Modeling Nonlinear Control Systems via Koopman Control Family: Universal Forms and Subspace Invariance Proximity

Masih Haseli [a] Jorge Cortés [a]

[a] *Department of Mechanical and Aerospace Engineering, University of California, San Diego, {mhaseli,cortes}@ucsd.edu*

Abstract

This paper introduces the Koopman Control Family (KCF), a mathematical framework for modeling general (not necessarily control-affine) discrete-time nonlinear control systems with the aim of providing a solid theoretical foundation for the use of Koopman-based methods in systems with inputs. We demonstrate that the concept of KCF captures the behavior of nonlinear control systems on a (potentially infinite-dimensional) function space. By employing a generalized notion of subspace invariance under the KCF, we establish a universal form for finite-dimensional models, which encompasses the commonly used linear, bilinear, and linear switched models as specific instances. In cases where the subspace is not invariant under the KCF, we propose a method for approximating models in general form and characterize the model's accuracy using the concept of invariance proximity. We end by discussing how the proposed framework naturally lends itself to data-driven modeling of control systems.

1 Introduction

The Koopman operator approach to dynamical systems has gained widespread attention in recent years. While traditional state-space methods for nonlinear systems rely on the description of system trajectories, the Koopman viewpoint provides an equivalent formulation of the system behavior using a linear operator acting on a vector space of functions. The Koopman operator framework yields beneficial algebraic constructions that can be leveraged for efficient computational learning and prediction. These benefits have motivated researchers to consider extending the framework to control systems. However, unlike the case of autonomous systems [1], this has proven difficult due to the fact that the role of input is fundamentally different from the state's role: without a priori knowledge of the input signal, there is not enough information to predict the system trajectories since the choice of input can drastically alter the system behavior. Our aim here is to provide a comprehensive mathematical framework for Koopman operator-based modeling of control systems.

Literature Review: The Koopman operator [24] provides an alternative description of nonlinear *autonomous* systems that encodes the system behavior through the evolution of functions (a.k.a., observables) belonging to a vector space. Even though the system might be nonlinear, the Koopman operator is always linear. This inherent linearity gives rise to favorable algebraic properties, leading to powerful tools to analyze complex dynamical systems [7, 33, 41] for which typical state-space and geometric methods are cumbersome. However, linearity comes at the expense of the infinite–dimensional nature of the operator. To make possible its direct and efficient implementation on digital computers, one needs to develop finite-dimensional descriptions for it. This generally relies on the concept of subspace invariance [6]. If a finite-dimensional subspace is invariant under the operator, then one can restrict the operator to the subspace and represent its action with a matrix given a chosen basis. This has led to a search for invariant subspaces through a variety of techniques, including the direct identification of eigenfunctions (which span invariant subspaces) [21, 27], optimization and neural network-based methods [22, 29, 36, 37, 48, 49, 53], and efficient algebraic searches [13, 14].

Even without a finite-dimensional invariant subspace available, one can still approximate the action of the Koopman operator on any finite-dimensional subspace via an orthogonal projection. Prominent data-driven methods in this category are Dynamic Mode Decomposition (DMD) [42, 50] and its variant, Extended Dynamic Mode Decomposition (EDMD) [26, 52]. In addition, [30] provides methods to handle time-varying systems. For such methods, criteria that must be balanced to choose the finite-dimensional space are the relevance of the dynamical information captured by the subspace and the accuracy of the resulting approximation. The work [16] provides a tight upper bound on the worst-case prediction error on a subspace, providing a measure of the quality of the subspace independently of the chosen basis. The works [28, 35] provide several error bounds for accuracy of DMD, EDMD, and extensions to Koopman-based control models. The work [15] provides an algebraic algorithm to approximate Koopman-invariant subspaces of an arbitrary finite-dimensional space with tunable accuracy.

The algebraic properties of the Koopman operator have been used in a myriad applications, including fluid dynamics [41], stability analysis [9, 32, 54, 55], reachability analysis [1, 2, 43, 51], safety-critical control [3, 10, 56], and robotics [5, 31, 44].

⋆ This work was supported by ONR Award N00014-23-1-2353 and NSF Award IIS-2007141.

[1] Consistent with the literature on the Koopman operator, we use the term "autonomous" to describe systems without input. Note that this terminology is different from the one used in classical system theory, which refers to time-invariant systems as autonomous.

The fact that the Koopman operator is only formally defined for autonomous systems has not been an obstacle for the development of many data-based methods inspired by it to construct low-complexity representations of control systems. Many advances do not directly require an operator-theoretic approach for the open control system, but instead rely on the Koopman operator of the unforced system and address control design based on control Lyapunov functions [34, 57] or feedback linearization [11]. A significant amount of attention has been devoted to deriving finite-dimensional forms by lifting to higher dimensions. Due to their simplicity, lifted linear models are the most popular in the literature [6, 25] and lead to highly efficient computational algorithms. Such models, however, are not capable of capturing certain structures, such as terms containing the products of inputs and states, which are prevalent in control-affine nonlinear systems. For these, the works [12, 19, 39, 47] propose the use of bilinear lifted models based on geometric arguments relying on the control-affine structure. The work [20] proposes a different lifted form based on invariant subspaces for the Koopman operator associated with the unforced system. An interesting alternative is to model the system by switching between several Koopman operators, each associated with the system under a different constant [4, 38] or piecewise constant [45, 46] input signal. The work [25] takes a different approach and considers the system behavior under all possible infinite input sequences, defining a Koopman operator for the control system on a function space whose members' domain is the Cartesian product of the state space and all possible input sequences. This is perhaps the most direct approach in terms of an operator-theoretic viewpoint for controls systems. However, given the reliance on infinite input sequences, working with finite-time restrictions should be done with care. Here, we take a different operator-theoretic approach to capture the behavior of control systems that we find easier to work with on finite-dimensional subspaces with only finitely many input steps available.

Statement of Contributions: Our goal here is to provide a solid theoretical framework to model general discrete-time nonlinear *control* systems based on Koopman operator theory. The starting point of our approach is the observation made in the literature that if the input is a constant signal, then the control system becomes an autonomous dynamics. Therefore, any Koopman-theoretic model for the control system must reduce to the conventional Koopman operator associated with the corresponding autonomous system. Motivated by this, we define the concept of *Koopman control family (KFC)* as the collection of all Koopman operators associated with constant-input autonomous dynamics derived from the control system. We show that the KCF can completely capture the control system behavior on a potentially infinite-dimensional function space. Since dealing with infinite-dimensional operators is computationally intractable, we provide a theoretical structure for finite-dimensional models whose construction is based on projection operators, analogous to the case of autonomous systems. To find a general finite-dimensional form for Koopman-based models for the control system, we rely on a generalized notion of subspace invariance. Specifically, we show that on a common-invariant subspace for the KCF, the finite-dimensional model always has a specific *"input-state separable"* form. Remarkably, the linear, bilinear, and switched linear models commonly used in the literature are all special cases of the input-state separable form. Since KCF contains uncountably many operators (given the infinite choices for the constant input signal), directly finding a common invariant subspace is challenging. To tackle this, we parametrize the KCF with one operator, termed augmented Koopman operator, and show that invariant subspaces under this augmented operator lead to common invariant subspaces for the KCF. As a result, the problem of working with uncountably many operators simplifies to working with a single linear operator on an extended function space. Similarly to the case of autonomous systems, finding an exact and informative finite-dimensional common invariant subspace for the KCF is generally challenging and in some cases might not even exist. To address this, we define the concept of invariance proximity under an operator, which enables us to extended our results to approximations on non-invariant subspaces and provide a bound on the accuracy of the resulting approximated models for the control system. Our final contribution shows how the results of the paper can directly be used in data-driven modeling applications.

Notation: The symbols $\mathbb{N}$, $\mathbb{R}$, and $\mathbb{C}$, represent the sets of natural, real, and complex numbers, respectively. Given $A \in \mathbb{C}^{m\times n}$, we denote its transpose, pseudo-inverse, conjugate transpose, Frobenius norm and range space by A^T, $A^\dagger$, A^H, $\|A\|_F$, and $\mathcal{R}(A)$, respectively. If A is square, A^{-1} and $\mathrm{Tr}(A)$ denote its inverse and trace respectively. When all eigenvalues of A are real, $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$ denote the smallest and largest eigenvalues of A. We use I_m and $\mathbf{0}_{m\times n}$ to denote the $m \times m$ identity matrix and $m \times n$ zero matrix (we drop the indices when appropriate). We denote the 2-norm of the vector $v \in \mathbb{C}^n$ by $\|v\|_2$. Given sets S_1 and S_2, their union and intersection are represented by $S_1 \cup S_2$ and $S_1 \cap S_2$, respectively. Also, $S_1 \subseteq S_2$ and $S_1 \subsetneq S_2$ mean that S_1 is a subset and proper subset of S_2, respectively. Given the vector space $\mathcal{V}$ defined on the field $\mathbb{C}$, $\dim \mathcal{V}$ denotes its dimension. Given a set $\mathcal{S} \subseteq \mathcal{V}$, $\mathrm{span}(\mathcal{S})$ is a vector space comprised of all linear combinations of elements in $\mathcal{S}$. Given vector spaces $\mathcal{V}_1$ and $\mathcal{V}_2$, we define $\mathcal{V}_1 + \mathcal{V}_2 := \{v_1 + v_2 \mid v_1 \in \mathcal{V}_1,\, v_2 \in \mathcal{V}_2\}$. Given functions f and g with appropriate domains and co-domains, $f \circ g$ denotes their composition. Let $f : A \times B \to C$ be a multivariable function yielding $f(a, b)$ for $(a, b) \in A \times B$. Then, $f\restriction_{b=b^*}\colon A \to C$ is defined as $f\restriction_{b=b^*}(a) := f(a, b^*)$ for $a \in A$. If F consists of multivariable functions of the form $f : A \times B \to C$, then $F\restriction_{b=b^*} := \{f\restriction_{b=b^*} \mid f \in F\}$. Given a positive measure μ on a set A and functions $f, g : A \to \mathbb{C}$, we define their L_2 inner product as $\langle f, g\rangle_{L_2(\mu)} := \int_A f(x)\bar{g}(x)d\mu(x)$, where $\bar{g}$ is the complex conjugate of g. The L_2 norm of f is defined as $\|f\|_{L_2(\mu)} = \sqrt{\langle f, f\rangle_{L_2(\mu)}}$. We drop the dependency on the measure μ when the context is clear.

2 Preliminaries

In this section, we review notions and results regarding the Koopman operator, Extended Dynamic Mode Decomposition, and the concept of consistency index.

2.1 Koopman Operator

Here, we briefly explain the Koopman operator associated with a dynamical system and its properties following the terminology in [7]. Consider a discrete-time system

$$x^+ = T(x), \tag{1}$$

with state space $\mathcal{X} \subseteq \mathbb{R}^n$, where $T : \mathcal{X} \to \mathcal{X}$ is the function describing the system behavior. Consider a linear function

space $\mathcal{F}$ (defined on the field $\mathbb{C}$) comprised of functions form $\mathcal{X}$ to $\mathbb{C}$ and assume it is closed under composition with T, i.e., $f \circ T \in \mathcal{F}$ for all $f \in \mathcal{F}$. We define the Koopman operator $\mathcal{K} : \mathcal{F} \to \mathcal{F}$ as

$$\mathcal{K}f = f \circ T. \tag{2}$$

It is easy to verify that (2) is a linear operator, i.e.,

$$\mathcal{K}(\alpha f + \beta g) = \alpha \mathcal{K}f + \beta \mathcal{K}g, \ \forall f, g \in \mathcal{F}, \ \forall \alpha, \beta \in \mathbb{C}. \tag{3}$$

The Koopman operator's action on a given function can be viewed as pushing forward the function values across all system trajectories by one time step. We can repeatedly apply the Koopman operator on a function $f \in \mathcal{F}$ to predict its evolution on any system trajectory $\{x(i)\}_{i=0}^{\infty}$ as

$$f(x(i)) = \mathcal{K}^i f(x(0)), \quad \forall i \in \mathbb{N}_0. \tag{4}$$

Since $\mathcal{K}$ is a linear operator, we can define its eigendecomposition. We say the function $\phi \in \mathcal{F}$ is an *eigenfunction* of $\mathcal{K}$ with *eigenvalue* λ if

$$\mathcal{K}\phi = \lambda\phi. \tag{5}$$

By comparing (4) and (5), one can see that Koopman eigenfunctions evolve linearly on system trajectories,

$$\phi(x(i)) = \lambda\phi(x(i-1)), \quad \forall i \in \mathbb{N}. \tag{6}$$

We refer to (6) as *temporal linear evolution* of eigenfunctions. This temporal linearity of eigenfunctions combined with the linearity (3) of the operator on the space $\mathcal{F}$ enables us to linearly predict function values on system trajectories. Specifically, given eigenfunctions $\{\phi_k\}_{k=1}^{N_k}$ with corresponding eigenvalues $\{\lambda_k\}_{k=1}^{N_k}$, one can write the evolution of the function $f = \sum_{k=1}^{N_k} c_k \phi_k$ on the system trajectories as

$$f(x(i)) = \sum_{k=1}^{N_k} c_k \lambda_k^i \phi_k(x(0)), \ \forall i \in \mathbb{N}_0.$$

This equation is of utmost importance since it provides a linear structure facilitating the prediction of nonlinear systems [25, 27] as well as learning the system behavior from data [14, 15, 41]. One should keep in mind that, in general, to capture the full state of the system, one might need $\mathcal{F}$ to be infinite dimensional since it must be closed under composition with T.

Next, we define the concept of subspace invariance under the Koopman operator. A subspace $\mathcal{G} \subseteq \mathcal{F}$ is *Koopman invariant* if $\mathcal{K}f \in \mathcal{G}$ for all $f \in \mathcal{G}$. Koopman eigenfunctions trivially span invariant subspaces.

Remark 2.1 *(Simplifying Notation For Vector-Valued Functions):* For convenience, we introduce some notation simplifying the operation of the Koopman operator on finite-dimensional spaces. Let $\Psi : \mathcal{X} \to \mathbb{C}^s$ be a vector-valued map represented as $\Psi(\cdot) = [\psi_1(\cdot), \ldots, \psi_s(\cdot)]^T$, where $\psi_i : \mathcal{X} \to \mathbb{C}$ for all $i \in \{1, \ldots, s\}$. We define the span of the elements of Ψ and action of Koopman operator on the elements of Ψ as

$$\text{span}(\Psi) := \text{span}(\{\psi_1, \ldots, \psi_s\}),$$

$$\mathcal{K}\Psi := [\mathcal{K}\psi_1, \ldots, \mathcal{K}\psi_s]^T = \Psi \circ T.$$

Given a finite-dimensional subspace $\mathcal{H} \subset \mathcal{F}$, we often describe a basis for it by a vector-valued map $\Phi : \mathcal{X} \to \mathbb{C}^{\dim(\mathcal{H})}$ satisfying $\mathcal{H} = \text{span}(\Phi)$. □

An important property of finite-dimensional Koopman-invariant subspaces is that one can capture the action of the operator by a matrix once a basis is chosen. Formally, given the invariant subspace $\mathcal{S} \subseteq \mathcal{F}$ with basis $\Psi : \mathcal{X} \to \mathbb{C}^{\dim(\mathcal{S})}$, there exists a unique matrix $K \in \mathbb{C}^{\dim(\mathcal{S}) \times \dim(\mathcal{S})}$ such that

$$\mathcal{K}\Psi = \Psi \circ T = K\Psi. \tag{7}$$

This equation combined with the linearity of the operator allow us to easily calculate the action of the operator on any function in $\mathcal{S}$. Formally, given any function $f \in \mathcal{S}$ with description $f(\cdot) = w^T \Psi(\cdot)$ where $w \in \mathbb{C}^{\dim(\mathcal{S})}$, one has

$$\mathcal{K}f = w^T K \Psi. \tag{8}$$

The concept of subspace invariance is of utmost importance since it allows us to operate on finite-dimensional subspaces and use numerical matrix computation for prediction, as laid out in equations (7)-(8).

Even if the subspace $\mathcal{S} \subseteq \mathcal{F}$ is *not* invariant under $\mathcal{K}$, it is still possible to use the notion of subspace invariance to *approximate* the action of $\mathcal{K}$ on $\mathcal{S}$. To achieve this, one usually utilizes $\mathcal{P}_{\mathcal{S}} : \mathcal{F} \to \mathcal{F}$, the orthogonal projection operator (given an inner product on $\mathcal{F}$) on $\mathcal{S}$. Observe that the space $\mathcal{S}$ is invariant under the operator $\mathcal{P}_{\mathcal{S}}\mathcal{K} : \mathcal{F} \to \mathcal{F}$; hence, equations (7)-(8) are valid when we substitute in them the Koopman operator $\mathcal{K}$ by $\mathcal{P}_{\mathcal{S}}\mathcal{K}$. Let K_{approx} be the matrix calculated by applying equation (7) to the operator $\mathcal{P}_{\mathcal{S}}\mathcal{K}$. Then, this matrix provides an approximation for the action of $\mathcal{K}$ on $\mathcal{S}$ as follows

$$\mathcal{K}\Psi = \Psi \circ T \approx \mathcal{P}_{\mathcal{S}}\mathcal{K}\Psi = K_{\text{approx}}\Psi. \tag{9}$$

Moreover, the analogous but approximated version of function prediction in (8) is given by

$$\mathcal{K}f \approx \mathcal{P}_{\mathcal{S}}\mathcal{K}f = w^T K_{\text{approx}}\Psi. \tag{10}$$

Remark 2.2 *(General Linear Form and Subspace Invariance):* When dealing with the action of the Koopman operator on finite-dimensional spaces, we use linear models that are either exact, cf. (7)-(8), or approximated, cf. (9)-(10). Note that in either case the model has the same form. It is in this sense that we say that the general finite-dimensional form of Koopman-based models is linear. Note that this general form is a consequence of the notion of subspace invariance (whether the subspace is actually Koopman-invariant or not). □

2.2 Extended Dynamic Mode Decomposition

In many engineering applications, the system dynamics is unknown and we only have access to data from the system trajectories. The Extended Dynamic Mode Decomposition (EDMD) method [52] reviewed here uses data to approximate the action of the Koopman operator on a given *finite-dimensional* space of functions.

Remark 2.3 *(Use of Real-valued Basis Functions in Data-Driven Applications):* All the systems in this paper are defined on state and input spaces with real-valued elements. Consequently, the Koopman operator's action on any pair of complex-conjugate functions leads to another complex-conjugate pair, which can be captured by a pair of real-valued functions. Hence, even though we develop our theory based on complex functions, in data-driven applications, we work with bases with real-valued elements to simplify the numerical operations, without loss of generality [2]. □

To specify the function space, EDMD uses a dictionary comprised of N_Ψ functions form $\mathcal{X}$ to $\mathbb{R}$. Formally, we define our dictionary as a vector-valued function

$$\Psi(\cdot) = [\psi_1(\cdot), \dots, \psi_{N_\Psi}(\cdot)]^T,$$

where $\psi_1, \dots, \psi_{N_\Psi} \in \mathcal{F}$ are the dictionary elements. To approximate the behavior of the Koopman operator (and therefore the system) on $\text{span}(\Psi)$, EDMD uses data snapshots from the trajectories in two matrices $X, X^+ \in \mathbb{R}^{n \times N}$ such that

$$x_i^+ = T(x_i),\ \forall i \in \{1, \dots, N\}, \tag{11}$$

where x_i and x_i^+ are the ith columns of matrices X and X^+ respectively. For convenience, we define the action of the dictionary on data matrix X (similarly for any data matrix) as

$$\Psi(X) = [\Psi(x_1), \Psi(x_2), \dots, \Psi(x_n)] \in \mathbb{R}^{N_\psi \times N}.$$

Note that based on (2) and (11), one can see $\Psi(X^+) = \Psi \circ T(X) = \mathcal{K}\Psi(X)$. Hence, the dictionary matrices $\Psi(X)$ and $\Psi(X^+)$ capture the behavior of the Koopman operator on $\text{span}(\Psi)$. EDMD approximates the action of the operator by solving a least-squares problem

$$\underset{K}{\text{minimize}} \|\Psi(X^+) - K\Psi(X)\|_F, \tag{12}$$

with the following closed-form solution

$$K_{\text{EDMD}} = \Psi(X^+)\Psi(X)^\dagger. \tag{13}$$

Throughout this paper, we make the following assumption.

Assumption 2.4 *(Full Rank Dictionary Matrices):* $\Psi(X)$ *and* $\Psi(X^+)$ *have full row rank.* □

Note that Assumption 2.4 implies that the element of Ψ are linearly independent. Also, it implies that data in X and X^+ are diverse enough to distinguish between functions in $\text{span}(\Psi)$. Moreover, if Assumption 2.4 holds, K_{EDMD} is the unique solution of (12).

The matrix K_{EDMD} captures relevant information about the system behavior and can be used to approximate the action of the Koopman operator on $\text{span}(\Psi)$. Formally, we define the EDMD predictor of $\mathcal{K}\Psi$ by EDMD as

$$\mathfrak{P}^{\text{EDMD}}_{\mathcal{K}\Psi} := K_{\text{EDMD}}\Psi. \tag{14}$$

[2] Given a vector-valued function Ψ with real-valued elements, $\text{span}(\Psi)$ contains complex-valued functions since we employ $\mathbb{C}$ as the underlying field.

Similarly, for an arbitrary function $f \in \text{span}(\Psi)$ with description $f(\cdot) = w^T\Psi(\cdot)$ for $w \in \mathbb{C}^{N_\Psi}$, one can define the EDMD predictor of $\mathcal{K}f$ as

$$\mathfrak{P}^{\text{EDMD}}_{\mathcal{K}f} := w^T K_{\text{EDMD}}\Psi. \tag{15}$$

The predictors (14)-(15) are special cases of the approximations in (9)-(10), where the orthogonal projection corresponds to the $L_2(\mu_X)$ inner product, with empirical measure

$$\mu_X = \frac{1}{N}\sum_{i=1}^{N} \delta_{x_i}, \tag{16}$$

where δ_{x_i} is the Dirac measure defined on the ith column of X (see e.g., [26]). The quality of predictors in (14)-(15) depends on the quality of $\text{span}(\Psi)$ in terms of being close to invariant under $\mathcal{K}$. If $\text{span}(\Psi)$ is invariant under $\mathcal{K}$, then the predictors in (14)-(15) are exact and match equations (7)-(8), respectively. Determining closeness to invariance necessitates an appropriate metric, which is the concept we review next.

2.3 Consistency Index Measures The Dictionary's Quality

We recall the concept of temporal forward-backward consistency to measure how close a dictionary span is to being Koopman invariant. Given a dictionary Ψ with real-valued elements and data matrices X, X^+, the *consistency index* [16] is

$$\mathcal{I}_C(\Psi, X, X^+) = \lambda_{\max}(I - K_F K_B), \tag{17}$$

where $K_F = \Psi(X^+)\Psi(X)^\dagger$ and $K_B = \Psi(X)\Psi(X^+)^\dagger$ are EDMD matrices applied forward and backward in time [3]. When the context is clear, we drop the arguments and use $\mathcal{I}_C$.

The intuition behind the consistency index is that when $\text{span}(\Psi)$ is Koopman-invariant, the forward and backward EDMD matrices K_F and K_B are inverse of each other. Otherwise, their product will deviate from the identity matrix, with the consistency index providing a measure for this deviation. The consistency index is easy to compute based on data and its value only depends on the vector space, not on the choice of particular basis. The following result states a key property of relevance to the ensuing discussion.

Theorem 2.5 *([16, Theorem 1]: $\sqrt{\mathcal{I}_C}$ Bounds the Relative L_2-norm error for EDMD's Koopman Predictions): Given Assumption 2.4 for dictionary Ψ, data matrices X, X^+, empirical measure* (16)*, consistency index* (17)*, and the predictor of EDMD defined in* (15)*, we have*

$$\sqrt{\mathcal{I}_C(\Psi, X, X^+)} = \max_{f \in \text{span}(\Psi)} \frac{\|\mathcal{K}f - \mathfrak{P}^{\text{EDMD}}_{\mathcal{K}f}\|_{L_2(\mu_X)}}{\|\mathcal{K}f\|_{L_2(\mu_X)}}.$$

The maximum above is taken over all functions leading to nonzero denominator (when the denominator is zero, the numerator is also zero and the prediction is exact). □

[3] Note that this definition is equivalent but different from [16, Definition 1]. The data matrices in this paper are transpose of the ones in [16]; however, this transposition does not affect the value of the consistency index.

Remark 2.6 *(Properties of the Consistency Index and Implications for Learning):* The consistency index (17) provides a notion of worst-case error bound for Koopman predictions on the vector space $\text{span}(\Psi)$. Note that $\mathcal{I}_C(\Psi, X, X^+) \in [0, 1]$ for all subspaces and data. Moreover, the index does not depend on the choice of basis for the subspace: for example, if Φ provides a different basis for $\text{span}(\Psi)$, we have $\mathcal{I}_C(\Psi, X, X^+) = \mathcal{I}_C(\Phi, X, X^+)$. Therefore, one can use the consistency index as an effective loss function for subspace learning [4]. We refer the reader to [16, Proposition 1 and Lemma 1] for more information. □

3 Motivation and Problem Statement

Consider the discrete-time control system

$$x^+ = \mathcal{T}(x, u), \quad x \in \mathcal{X} \subseteq \mathbb{R}^n,\ u \in \mathcal{U} \subseteq \mathbb{R}^m, \tag{18}$$

where x and u are the state and input vectors, $\mathcal{X}$ and $\mathcal{U}$ are the state and input spaces, and $\mathcal{T} : \mathcal{X} \times \mathcal{U} \to \mathcal{X}$ is the function describing the system dynamics. Note that no special structure (e.g., control affine) is assumed on the system (18). Our goal is to provide a Koopman operator theory description of the nonlinear control system. The challenge for extending the concept of the Koopman operator to systems with inputs is that unlike the autonomous system (1), the behavior of the control system (18) cannot be determined without knowledge of the input sequence [5]. Here, we aim to provide a rigorous mathematical description of how to employ the Koopman operator for control systems in both infinite- and finite-dimensional cases, and articulate its application in data-driven modeling of control systems. We formalize the problem next.

Problem 3.1 *(Challenges Regarding the Extension of Koopman Theory to Control Systems): We aim to provide a mathematical framework based on the Koopman operator to*

- *(a) capture the behavior of control system* (18)*;*
- *(b) provide a generalized notion of subspace invariance, leading to a general form for finite-dimensional Koopman-based models that*
  - *(i) encompasses commonly used linear, bilinear, and linear switched control models;*
  - *(ii) is consistent with the linear form for autonomous systems in* (7) *and* (9)*: if we set the input to be constant (which yields an autonomous system), the finite-dimensional form should reduce to the lifted linear model in* (7) *and* (9)*;*
- *(c) evaluate the accuracy of such finite-dimensional models;*
- *(d) use for data-driven identification of control systems.*

[4] The residual error of EDMD $\|\Psi(X^+) - K_{\text{EDMD}}\Psi(X)\|_F$ depends on the choice of basis and is not suitable for measuring quality of $\text{span}(\Psi)$. In fact, it is easy to show [16, Example 1] that, if $\text{span}(\Psi)$ is not invariant but contains one exact eigenfunction, then one can find a linear transformation on the dictionary to make the residual error *arbitrarily* close to zero.

[5] Given an infinite input sequence, one can determine [25] the system behavior completely and define a Koopman operator for it. Moreover, if one closes the loop by means of feedback, the system takes the autonomous form (1) and hence has a well-defined Koopman operator, see e.g., [21].

4 Koopman Control Family and General Form for Finite-Dimensional Models

Here, we take the first step towards extending Koopman theory to the control system case and providing a generalized notion of subspace invariance. As we show below, this ultimately leads to a finite-dimensional model that is the extension of the linear form in (7). We start from the observation that, if we fix the input as a constant, we get an autonomous system in the form of (1) which admits a well-defined Koopman operator. Motivated by this idea, one can model the system (18) by switching between constant input autonomous systems [6]. Formally, consider the family of autonomous systems created by setting the input as a constant signal

$$x^+ = \mathcal{T}_{u^*}(x) := \mathcal{T}(x, u \equiv u^*), \quad u^* \in \mathcal{U}. \tag{19}$$

Note that any trajectory $\{x_k\}_{k=0}^{L} \subset \mathcal{X}$ of system (18) generated with input sequence $\{u_k\}_{k=0}^{L-1} \subset \mathcal{U}$, can be generated by applying the autonomous systems $\mathcal{T}_{u_k}, k \in \{0, \dots, L-1\}$ subsequently on the initial condition x_0. Hence, we have

$$x_k = \mathcal{T}_{u_{k-1}} \circ \cdots \circ \mathcal{T}_{u_0}(x_0) \tag{20a}$$
$$= \mathcal{T}_{u_{k-1}}(x_{k-1}), \quad k \in \{1, \dots, L\}. \tag{20b}$$

Noting that the members of the family $\{\mathcal{T}_{u^*}\}_{u^* \in \mathcal{U}}$ are all autonomous, we can define Koopman operators for each of them as in (2), leading to the following definition.

Definition 4.1 *(Koopman Control Family (KCF)): Let $\mathcal{F}$ be a vector space (over the field $\mathbb{C}$) of complex-valued functions with domain $\mathcal{X}$ that is closed under composition with members of $\{\mathcal{T}_{u^*}\}_{u^* \in \mathcal{U}}$. The associated* ***Koopman control family (KCF)*** *is the family of operators $\{\mathcal{K}_{u^*} : \mathcal{F} \to \mathcal{F}\}_{u^* \in \mathcal{U}}$ where, for each $u^* \in \mathcal{U}$, $\mathcal{K}_{u^*}$ defined by $\mathcal{K}_{u^*} f = f \circ \mathcal{T}_{u^*}$, for all $f \in \mathcal{F}$, is the Koopman operator corresponding to the dynamics $\mathcal{T}_{u^*}$.* □

Similarly to the multi-step prediction (4) under the Koopman operator of an autonomous system, one can use (20a) and the definition of KCF to provide a similar identity for the non-autonomous case,

$$f(x_k) = \mathcal{K}_{u_0}\mathcal{K}_{u_1} \dots \mathcal{K}_{u_{k-1}} f(x_0), \quad \forall f \in \mathcal{F}.$$

Note that the identity above is *exact* and general and can be utilized for all trajectories of (18).

Even though a KCF on an appropriate function space can completely capture the behavior of control system (18), the infinite-dimensional nature of the function space $\mathcal{F}$ can make its implementation on digital computers impossible. To address this issue, we need finite-dimensional representations for KCF. A simple guiding observation in this regard is that if we have an exact finite-dimensional representation and fix the input to be constant, then the system is autonomous and the model should reduce to a linear finite-dimensional case similar to equation (7). This leads us to the concept of common invariant subspaces under KCF.

Definition 4.2 *(Common Invariant Subspaces Under the Koopman Control Family): The space $\mathcal{L} \subseteq \mathcal{F}$ is a* ***common***

[6] The idea of modeling a control system via constant input systems has already been considered several times in the literature [38, 45, 46].

invariant subspace *under the KCF if* $\mathcal{K}_{\bar{u}} f \in \mathcal{L}$*, for all* $\mathcal{K}_{\bar{u}} \in \{\mathcal{K}_{u^*}\}_{u^* \in \mathcal{U}}$ *and all* $f \in \mathcal{L}$. □

Finite-dimensional common invariant subspaces under the KCF $\{\mathcal{K}_{u^*}\}_{u^* \in \mathcal{U}}$ are of utmost importance because the action of all its members on such subspaces can be captured *exactly* by matrices. This provides a general framework for treating control systems. Next, we show that finite-dimensional common invariant subspaces under KCF lead to a universal form of models that can be viewed as a generalization of (7) to the case of control systems [7].

Theorem 4.3 *(General Form on Common Invariant Subspaces:* ***Input-State Separable****): The Koopman control family has a finite-dimensional (of dimension* s*) common invariant subspace if and only if there exist functions* $\Psi : \mathcal{X} \to \mathbb{C}^s$ *and* $A : \mathcal{U} \to \mathbb{C}^{s \times s}$ *such that for all* $(x, u) \in \mathcal{X} \times \mathcal{U}$*,*

$$\Psi(x^+) = \Psi \circ \mathcal{T}(x, u) = \mathcal{A}(u)\Psi(x). \tag{21}$$

In this formulation, the common invariant subspace under the KCF is described by $\operatorname{span}(\Psi)$*.*

PROOF. ($\Rightarrow$) : Let $\mathcal{S} \subset \mathcal{F}$, with $\dim \mathcal{S} = s$, be a common invariant subspace of the Koopman control family $\{\mathcal{K}_{u^*}\}_{u^* \in \mathcal{U}}$. Let functions $\{\psi_1, \ldots, \psi_s\}$ be a basis for $\mathcal{S}$ and define the vector-valued function $\Psi : \mathcal{X} \to \mathbb{C}^s$ as $\Psi(x) = [\psi_1(x), \ldots, \psi_s(x)]^T$ for all $x \in \mathcal{X}$. Since $\mathcal{S} = \operatorname{span}(\Psi)$ is invariant under the KCF, for *each* $u^* \in \mathcal{U}$, there exists a matrix $K_{u^*} \in \mathbb{C}^{s \times s}$ (which represents the action of operator $\mathcal{K}_{u^*}$ in KCF on subspace $\mathcal{S}$ with respect to basis Ψ) such that

$$\mathcal{K}_{u^*} \Psi(x) = K_{u^*} \Psi(x), \ \forall x \in \mathcal{X}, \tag{22}$$

where we have used (7) and the notation in Remark 2.1. Define then the matrix-valued function $\mathcal{A} : \mathcal{U} \to \mathbb{C}^{s \times s}$ as

$$\mathcal{A}(u^*) = K_{u^*},$$

for each $u^* \in \mathcal{U}$. Noting that equation (22) holds for all $u^* \in \mathcal{U}$, one can use the definition of $\mathcal{A} : \mathcal{U} \to \mathbb{C}^{s \times s}$ and write

$$\Psi \circ \mathcal{T}(x, u^*) = K_{u^*} \Psi(x) = \mathcal{A}(u^*)\Psi(x), \ \forall x \in \mathcal{X}, \ \forall u^* \in \mathcal{U}.$$

Noting that the equation above holds for all $u^* \in \mathcal{U}$, one can do a change of symbol (u^* to u), leading to (21).

($\Leftarrow$) : Assume equation (21) holds. Hence, for all $u^* \in \mathcal{U}$,

$$\Psi \circ \mathcal{T}(x, u \equiv u^*) = \Psi \circ \mathcal{T}_{u^*}(x) = \mathcal{A}(u \equiv u^*)\Psi(x), \forall x \in \mathcal{X}.$$

Given that $\mathcal{A}(u \equiv u^*)$ is a constant matrix, for any function $f \in \operatorname{span}(\Psi)$ in the form of $f = v_f^T \Psi$ with $v_f \in \mathbb{C}^s$, we have

$$\mathcal{K}_{u^*} f = f \circ \mathcal{T}_{u^*} = v_f^T \Psi \circ \mathcal{T}_{u^*} = v_f^T \mathcal{A}(u \equiv u^*)\Psi \in \operatorname{span}(\Psi).$$

This equality holds for all $u^* \in \mathcal{U}$; hence, $\operatorname{span}(\Psi)$ is invariant under the Koopman control family $\{\mathcal{K}_{u^*}\}_{u^* \in \mathcal{U}}$. □

[7] Similarly to the case of autonomous systems (cf. Remark 2.2 and its preceding discussions), we rely on a notion of subspace invariance to find a general finite-dimensional form. In the following sections, we also rigorously investigate approximations on non-invariant subspaces.

The input-state separable form (21) (note the composition on the left and the matrix product on the right) can be viewed as a generalization of (7), which describes the exact action of the Koopman operator on an invariant subspace. Next, we discuss the special structure of the input-state separable models and their closed-form solutions.

Remark 4.4 *(Linearity in Lifted State and Closed-Form Solution of Input-State Separable Models):* By defining the lifting map $x \mapsto z := \Psi(x)$, one can rewrite the input-state separable form in (21) as $z^+ = \mathcal{A}(u)z$. This system is linear in the lifted state z and has a closed-from solution: given initial condition x_0 and input sequence $\{u_i\}_{i=0}^{\infty}$, one can write

$$z_k = \Big(\prod_{i=0}^{k-1} \mathcal{A}(u_i) \Big) z_0, \quad \forall k \in \mathbb{N},$$

where $\prod_{i=0}^{k-1} \mathcal{A}(u_i) := \mathcal{A}(u_{k-1})\mathcal{A}(u_{k-2}) \cdots \mathcal{A}(u_0)$ and $z_0 = \Psi(x_0)$. □

It is important to note that the condition in Theorem 4.3 is *necessary and sufficient (and hence cannot be relaxed)*; therefore the input-state separable form is general. In fact, as we show next, it provides a mathematical framework encompassing common Koopman-inspired descriptions of the control system (18). It is easy to see that the linear switched systems used in [38] are a special case of the input-state separable form where the input space $\mathcal{U}$ contains finitely many elements. We formalize this observation in the following result that follows directly from the definition of input-state separable form.

Lemma 4.5 *(Linear Switched Form is a Special Case of Input-State Separable Form): For system* (18)*, let* $\mathcal{U} = \{u_1, \ldots, u_l\}$ *and assume the system has a lifted linear switched representation of the form*

$$\Psi(x^+) = A_u \Psi(x), \ A_u \in \{A_{u_1}, \ldots, A_{u_l}\} \subset \mathbb{R}^{N_\Psi \times N_\Psi}, \tag{23}$$

where $\Psi : \mathcal{X} \to \mathbb{R}^{N_\Psi}$ *with* $N_\Psi \in \mathbb{N}$ *and* $u \in \mathcal{U}$*. Then,* $\operatorname{span}(\Psi)$ *is a finite-dimensional common invariant subspace under the KCF associated with the system and* (23) *is an input-state separable representation.* □

Next, we show that the commonly used linear and bilinear Koopman-based models are also special cases of the input-state separable form.

Lemma 4.6 *(Linear and Bilinear Forms are Special Cases of Input-State Separable Form): Assume the system* (18) *has a finite-dimensional lifted representation of the form*

$$\psi(x^+) = A\psi(x) + \sum_{i=1}^{m} B_i \psi(x) u_i + Cu, \tag{24}$$

where $\psi : \mathcal{X} \to \mathbb{R}^{N_\psi}$ *with* $N_\psi \in \mathbb{N}$*. Moreover,* $A, B_i \in \mathbb{R}^{N_\psi \times N_\psi}$ *for* $i \in \{1, \ldots, m\}$ *and* $C \in \mathbb{R}^{N_\psi \times m}$ *where* m *is the dimension of the input vector. Then* $\operatorname{span}(\psi) + \operatorname{span}(1_{\mathcal{X}})$ *is a finite-dimensional common invariant subspace under the KCF associated with the system* [8]*, which has the input-state*

[8] Here, $1_{\mathcal{X}} : \mathcal{X} \to \mathbb{C}$ is the constant function defined by $1_{\mathcal{X}}(x) = 1$ for all $x \in \mathcal{X}$.

separable representation

$$\begin{bmatrix} \psi(x^+) \\ 1_{\mathcal{X}}(x^+) \end{bmatrix} = \begin{bmatrix} A + \sum_{i=1}^m u_i B_i & Cu \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \psi(x) \\ 1_{\mathcal{X}}(x) \end{bmatrix}. \tag{25}$$

PROOF. Using the constant function $1_{\mathcal{X}}$, one can rewrite the dynamics (24) as (25), which is in input-state separable form. Therefore, based on Theorem 4.3, $\text{span}(\psi) + \text{span}(1_{\mathcal{X}})$ is a finite-dimensional common invariant subspace under the KCF associated with the system. □

Remark 4.7 *(Existence of Linear or Bilinear Forms Implies Common Invariant Subspaces of KCF):* Lemma 4.6 shows that if a system has a linear or bilinear lifted form, then its associated KCF has a common invariant subspace. However, the converse does not hold, as corroborated by the necessary and sufficient condition in Theorem 4.3. Therefore, for a system to have a linear or bilinear lifted form, stronger conditions than the existence of common invariant subspace under KCF are required. □

Note that linear and bilinear models are special cases of the model in (24). Therefore, the input-state separable model captures these important special cases.

Example 4.8 *(Input-State Separable Form):* Consider

$$\begin{aligned} x_1^+ &= ax_1 + bu \\ x_2^+ &= cx_2 + dx_1^2 + ex_1u + fu + g\sin(u) + h \end{aligned} \tag{26}$$

where x_1, x_2 are the state variables and u is the input. The system has the input-state separable form

$$\begin{bmatrix} x_1 \\ x_2 \\ x_1^2 \\ 1 \end{bmatrix}^+ = \begin{bmatrix} a & 0 & 0 & bu \\ eu & c & d & fu + g\sin(u) + h \\ 2abu & 0 & a^2 & b^2u^2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_1^2 \\ 1 \end{bmatrix}. \tag{27}$$

Note that for any constant $u \equiv u^*$, the system turns into an exact lifted linear form on a Koopman invariant subspace (compare with the linear switched model in [38] and Lemma 4.5). If $b = g = 0$, (27) turns into the following bilinear form (cf. Lemma 4.6)

$$\begin{bmatrix} x_1 \\ x_2 \\ x_1^2 \\ 1 \end{bmatrix}^+ = \begin{bmatrix} a & 0 & 0 & 0 \\ 0 & c & d & h \\ 0 & 0 & a^2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_1^2 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ e & 0 & 0 & f \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_1^2 \\ 1 \end{bmatrix} u.$$

If in addition we have $e = 0$, the previous equation can be written in linear form. □

So far, we have established the KCF modeling can completely capture the behavior of the control system (18). Moreover, we have found the general form of finite-dimensional models on the common invariant subspaces of KCF. However, given that, in general, the KCF contains uncountably many linear operators, one needs to find tractable ways to find or approximate finite-dimensional common invariant subspaces under the KCF. We tackle this task in the following sections.

5 Parameterizing the Koopman Control Family

Here we provide a way to parametrize a Koopman Control Family via a single linear operator defined on an augmented function space. This allows us to provide an equivalent characterization for a common invariant subspace under the KCF.

5.1 Augmented Koopman Operator

To parametrize the KCF, we first parametrize the family of constant input systems in (19) as the following augmented dynamical system

$$\begin{bmatrix} x \\ u \end{bmatrix}^+ = \begin{bmatrix} \mathcal{T}(x,u) \\ u \end{bmatrix}.$$

For convenience, we define the following tuple notation for the system above

$$(x^+, u^+) = \mathcal{T}^{\text{aug}}(x,u) := (\mathcal{T}(x,u), u), \tag{28}$$

for $(x,u) \in \mathcal{X} \times \mathcal{U}$. Note that in (28), u is a part of the state vector and not an input. The next result shows that this augmented system captures the behavior of all members of constant-input systems defined in (19).

Lemma 5.1 *(Augmented System Parametrizes the Constant-Input Family): For the augmented system* (28)*, the following hold:*

(a) the set $\mathcal{X} \times \{u^*\}$ is forward invariant under (28) *for all $u^* \in \mathcal{U}$;*

(b) for any $u^* \in \mathcal{U}$, let $\{x_i\}_{i=1}^\infty$ be a trajectory of $\mathcal{T}_{u^*}$ in (19) *with initial condition $x_0 \in \mathcal{X}$ and let $\{(x_i^{\text{aug}}, u_i^{\text{aug}})\}_{i=1}^\infty$ be a trajectory of $\mathcal{T}^{\text{aug}}$ starting from $(x_0^{\text{aug}}, u^*) \in \mathcal{X} \times \mathcal{U}$ with $x_0^{\text{aug}} = x_0$. Then, $x_i = x_i^{\text{aug}}$ for all $i \in \mathbb{N}$.* □

The proof of Lemma 5.1 directly follows from the definition of system (28) and is omitted for space reasons. As a result of Lemma 5.1(a), if we restrict the state space of (28) to $\mathcal{X} \times \{u^*\}$ for any $u^* \in \mathcal{U}$, we get a well-defined dynamics, which we denote by $\mathcal{T}^{\text{aug}} \restriction_{\mathcal{X} \times \{u^*\}}$. Moreover, by Lemma 5.1(b), $\mathcal{T}^{\text{aug}} \restriction_{\mathcal{X} \times \{u^*\}}$ captures the behavior of $\mathcal{T}_{u^*}$ for all $u^* \in \mathcal{U}$. It is in this sense that we say that $\mathcal{T}^{\text{aug}}$ on the state space $\mathcal{X} \times \mathcal{U}$ parametrizes the family of constant-input systems $\{\mathcal{T}_{u^*}\}_{u^* \in \mathcal{U}}$.

Since the augmented system is an autonomous dynamics in the form of (1), we can define a Koopman operator as given in (2). Appropriately defined, this operator would encompass the KCF's information, as supported by Lemma 5.1, which connects the augmented system (28) to the constant-input systems (19). Nonetheless, before making this connection, we must first bridge the gap between the state-space of constant-input systems ($\mathcal{X}$) and that of the augmented system ($\mathcal{X} \times \mathcal{U}$), and define a proper function space. To do this, we first provide the following definition.

Definition 5.2 *(Control-Independent Extension of Functions in $\mathcal{F}$ to Domain $\mathcal{X} \times \mathcal{U}$): Given the function $\phi \in \mathcal{F}$ where $\phi : \mathcal{X} \to \mathbb{C}$, we define its* ***control-independent extension*** *to the domain $\mathcal{X} \times \mathcal{U}$ as $\phi_e : \mathcal{X} \times \mathcal{U} \to \mathbb{C}$,*

$$\phi_e(x,u) = \phi(x) 1_{\mathcal{U}}(u), \quad \forall (x,u) \in \mathcal{X} \times \mathcal{U},$$

where $1_{\mathcal{U}} : \mathcal{U} \to \mathbb{C}$ is defined as $1_{\mathcal{U}}(u) = 1$ for all $u \in \mathcal{U}$. Similarly, for a vector-valued function $\Phi(x) =$

$[\phi_1(x), \ldots, \phi_n(x)]^T$, *where* $\phi_i \in \mathcal{F}$ *for all* $i \in \{1, \ldots, n\}$, *we define* $\Phi_e(x, u) = [\phi_1(x)1_{\mathcal{U}}(u), \ldots, \phi_n(x)1_{\mathcal{U}}(u)]^T$. □

One could equivalently define the control-independent extension as $\phi_e(x, u) = \phi(x)$ for $(x, u) \in \mathcal{X} \times \mathcal{U}$. However, the structure of Definition 5.2 is consistent with input-state separable forms, which is particularly convenient in our forthcoming theoretical analysis. Next, we state straightforward but useful properties of control-independent extensions that follow from the definition.

Lemma 5.3 *(Control-Independent Extensions' Properties): Let* $\phi : \mathcal{X} \to \mathbb{C}$ *and* $\Phi : \mathcal{X} \to \mathbb{C}^n$, *and let* $I_{\mathcal{U}}^{n\times n} : \mathcal{U} \to \mathbb{C}^{n\times n}$ *be a constant function returning the identity matrix, i.e.,* $I_{\mathcal{U}}^{n\times n}(u) = I_{n\times n}$ *for all* $u \in \mathcal{U}$. *Then, for all* $(x, u) \in \mathcal{X} \times \mathcal{U}$,

(a) $\phi_e(x, u) = \phi(x)$ *and* $\Phi_e(x, u) = \Phi(x)$;
(b) $\Phi_e(x, u) = I_{\mathcal{U}}^{n\times n}(u)\Phi(x)$;
(c) for all $f \in \text{span}(\Phi)$ *with description* $f = v_f^T\Phi$ *where* $v_f \in \mathbb{C}^n$, *we have* $f_e = v_f^T\Phi_e$. □

We next define a proper function space for the Koopman operator associated with the augmented system (28).

Definition 5.4 *(Function Space and Koopman Operator for* $\mathcal{T}^{\text{aug}}$*): Let* $\mathcal{F}^{\text{aug}}$ *be a linear space (on the field* $\mathbb{C}$*) of complex-valued functions with domain* $\mathcal{X} \times \mathcal{U}$ *such that*

(a) is closed under composition with $\mathcal{T}^{\text{aug}}$;
(b) contains $f \circ \mathcal{T}$ *for all* $f \in \mathcal{F}$;
(c) contains the control-independent extension f_e *for all* $f \in \mathcal{F}$;
(d) for all $u^* \in \mathcal{U}$, $\mathcal{F}^{\text{aug}}\restriction_{u=u^*} = \mathcal{F}$.

Then, the ***augmented Koopman operator****,* $\mathcal{K}^{\text{aug}} : \mathcal{F}^{\text{aug}} \to \mathcal{F}^{\text{aug}}$ *defined as*

$$\mathcal{K}^{\text{aug}} g = g \circ \mathcal{T}^{\text{aug}}, \quad \forall g \in \mathcal{F}^{\text{aug}}, \tag{29}$$

encodes the behavior of the augmented system (28). □

Note that, as long as we allow the function spaces $\mathcal{F}$ and $\mathcal{F}^{\text{aug}}$ to be infinite-dimensional, the conditions in Definition 5.4 are easy to satisfy. The choice of $\mathcal{F}$ and $\mathcal{F}^{\text{aug}}$ depends on $\mathcal{T}$ and $\mathcal{T}^{\text{aug}}$ and the choices are not unique. Here, we provide a few generic examples for $\mathcal{F}$ and $\mathcal{F}^{\text{aug}}$:

- one of the simplest examples is to choose $\mathcal{F}$ and $\mathcal{F}^{\text{aug}}$ as the spaces of bounded complex-valued functions on domains $\mathcal{X}$ and $\mathcal{X} \times \mathcal{U}$, resp.;
- if $\mathcal{T}$ in (18) is continuous in both variables (e.g., with respect to usual metrics inherited from $\mathbb{R}^n$ and $\mathbb{R}^m$), then $\mathcal{F}$ and $\mathcal{F}^{\text{aug}}$ can be chosen as the spaces of continuous complex-valued functions on domains $\mathcal{X}$ and $\mathcal{X} \times \mathcal{U}$, resp.;
- if $\mathcal{T}$ is polynomial in both variables, then $\mathcal{F}$ and $\mathcal{F}^{\text{aug}}$ can be chosen as the spaces of all polynomials with complex coefficients on $\mathcal{X}$ and $\mathcal{X} \times \mathcal{U}$, resp.

It is also important to note that, at this point of the exposition, there are no requirements on these function spaces having inner products, norms, or metrics.

5.2 Augmented Koopman Operator Parametrizes the Koopman Control Family

Here, we investigate the connection between the augmented operator and the KCF, and the implications for the search of common invariant subspaces for the KCF. The next result shows how $\mathcal{K}^{\text{aug}}$ parameterizes the KCF $\{\mathcal{K}_{u^*}\}_{u^*\in\mathcal{U}}$.

Lemma 5.5 *(*$\mathcal{K}^{\text{aug}}$ *Parametrizes the KCF): Let* $f \in \mathcal{F}^{\text{aug}}$. *Then for all* $u^* \in \mathcal{U}$ *we have* $(\mathcal{K}^{\text{aug}} f)\restriction_{u=u^*} = \mathcal{K}_{u^*}(f\restriction_{u=u^*})$. □

The proof of Lemma 5.5 directly follows from the definition of $\mathcal{K}^{\text{aug}}$ and is omitted for space reasons. Lemma 5.5 establishes the important fact that the action of $\mathcal{K}^{\text{aug}}$ on $\mathcal{F}^{\text{aug}}$ completely captures the effect of $\mathcal{K}_{u^*}$ on $\mathcal{F}^{\text{aug}}\restriction_{u=u^*} = \mathcal{F}$ (cf. Definition 5.4) for all $u^* \in \mathcal{U}$. This shows that $\mathcal{K}^{\text{aug}}$ can be viewed as a parametrization of the KCF, i.e., by knowing the effect of $\mathcal{K}^{\text{aug}}$ on $\mathcal{F}^{\text{aug}}$, one can calculate the effect of all (potentially uncountably many) members of the KCF. The next result shows how the augmented Koopman operator can capture relevant information regarding the evolution of functions in $\mathcal{F}$ under the trajectories of the control system (18).

Lemma 5.6 *(Augmented Koopman Operator Predicts the Functions Evolutions on System Trajectories): Let* $f \in \mathcal{F}$ *and denote by* $f \circ \mathcal{T} \in \mathcal{F}^{\text{aug}}$ *the function created by pushing the values of* f *one time-step forward through the trajectories of* $\mathcal{T}$. *Let* f_e *be the control-independent extension of* f *to* $\mathcal{X} \times \mathcal{U}$. *Then,* $f \circ \mathcal{T} = \mathcal{K}^{\text{aug}} f_e$. □

The proof of Lemma 5.6 directly follows from the definition of $\mathcal{K}^{\text{aug}}$ and Lemma 5.3(a), and is omitted for space reasons. Lemma 5.6 provides a crucial tool to analyze the behavior of functions in $\mathcal{F}$ on the trajectories of the control system (18) (note the similarity of the composition $f \circ \mathcal{T}$ with the definition of the Koopman operator (2) for autonomous systems). In this result, observe that even though $\mathcal{K}^{\text{aug}}$ is the Koopman operator associated with (28), its action on control-independent function extensions leads to the prediction of the function values on trajectories of the actual control system (18).

The next result provides a link between the invariant subspaces of $\mathcal{K}^{\text{aug}}$ and common invariant subspaces of the KCF.

Proposition 5.7 *(Invariant Subspaces of* $\mathcal{K}^{\text{aug}}$ *Characterize Common Invariant Subspaces for the KCF): Let* $\mathcal{S} \subseteq \mathcal{F}^{\text{aug}}$ *be an invariant subspace under* $\mathcal{K}^{\text{aug}}$. *Then,*

(a) for all $u^* \in \mathcal{U}$, $\mathcal{S}\restriction_{u=u^*}$ *is an invariant subspace of* $\mathcal{K}_{u^*}$;
(b) if $\mathcal{S}\restriction_{u=u_1} = \mathcal{S}\restriction_{u=u_2}$ *for all* $u_1, u_2 \in \mathcal{U}$, *then* $\mathcal{S}\restriction_{u=u^*}$ *(for any* $u^* \in \mathcal{U}$*) is a common invariant subspace under the Koopman control family* $\{\mathcal{K}_{u^*}\}_{u^*\in\mathcal{U}}$.

PROOF. (a) First note that $\mathcal{S}\restriction_{u=u^*}$ is a vector space for all $u^* \in \mathcal{U}$. Given any $u^* \in \mathcal{U}$, consider an arbitrary function $g \in \mathcal{S}\restriction_{u=u^*}$. By definition of $\mathcal{S}\restriction_{u=u^*}$, there exists a function $\tilde{g} \in \mathcal{S}$ such that $g = \tilde{g}\restriction_{u=u^*}$ (note that $\tilde{g}$ might not be unique). By Lemma 5.5, one can write

$$\mathcal{K}_{u^*} g = \mathcal{K}_{u^*}(\tilde{g}\restriction_{u=u^*}) = (\mathcal{K}^{\text{aug}}\tilde{g})\restriction_{u=u^*} \in \mathcal{S}\restriction_{u=u^*},$$

where we have used the fact that $\mathcal{K}^{\text{aug}}\tilde{g} \in \mathcal{S}$ because $\mathcal{S}$ is invariant under $\mathcal{K}^{\text{aug}}$. Therefore, $\mathcal{S}\restriction_{u=u^*}$ is an invariant subspace of $\mathcal{K}_{u^*}$.

(b) This is a direct consequence of part (a) and the definition of common invariant subspace for the KCF. □

Proposition 5.7 provides a tool for the identification of common invariant subspaces under KCF based on the invariant subspaces of the augmented Koopman operator. However, checking the condition in Proposition 5.7(b) requires one to compare different vector spaces, which can be cumbersome. In the following section, we provide more direct conditions

that can be checked easily and lead to input-state separable models, as laid out in Theorem 4.3.

6 Input-State Separable Forms via the Augmented Koopman Operator

Here, we aim to build on Proposition 5.7 and Theorem 4.3 to provide more specific practical criteria to identify common invariant subspaces of the KCF and derive input-state separable models. Based on Theorem 4.3 we know that on a common invariant subspace, function composition with $\mathcal{T}$ leads to functions that can be written as a linear combination of separable functions in x and u. For convenience, we provide the following definition.

Definition 6.1 *(Input-State Separable Functions and Their Linear Combinations): A function $f \in \mathcal{F}^{\text{aug}}$ is **input-state separable** if there exist $g : \mathcal{U} \to \mathbb{C}$ and $h : \mathcal{X} \to \mathbb{C}$ such that $f(x,u) = g(u)h(x)$ for all $x \in \mathcal{X}$ and $u \in \mathcal{U}$. A function J is an **input-state separable combination** (or **separable combination** for short) if it can be written as a* finite *linear combination of input-state separable functions.* □

Next, we show a property of the bases for spaces of separable combinations.

Proposition 6.2 *(Spaces of Separable Combinations Have Separable Bases): Let $\mathcal{S} \subseteq \mathcal{F}^{\text{aug}}$ be a finite-dimensional (of dimension $s \in \mathbb{N}$) subspace comprised of input-state separable combinations. Then, for any arbitrary basis $\{\phi_1, \ldots, \phi_s\}$ of $\mathcal{S}$, the vector-valued function $\Phi(x,u) = [\phi_1(x,u), \ldots, \phi_s(x,u)]^T$ can be decomposed as the product of two functions as follows*

$$\Phi(x,u) = G(u)H(x),\ \forall (x,u) \in \mathcal{X} \times \mathcal{U}. \tag{30}$$

where $G : \mathcal{U} \to \mathbb{C}^{s \times l}$ and $H : \mathcal{X} \to \mathbb{C}^l$ for some $l \in \mathbb{N}$. □

The proof is available online at [17]. With this result in place, we can show how to obtain a common invariant subspace of the KCF using the invariant subspaces of the augmented Koopman operator.

Theorem 6.3 *(Rank Condition for Identification of Common Invariant Subspaces of KCF via $\mathcal{K}^{\text{aug}}$): Let $\mathcal{S} \subseteq \mathcal{F}^{\text{aug}}$ be a finite-dimensional (of dimension $s \in \mathbb{N}$) subspace comprised of input-state separable combinations that is invariant under $\mathcal{K}^{\text{aug}}$ and let $\Phi(x,u) = G(u)H(x)$ be a decomposition of a basis for $\mathcal{S}$, where $G : \mathcal{U} \to \mathbb{C}^{s \times l}$ and $H : \mathcal{X} \to \mathbb{C}^l$. If $G(u)$ has full column rank for all $u \in \mathcal{U}$, then the space $\mathcal{H} = \text{span}(H)$ is a common invariant subspace for the KCF.*

PROOF. Since $\mathcal{S}$ is a finite-dimensional invariant subspace under $\mathcal{K}^{\text{aug}}$, given the basis Φ, one can represent the action of $\mathcal{K}^{\text{aug}}$ on $\mathcal{S}$ by a matrix $A \in \mathbb{C}^{s \times s}$ as

$$\mathcal{K}^{\text{aug}}\Phi = A\Phi, \tag{31}$$

where we have used the compact notation in Remark 2.1. Using this and the decomposition $\Phi(x,u) = G(u)H(x)$,

$$\mathcal{K}^{\text{aug}}\big(G(\cdot)H(\cdot)\big) = A\,G(\cdot)H(\cdot).$$

With this compact description, in order to invoke Proposition 5.7(b), we need to show that $\big[\text{span}(G(u)H(\cdot))\big]\restriction_{u=u^*} = \text{span}(G(u^*)H)$ is the same for all $u^* \in \mathcal{U}$. We show this by establishing

$$\text{span}(G(u^*)H) = \text{span}(H),\ \forall u^* \in \mathcal{U}. \tag{32}$$

To show the inclusion from left to right, consider $g : \mathcal{X} \to \mathbb{C}$ with $g \in \text{span}(G(u^*)H)$. Hence, there is a vector $v_g \in \mathbb{C}^s$ such that $g(\cdot) = v_g^T G(u^*)H(\cdot)$. Defining $w_g = G(u^*)^T v_g \in \mathbb{C}^l$, one can write $g(\cdot) = w_g^T H(\cdot) \in \text{span}(H)$, proving

$$\text{span}(G(u^*)H) \subseteq \text{span}(H),\ \forall u^* \in \mathcal{U}. \tag{33}$$

To prove the inclusion from right to left, consider $p : \mathcal{X} \to \mathbb{C}$ with $p \in \text{span}(H)$. Hence, there is a vector $v_p \in \mathbb{C}^l$ such that $p(\cdot) = v_p^T H(\cdot)$. For a given u^*, we need to show that there exists a vector $w_p \in \mathbb{C}^s$ such that $p(\cdot) = w_p^T G(u^*)H(\cdot)$. In other words, we have to show the following linear equation holds for some $w_p \in \mathbb{C}^s$

$$G(u^*)^T w_p = v_p. \tag{34}$$

Given that $G(u^*)$ has full column rank, equation (34) always have at least one solution, which might not be unique. Therefore, $p(\cdot) = w_p^T G(u^*)H(\cdot) \in \text{span}(G(u^*)H)$ and consequently

$$\text{span}(H) \subseteq \text{span}(G(u^*)H),\ \forall u^* \in \mathcal{U}. \tag{35}$$

Combining (33) and (35) yields the subspace equality (32). By Proposition 5.7(b), we conclude that $\text{span}(H) = \mathcal{H}$ is a common invariant subspace for the KCF. □

Theorem 6.3 provides an algebraic rank condition that is far easier to check than the condition in Proposition 5.7.

Remark 6.4 *(A Note on Rank Condition in Theorem 6.3):* In Theorem 6.3, if the matrix $G(u)$ is column-rank deficient only for some $u \in \mathcal{U}$, one might be able to use the result with a slight relaxation. In particular, define

$$\tilde{\mathcal{U}} := \{u \in \mathcal{U} \mid G(u) \text{ has full column rank}\}.$$

If the control system (18) exhibits favorable control properties, e.g., controllability, reachability, or stabilizability, etc., on $\tilde{\mathcal{U}}$, then one can restrict the input space to $\tilde{\mathcal{U}}$ and apply Theorem 6.3. A notable example of this restriction is the case of switched linear modeling, see e.g., [38], that only requires $\tilde{\mathcal{U}}$ to contain finitely many predetermined inputs. □

Example 6.5 *(Revisiting Example 4.8 – Invariant Subspace for $\mathcal{K}^{\text{aug}}$):* For the system (26), one can derive a lifted linear form on an invariant subspace of $\mathcal{K}^{\text{aug}}$ as

$$\begin{bmatrix} x_1 \\ x_2 \\ x_1^2 \\ 1 \\ x_1 u \\ u \\ u^2 \\ \sin(u) \end{bmatrix}^+ = \begin{bmatrix} a & 0 & 0 & 0 & 0 & b & 0 & 0 \\ 0 & c & d & h & e & f & 0 & g \\ 0 & 0 & a^2 & 0 & 2ab & 0 & b^2 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & a & 0 & b & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_1^2 \\ 1 \\ x_1 u \\ u \\ u^2 \\ \sin(u) \end{bmatrix}.$$

Note that the evolution is based on the augmented system (28), which does not evolve u. One can decompose the basis $\Phi(x,u) = [x_1, x_2, x_1^2, 1, x_1u, u, u^2, \sin(u)]^T$ as $\Phi(x,u) = G(u)H(x)$, with $G(u) = [I_{4\times 4}, \tilde{G}(u)^T]^T$ where

$$H(x) = \begin{bmatrix} x_1 \\ x_2 \\ x_1^2 \\ 1 \end{bmatrix} \text{ and } \tilde{G}(u) = \begin{bmatrix} u & 0 & 0 & 0 \\ 0 & 0 & 0 & u \\ 0 & 0 & 0 & u^2 \\ 0 & 0 & 0 & \sin(u) \end{bmatrix}.$$

In this decomposition, the rank condition in Theorem 6.3 holds (note the presence of $I_{4\times 4}$ in $G(u)$). Hence, $\text{span}(H)$ is a common invariant subspace for the KCF, which is in agreement with Example 4.8. □

According to Theorem 4.3, a common invariant subspace for the KCF comes with an associated input-state separable model for the control system (18). The next result specifies how to obtain it under the conditions of Theorem 6.3.

Proposition 6.6 *(Deriving Input-State Separable Models using Invariant Subspaces of $\mathcal{K}^{\text{aug}}$): Let $\mathcal{S} \subseteq \mathcal{F}^{\text{aug}}$ be a finite-dimensional (of dimension $s \in \mathbb{N}$) subspace comprised of input-state separable combinations that is invariant under $\mathcal{K}^{\text{aug}}$ and $\Phi : \mathcal{X} \times \mathcal{U} \to \mathbb{C}^s$ a basis of $\mathcal{S}$. Let $A \in \mathbb{C}^{s\times s}$ be*[9] *such that $\mathcal{K}^{\text{aug}}\Phi = A\Phi$. Let $\Phi(x,u) = G(u)H(x)$ be a decomposition of a basis for $\mathcal{S}$, where $G : \mathcal{U} \to \mathbb{C}^{s\times l}$ and $H : \mathcal{X} \to \mathbb{C}^l$. If $G(u)$ has full column rank for all $u \in \mathcal{U}$, then the matrix-valued map $\mathcal{A} : \mathcal{U} \to \mathbb{C}^{l\times l}$ given by*

$$\mathcal{A}(u) = G(u)^\dagger AG(u) = \left(G(u)^H G(u)\right)^{-1} G(u)^H AG(u),$$

turns the common-invariant subspace $\mathcal{H} = \text{span}(H)$ for the KCF into the input-state separable form of Theorem 4.3, i.e., for all $(x,u) \in \mathcal{X} \times \mathcal{U}$

$$H(x^+) = H \circ \mathcal{T}(x,u) = \mathcal{A}(u)H(x).$$

PROOF. Using the definition of $\mathcal{T}^{\text{aug}}$, cf. equation (28), one can write $\mathcal{K}^{\text{aug}}\Phi(x,u) = \Phi \circ \mathcal{T}^{\text{aug}}(x,u) = \Phi(\mathcal{T}(x,u),u) = A\Phi(x,u)$ for all $(x,u) \in \mathcal{X} \times \mathcal{U}$. Now, by using $\Phi(x,u) = G(u)H(x)$,

$$G(u)H(\mathcal{T}(x,u)) = AG(u)H(x), \quad \forall (x,u) \in \mathcal{X} \times \mathcal{U}.$$

Keeping in mind that $G(u)$ has full column rank, one can multiply both sides from the left by $G(u)^\dagger = \left(G(u)^H G(u)\right)^{-1} G(u)^H$, use $G(u)^\dagger G(u) = I$, and reorder the terms to write

$$H \circ \mathcal{T}(x,u) = H(\mathcal{T}(x,u)) = G(u)^\dagger AG(u)H(x) = \mathcal{A}(u)H(x),$$

for all $(x,u) \in \mathcal{X} \times \mathcal{U}$. □

Theorem 6.3 and Proposition 6.6 provide us with a way to leverage the augmented Koopman operator $\mathcal{K}^{\text{aug}}$ to identify common invariant subspaces for the KCF and derive input-state separable models for the control system.

[9] The existence of this matrix is a direct consequence of the fact that $\mathcal{S}$ is invariant under $\mathcal{K}^{\text{aug}}$.

7 Input-State Separable Forms on Normal Spaces

In this section, we turn our attention to a special case of subspaces that are of practical significance. This focus is motivated by examining the rank condition on $G(u)$ presented in Theorem 6.3, and observing that the matrix-valued function G constitutes an element of the basis description for subspace $\mathcal{S}$. It becomes then clear that this condition specifies a structural characteristic of the subspace $\mathcal{S}$ and its basis, which is independent of the operator $\mathcal{K}^{\text{aug}}$. Therefore, here we study a specific class of subspaces and their bases that always satisfy the rank condition in Theorem 6.3.

Definition 7.1 *(Vector-Valued Function of Separable Combinations in Normal Form and Normal Spaces): Let $\Phi : \mathcal{X} \times \mathcal{U} \to \mathbb{C}^s$ be a vector-valued function of separable combinations. Moreover, let the set of elements of Φ be linearly independent. Φ is in* ***normal form*** *if it has a decomposition as one of the following:*

$$\Phi(x,u) = \begin{bmatrix} I_{\mathcal{U}}^{l\times l}(u) \\ \tilde{G}(u) \end{bmatrix} H(x), \qquad s > l, \tag{36a}$$

$$\Phi(x,u) = I_{\mathcal{U}}^{l\times l}(u)H(x), \qquad s = l, \tag{36b}$$

where $H : \mathcal{X} \to \mathbb{C}^l$ and $\tilde{G} : \mathcal{U} \to \mathbb{C}^{(s-l)\times l}$ for some $l \leq s$ and the elements of H are linearly independent. Moreover, $I_{\mathcal{U}}^{l\times l} : \mathcal{U} \to \mathbb{C}^{l\times l}$ is the constant identity function, $I_{\mathcal{U}}^{l\times l}(u) \equiv I$. A finite-dimensional space of separable combinations is ***normal*** *if it has a basis that can be written as a vector-valued function of normal form.* □

From Definition 7.1, it is clear that a basis in normal form satisfies the rank condition in Theorem 6.3. The next result shows a useful property of normal spaces.

Proposition 7.2 *(Normal Spaces Capture Control-Independent Functions): Let $\mathcal{S} \subset \mathcal{F}^{\text{aug}}$ be a finite-dimensional space of input-state separable combinations and let $\Phi(x,u) = G(u)H(x)$ be a decomposition of a basis for $\mathcal{S}$, where $G : \mathcal{U} \to \mathbb{C}^{s\times l}$ and $H : \mathcal{X} \to \mathbb{C}^l$ for some $l \leq s$ (here, $s \in \mathbb{N}$ is the dimension of $\mathcal{S}$). Then, $\mathcal{S}$ is normal if and only if $h_e \in \mathcal{S}$ (cf. Definition 5.2) for all $h \in \text{span}(H)$.*

PROOF. ($\Rightarrow$) : Since $\mathcal{S}$ is normal, it has a basis with one of the following forms:

$$\hat{\Phi}(x,u) = \begin{bmatrix} I_{\mathcal{U}}^{l\times l}(u) \\ \tilde{G}(u) \end{bmatrix} \tilde{H}(x), \qquad \text{if } s > l,$$

$$\hat{\Phi}(x,u) = I_{\mathcal{U}}^{l\times l}(u)\tilde{H}(x), \qquad \text{if } s = l, \tag{37}$$

where $\text{span}(\tilde{H}) = \text{span}(H)$ (this is a direct consequence of the fact that Φ and $\hat{\Phi}$ are bases for the same subspace). Then, for every $h \in \text{span}(H)$, there exists a vector $w_h \in \mathbb{C}^l$, such that $h(\cdot) = w_h^T \tilde{H}(\cdot)$. Based on (37), one can write

$$h_e(x,u) = [w_h^T, 0_{1\times(s-l)}]\hat{\Phi}(x,u) \in \mathcal{S}, \qquad \text{if } s > l,$$

$$h_e(x,u) = w_h^T \hat{\Phi}(x,u) \in \mathcal{S}, \qquad \text{if } s = l.$$

This concludes the proof of this part.

($\Leftarrow$) : Let $H(\cdot) = [h_1(\cdot), \ldots, h_l(\cdot)]^T$. By hypothesis, we have $h_i(x)1_{\mathcal{U}}(u) \in \mathcal{S}$ for all $i \in \{1, \ldots, l\}$. As a result, there exist

vectors $\{w_1, \ldots, w_l\} \subset \mathbb{C}^s$ such that for all $(x,u) \in \mathcal{X} \times \mathcal{U}$ we have

$$h_i(x)1_{\mathcal{U}}(u) = w_i^T G(u)H(x), \ \forall i \in \{1, \ldots, l\}. \tag{38}$$

Let $W = [w_1, \ldots, w_l]^T \in \mathbb{C}^{l \times s}$ and consider two cases:

Case (i): Suppose $s = l$. Define the vector-valued function $\tilde{\Phi}(\cdot,\cdot) = W\Phi(\cdot,\cdot)$. This function can be written as

$$\tilde{\Phi}(x,u) = WG(u)H(x) = I_{\mathcal{U}}^{l\times l}(u)H(x), \quad \forall (x,u) \in \mathcal{X} \times \mathcal{U},$$

where we have used (38). Therefore $\tilde{\Phi}$ is a normal-form basis and hence $\mathcal{S}$ is normal.

Case (ii) Suppose $s > l$ and decompose $G(u) = [G_1^T(u), G_2^T(u)]^T$, where $G_1 : \mathcal{U} \to \mathbb{C}^{l\times l}$ and $G_2 : \mathcal{U} \to \mathbb{C}^{(s-l)\times l}$. Define the vector-valued function $\hat{\Phi} : \mathcal{X} \times \mathcal{U} \to \mathbb{C}^{s\times l}$,

$$\begin{aligned}\hat{\Phi}(x,u) = \begin{bmatrix} W \\ B \end{bmatrix} \Phi(x,u) &= \begin{bmatrix} W \\ B \end{bmatrix} \begin{bmatrix} G_1(u) \\ G_2(u) \end{bmatrix} H(x) \\ &= \begin{bmatrix} I_{\mathcal{U}}^{l\times l}(u) \\ G_2(u) \end{bmatrix} H(x),\end{aligned}$$

where $B = [0_{(s-l)\times l}, I_{(s-l)\times(s-l)}] \in \mathbb{C}^{(s-l)\times s}$, and in the third equality we have used (38). Therefore, $\hat{\Phi}$ is a normal-form basis and hence $\mathcal{S}$ is normal. □

Proposition 7.2 reveals a useful property of normal spaces that allows us to directly predict the evolution of functions in $\mathcal{F}$ under the system trajectories by applying $\mathcal{K}^{\text{aug}}$ on control-independent extensions through Lemma 5.6, as we explain next.

Theorem 7.3 *(Identification of Common Invariant Subspaces of the KCF and Input-State Separable Forms on Normal Spaces): Let $\mathcal{S} \subset \mathcal{F}^{\text{aug}}$ be a finite-dimensional normal space of input-state separable combinations that is invariant under $\mathcal{K}^{\text{aug}}$. Let $\Phi(x,u) = G(u)H(x)$ be a decomposition of a basis for $\mathcal{S}$ where $G : \mathcal{U} \to \mathbb{C}^{s\times l}$ and $H : \mathcal{X} \to \mathbb{C}^l$ for some $l \leq s$ (here, $s \in \mathbb{N}$ is the dimension of $\mathcal{S}$). Then,*

(a) $\text{span}(H) \subset \mathcal{F}$ *is a common invariant subspace under the Koopman control family $\{\mathcal{K}_{u^*}\}$;*

(b) for all $h \in \text{span}(H)$ and for all $(x,u) \in \mathcal{X} \times \mathcal{U}$, it holds that $h(x^+) = h \circ \mathcal{T}(x,u) = \mathcal{K}^{\text{aug}} h_e(x,u)$;

(c) without loss of generality, assume Φ is in normal form, i.e., $G(u) = I_{\mathcal{U}}^{l\times l}(u)$ if $l = s$ or $G(u) = [I_{\mathcal{U}}^{l\times l}(u)^T, \tilde{G}(u)^T]^T$ if $s > l$. Moreover, let $A \in \mathbb{C}^{s\times s}$ be a matrix such that $\mathcal{K}^{\text{aug}}\Phi = A\Phi$ (note that A exists because $\mathcal{S}$ is invariant under $\mathcal{K}^{\text{aug}}$). If $s > l$, consider the block-decomposition of A,

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

where $A_{11} \in \mathbb{C}^{l\times l}$, $A_{12} \in \mathbb{C}^{l\times(s-l)}$, $A_{21} \in \mathbb{C}^{(s-l)\times l}$, and $A_{22} \in \mathbb{C}^{(s-1)\times(s-l)}$. Then, the associated input-state separable dynamics can be written as

$$H(x^+) = H \circ \mathcal{T}(x,u) = \mathcal{A}(u)H(x), \tag{39}$$

where, for each $u \in \mathcal{U}$,

$$\begin{aligned}\mathcal{A}(u) &= A_{11} + A_{12}\tilde{G}(u), && \text{if } s > l. \\ \mathcal{A}(u) &= A\, I_{\mathcal{U}}^{l\times l}(u) = A, && \text{if } s = l.\end{aligned}$$

PROOF. (a) Since $\mathcal{S}$ is normal, one can do a linear transformation of the basis $\Phi(x,u) = G(u)H(x)$ to put it in normal form. Hence, there is a nonsingular square matrix E, such that $E\Phi(x,u) = EG(u)H(x)$ is in normal form. Therefore, by Definition 7.1, $EG(u)$ has full column rank for all $u \in \mathcal{U}$. Since E is nonsingular, we deduce that $G(u)$ has full column rank for all $u \in \mathcal{U}$. As a result, we can invoke Theorem 6.3 to deduce that $\text{span}(H) \subset \mathcal{F}$ is a common invariant subspace under the KCF.

(b) This part is the direct consequence of Lemma 5.6.

(c) Using the definition of $\mathcal{T}^{\text{aug}}$, cf. equation (28), one can write $\Phi \circ \mathcal{T}^{\text{aug}}(x,u) = \Phi(\mathcal{T}(x,u),u) = A\Phi(x,u)$ for all $(x,u) \in \mathcal{X} \times \mathcal{U}$. Now, using $\Phi(x,u) = G(u)H(x)$,

$$G(u)H(x^+) = G(u)H(\mathcal{T}(x,u)) = AG(u)H(x). \tag{40}$$

The case $s = l$ is trivial since $G(u)$ is an identity map. For the case $s > l$, the proof directly follows by multiplying both sides of (40) from the left by the matrix $[I_{l\times l}, 0_{l\times(s-l)}]$ and using the decompositions of $G(u)$ and A. □

Example 7.4 *(Examples 4.8 and 6.5 Revisited):* The basis decomposition in Example 6.5 is in normal form. One can readily use the formula in Theorem 7.3(c) with this decomposition to calculate the input-state separable form (27). □

Theorem 7.3 has significant practical implications: not only it connects the invariant subspaces of $\mathcal{K}^{\text{aug}}$ to common invariant subspaces of the KCF, but more importantly, unlike Proposition 6.6, it provides a direct way of predicting the evolution of observables in $\mathcal{F}$ under the control system based on the application of $\mathcal{K}^{\text{aug}}$ on control-independent extensions. This direct computation does not require taking a pseudo-inverse (cf. Proposition 6.6) and is helpful to find accuracy bounds when we have to approximate invariant subspaces of $\mathcal{K}^{\text{aug}}$, as we explain next.

8 Non-Invariant Subspaces, Invariance Proximity, and Approximation Error

In the sections above we have provided results connecting the finite-dimensional invariant subspaces of $\mathcal{K}^{\text{aug}}$ to common invariant subspaces of the Koopman control family $\{\mathcal{K}_{u^*}\}_{u^*\in\mathcal{U}}$, and how these can be used in predicting the evolution of functions on the common invariant subspace under the trajectories of the control system. In practice, however, finding exact invariant subspaces that capture proper information is an arduous task and one might need to settle for approximations on non-invariant subspaces. In such case, three fundamental questions immediately arise:

(Q1) How can we measure the closeness of a subspace to being invariant?

(Q2) How does this measure characterize the approximation error of the action of the operator on a non-invariant subspace?

(Q3) How do the previous results regarding the prediction of observables on the trajectories of the control system extend to the case of non-invariant subspaces?

These are the questions we tackle in this section. To determine whether a finite-dimensional subspace $\mathcal{S} \subset \mathcal{F}^{\text{aug}}$ is invariant under $\mathcal{K}^{\text{aug}}$ we only need the concept of set inclusion. However, in order to quantify how close to invariant a subspace is, we need to be able to measure angles, lengths, and distances. Therefore, we equip the space $\mathcal{F}^{\text{aug}}$ with an inner product, that induces a norm and, in turn, a metric [10].

Definition 8.1 *(Inner Product, Norm, and Metric on $\mathcal{F}^{\text{aug}}$): An arbitrary inner product*[11] $\langle \cdot, \cdot \rangle : \mathcal{F}^{\text{aug}} \times \mathcal{F}^{\text{aug}} \to \mathbb{C}$ *on $\mathcal{F}^{\text{aug}}$ induces a norm $\| \cdot \| : \mathcal{F}^{\text{aug}} \to [0, \infty)$ and a metric* $\text{dist} : \mathcal{F}^{\text{aug}} \times \mathcal{F}^{\text{aug}} \to [0, \infty)$ *as*

$$\|f\| = \sqrt{\langle f, f \rangle}, \qquad \text{dist}(f, g) = \|f - g\|. \qquad \square$$

Since we work with a finite-dimensional subspace that is not necessarily invariant under the operator, we have to approximate the action of the operator on the subspace. This approximation is generally done by performing an orthogonal projection on the subspace, as explained next.

Definition 8.2 *(Linear Predictors on Finite-Dimensional Subspaces): Consider the finite-dimensional subspace $\mathcal{S} \subset \mathcal{F}^{\text{aug}}$ and let $\mathcal{P}_{\mathcal{S}} : \mathcal{F}^{\text{aug}} \to \mathcal{F}^{\text{aug}}$ be the orthogonal projection operator*[12] *on $\mathcal{S}$. We define the predictor for the function $\psi \in \mathcal{F}^{\text{aug}}$ on $\mathcal{S}$ as*

$$\psi \approx \mathfrak{P}_{\psi}^{\mathcal{S}} := \mathcal{P}_{\mathcal{S}} \psi.$$

For a vector-valued function $\Psi = [\psi_1, \ldots, \psi_n]^T$, where $\psi_i \in \mathcal{F}^{\text{aug}}$ for $i \in \{1, \ldots, n\}$, we define the linear predictor $\Psi \approx \mathfrak{P}_{\Psi}^{\mathcal{S}} := [\mathfrak{P}_{\psi_1}^{\mathcal{S}}, \ldots, \mathfrak{P}_{\psi_n}^{\mathcal{S}}]$. We remove the superscript $\mathcal{S}$ when the choice of subspace is clear from the context. $\square$

The properties of the operator $\mathcal{P}_{\mathcal{S}}$ lead to useful properties of the linear predictors defined in Definition 8.2.

Lemma 8.3 *(Properties of Linear Predictors): Linear predictors on the finite-dimensional subspace $\mathcal{S} \subset \mathcal{F}^{\text{aug}}$ satisfy:*

(a) $\mathfrak{P}_f \in \mathcal{S}$ is the best approximation for $f \in \mathcal{F}^{\text{aug}}$ on $\mathcal{S}$, i.e., $\|f - \mathfrak{P}_f\| \leq \|f - g\|$ for all $g \in \mathcal{S}$;
(b) $\mathfrak{P}_{c_1 f_1 + c_2 f_2} = c_1 \mathfrak{P}_{f_1} + c_2 \mathfrak{P}_{f_2}$ for all $f_1, f_2 \in \mathcal{F}^{\text{aug}}$ and $c_1, c_2 \in \mathbb{C}$;
(c) let Ψ be a vector-valued function with $\text{span}(\Psi) \subset \mathcal{F}^{\text{aug}}$ and let $f = v_f^T \Psi$, where v_f is a complex vector of appropriate size. Then, $\mathfrak{P}_f = v_f^T \mathfrak{P}_{\Psi}$. $\square$

The proof of Lemma 8.3 is a direct consequence of the properties of orthogonal projections and is omitted for space reasons. Lemma 8.3(a) states that the predictor defined in Definition 8.2 is the best predictor on the subspace: in this sense, we use the notation $f \approx \mathfrak{P}_f$ when we aim to emphasize that we approximate f with $\mathfrak{P}_f$.

[10] Even though we aim to approximate a common invariant subspace $\mathcal{H} \subset \mathcal{F}$ under the Koopman control family, our end goal is to predict the evolution of observables under the system trajectories, i.e., we aim to predict $h(x^+) = h \circ \mathcal{T}(x, u)$ for all $h \in \mathcal{H}$ and $(x, u) \in \mathcal{X} \times \mathcal{U}$. Since $h \circ \mathcal{T} \in \mathcal{F}^{\text{aug}}$, we need to reason with $\mathcal{F}^{\text{aug}}$.

[11] Since we are working with finite-dimensional subspaces, we do not require the inner product space $\mathcal{F}^{\text{aug}}$ to be complete (Hilbert) or separable.

[12] Given an orthonormal basis $\{e_1, \ldots, e_n\}$ for $\mathcal{S}$, one can calculate the orthogonal projection of $g \in \mathcal{F}^{\text{aug}}$ on $\mathcal{S}$ by $\mathcal{P}_{\mathcal{S}}(g) = \sum_{i=1}^{n} \langle g, e_i \rangle e_i$.

We next use the linear predictors to approximate the action of the operator $\mathcal{K}^{\text{aug}}$ on a non-invariant finite-dimensional subspace and provide a matrix notation for it.

Lemma 8.4 *(Approximating an Operator's Action using Linear Predictors): Any finite-dimensional subspace $\mathcal{S} \subset \mathcal{F}^{\text{aug}}$ is invariant under $\mathcal{P}_{\mathcal{S}} \mathcal{K}^{\text{aug}}$. Let $\Phi : \mathcal{X} \times \mathcal{U} \to \mathbb{C}^s$ be a basis for $\mathcal{S}$ and let $\tilde{A} \in \mathbb{C}^{s \times s}$ be a matrix such that $\mathcal{P}_{\mathcal{S}} \mathcal{K}^{\text{aug}} \Phi = \tilde{A} \Phi$. Then,*

(a) $\mathfrak{P}_{\mathcal{K}^{\text{aug}} \Phi} = \tilde{A} \Phi$;
(b) for $f \in \mathcal{S}$ with description $f = v_f^T \Phi$, where $v_f \in \mathbb{C}^s$, we have $\mathfrak{P}_{\mathcal{K}^{\text{aug}} f} = v_f^T \tilde{A} \Phi$. $\square$

Note the parallelism of Lemma 8.4 with (9)-(10). Its proof is a direct consequence of the linearity of $\mathcal{K}^{\text{aug}}$ and Lemma 8.3, and is omitted for space reasons. The prediction error associated with the predictors in Lemma 8.4 directly depends on how close to invariant the space is under the operator $\mathcal{K}^{\text{aug}}$. To capture this, we define the concept of invariance proximity under an operator.

Definition 8.5 *(Invariance Proximity): The **invariance proximity** of a finite-dimensional subspace $\mathcal{S} \subset \mathcal{F}^{\text{aug}}$ under the operator $\mathcal{K}^{\text{aug}}$, denoted $I_{\mathcal{K}^{\text{aug}}}(\mathcal{S})$, is*

$$I_{\mathcal{K}^{\text{aug}}}(\mathcal{S}) = \sup_{f \in \mathcal{S}, \|\mathcal{K}^{\text{aug}} f\| \neq 0} \frac{\|\mathcal{K}^{\text{aug}} f - \mathfrak{P}_{\mathcal{K}^{\text{aug}} f}\|}{\|\mathcal{K}^{\text{aug}} f\|}. \qquad \square$$

Invariance proximity measures the worst-case relative error of approximation by projecting the action of $\mathcal{K}^{\text{aug}}$ on $\mathcal{S}$ and provides an answer to Q2 above. Invariance proximity does not depend on the specific basis for the subspace, and is instead a property of the linear space $\mathcal{S}$ and the operator $\mathcal{K}^{\text{aug}}$.

Proposition 8.6 *(Properties of Invariance Proximity): Given a finite-dimensional subspace $\mathcal{S} \subset \mathcal{F}^{\text{aug}}$,*

(a) $I_{\mathcal{K}^{\text{aug}}}(\mathcal{S}) \in [0, 1]$;
(b) $I_{\mathcal{K}^{\text{aug}}}(\mathcal{S}) = 0$ if[13] *$\mathcal{S}$ is invariant under $\mathcal{K}^{\text{aug}}$.*

PROOF. (a) Let $f \in \mathcal{F}^{\text{aug}}$ with $\|\mathcal{K}^{\text{aug}} f\| \neq 0$. Noting that $\mathfrak{P}_{\mathcal{K}^{\text{aug}} f} = \mathcal{P}_{\mathcal{S}} \mathcal{K}^{\text{aug}} f$ is an orthogonal projection on $\mathcal{S}$, we can decompose $\mathcal{K}^{\text{aug}} f$ as $\mathcal{K}^{\text{aug}} f = \mathfrak{P}_{\mathcal{K}^{\text{aug}} f} + e$, where $\langle \mathfrak{P}_{\mathcal{K}^{\text{aug}} f}, e \rangle = 0$. Using the definition of the norm induced by the inner product then yields $\|\mathcal{K}^{\text{aug}} f\|^2 = \|\mathfrak{P}_{\mathcal{K}^{\text{aug}} f}\|^2 + \|e\|^2$. Therefore, $\|e\| \leq \|\mathcal{K}^{\text{aug}} f\|$ and we can write

$$\frac{\|\mathcal{K}^{\text{aug}} f - \mathfrak{P}_{\mathcal{K}^{\text{aug}} f}\|}{\|\mathcal{K}^{\text{aug}} f\|} = \frac{\|e\|}{\|\mathcal{K}^{\text{aug}} f\|} \leq 1.$$

Since this inequality holds for all functions $f \in \mathcal{F}^{\text{aug}}$ where $\mathcal{K}^{\text{aug}} f \neq 0$, we deduce $I_{\mathcal{K}^{\text{aug}}}(\mathcal{S}) \leq 1$. Moreover, by definition of $I_{\mathcal{K}^{\text{aug}}}(\mathcal{S})$ and the fact that norms are nonnegative, we conclude $I_{\mathcal{K}^{\text{aug}}}(\mathcal{S}) \geq 0$, completing the proof.

(b) If $\mathcal{S}$ is invariant under $\mathcal{K}^{\text{aug}}$, we have $\mathcal{K}^{\text{aug}} f \in \mathcal{S}$ and therefore $\|\mathcal{K}^{\text{aug}} f - \mathfrak{P}_{\mathcal{K}^{\text{aug}} f}\| = 0$ for all $f \in \mathcal{S}$. Hence, $I_{\mathcal{K}^{\text{aug}}}(\mathcal{S}) = 0$. $\square$

[13] The converse also holds if $\|f - g\| = 0$ implies $f = g$ everywhere. This might not hold for typical norms on function spaces that operate on equivalence classes and allow for violations of equality on measure-zero sets.

Proposition 8.6 means that invariance proximity provides an answer to Q1 above. The next result extends to non-invariant subspaces the results on prediction of the evolution of functions in $\mathcal{F}$ under the control system (18), providing an answer to Q3.

Theorem 8.7 *(Approximate Input-State Separable Form and Accuracy Bound): Let $\mathcal{S} \subset \mathcal{F}^{\text{aug}}$ be a finite-dimensional normal subspace comprised of input-state separable combinations. Let $\Phi(x,u) = G(u)H(x)$ be a decomposition of a basis for $\mathcal{S}$ where $G : \mathcal{U} \to \mathbb{C}^{s\times l}$ and $H : \mathcal{X} \to \mathbb{C}^l$ for some $l \leq s$ (here, $s \in \mathbb{N}$ is the dimension of $\mathcal{S}$). Let H_e and h_e be the control-independent extensions of H and $h \in \text{span}(H)$ respectively. Then,*

(a) $\mathfrak{P}_{h\circ\mathcal{T}} = \mathfrak{P}_{\mathcal{K}^{\text{aug}}h_e}$ for all $h \in \text{span}(H)$, and $\mathfrak{P}_{H\circ\mathcal{T}} = \mathfrak{P}_{\mathcal{K}^{\text{aug}}H_e}$;

(b) without loss of generality, assume Φ is in normal form, i.e., $G(u) = I^{l\times l}_{\mathcal{U}}(u)$ if $l = s$ or $G(u) = [I^{l\times l}_{\mathcal{U}}(u)^T, \tilde{G}(u)^T]^T$ if $s > l$. Moreover, let $\tilde{A} \in \mathbb{C}^{s\times s}$ be a matrix such that $\mathcal{P}_{\mathcal{S}}\mathcal{K}^{\text{aug}}\Phi = \tilde{A}\Phi$ (note that $\tilde{A}$ exists because $\mathcal{S}$ is invariant under $\mathcal{P}_{\mathcal{S}}\mathcal{K}^{\text{aug}}$). If $s > l$, consider the block-decomposition of $\tilde{A}$,

$$\tilde{A} = \begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ \tilde{A}_{21} & \tilde{A}_{22} \end{bmatrix},$$

where $\tilde{A}_{11} \in \mathbb{C}^{l\times l}$, $\tilde{A}_{12} \in \mathbb{C}^{l\times(s-l)}$, $\tilde{A}_{21} \in \mathbb{C}^{(s-l)\times l}$, and $\tilde{A}_{22} \in \mathbb{C}^{(s-1)\times(s-l)}$. Then, the associated approximate input-state separable dynamics can be written as

$$H(x^+) = H \circ \mathcal{T}(x,u) \approx \mathfrak{P}_{H\circ\mathcal{T}}(x,u) = \mathcal{A}(u)H(x), \tag{41}$$

where, for each $u \in \mathcal{U}$,

$$\mathcal{A}(u) = \tilde{A}_{11} + \tilde{A}_{12}\tilde{G}(u), \qquad \text{if } s > l.$$
$$\mathcal{A}(u) = \tilde{A}\, I^{l\times l}_{\mathcal{U}}(u) = \tilde{A}, \qquad \text{if } s = l.$$

(c) for all $h \in \text{span}(H)$ with description $h = v_h^T H$, $v_h \in \mathbb{C}^l$,

$$h(x^+) = h \circ \mathcal{T}(x,u) \approx \mathfrak{P}_{h\circ\mathcal{T}}(x,u) = v_h^T \mathcal{A}(u)H(x);$$

(d) for all $h \in \text{span}(H)$ with $\|h \circ \mathcal{T}\| \neq 0$, the predictor's relative error is bounded by the invariance proximity of $\mathcal{S}$ under $\mathcal{K}^{\text{aug}}$,

$$\frac{\|h \circ \mathcal{T} - \mathfrak{P}_{h\circ\mathcal{T}}\|}{\|h \circ \mathcal{T}\|} \leq I_{\mathcal{K}^{\text{aug}}}(\mathcal{S}).$$

PROOF. (a) By Definition 8.2, $\mathfrak{P}_{h\circ\mathcal{T}} = \mathcal{P}_{\mathcal{S}}(h \circ \mathcal{T})$. Using Lemma 5.6, we have $h \circ \mathcal{T} = \mathcal{K}^{\text{aug}}h_e$. Hence, $\mathfrak{P}_{h\circ\mathcal{T}} = \mathcal{P}_{\mathcal{S}}\mathcal{K}^{\text{aug}}h_e = \mathfrak{P}_{\mathcal{K}^{\text{aug}}h_e}$. The statement regarding H follows directly by applying this to each element of the equality $\mathfrak{P}_{H\circ\mathcal{T}} = \mathfrak{P}_{\mathcal{K}^{\text{aug}}H_e}$.

(b) We need to prove the rightmost equality in (41), since the rest follow directly from their definitions. From part (a), and using the vector-valued notation in Remark 2.1, we have

$$\mathfrak{P}_{H\circ\mathcal{T}} = \mathfrak{P}_{\mathcal{K}^{\text{aug}}H_e} = \mathcal{P}_{\mathcal{S}}\mathcal{K}^{\text{aug}}H_e. \tag{42}$$

For the case $s = l$, we use Lemma 5.3(b) to write $\Phi(x,u) = I^{l\times l}_{\mathcal{U}}(u)H(x) = H_e(x,u)$. Hence, noting that $\mathcal{P}_{\mathcal{S}}\mathcal{K}^{\text{aug}}\Phi = \tilde{A}\Phi$, we have $\mathcal{P}_{\mathcal{S}}\mathcal{K}^{\text{aug}}H_e(x,u) = \tilde{A}H_e(x,u) = \tilde{A}I^{l\times l}_{\mathcal{U}}(u)H(x)$. Using (42), we can write $\mathfrak{P}_{H\circ\mathcal{T}}(x,u) = \tilde{A}H_e(x,u) = \tilde{A}I^{l\times l}_{\mathcal{U}}(u)H(x)$, which completes the proof.

Next, we turn our attention to the case $s > l$. Using Lemma 5.3(b), one can write

$$\Phi(x,u) = [H_e(x,u)^T, (\tilde{G}(u)H(x))^T]^T. \tag{43}$$

Multiplying both sides of $\mathcal{P}_{\mathcal{S}}\mathcal{K}^{\text{aug}}\Phi = \tilde{A}\Phi$ from the left by $W = [I_{l\times l}, 0_{l\times(s-l)}]$, and using (43), the decomposition of $\tilde{A}$, the properties of the vector-valued notation in Remark 2.1, and the linearity of the operator $\mathcal{P}_{\mathcal{S}}\mathcal{K}^{\text{aug}}$, one can write

$$\begin{aligned}\mathcal{P}_{\mathcal{S}}\mathcal{K}^{\text{aug}}H_e &= W\mathcal{P}_{\mathcal{S}}\mathcal{K}^{\text{aug}}\Phi = W\tilde{A}\Phi \\ &= (\tilde{A}_{11}I^{l\times l}_{\mathcal{U}}(u) + \tilde{A}_{12}\tilde{G}(u))H(x).\end{aligned}$$

The statement then follows from equation (42) and the fact that $I^{l\times l}_{\mathcal{U}}(u) = I$ for all $u \in \mathcal{U}$.

(c) We need to prove the rightmost equality $\mathfrak{P}_{h\circ\mathcal{T}}(x,u) = v_h^T\mathcal{A}(u)H(x)$, since the rest follow directly from their definitions. By hypothesis $h \circ \mathcal{T} = v_h^T H \circ \mathcal{T}$; hence, from Lemma 8.3(c), we have $\mathfrak{P}_{h\circ\mathcal{T}} = v_h^T\mathfrak{P}_{H\circ\mathcal{T}}$. The result then follows from (41).

(d) By Proposition 7.2, and using the definition of invariance proximity, for all $h \in \text{span}(H)$ with $\|h \circ \mathcal{T}\| \neq 0$, one can write

$$\frac{\|\mathcal{K}^{\text{aug}}h_e - \mathfrak{P}_{\mathcal{K}^{\text{aug}}h_e}\|}{\|\mathcal{K}^{\text{aug}}h_e\|} \leq I_{\mathcal{K}^{\text{aug}}}(\mathcal{S}).$$

The statement then follows from the fact that $\mathcal{K}^{\text{aug}}h_e = h\circ\mathcal{T}$ (cf. Lemma 5.6) and part (a). □

This result can be viewed as an analog of Theorem 7.3 for non-invariant subspaces. Theorem 8.7 is a central result, as it allows to approximate models in the input-state separable form (cf. Theorem 4.3) by approximating a single normal invariant subspace of $\mathcal{K}^{\text{aug}}$, which is significantly easier than working with the KCF directly. Moreover, the concept of invariance proximity provides a bound for approximation errors on the entire subspace. This has important implications for the validity and approximation accuracy of common Koopman-inspired descriptions of the control system (18), cf. Lemmas 4.5 and 4.6.

9 Implications for Robust Data-driven Learning

In this section we illustrate how the results of the paper can be used in data-driven modeling of control systems. We provide an algorithmic description that specifies how to process the data, the choice of inner product space, and the formulation for the dictionary learning.

9.1 Gathering Data for the Augmented Koopman Operator

Our strategy for learning relies on using Theorem 8.7 to approximate an input-state separable form and bound the prediction error for all functions in the identified subspace. This result employs the augmented Koopman operator associated

with the augmented system (28) and, instead, we can only collect trajectory data from the original control system (18). This mismatch can be easily reconciled as we explain next.

Let $\{x_i\}_{i=1}^N \subset \mathcal{X}$, $\{u_i\}_{i=1}^N \subset \mathcal{U}$, and $\{x_i^+\}_{i=1}^N \subset \mathcal{X}$ be state and input data from trajectories of system (18) such that

$$x_i^+ = \mathcal{T}(x_i, u_i), \quad \forall i \in \{1, \dots, N\}. \tag{44}$$

A close look at the definition of $\mathcal{T}^{\text{aug}}$ in (28) reveals that it does not alter the input signal, i.e., if we apply it on the state-input pair x_i, u_i for all $i \in \{1, \dots, N\}$, we get $\mathcal{T}^{\text{aug}}(x_i, u_i) = (\mathcal{T}(x_i, u_i), u_i) = (x_i^+, u_i)$. Therefore, we already have access to all the information $\mathcal{T}^{\text{aug}}$ generates: the first element returned by $\mathcal{T}^{\text{aug}}$ is exactly the action of the control system $\mathcal{T}$ that we have measured in (44) and the second element is exactly the input (without any change) to $\mathcal{T}$, again measured in (44). For convenience, we gather these data snapshots for $\mathcal{T}^{\text{aug}}$ in snapshot matrices as follows

$$\begin{aligned} &X = [x_1, \dots, x_N] \in \mathbb{R}^{n \times N}, \quad X^+ = [x_1^+, \dots, x_N^+] \in \mathbb{R}^{n \times N}, \\ &U = [u_1, \dots, u_N] \in \mathbb{R}^{m \times N}, \quad U^+ = U. \end{aligned} \tag{45}$$

Note that even though matrix U^+ does not capture additional information we have created it, since it is a part of the corresponding state for $\mathcal{T}^{\text{aug}}$. To apply existing numerical methods such as EDMD on $\mathcal{K}^{\text{aug}}$, we gather the augmented state snapshots of $\mathcal{T}^{\text{aug}}$ as

$$\begin{aligned} &Z = [X^T, U^T]^T \in \mathbb{R}^{(n+m) \times N}, \\ &Z^+ = [(X^+)^T, (U^+)^T]^T \in \mathbb{R}^{(n+m) \times N}. \end{aligned} \tag{46}$$

9.2 Choice of Inner Product Space

The results in the previous sections can be used for subspace learning on any arbitrary inner product space. Here we focus on the most popular inner product space in the literature that is used for the EDMD method [26, 52]. Consider the empirical measure μ_Z defined by

$$\mu_Z = \frac{1}{N} \sum_{i=1}^{N} \delta_{z_i}, \tag{47}$$

where δ_{z_i} is the Dirac measure at point z_i, the ith column of matrix Z defined in (46). We then choose the space $L_2(\mu_Z)$ comprised of functions on the domain $\mathcal{X} \times \mathcal{U}$. Under this choice, given any basis $\Phi : \mathcal{X} \times \mathcal{U} \to \mathbb{R}^s$ with *real-valued* elements (cf. Remark 2.3) for the finite-dimensional (with dimension s) normal subspace $\mathcal{S}$, the matrix $\tilde{A}$ in the hypotheses of Theorem 8.7 is the EDMD solution applied on dictionary Φ and data in (46) (cf. Section 2.2), i.e.,

$$\tilde{A} = \Phi(Z^+)\Phi(Z)^\dagger. \tag{48}$$

Moreover, under the condition that $\Phi(Z)$ and $\Phi(Z^+)$ have full row rank, the invariance proximity turns into the square root of the consistency index (cf. Section 2.3) and has the following closed-form expression

$$\begin{aligned} I_{\mathcal{K}^{\text{aug}}}(\mathcal{S}) &= \sqrt{\mathcal{I}_C(\Phi, Z, Z^+)} \\ &= \sqrt{\lambda_{\max}\big(I - \Phi(Z^+)\Phi(Z)^\dagger \Phi(Z)\Phi(Z^+)^\dagger\big)}. \end{aligned} \tag{49}$$

We use (49) to formulate an optimization-based learning problem for modeling the control system.

9.3 Optimization-Based Subspace Learning

Based on Theorem 8.7(d), the invariance proximity determines the accuracy of the model provided on a given normal subspace. Hence, we formulate an optimization problem to find an accurate model by minimizing the invariance proximity over a parametric family of normal spaces with basis Φ in normal form (36) as

$$\underset{\Phi \in \text{PF}}{\text{minimize}}\ I_{\mathcal{K}^{\text{aug}}}(\mathcal{S}) \quad \Leftrightarrow \quad \underset{\Phi \in \text{PF}}{\text{minimize}}\ \sqrt{\mathcal{I}_C(\Phi, Z, Z^+)}\,, \tag{50}$$

where PF is the parametric family of choice (e.g., neural networks, polynomials), $\mathcal{S} = \text{span}(\Phi)$, and one can use the closed-form solution of the invariance proximity in (49). Note that depending on the choice of the parametric family, the optimization problem (50) is generally non-convex.

We make the following observations regarding the optimization problem (50) and its properties:

Alternative formulation for efficiency and numerical resiliency to finite-precision errors Using the closed-form expression for invariance proximity in (49) requires calculating the maximum eigenvalue of $M_C = I - \Phi(Z^+)\Phi(Z)^\dagger\Phi(Z)\Phi(Z^+)^\dagger$. This matrix has spectrum in $[0, 1]$, cf. [16, Lemma 1]. Many software packages for finding maximum eigenvalues rely on iterative methods that are sensitive to the separation between the largest and second largest eigenvalues. To avoid numerical issues, one can use $\text{Tr}(M_C)$ instead of $\lambda_{\max}(M_C)$, as justified by

$$\frac{1}{s}\,\text{Tr}(M_C) \le \lambda_{\max}(M_C) \le \text{Tr}(M_C),$$

where s is the dimension of M_C. Note that the inequalities follow from the fact that the spectrum of M_C belongs to $[0, 1]$.

Equivalence to robust minimax problem Based on Theorem 2.5, the optimization problem (50) is equivalent to the following robust minimax problem

$$\underset{\Phi \in \text{PF}}{\text{minimize}} \max_{f \in \mathcal{S}, \|\mathcal{K}^{\text{aug}} f\|_{L_2(\mu_Z)} \neq 0} \frac{\|\mathcal{K}^{\text{aug}} f - \mathfrak{P}_{\mathcal{K}^{\text{aug}} f}\|_{L_2(\mu_Z)}}{\|\mathcal{K}^{\text{aug}} f\|_{L_2(\mu_Z)}},$$

where $\mathcal{S} = \text{span}(\Phi)$ and μ_Z is defined in (47). This equivalence makes it clear that optimization (50) minimizes the worst-case error on the subspace, does not depend on the choice of basis, and is not sensitive to the scaling of variables.

For the readers' convenience, Algorithm 1 summarizes the steps described above to learn input-state separable models.

Example 9.1 *(DC Motor with Nonlinear Multiplicative Input Injection):* Consider the following equation for a DC mo-

Algorithm 1 Learning Input-State Separable Models

1: Gather data according to (45)-(46)
2: Choose a parametric family of normal dictionaries (e.g., neural networks, polynomials) with real-valued elements in the form (36)
3: Obtain Φ^* by solving (50) using the closed-form expression for the cost in (49)
4: Calculate $\tilde{A} = \Phi^*(Z^+)\Phi^*(Z)^\dagger$
5: Find the input-state separable form via Theorem 8.7(b)

tor [14]

$$\begin{aligned} \dot{x}_1 &= -(R_a/L_a)x_1 - (k_m/L_a)x_2 f(u) + u_a/L_a, \\ \dot{x}_2 &= -(B/J)x_2 + (k_m/J)x_1 f(u) - \tau_l/J, \end{aligned} \tag{51}$$

where x_1 is the armature current, x_2 is the angular velocity, $x = [x_1, x_2]^T$ is the state vector and u is the input. The value of the parameters are $R_a = 12.345$, $L_a = 0.314$, $k_m = 0.253$, $u_a = 60$, $B = 0.00732$, $\tau_l = 1.47$, and $J = 0.00441$. We consider two choices for f: (i) $f(u) = 2\tanh(u)$ (saturated input) and (ii) $f(u) = 2\tanh(u\cos(u))$ (saturated non-monotone input).

Data: We run 10^4 experiments with constant inputs and length $50\,ms$ with uniformly selected initial conditions from the normal operating range $[-5, 15] \times [-250, 125]$ and inputs from $[-4, 4]$. We sample the trajectories with time step $\Delta t = 5\,ms$, resulting in a total of 10^5 data snapshots. Out of this data set, we select half as the training data set and the rest as the test data set.

Parametric Families for Comparison: Our aim is to compare the effectiveness of our methods with widely used lifted-linear (an extension of DMD with control [40]) and bilinear Koopman-inspired forms as

$$\begin{aligned} \Psi_l(x^+) &= A_l\Psi_l(x) + B_l u, \\ \Psi_b(x^+) &= A_b\Psi_b(x) + B_b\Psi_b(x)u, \end{aligned}$$

We use data to learn all the models with dimension four. For the input-state separable model, in the normal basis (36), we set the dimension of normal space $s = 20$ and the dimension of the input-state separable model as $l = 4$. We model the functions $H(x)$ and $\tilde{G}(u)$ in (36) by two residual neural networks [18] comprised of 5 residual blocks each with 64 neuron per hidden layer and ReLU activation functions. We also fix the first two elements of $H(x)$ to be the state vector corresponding to the system. For the linear and bilinear models, we set the functions $\Psi_l(x)$ and $\Psi_b(x)$ to be the same type of neural network used for $H(x)$ in the input-state separable model. To learn the input-state separable model we use Algorithm 1. We train the neural networks for lifted linear and bilinear models by minimizing the following typical least norm residual errors

$$\begin{aligned} &\underset{\Psi_l \in \mathrm{PF}}{\text{minimize}} \quad \|\Psi_l(X^+) - A_l^*\Psi_l(X) - B_l^* U\|_F, \\ &\underset{\Psi_b \in \mathrm{PF}}{\text{minimize}} \quad \|\Psi_b(X^+) - A_b^*\Psi_b(X) - B_b^*\Psi_b(X)\cdot U\|_F, \end{aligned} \tag{52}$$

[14] This system is a modified version of the experimental study [8] which has been used as an example in [25]. We have added the nonlinear function $f : \mathcal{U} \to \mathbb{R}$ in the mechanism generating the field current.

where $\Psi_b(X)\cdot U$ denotes the column-wise product of $\Psi_b(X)$ and U. Moreover, A_l^* and B_l^* are the best parameters by minimizing the same cost function over A_l and B_l instead of Ψ_l (see e.g., [25, Section 4]). A_b^* and B_b^* are computed similarly. Note that the optimization problems in (52) are solved in Ψ_l and Ψ_b and are nonconvex.

Training and Practical Considerations: We randomly initialize the neural networks. The networks for $H(x)$ in the input state separable model, and Ψ_l and Ψ_b start from the same initial weights and biases. To make sure all variables are in the same scale and one does not dominate the others, we do a change of coordinates by $x_1 \mapsto 0.1\,x_1$, $x_2 \mapsto 0.004\,x_2$ and $u \mapsto 0.25\,u$. We scale back to the original coordinates after training. One also might benefit from regularization in case of overfitting. However, it should be noted that the robust minimax problem used in our method is resilient to overfitting since it considers all uncountably many functions in the vector space and we have not used any regularization. To train the neural networks, we use the Adam method [23] with batch size of 200. We train the networks for 500 epochs while decreasing the learning rate linearly from 5×10^{-4} to 10^{-6}. Finally, we use the formula in Theorem 8.7(b) to build an input-state separable model based on the augmented operator.

Evaluation and Comparison: To evaluate the accuracy of models, we create a piecewise constant random input with time step Δt for 600 time steps (or 3 seconds) and compare the learned models' response with the actual system trajectories generated by (51). Figures 1 shows the generated input signal used for comparison. Figure 2 shows the angular velocity of the motor with nonlinear input injection $f(u) = 2\tanh(u)$ compared to predictions derived by the input-state separable, lifted linear, and lifted bilinear models. Figure 3 depicts the same comparison for the case where $f(u) = 2\tanh(u\cos(u))$ in system (51). Clearly, in both cases $f(u) = 2\tanh(u)$ and $f(u) = 2\tanh(u\cos(u))$ the input-state separable model outperforms the other methods. Moreover, by comparing Figure 2 with Figure 3, one can see that the more nonlinear the system is in the input, the less accurate the lifted linear and bilinear models become (even for short-term predictions).

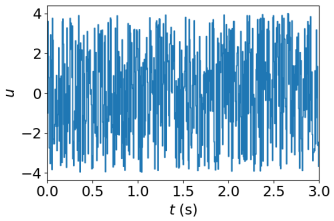

Fig. 1. The piecewise constant random signal used for comparing different models with $\Delta t = 5\,ms$ and length of 600 steps.

□

10 Conclusions

We have presented the notion of Koopman Control Family (KCF), a theoretical framework for modeling general nonlinear control systems. We have shown that the KCF can fully characterize the behavior of a control system on a (potentially infinite-dimensional) function space. To build finite-dimensional models, we have introduced a generalized notion of subspace invariance, leading to a universal finite-dimensional form which we refer to as *input-state separable*.

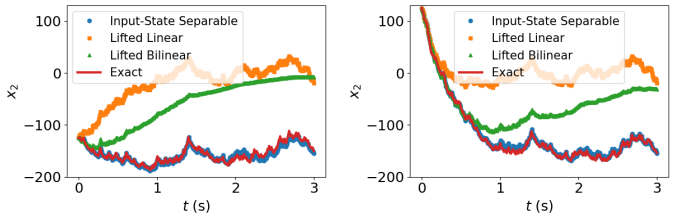

Fig. 2. The angular velocity, x_2, of the DC motor in (51) with $f(u) = 2\tanh(u)$ and predictions by input-state separable (our method), lifted linear, and lifted bilinear models. The trajectories start from two initial conditions $x_0 = [0, -125]$ (left), and $x_0 = [0, 125]$ (right).

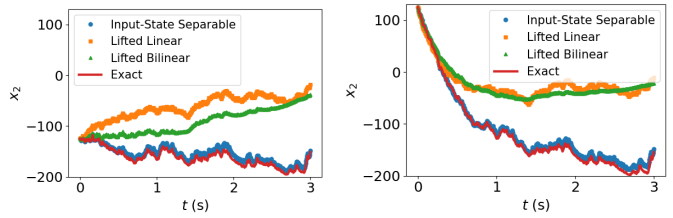

Fig. 3. The angular velocity, x_2, of the DC motor in (51) with $f(u) = 2\tanh(u\cos(u))$ and predictions by input-state separable (our method), lifted linear, and lifted bilinear models. The trajectories start from two initial conditions $x_0 = [0, -125]$ (left), and $x_0 = [0, 125]$ (right).

Remarkably, the commonly-used lifted linear, bilinear, and switched linear models are all special cases of the input-state separable form. We have provided a complete theoretical analysis accompanied by discussions on usage in data-driven applications. Future work will build on the results of the paper to develop strategies for control design, such as using the closed-form solution of the input-state separable models to provide computational gains and performance guarantees for model predictive control as well as extending switching-based linear control designs to the case of uncountable and potentially unbounded input sets. We also aim to build on the proposed framework to determine reachable and control-invariant sets. We also aim to explore additional structures that the KCF might enjoy for special classes of nonlinear systems such as control-affine and monotone systems.

References

[1] S. Bak, S. Bogomolov, P. S. Duggirala, A. R. Gerlach, and K. Potomkin. Reachability of black-box nonlinear systems after Koopman operator linearization. *IFAC-PapersOnLine*, 54(5):253–258, 2021.

[2] H. Balim, A. Aspeel, Z. Liu, and N. Ozay. Koopman-inspired implicit backward reachable sets for unknown nonlinear systems. *IEEE Control Systems Letters*, 7:2245–2250, 2023.

[3] M. Black and D. Panagou. Safe control design for unknown nonlinear systems with Koopman-based fixed-time identification. *arXiv preprint arXiv:2212.00624*, 2022.

[4] M. Blischke and J. P. Hespanha. Learning switched Koopman models for control of entity-based systems. In *IEEE Conf. on Decision and Control*, pages 6006–6013, Singapore, Singapore, 2023. IEEE.

[5] D. Bruder, X. Fu, R. B. Gillespie, C. D. Remy, and R. Vasudevan. Data-driven control of soft robots using Koopman operator theory. *IEEE Transactions on Robotics*, 37(3):948–961, 2020.

[6] S. L. Brunton, B. W. Brunton, J. L. Proctor, and J. N. Kutz. Koopman invariant subspaces and finite linear representations of nonlinear dynamical systems for control. *PLOS One*, 11(2):1–19, 2016.

[7] M. Budišić, R. Mohr, and I. Mezić. Applied Koopmanism. *Chaos*, 22(4):047510, 2012.

[8] S. Daniel-Berhe and H. Unbehauen. Experimental physical parameter estimation of a thyristor driven DC-motor using the HMF-method. *Control Engineering Practice*, 6(5):615–626, 1998.

[9] S. A. Deka, A. M. Valle, and C. J. Tomlin. Koopman-based neural Lyapunov functions for general attractors. In *IEEE Conf. on Decision and Control*, pages 5123–5128, Cancun, Mexico, December 2022.

[10] C. Folkestad, Y. Chen, A. D. Ames, and J. W. Burdick. Data-driven safety-critical control: Synthesizing control barrier functions with Koopman operators. *IEEE Control Systems Letters*, 5(6):2012–2017, 2020.

[11] D. Gadginmath, V. Krishnan, and F. Pasqualetti. Data-driven feedback linearization using the Koopman generator. *arXiv preprint arXiv:2210.05046*, 2022.

[12] D. Goswami and D. A. Paley. Bilinearization, reachability, and optimal control of control-affine nonlinear systems: A Koopman spectral approach. *IEEE Transactions on Automatic Control*, 67(6):2715–2728, 2022.

[13] M. Haseli and J. Cortés. Parallel learning of Koopman eigenfunctions and invariant subspaces for accurate long-term prediction. *IEEE Transactions on Control of Network Systems*, 8(4):1833–1845, 2021.

[14] M. Haseli and J. Cortés. Learning Koopman eigenfunctions and invariant subspaces from data: Symmetric Subspace Decomposition. *IEEE Transactions on Automatic Control*, 67 (7):3442–3457, 2022.

[15] M. Haseli and J. Cortés. Generalizing dynamic mode decomposition: balancing accuracy and expressiveness in Koopman approximations. *Automatica*, 153:111001, 2023.

[16] M. Haseli and J. Cortés. Temporal forward-backward consistency, not residual error, measures the prediction accuracy of Extended Dynamic Mode Decomposition. *IEEE Control Systems Letters*, 7:649–654, 2023.

[17] M. Haseli and J. Cortés. Modeling nonlinear control systems via Koopman control family: universal forms and subspace invariance proximity. *Preprint. Available at* `https://arxiv.org/abs/2307.15368`, 2023.

[18] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[19] B. Huang, X. Ma, and U. Vaidya. Feedback stabilization using Koopman operator. In *IEEE Conf. on Decision and Control*, pages 6434–6439, Miami Beach, FL, December 2018.

[20] L. C. Iacob, R. Tóth, and M. Schoukens. Koopman form of nonlinear systems with inputs. *arXiv preprint arXiv:2207.12132*, 2022.

[21] E. Kaiser, J. N. Kutz, and S. L. Brunton. Data-driven discovery of Koopman eigenfunctions for control. *Machine Learning: Science and Technology*, 2(3):035023, 2021.

[22] M. Khosravi. Representer theorem for learning Koopman operators. *IEEE Transactions on Automatic Control*, 2023.

[23] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. In *International Conference for Learning Representations*, San Diego, CA, USA, May 2015.

[24] B. O. Koopman. Hamiltonian systems and transformation in Hilbert space. *Proceedings of the National Academy of Sciences*, 17(5):315–318, 1931.

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

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

[27] M. Korda and I. Mezic. Optimal construction of Koopman eigenfunctions for prediction and control. *IEEE Transactions on Automatic Control*, 65(12):5114–5129, 2020.

[28] H. Lu and D. M. Tartakovsky. Prediction accuracy of dy-

namic mode decomposition. *SIAM Journal on Scientific Computing*, 42(3):A1639–A1662, 2020.
[29] B. Lusch, J. N. Kutz, and S. L. Brunton. Deep learning for universal linear embeddings of nonlinear dynamics. *Nature Communications*, 9(1):1–10, 2018.
[30] S. Maćešić, N. Črnjarić Žic, and I. Mezić. Koopman operator family spectrum for nonautonomous systems. *SIAM Journal on Applied Dynamical Systems*, 17(4):2478–2515, 2018. doi: 10.1137/17M1133610.
[31] G. Mamakoukas, I. Abraham, and T. D. Murphey. Learning stable models for prediction and control. *IEEE Transactions on Robotics*, 39(3):2255–2275, 2023.
[32] A. Mauroy and I. Mezić. Global stability analysis using the eigenfunctions of the Koopman operator. *IEEE Transactions on Automatic Control*, 61(11):3356–3369, 2016.
[33] I. Mezić. Spectral properties of dynamical systems, model reduction and decompositions. *Nonlinear Dynamics*, 41(1-3):309–325, 2005.
[34] A. Narasingam, S. H. Son, and J. S. Kwon. Data-driven feedback stabilisation of nonlinear systems: Koopman-based model predictive control. *International Journal of Control*, 0(0):1–12, 2022. doi: 10.1080/00207179.2021.2013541.
[35] F. Nüske, S. Peitz, F. Philipp, M. Schaller, and K. Worthmann. Finite-data error bounds for Koopman-based prediction and control. *Journal of Nonlinear Science*, 33(1):14, 2023.
[36] S. E. Otto and C. W. Rowley. Linearly recurrent autoencoder networks for learning dynamics. *SIAM Journal on Applied Dynamical Systems*, 18(1):558–593, 2019.
[37] S. Pan, N. Arnold-Medabalimi, and K. Duraisamy. Sparsity-promoting algorithms for the discovery of informative Koopman-invariant subspaces. *Journal of Fluid Mechanics*, 917, 2021.
[38] S. Peitz and S. Klus. Koopman operator-based model reduction for switched-system control of PDEs. *Automatica*, 106:184–191, 2019.
[39] S. Peitz, S. E. Otto, and C. W. Rowley. Data-driven model predictive control using interpolated Koopman generators. *SIAM Journal on Applied Dynamical Systems*, 19(3):2162–2193, 2020.
[40] J. L. Proctor, S. L. Brunton, and J. N. Kutz. Dynamic mode decomposition with control. *SIAM Journal on Applied Dynamical Systems*, 15(1):142–161, 2016.
[41] C. W. Rowley, I. Mezić, S. Bagheri, P. Schlatter, and D. S. Henningson. Spectral analysis of nonlinear flows. *Journal of Fluid Mechanics*, 641:115–127, 2009.
[42] P. J. Schmid. Dynamic mode decomposition of numerical and experimental data. *Journal of Fluid Mechanics*, 656: 5–28, 2010.
[43] W. Sharpless, N. Shinde, M. Kim, Y. T. Chow, and S. Herbert. Koopman-Hopf Hamilton-Jacobi reachability and control. *arXiv preprint arXiv:2303.11590*, 2023.
[44] L. Shi and K. Karydis. Enhancement for robustness of Koopman operator-based data-driven mobile robotic systems. In *IEEE Int. Conf. on Robotics and Automation*, pages 2503–2510, 2021.
[45] A. Sootla and D. Ernst. Pulse-based control using Koopman operator under parametric uncertainty. *IEEE Transactions on Automatic Control*, 63(3):791–796, 2017.
[46] A. Sootla, A. Mauroy, and D. Ernst. Optimal control formulation of pulse-based control using Koopman operator. *Automatica*, 91:217–224, 2018.
[47] R. Strässer, J. Berberich, and F. Allgöwer. Robust data-driven control for nonlinear systems using the Koopman operator. *IFAC-PapersOnLine*, 56(2):2257–2262, 2023.
[48] M. Sznaier. A data driven, convex optimization approach to learning Koopman operators. In *Learning for Dynamics and Control Conference*, pages 436–446. PMLR, 2021.
[49] N. Takeishi, Y. Kawahara, and T. Yairi. Learning Koopman invariant subspaces for dynamic mode decomposition. In *Conference on Neural Information Processing Systems*, pages 1130–1140, 2017.
[50] J. H. Tu, C. W. Rowley, D. M. Luchtenburg, S. L. Brunton, and J. N. Kutz. On dynamic mode decomposition: theory and applications. *Journal of Computational Dynamics*, 1 (2):391–421, 2014.
[51] B. Umathe, D. Tellez-Castro, and U. Vaidya. Reachability analysis using spectrum of Koopman operator. *IEEE Control Systems Letters*, 7:595–600, 2022.
[52] M. O. Williams, I. G. Kevrekidis, and C. W. Rowley. A data-driven approximation of the Koopman operator: Extending dynamic mode decomposition. *Journal of Nonlinear Science*, 25(6):1307–1346, 2015.
[53] E. Yeung, S. Kundu, and N. Hodas. Learning deep neural network representations for Koopman operators of nonlinear dynamical systems. In *American Control Conference*, pages 4832–4839, Philadelphia, PA, July 2019.
[54] B. Yi and I. R. Manchester. On the equivalence of contraction and Koopman approaches for nonlinear stability and control. In *IEEE Conf. on Decision and Control*, pages 4609–4614, 2021.
[55] C. M. Zagabe and A. Mauroy. Switched nonlinear systems in the Koopman operator framework: Toward a Lie-algebraic condition for uniform stability. In *European Control Conference*, pages 281–286. IEEE, 2021.
[56] V. Zinage and E. Bakolas. Neural Koopman control barrier functions for safety-critical control of unknown nonlinear systems. In *American Control Conference*, pages 3442–3447. IEEE, 2023.
[57] V. Zinage and E. Bakolas. Neural Koopman Lyapunov control. *Neurocomputing*, 527:174–183, 2023.

Masih Haseli received the B.Sc. and M.Sc. degrees in electrical engineering from the Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran, in 2013 and 2015, respectively. He also received the Ph.D. degree in Engineering Sciences (Mechanical Engineering) from the University of California San Diego, CA, USA, in 2022. He is currently a postdoctoral researcher with the Department of Mechanical and Aerospace Engineering, University of California, San Diego, CA, USA. His research interests include system identification, nonlinear systems, network systems, data-driven modeling and control, and distributed and parallel computing. Dr. Haseli is the recipient of the Bronze Medal of the 2014 Iran National Mathematics Competition and the Best Student Paper Award of the 2021 American Control Conference.

Jorge Cortés received the Licenciatura degree in mathematics from Universidad de Zaragoza, Spain, in 1997, and the Ph.D. degree in engineering mathematics from Universidad Carlos III de Madrid, Spain, in 2001. He held postdoctoral positions with the University of Twente, The Netherlands, and the University of Illinois at Urbana-Champaign, USA. He was an Assistant Professor with the Department of Applied Mathematics and Statistics, UC Santa Cruz, USA. He is a Professor in the Department of Mechanical and Aerospace Engineering, UC San Diego, USA. He is a Fellow of IEEE, SIAM, and IFAC. His research interests include distributed control and optimization, network science, nonsmooth analysis, reasoning and decision making under uncertainty, network neuroscience, and multi-agent coordination in robotic, power, and transportation networks.