_i\|^2}$, $i = 1, 2, 3$, $n_\psi = 5$, $c_1 = (-0.644, -1.09)$, $c_2 = (-0.99, 0.76)$, and $c_3 = (-0.26, -1.48)$ are the kernel centers generated with UDR numbers. The parameters of the linear predictor have been computed according to (6) and (7).

The penalty matrices $\tilde{Q}$ and R have been selected as $\tilde{Q} = I_5$, $R = 0.1$. The terminal penalty matrix, the sets $\bar{\mathcal{W}}$, $\mathcal{V}$, the gain matrix K, the robust invariant set, and the terminal constraint $\mathcal{S}_f$ have been computed similar to that in Section 4.1. The prediction horizon has been chosen as $N = 10$.

The proposed r-KMPC has been implemented with $x_0 = (0.2, 1)$. The simulation results are reported in Figures 4-6. It can be seen from Figures 4 and 5 that the robustness of the closed-loop control system, the satisfaction of state and control constraints as well as the convergence of the nominal system are verified. Also, under no external perturbation, the state and control of r-KMPC converge asymptotically to the origin, see Figure 6.

Numerical comparison with KMPC [20]: The comparison has been performed similar to that in Section 4.1, from the simulation

Table 2
Cumulative cost comparison: "S" and "SW" stand for "sinusoidal" and "Step-wise" respectively; "InvQuad" stands for "Inverse quadratic".

	Algorithm		R-KMPC	KMPC [20]		
			$n_\psi = 5$	$n_\psi = 5$	$n_\psi = 15$	$n_\psi = 25$
Cost J	Gaussian kernel	Nominal	175	434	228	203
		S-noise	333	695	407	367
		UDR-noise	191	570	235	189
		SW-noise	201	763	255	214
	InvQuad kernel	Nominal	173	512	252	–
		S-noise	319	871	432	–
		UDR-noise	236	546	276	–
		SW-noise	248	484	261	–

results in Figure 6, one can observe the diverging trends of the state and control in KMPC and asymptotic convergence of the state and control in our approach. The numerical comparison in terms of the cumulative costs of both approaches in Table 2 show that the proposed r-KMPC results in a lower cost consumption compared with KMPC [20] even using a greater number of observable functions.

4.3 Regulation of a non-affine system

Consider a non-affine system [41], whose continuous-time model is

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ x_1^2 + 0.15u^3 + 0.1(1 + x_2^2)u + sin(0.1u) \end{bmatrix} + w_o, \tag{35}$$

where $x = (x_1, x_2)$ are the state variables, u is the control, $w_o = \sin(10\pi t)$. The constraints to be enforced are $-2.5 \leq x \leq 2.5$, $-25 \leq u \leq 25$.

To obtain the Koopman model for control, model (35) is discretized with $T = 0.005s$. The data samples have been obtained with $M = 5 \cdot 10^4$ under a UDR control policy. The observable functions have been constructed as $\Psi(x) = (x, \psi_1(x), \psi_2(x), \psi_3(x)) -$
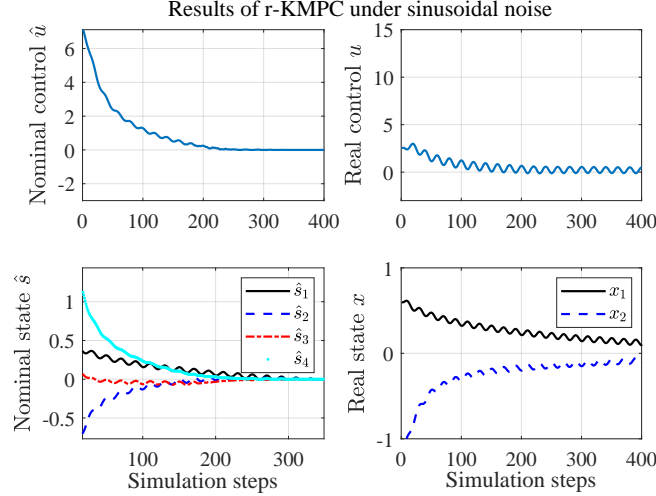
Fig. 7. The state and control of the controlled nonaffine system with r-KMPC under sinusoidal noise.
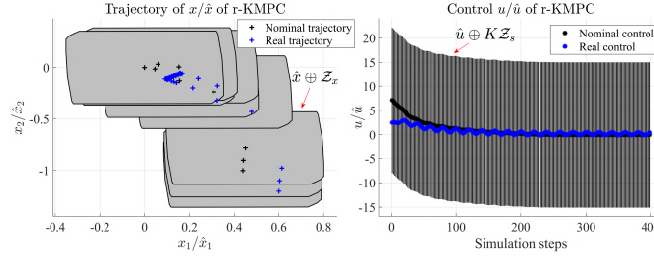

Fig. 8. The state and control tubes of the controlled non-affine system under sinusoidal noise. In the left (right) panel, the real trajectory, marked with blue + ($*$), lies in the grey tube, i.e., $\mathcal{Z}_x$ ($K\mathcal{Z}_s$), centered at the nominal one, marked with black + ($*$).
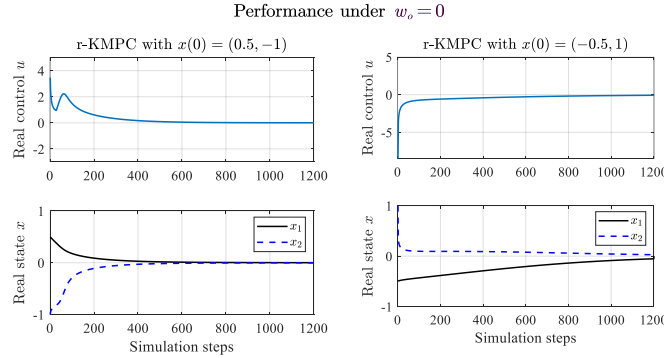

Fig. 9. Performance of the controlled nona-ffine system with r-KMPC under $w_o = 0$. The state and control converge asymptotically to the origin.

$(\mathbf{0}, \psi_1(0), \psi_2(0), \psi_3(0))$, where $\psi_i(x) = \|x - c_i\| \log(\|x - c_i\|)$, $i = 1, 2, 3$ are polyharmonic kernels, $n_\psi = 5$, c_i, $i = 1, 2, 3$ are the kernel centers generated with UDR numbers. The parameters of the Koopman model have been computed according to (6) and (7).

To control model (35) with the r-KMPC, the penalty matrices $\tilde{Q}$, R have been selected as $\tilde{Q} = I_5$, $R = 0.1$, and the prediction horizon has been chosen as $N = 30$. The terminal penalty matrix, the sets $\bar{\mathcal{W}}$, $\mathcal{V}$, the gain matrix K, the robust invariant set, and the terminal constraint $\mathcal{S}_f$ have been computed similar to that in Section 4.1. The proposed r-KMPC has been implemented with $x_0 = (0.6, -1.2)$. The simulation results in Figures 7 and 8 illustrate that the robustness of the closed-loop control system, the satisfaction of state and control constraints as well as the convergence of the nominal system are verified. Also, under no external perturbation, the state and control of r-KMPC converge asymptotically to the origin, see Figure 9.

15

5 Conclusions

In this paper, we proposed a robust tube-based model predictive control scheme, i.e., r-KMPC, for nonlinear systems with additive disturbances as well as state and control constraints. The proposed r-KMPC relies upon a lifted global linear model built using Koopman operators. The closed-loop robust controller is composed of a nonlinear control action computed by an online linear MPC using the nominal Koopman model and a nonlinear static state-feedback policy. The robustness of the closed-loop system was verified under internal modeling errors and exogenous disturbances. Moreover, the asymptotic stability of the controlled system under no exogenous disturbance was proven under mild assumptions. Simulation results verified the effectiveness of the proposed approach.

A   Statistical validations of sets $\bar{\mathcal{W}}$ and $\mathcal{V}$

The derivation of $\bar{\mathcal{W}}$ and $\mathcal{V}$ is not straightforward since they show dependencies on $\Psi(x)$ and the datasets used in (6) (or (14)) and (7). In the following, we show that the above sets can be estimated using statistic learning theory. As described in [17], the true risk of a learning problem can be estimated according to the law of large numbers. This relies upon a relaxed risk evaluation condition developed by Hoeffding (1963):

$$P\left(\left|\hat{G}(f_l) - G(f_l)\right| \geq \epsilon\right) \leq \delta_r \tag{A.1}$$

where $\delta_r := 2\exp\left(-2L\epsilon^2\right)$, $\hat{G}(f_l)$ and $G(f_l)$ are the empirical and true risk respectively, f_l is the learned function, and L is the number of samples. For fixed f_l, provided a large number of datasets, the empirical risk is a good estimate of the true one.

Loss function for validation: Let L tuples $(u_i, \hat{w}_{o,i}, x_i, x_i^+)$ be generated under Assumption 1. Let $\bar{w}_i = \Psi(x_i^+) - \hat{s}_i^+$ where $\hat{s}_i^+$ is computed by (8) with $\hat{s} = \Psi(x_i)$ and $\hat{u} = u_i$ and let $v_i = x_i - C\Psi(x_i)$. We define the loss function as

$$\ell_\star(u_i, x_i, x_i^+) = \begin{cases} 0 & \text{if } g_\star = 1 \\ 1 & \text{otherwise} \end{cases} \tag{A.2}$$

where $\star = \bar{w}, v$ in turns, $g_{\bar{w}_i} = 1$ stands for $\bar{w}_i \in \bar{\mathcal{W}}$ and $g_{v_i} = 1$ stands for $v_i \in \mathcal{V}$.

With (A.2), for any $\star = \bar{w}, v$, the empirical risk of the learned linear predictor is $\hat{G}_\star = \frac{1}{L}\sum_{i=1}^{L}\ell_\star(u_i, x_i, x_i^+)$. In view of (A.1), $P_\star(\ell = 0) = G_\star \leq \hat{G}_\star + \epsilon_\star$. For the validation of uncertainty sets, we choose $\bar{G}_\star \geq G_\star$ and $\delta_r > 0$, then we check if for the L samples that [17]

$$\bar{G}_\star \geq \hat{G}_\star + \epsilon_\star \tag{A.3}$$

with a confidence level of $1 - \delta_r$, where $\epsilon_\star = \sqrt{-\log\left(0.5\delta_r\right)/2L}$.

The computation steps of $\bar{\mathcal{W}}$ and $\mathcal{V}$ are given in Algorithm 2. The rationale behind the uncertainty scaling in step 6)

Algorithm 2 Computation of $\bar{\mathcal{W}}$ and $\mathcal{V}$

---

1) Set initial sets $\bar{\mathcal{W}}_I$ and $\mathcal{V}_I$;
2) Set $\bar{G}_\star \geq G_\star$ and δ_r, we calculate $\hat{G}_\star$ with (A.2) and $\epsilon_\star = \sqrt{-\log\left(0.5\delta_r\right)/2L}$, for $\star = \bar{w}, v$;
if Condition (A.3) is satisfied for $\star = \bar{w}$ (or $\star = v$) **then**
    3) Return $\bar{\mathcal{W}}_I$ (or $\mathcal{V}_I$);
else
    4) Increase L or enlarge the size of $\bar{\mathcal{W}}_I$ (or $\bar{\mathcal{V}}_I$);
    5) Go back to step 2);
end if
6) Set $\bar{\mathcal{W}} = \gamma_w \bar{\mathcal{W}}_I$ and $\mathcal{V} = \gamma_v \mathcal{V}_I$ where $\gamma_w, \gamma_v > 1$.

---

of Algorithm 2 is given in the following proposition. Denote $\mathcal{D}_h = \{h_i\}_{i=1}^{L}$, $h_i = x_i, u_i, \hat{w}_{o,i}$ in turns; define $d_x = \max_{x,y \in \mathcal{D}_x, y \in x_\mathcal{N}}\|x - y\|$, $d_u = \max_{u \in \mathcal{D}_u, y \in y_\mathcal{N}}\|u - y\|$, $d_{\hat{w}} = \max_{w \in \mathcal{D}_{w_o}, y \in u_\mathcal{N}}\|w - y\|$, where $z_\mathcal{N}$ is the collection of neighbor point of z for $z = x, u, w$. Let $d_{\delta w}$ be the maximal estimation deviation of the real disturbance and the estimated one. The following proposition can be stated.

Proposition 6 (Uncertainty set scaling) *The real uncertainty sets* $\bar{\mathcal{W}}, \mathcal{V}$ *satisfy* $\bar{\mathcal{W}} \subseteq \bar{\mathcal{W}}_I \oplus \Delta\mathcal{W}$, $\mathcal{V} \subseteq \mathcal{V}_I \oplus \Delta\mathcal{V}$, *where* $\Delta\mathcal{W} = \{w \in \mathbb{R}^{n_\psi} \mid \|w\| \leq L_s L_\Psi d_x + L_u d_u + L_{\delta w} d_{\delta w} + L_{\hat{w}} d_{\hat{w}}\}$, $\Delta\mathcal{V} = \{v \in \mathbb{R}^n \mid \|v\| \leq L_v d_x\}$, L_Ψ *and* L_v *are the local Lipschitz constants of* $\Psi(x)$ *and* $x - \Psi^{-1}$ *respectively.*

Proof. Let $w_1 \in \bar{\mathcal{W}}_I \ominus D\hat{\mathcal{W}}_o$, $w_2 \in \bar{\mathcal{W}} \backslash (\bar{\mathcal{W}}_I \ominus D\hat{\mathcal{W}}_o)$. In view of (13), one has that $\|w_1 - w_2\| \leq L_s L_\Psi d_x + L_u d_u + L_{\delta w} d_{\delta w} + L_{\hat{w}} d_{\hat{w}}$ where the second inequality follows from the Lipschitz continuity of $\Psi(x)$. Hence, $\mathcal{W} \subseteq \bar{\mathcal{W}}_I \oplus \Delta \mathcal{W}$. Likewise, let $v_1 \in \mathcal{V}_I$, $v_2 \in \mathcal{V} \backslash \mathcal{V}_I$, then one has that $\|v_1 - v_2\| \leq L_v d_x$. Hence, $\mathcal{V} \subseteq \mathcal{V}_I \oplus \Delta \mathcal{V}$. $\qquad \square$

References

[1] David Q Mayne, James B Rawlings, Christopher V Rao, and Pierre OM Scokaert. Constrained model predictive control: Stability and optimality. *Automatica*, 36(6):789–814, 2000.

[2] S Joe Qin and Thomas A Badgwell. A survey of industrial model predictive control technology. *Control Engineering Practice*, 11(7):733–764, 2003.

[3] James Blake Rawlings and David Q Mayne. *Model predictive control: Theory and design*. Nob Hill Pub., 2009.

[4] Alberto Bemporad, Francesco Borrelli, and Manfred Morari. Min-max control of constrained uncertain discrete-time linear systems. *IEEE Transactions on Automatic Control*, 48(9):1600–1606, 2003.

[5] David Q Mayne, María M Seron, and SV Raković. Robust model predictive control of constrained linear systems with bounded disturbances. *Automatica*, 41(2):219–224, 2005.

[6] David Q Mayne, Erric C Kerrigan, EJ Van Wyk, and Paola Falugi. Tube-based robust nonlinear model predictive control. *International Journal of Robust and Nonlinear Control*, 21(11):1341–1353, 2011.

[7] Dario Piga, Simone Formentin, and Alberto Bemporad. Direct data-driven control of constrained systems. *IEEE Transactions on Control Systems Technology*, 26(4):1422–1429, 2017.

[8] Andrea Carron, Elena Arcari, Martin Wermelinger, Lukas Hewing, Marco Hutter, and Melanie N Zeilinger. Data-driven model predictive control for trajectory tracking with a robotic arm. *IEEE Robotics and Automation Letters*, 4(4):3758–3765, 2019.

[9] Enrico Terzi, Lorenzo Fagiano, Marcello Farina, and Riccardo Scattolini. Learning-based predictive control for linear systems: A unitary approach. *Automatica*, 108:108473, 2019.

[10] Sebastian Peitz, Samuel E Otto, and Clarence W Rowley. Data-driven model predictive control using interpolated Koopman generators. *SIAM Journal on Applied Dynamical Systems*, 19(3):2162–2193, 2020.

[11] Monimoy Bujarbaruah, Xiaojing Zhang, Marko Tanaskovic, and Francesco Borrelli. Adaptive MPC under time varying uncertainty: Robust and stochastic. *arXiv preprint arXiv:1909.13473*, 2019.

[12] Johannes Köhler, Elisa Andina, Raffaele Soloperto, Matthias A Müller, and Frank Allgöwer. Linear robust adaptive model predictive control: Computational complexity and conservatism. In *2019 IEEE 58th Conference on Decision and Control (CDC)*, pages 1383–1388. IEEE, 2019.

[13] Matthias Lorenzen, Frank Allgöwer, and Mark Cannon. Adaptive model predictive control with robust constraint satisfaction. *IFAC-PapersOnLine*, 50(1):3313–3318, 2017.

[14] Lorenzo Fagiano, Georg Schildbach, Marko Tanaskovic, and Manfred Morari. Scenario and adaptive model predictive control of uncertain systems. *IFAC-PapersOnLine*, 48(23):352–359, 2015.

[15] Ugo Rosolia and Francesco Borrelli. Learning model predictive control for iterative tasks. a data-driven control framework. *IEEE Transactions on Automatic Control*, 63(7):1883–1896, 2017.

[16] D Limon, J Calliess, and Jan Marian Maciejowski. Learning-based nonlinear model predictive control. *IFAC-PapersOnLine*, 50(1):7769–7776, 2017.

[17] Michael Hertneck, Johannes Köhler, Sebastian Trimpe, and Frank Allgöwer. Learning an approximate model predictive controller with guarantees. *IEEE Control Systems Letters*, 2(3):543–548, 2018.

[18] Hassan Arbabi, Milan Korda, and Igor Mezic. A data-driven Koopman model predictive control framework for nonlinear flows. *arXiv preprint arXiv:1804.05291*, 2018.

[19] Milan Korda and Igor Mezić. On convergence of extended dynamic mode decomposition to the Koopman operator. *Journal of Nonlinear Science*, 28(2):687–710, 2018.

[20] Milan Korda and Igor Mezić. Linear predictors for nonlinear dynamical systems: Koopman operator meets model predictive control. *Automatica*, 93:149–160, 2018.

[21] Yingzhao Lian, Renzi Wang, and Colin N Jones. Koopman based data-driven predictive control. *arXiv preprint arXiv:2102.05122*, 2021.

[22] Yiqiang Han, Wenjian Hao, and Umesh Vaidya. Deep learning of Koopman representation for control. In *2020 59th IEEE Conference on Decision and Control (CDC)*, pages 1890–1895. IEEE, 2020.

[23] Abhinav Narasingam and Joseph Sang-Il Kwon. Data-driven feedback stabilization of nonlinear systems: Koopman-based model predictive control. *arXiv preprint arXiv:2005.09741*, 2020.

[24] Matthew O Williams, Ioannis G Kevrekidis, and Clarence W Rowley. A data–driven approximation of the Koopman operator: Extending dynamic mode decomposition. *Journal of Nonlinear Science*, 25(6):1307–1346, 2015.

[25] Sang Hwan Son, Abhinav Narasingam, and Joseph Sang-Il Kwon. Handling plant-model mismatch in Koopman Lyapunov-based model predictive control via offset-free control framework. *arXiv preprint arXiv:2010.07239*, 2020.

[26] Andrea Bacciotti. The local stabilizability problem for nonlinear systems. *IMA Journal of Mathematical Control and Information*, 5(1):27–39, 1988.

[27] Frank H Clarke. Generalized gradients and applications. *Transactions of the American Mathematical Society*, 205:247–262, 1975.

[28] Francis Clarke. On the inverse function theorem. *Pacific Journal of Mathematics*, 64(1):97–102, 1976.

[29] Stefan Klus, Feliks Nüske, and Boumediene Hamzi. Kernel-based approximation of the Koopman generator and schrödinger operator. *Entropy*, 22(7):722, 2020.

[30] Stefan Klus, Feliks Nüske, Sebastian Peitz, Jan-Hendrik Niemann, Cecilia Clementi, and Christof Schütte. Data-driven approximation of the Koopman generator: Model reduction, system identification, and control. *Physica D: Nonlinear Phenomena*, 406:132416, 2020.

[31] Matthew O Williams, Maziar S Hemati, Scott TM Dawson, Ioannis G Kevrekidis, and Clarence W Rowley. Extending data-driven Koopman analysis to actuated systems. *IFAC-PapersOnLine*, 49(18):704–709, 2016.

[32] Joshua L Proctor, Steven L Brunton, and J Nathan Kutz. Generalizing koopman theory to allow for inputs and control. *SIAM Journal on Applied Dynamical Systems*, 17(1):909–930, 2018.

[33] Amit Surana and Andrzej Banaszuk. Linear observer synthesis for nonlinear systems using Koopman operator framework. *IFAC-PapersOnLine*, 49(18):716–723, 2016.

[34] Carl Folkestad, Daniel Pastor, Igor Mezic, Ryan Mohr, Maria Fonoberova, and Joel Burdick. Extended dynamic mode decomposition with learned Koopman eigenfunctions for prediction and control. In *2020 American Control Conference (ACC)*, pages 3906–3913. IEEE, 2020.

[35] Samuel E Otto and Clarence W Rowley. Koopman operators for estimation and control of dynamical systems. *Annual Review of Control, Robotics, and Autonomous Systems*, 4:59–87, 2021.

[36] Sebastian Peitz and Stefan Klus. Koopman operator-based model reduction for switched-system control of PDEs. *Automatica*, 106:184–191, 2019.

[37] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[38] Francesca Albertini and Domenico D'Alessandro. Remarks on the observability of nonlinear discrete time systems. In *System Modelling and Optimization*, pages 155–162. Springer, 1996.

[39] Marcello Farina, Xinglong Zhang, and Riccardo Scattolini. A hierarchical multi-rate MPC scheme for interconnected systems. *Automatica*, 90:38–46, 2018.

[40] J. Löfberg. Yalmip : A toolbox for modeling and optimization in matlab. In *In Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004.

[41] Shuzhi Sam Ge, Chang Chieh Hang, and Tao Zhang. Adaptive neural network control of nonlinear systems by state and output feedback. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 29(6):818–828, 1999.