

Article

Not peer-reviewed version

Correct Degree Selection for Koopman Mode Decomposition

[Kilho Shin](#) * and [Shodai Asaoka](#)

Posted Date: 18 December 2025

doi: 10.20944/preprints202512.1616.v1

Keywords: Koopman mode decomposition; uniquely feasible degree; Hankel matrix; spectral decomposition; model selection; dynamical systems; time series analysis



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Correct Degree Selection for Koopman Mode Decomposition

Kilho Shin ¹  and Shodai Asaoka ²

¹ Gakushuin University

² Kyushu Institute of Technology

* Correspondence: kilhoshin314@gmail.com

Abstract

Fourier Decomposition (FD) and Koopman Mode Decomposition (KMD) are important tools for time series data analysis, applied across a broad spectrum of applications. Both aim to decompose time series functions into superpositions of countably many wave functions, with strikingly similar mathematical foundations. These methodologies derive from the linear decomposition of functions within specific function spaces: FD uses a fixed basis of sine and cosine functions, while KMD employs eigenfunctions of the Koopman linear operator. A notable distinction lies in their scope: FD is confined to periodic functions, while KMD can decompose functions into exponentially amplifying or damping waveforms, making it potentially better suited for describing phenomena beyond FD's capabilities. However, practical applications of KMD often show that despite accurate approximation of training data, its prediction accuracy is limited. This paper clarifies that this issue is closely related to the number of wave components used in decomposition, referred to as the *degree* of a KMD. Existing methods use predetermined, arbitrary, or ad hoc values for this degree. We demonstrate that using a degree different from a uniquely determined value for the data allows infinite KMDs to accurately approximate training data, explaining why current methods, which select a single KMD from these candidates, struggle with prediction accuracy. Furthermore, we introduce mathematically supported algorithms to determine the correct degree. Simulations verify that our algorithms can identify the right degrees and generate KMDs that can make accurate predictions, even with noisy data.

Keywords: Koopman mode decomposition; uniquely feasible degree; Hankel matrix; spectral decomposition; model selection; dynamical systems; time series analysis

1. Introduction

A wide range of natural and social phenomena are observed as superpositions of multiple nonlinear elemental processes. For example, recorded audio signals typically include not only the target speech—for instance, a conversation between individuals—but also various environmental noise components. Similarly, variations in the geomagnetic field arise from both internal processes, such as temporal fluctuations in the Earth's main magnetic field, and external perturbations, such as solar flares and solar wind. Decomposing such observations into constituent processes and extracting only the components relevant to the study is a fundamental procedure in scientific research. Among such methods, frequency analysis—where time-series data are decomposed into countably (often finitely) many frequency components—plays a central role in data science.

A foundational principle across the natural sciences involves reducing nonlinear phenomena to linear problems, enabling analysis via linear algebra. For instance, kernel methods in machine learning embed data into high-dimensional (often infinite-dimensional Hilbert) spaces to facilitate linear solutions. Likewise, neural networks—which approximate arbitrary continuous (and thus potentially nonlinear) functions via linear combinations followed by nonlinear activations—rely on efficient linear transformations during training. The backpropagation algorithm, essential for learning from large-scale data, exemplifies this reliance.

A well-established and widely applied technique based on this principle is **Fourier decomposition**, which forms the foundation of frequency analysis across a wide range of fields. Grounded in functional analysis, it represents a function as a linear combination of frequency components, typically expressed in terms of an orthonormal basis of trigonometric functions. This decomposition facilitates tasks such as signal characterization and noise reduction and has found broad applications in speech recognition and compression, image processing, radar and sonar analysis, time-series forecasting, and medical imaging.

Koopman Mode Decomposition (KMD) has recently attracted considerable theoretical attention as a powerful extension of Fourier decomposition. Its origin traces back to the 1931 work of B.O. Koopman, who formulated a representation of nonlinear dynamical systems through linear operators acting on function spaces—now referred to as *Koopman operators*. The theoretical foundation of this framework was subsequently formalized, and beginning in the 1990s, research by I. Mezić and collaborators renewed interest in its potential to reveal latent dynamics in nonlinear systems. The development of Dynamic Mode Decomposition (DMD) by P.J. Schmid ignited a new wave of research and led to advanced extensions such as Extended DMD (EDMD), which enable practical estimation of Koopman spectral components from data beyond the original limitations of DMD. Although KMD often provides accurate representations of observed data, it has been noted that its predictive accuracy may deteriorate under certain conditions.

The present study aims to address this limitation by identifying the sources of prediction error in KMD and proposing efficient algorithms to extract those Koopman modes that, if present, are the only viable candidates for accurate forecasting.

To illustrate Koopman Mode Decomposition and our contributions, we begin by recalling the concept of Fourier Decomposition (FD). Let $f(t)$ be a 2π -periodic, complex-valued function in L^2 . That is, $f(t + 2\pi) = f(t)$ for all t , and

$$\int_{-\pi}^{\pi} |f(t)|^2 dt < \infty.$$

The space of such functions forms a Hilbert space, where the inner product between $g(t)$ and $h(t)$ is defined by

$$\langle g, h \rangle = \int_{-\pi}^{\pi} g(t) \cdot \overline{h(t)} dt,$$

with \bar{x} denoting the complex conjugate of x . Then, $f(t)$ admits the decomposition:

$$f(t) = \sum_{n=-\infty}^{\infty} c_n e^{int}. \quad (1)$$

This decomposition is justified by the fact that the family

$$\left\{ \frac{1}{\sqrt{2\pi}} e^{int} \mid n \in \mathbb{Z} \right\}$$

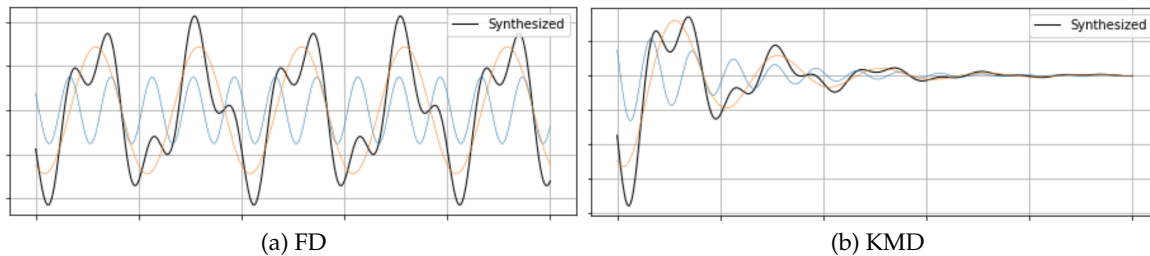
forms a countable orthonormal basis for the Hilbert space of 2π -periodic functions in $L^2([0, 2\pi])$. The convergence in Equation (1) is understood in the L^2 -norm. Although such convergence does not imply pointwise convergence, *Riesz's theorem* [1] [Theorem 3.12] guarantees that a subsequence of the partial sums converges pointwise almost everywhere.

Koopman Mode Decomposition (KMD) [2] is similar to FD in that it expresses a function as a sum of oscillatory components. However, unlike FD, KMD allows for exponentially growing or decaying components. Hence, if a KMD of $f(t)$ exists, it takes the form:

$$f(t) = \sum_{n=0}^{\infty} c_n \lambda_n^t, \quad (2)$$

where $\lambda_n^t = |\lambda_n|^t e^{i \arg(\lambda_n)t}$. Thus, unless $|\lambda_n| = 1$, each term represents an exponentially growing (if $|\lambda_n| > 1$) or decaying (if $|\lambda_n| < 1$) wave (see Figure 1). The λ_n constitute a countable subset of the spectrum of the so-called *Koopman operator* [2].

KMD is expected to provide a more flexible framework for representing diverse phenomena and has been applied in a wide array of domains, including: fluid dynamics [3–6], chaotic systems [7], neuroscience [8], plasma physics [9–11], sports analytics [12], robotics [13], and video processing [14].



The target function (shown in black) is a superposition of two trigonometric functions (represented in orange and blue). The amplitudes of these trigonometric functions are: (a) Constant for the FD and (b) Decreasing over time for KMD.

Figure 1. Comparison between FD and KMD.

In practical settings, both FD and KMD rely on a finite number of observations. Without restricting the summations in Equation (1) and Equation (2) to finitely many terms, the decomposition becomes ill-posed. We therefore approximate the function by a finite superposition of ℓ oscillatory components, where ℓ is called the *degree* of the decomposition.

In the case of the Discrete Fourier Transform (DFT), we assume observations $f(t_0), \dots, f(t_{T-1})$ at $t_k = \frac{2\pi k}{T}$, for $k = 0, \dots, T-1$. Since $e^{imt_k} = e^{int_k}$ whenever $m \equiv n \pmod{T}$, the problem of finding the coefficients c_0, \dots, c_{ℓ} reduces to solving the linear system:

$$(f(t_0), \dots, f(t_{T-1})) = (c_0 \dots c_{T-1}) \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & e^{i\frac{2\pi}{T}} & \dots & e^{i\frac{2\pi(T-1)}{T}} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & e^{i\frac{2\pi(T-1)}{T}} & \dots & e^{i\frac{2\pi(T-1)^2}{T}} \end{bmatrix}. \quad (3)$$

This system has a unique solution, as the coefficient matrix is a square Vandermonde matrix over distinct T -th roots of unity.

In general, an $m \times n$ matrix

$$\mathcal{V}_{n|a_1, \dots, a_m} = \begin{bmatrix} 1 & a_1 & a_1^2 & \dots & a_1^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_m & a_m^2 & \dots & a_m^{n-1} \end{bmatrix},$$

is referred to as a *Vandermonde* matrix, whose determinant when $m = n$ is given by

$$\det \mathcal{V}_{m|a_1, \dots, a_m} = \prod_{i>j} (a_i - a_j).$$

The square matrix on the right-hand side of Equation (3) is a Vandermonde matrix, and Equation (3) can be restated as

$$(x_0 \dots x_{T-1}) = (c_0 \dots c_{\ell-1}) \mathcal{V}_{T|\alpha^0, \alpha^1, \dots, \alpha^{\ell-1}}, \quad (4)$$

where $\alpha = e^{2\pi i/T}$ is a primitive T -th root of unity.

By the aforementioned invertibility of the Vandermonde matrix generated by distinct points $\alpha^0, \dots, \alpha^{T-1}$, Equation (3) admits a unique solution:

$$(c_0 \dots c_{T-1}) = \frac{1}{T} (x_0 \dots x_{T-1}) \mathcal{V}_{T|\alpha^0, \alpha^1, \dots, \alpha^{T-1}}^* \quad (5)$$

Here, \mathbb{M}^* denotes the conjugate transpose (i.e., Hermitian transpose) of a matrix \mathbb{M} .

In contrast, the KMD problem can be formulated analogously as:

$$[x_0 \dots x_{T-1}] = [m_1 \dots m_\ell] \mathcal{V}_{T|\lambda_1, \dots, \lambda_\ell} \quad (6)$$

with the following distinctions:

1. Each observable x_t is an m -dimensional vector. We denote the matrix of observations by $\mathbb{X} = [x_0 \dots x_{T-1}]$.
2. The eigenvalues $\lambda_1, \dots, \lambda_\ell$ and the corresponding modes m_1, \dots, m_ℓ are unknown and must be determined.
3. The choice $\ell = T$, which is required for DFT, is entirely unsuitable for KMD: for any distinct set of $\lambda_1, \dots, \lambda_T$, there always exists a corresponding set of modes m_1, \dots, m_T such that Equation (6) holds exactly.

Despite the increased complexity of the KMD problem, several numerical methods exist to solve Equation (6) for a given degree ℓ , including: Dynamic Mode Decomposition (DMD), which is typically applicable when $\ell = \text{rank } \mathbb{X}$; the Arnoldi method, applicable when $\ell = T - 1$; and the vector Prony method, which allows arbitrary ℓ . These methods yield approximate solutions minimizing the residual sum of squares (RSS), especially in the presence of observation noise.

However, what remained unresolved was how to determine an optimal degree ℓ . We illustrate its importance through the following example, highlighting the predictive risk of inappropriate choice.

Example 1. Consider one-dimensional observables given by $\mathbb{X} = [1 \ 1 \ 1 \ 1 \ 3 \ 5 \ 7]$. It is readily verified that Equation (6) admits no solution if $\ell \leq 3$. For $\ell = 4$, the roots $\lambda_1, \dots, \lambda_4$ of the following equation

$$f(x; \alpha) = x^4 - x^3 - x - 1 + \alpha(x - 1) = 0 \quad (7)$$

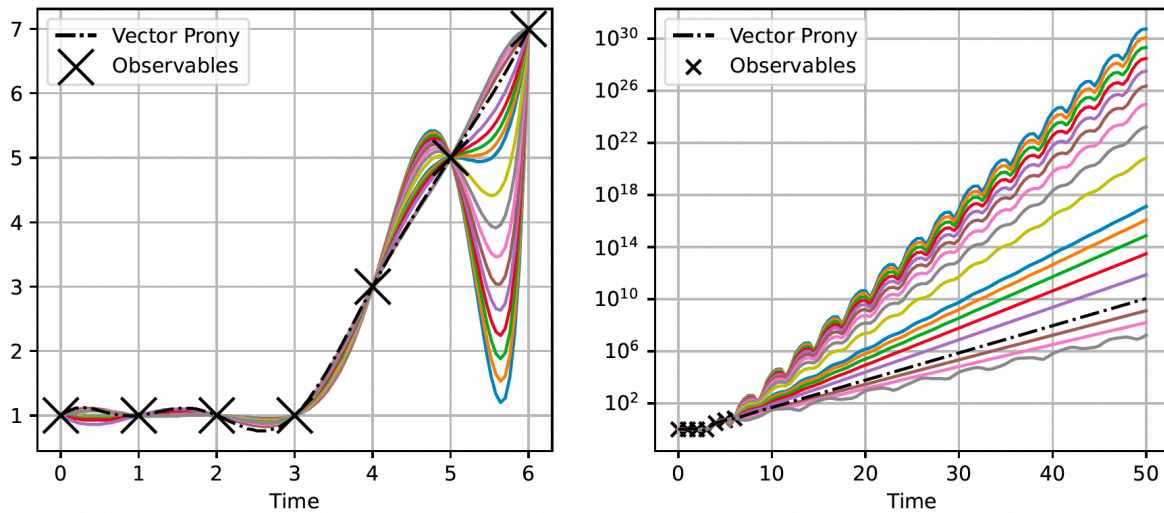
uniquely determine the modes m_1, \dots, m_4 such that Equation (6) is satisfied, thereby yielding a valid KMD for any value of the parameter α . As illustrated in Figure 2, all such KMDs exactly reproduce the observed sequence \mathbb{X} for $t = 0, \dots, 6$, yet their extrapolations for $t \geq 7$ differ significantly.

This example highlights a key issue: even if an algorithm happens to return a single quartic KMD, it is merely one among infinitely many KMDs that fit the observed data. Consequently, the forecast made by such a KMD is almost certainly different from the ground truth, and the chance of accurate prediction is negligibly small.

Thus, for the sake of predictive accuracy, it is crucial to select ℓ such that it is *uniquely feasible*, defined as follows:

Definition 1. Given an observable matrix \mathbb{X} , a degree ℓ is said to be feasible if there exists at least one solution $(\lambda_1, \dots, \lambda_\ell; m_1, \dots, m_\ell)$ to Equation (6). Moreover, if this solution is unique, then ℓ is said to be uniquely feasible.

This paper develops a theoretical framework for uniquely feasible degrees and, based on this foundation, proposes efficient and practical algorithms to determine whether a given set of observables \mathbb{X} admits a uniquely feasible degree—and if so, to identify it. We also demonstrate through simulations that the KMD selected by our algorithms can yield highly accurate predictions.



Each KMD exactly fits the 7 observed values (left), while their extrapolations for $t \geq 7$ differ substantially (right). The black dash-dot curve indicates the KMD obtained by the vector Prony method.

Figure 2. Illustration of multiple valid KMDs with divergent predictions.

2. Theoretical Frameworks Underlying Koopman Mode Decomposition

A key significance of Koopman Mode Decomposition (KMD) is its ability to analyze the dynamics of a nonlinear system using only methods from linear algebra. In this section, we provide a brief review of the theoretical framework of KMD, which bridges nonlinear dynamics and linear algebra.

2.1. Temporal Transition of States and Semigroup Property

Let Z denote a (possibly unobservable) state space. Under a deterministic assumption, once a state $\zeta \in Z$ is observed at some time, the state of the system after an elapsed time $t \geq 0$ is uniquely determined and is denoted by ζ_t . Accordingly, the temporal evolution of the system is described by the mapping

$$\widehat{\zeta} : Z \times [0, \infty) \ni (\zeta, t) \mapsto \zeta_t \in Z.$$

In the discrete-time setting, $\widehat{\zeta}$ is instead defined on $Z \times (\{0\} \cup \mathbb{N})$, which can be regarded as a special case of the continuous-time formulation.

While the notation $\widehat{\zeta}(\zeta, t)$ emphasizes the bivariate nature of the mapping, the map $t \mapsto \zeta_t$ is essentially regarded as a univariate function of t , with the initial state ζ fixed. The deterministic assumption also requires the identity $\zeta_{s+t} = (\zeta_s)_t$, which is equivalently expressed as

$$\widehat{\zeta}(\zeta, s+t) = \widehat{\zeta}(\widehat{\zeta}(\zeta, s), t)$$

for all $s, t \geq 0$. This implies that if we define $\sigma^t := \widehat{\zeta}(\cdot, t) : Z \rightarrow Z$ for $t \geq 0$, then the family

$$\{\sigma^t : Z \rightarrow Z \mid t \in [0, \infty)\}$$

forms a one-parameter semigroup under composition, that is,

$$\sigma^{s+t} = \sigma^t \circ \sigma^s$$

holds for all $s, t \geq 0$.

2.2. Koopman Operator

We denote the space of \mathbb{C} -valued functions defined over Z by \mathbb{C}^Z . The function space \mathbb{C}^Z forms a \mathbb{C} -algebra, equipped with addition, multiplication, and scalar multiplication, defined as follows for $f, g \in \mathbb{C}^Z$ and $z \in \mathbb{C}$:

$$(f + g)(\zeta) = f(\zeta) + g(\zeta), \quad (fg)(\zeta) = f(\zeta)g(\zeta), \quad \text{and} \quad (zf)(\zeta) = zf(\zeta)$$

In particular, \mathbb{C}^Z is a vector space over \mathbb{C} .

The *Koopman operator* parameterized by time t is defined as

$$U^t : \mathbb{C}^Z \ni f \mapsto f \circ \sigma^t \in \mathbb{C}^Z.$$

It is straightforward to verify that the Koopman operator is a \mathbb{C} -algebra homomorphism and, in particular, a linear operator. Furthermore, we have:

Proposition 1. *The collection of Koopman operators $\{U^t \mid t \in [0, \infty)\}$ forms a one-parameter semigroup. That is, $U^{s+t} = U^t \circ U^s$ holds for any $s, t \in [0, \infty)$.*

2.3. Koopman Generator

In general, when a one-parameter semigroup $\mathcal{T} = \{T^t \mid t \in [0, \infty)\}$ is defined on a Banach space \mathcal{B} , it is said to be *strongly continuous* if, for every $x \in \mathcal{B}$, the following norm convergence holds:

$$\lim_{t \downarrow 0} \|T^t x - x\| = 0.$$

A strongly continuous one-parameter semigroup \mathcal{T} has several important properties [15] [Chapter 13]:

- Each T^t is bounded; that is, the operator norm $\|T^t\|$ is well-defined. More precisely, there exist constants $M \geq 1$ and $\omega \in \mathbb{R}$ such that $\|T^t\| \leq Me^{\omega t}$ for all $t \geq 0$.
- The set $D(A)$ of all $x \in \mathcal{B}$ for which

$$Ax := \lim_{t \downarrow 0} \frac{T^t x - x}{t}$$

exists is a dense linear subspace of \mathcal{B} , and A is a closed linear operator with domain $D(A)$. This operator A is called the *infinitesimal generator* of $\{T^t\}$.

- If $\{T^t\}$ is bounded, i.e., $\sup_{t \geq 0} \|T^t\| < \infty$, then $D(A) = \mathcal{B}$.
- For every $x \in D(A)$, the derivative

$$\frac{d}{dt} T^t x = \lim_{\tau \downarrow 0} \frac{T^{t+\tau} x - T^t x}{\tau}$$

exists, and we have

$$\frac{d}{dt} T^t x = T^t Ax = AT^t x$$

for all $t \geq 0$ and $x \in D(A)$.

- If $Ax = \lambda x$ for some $x \neq 0$ and $\lambda \in \mathbb{C}$ (that is, λ is an eigenvalue of A), then

$$T^t x = e^{\lambda t} x.$$

When we say that the Koopman operator semigroup $\mathcal{U} = \{U^t \mid t \in [0, \infty)\}$ is *strongly continuous*, we assume that it acts on a Banach space \mathcal{B} whose elements can be regarded as functions in \mathbb{C}^Z in some way (e.g., $\mathcal{B} \subseteq \mathbb{C}^Z$), and that for each $f \in \mathcal{B}$ the limit

$$\lim_{t \downarrow 0} U^t f = f$$

holds in the norm of \mathcal{B} . Pointwise convergence of functions is a more primitive notion, and although these two modes of convergence are generally independent, they are closely related in certain settings.

- Let X be a compact topological space and let $C(X)$ denote the space of continuous functions on X . Since every continuous function on a compact space is bounded, we may equip $C(X)$ with the supremum norm:

$$\|f\|_\infty := \sup_{x \in X} |f(x)|.$$

With this norm, $C(X)$ is a Banach space. In this setting, convergence in norm is equivalent to uniform convergence, and in particular, uniform convergence implies pointwise convergence.

- Let (X, Σ, μ) be a measure space, and let $\mathcal{B} = L^p(\mu)$ for $1 \leq p < \infty$. Elements of $L^p(\mu)$ are equivalence classes of measurable functions that are equal almost everywhere. Thus, any statement about pointwise convergence should be interpreted in terms of representatives of these equivalence classes, that is, convergence almost everywhere. If a sequence $\{f_n\}_{n \in \mathbb{N}} \subset L^p(\mu)$ satisfies

$$\lim_{n \rightarrow \infty} \|f_n - f\|_{L^p} = 0,$$

then there exists a subsequence that converges to f pointwise almost everywhere. This follows from the completeness of L^p spaces and is sometimes referred to as a version of the Riesz convergence theorem (see [1] [Theorem 3.12]).

If the Koopman operator semigroup $\{U^t \mid t \in [0, \infty)\}$ is bounded, then its infinitesimal generator, referred to as the *Koopman generator*, and defined by

$$\mathcal{K}f = \lim_{t \downarrow 0} \frac{U^t f - f}{t}, \quad f \in D(\mathcal{K}), \quad (8)$$

is defined on the entire Banach space.

We next consider the case in which the Koopman operators are bounded. Let (Z, Σ, μ) be a measure space, and suppose that for each $t \geq 0$ the map $\sigma^t : Z \rightarrow Z$ is measurable. We examine the boundedness of the associated Koopman operator U^t acting on $L^p(\mu)$.

A sufficient condition for U^t to be bounded on $L^p(\mu)$ is that σ^t is *non-expansive* with respect to μ , meaning that

$$\mu((\sigma^t)^{-1}(A)) \leq \mu(A) \quad \text{for all } A \in \Sigma.$$

In this case, for any $f \in L^p(\mu)$,

$$\|U^t f\|_{L^p(\mu)}^p = \int_Z |f(\sigma^t(z))|^p d\mu(z) = \int_Z |f(z)|^p d((\sigma^t)_* \mu)(z),$$

where $(\sigma^t)_* \mu$ denotes the pushforward of μ by σ^t . If σ^t is non-expansive, then $(\sigma^t)_* \mu \leq \mu$ (as measures), and hence

$$\|U^t f\|_{L^p(\mu)} \leq \|f\|_{L^p(\mu)},$$

so in particular

$$\sup_{t \geq 0} \|U^t\|_{L^p(\mu) \rightarrow L^p(\mu)} \leq 1 < \infty.$$

(For $p = \infty$, the same argument shows $\|U^t\|_{L^\infty \rightarrow L^\infty} \leq 1$.)

Conversely, if σ^t is *expansive*, i.e., there exists $A \in \Sigma$ such that

$$\mu((\sigma^t)^{-1}(A)) > \mu(A),$$

then the family $\{U^t\}_{t \geq 0}$ may fail to be uniformly bounded in t (even though each fixed U^t can still be bounded).

In many applications, especially in ergodic theory and dynamical systems, σ^t is assumed to be *measure-preserving*, that is,

$$\mu((\sigma^t)^{-1}(A)) = \mu(A) \quad \text{for all } A \in \Sigma.$$

This implies $(\sigma^t)_*\mu = \mu$, and hence U^t acts as an isometry on $L^p(\mu)$ for every $1 \leq p \leq \infty$. On $L^2(\mu)$, the Koopman operators are therefore *unitary*. If, in addition, $\{\sigma^t\}_{t \in \mathbb{R}}$ forms a measure-preserving *flow* (so that $\{U^t\}_{t \in \mathbb{R}}$ is a strongly continuous unitary group), then by Stone's theorem [1] [Theorem 13.40] the Koopman generator \mathcal{K} is *skew-adjoint*, i.e., $\mathcal{K}^* = -\mathcal{K}$. Since unitary and skew-adjoint operators are normal, the spectral theorem applies and provides the functional-analytic foundation for the Koopman Mode Decomposition.

2.4. Koopman Mode Decomposition and Spectral Theorem

To introduce the Koopman mode decomposition, we assume that the Koopman operator semigroup $\{U^t \mid t \in [0, \infty)\}$ defined on a Banach space \mathcal{B} is strongly continuous in the norm of \mathcal{B} and induces a Koopman generator defined on the entire \mathcal{B} . For example, this holds when the semigroup $\{U^t\}$ is bounded.

Let $\sigma_p(\mathcal{K}) \subset \mathbb{C}$ denote the point spectrum of \mathcal{K} , and let $V_\lambda \subseteq \mathcal{B}$ be the eigenspace corresponding to $\lambda \in \sigma_p(\mathcal{K})$. When f belongs to the completion in \mathcal{B} of the linear span of $\bigcup_{\lambda \in \sigma_p(\mathcal{K})} V_\lambda$, that is, when

$$f = \sum_{\lambda \in \sigma_p(\mathcal{K})} \phi_\lambda, \quad \phi_\lambda \in V_\lambda,$$

holds in the norm of \mathcal{B} , only countably many ϕ_λ are nonzero. We denote the corresponding eigenvalues by $\{\lambda_n \mid n = 1, 2, \dots\}$. Then the Koopman mode decomposition of f is expressed as

$$f = \sum_{n=1}^N \phi_{\lambda_n}, \quad N \in \mathbb{N} \cup \{\infty\},$$

and the following relations hold:

$$\mathcal{K}f = \sum_{n=1}^N \lambda_n \phi_{\lambda_n}, \quad U^t f = \sum_{n=1}^N e^{\lambda_n t} \phi_{\lambda_n}, \quad t \in [0, \infty).$$

If, in addition, every element of the one-parameter semigroup $\{\sigma^t \mid t \in [0, \infty)\}$ is measure-preserving on a measure space (Z, Σ, μ) , then the Koopman operators U^t and the Koopman generator \mathcal{K} defined on $L^2(Z)$ are unitary and skew-adjoint, respectively; that is, they are normal linear operators defined on the entire \mathcal{B} . Hence, the Koopman mode decomposition can be understood in the context of the spectral theorem.

The spectral theorem asserts that a normal operator T defined on a Hilbert space \mathcal{H} can be represented as

$$T = \int_{\sigma(T)} \lambda dE(\lambda),$$

where E is a *projection-valued measure*, which plays the role of a Borel measure defined on the Borel σ -algebra \mathbb{B} of $\sigma(T) \subseteq \mathbb{C}$. For each $B \in \mathbb{B}$, the value $E(B)$ is an orthogonal projection operator on \mathcal{H} rather than a real number, and the measure E satisfies the following properties:

- **Orthogonality:** $E(B_1)E(B_2) = E(B_1 \cap B_2)$.

- **Countable additivity:** For any countable mutually disjoint family $\{B_i\}_{i \in I} \subseteq \mathbb{B}$,

$$E\left(\bigcup_{i \in I} B_i\right) = \sum_{i \in I} E(B_i),$$

where the convergence is in the strong operator topology.

Based on the projection-valued measure E , the integral of a measurable function F over \mathbb{C} is defined as

$$F(T) = \int_{\sigma(T)} F(\lambda) dE(\lambda),$$

in complete analogy with the Lebesgue integral.

This integral representation is essential, because the spectrum $\sigma(T)$ may be distributed continuously in \mathbb{C} . Furthermore, letting $\sigma_d(T) \subseteq \sigma_p(T)$ denote the set of all isolated eigenvalues of T , we can express the decomposition as

$$F(T) = \sum_{\lambda \in \sigma_d(T)} F(\lambda) E(\{\lambda\}) + \int_{\sigma(T) \setminus \sigma_d(T)} F(\lambda) dE(\lambda), \quad (9)$$

where each $E(\{\lambda\})$ for $\lambda \in \sigma_d(T)$ coincides with the orthogonal projection onto the eigenspace of λ .

For $f \in L^2(Z)$ —an equivalence class of functions equal almost everywhere—the Koopman mode decomposition of f and the actions of the Koopman generator \mathcal{K} and the Koopman operators U^t are obtained by setting $T = \mathcal{K}$ and assuming that the integral part of Equation (9) vanishes (or is negligible). Since $E(\{\lambda\})f$ is nonzero only for a countable subset of $\sigma_d(\mathcal{K})$, we label the corresponding eigenvalues as $\{\lambda_n\}_{n=1}^N$ for $N \in \mathbb{N} \cup \{\infty\}$, and write $\phi_{\lambda_n} := E(\{\lambda_n\})f$.

- For $F(x) = 1$, we obtain the Koopman mode decomposition of f :

$$f = \sum_{\lambda \in \sigma_d(T)} E(\{\lambda\})f = \sum_{n=1}^N \phi_{\lambda_n}.$$

- For $F(x) = x$, we obtain the action of \mathcal{K} :

$$\mathcal{K}f = \sum_{\lambda \in \sigma_d(T)} \lambda E(\{\lambda\})f = \sum_{n=1}^N \lambda_n \phi_{\lambda_n}.$$

- For $F(x) = e^{tx}$, we obtain the action of U^t :

$$e^{t\mathcal{K}}f = \sum_{\lambda \in \sigma_d(T)} e^{t\lambda} E(\{\lambda\})f = \sum_{n=1}^N e^{t\lambda_n} \phi_{\lambda_n}.$$

3. Discrete Koopman Mode Decomposition

The objective of Discrete Koopman Mode Decomposition (DKMD) is analogous to that of the Discrete Fourier Transform (DFT): to obtain a decomposition that fits a given finite sequence of samples of an unknown function. We begin by recalling the formulation of the DFT.

3.1. DFT and Vandermonde Matrix

Let (x_0, \dots, x_{T-1}) denote the values of an unknown function observed at times $t = 0, 1, \dots, T-1$. The DFT seeks coefficients $c_0, c_{\pm 1}, c_{\pm 2}, \dots$ such that

$$x_t = \sum_{n=-\infty}^{\infty} c_n e^{i2\pi n \frac{t}{T}}, \quad t = 0, \dots, T-1.$$

Since this is an underdetermined system with infinitely many unknowns, we restrict to a finite sum. Moreover, since $n \equiv n' \pmod{T}$ implies $e^{i2\pi n \frac{t}{T}} = e^{i2\pi n' \frac{t}{T}}$, we may assume $n = 0, 1, \dots, T-1$. Let $\omega := e^{i\frac{2\pi}{T}}$. Then

$$x_t = \sum_{n=0}^{T-1} c_n \omega^{nt}, \quad t = 0, \dots, T-1.$$

In matrix form, we can write

$$\begin{pmatrix} x_0 & x_1 & \dots & x_{T-1} \end{pmatrix} = \begin{pmatrix} c_0 & c_1 & \dots & c_{T-1} \end{pmatrix} \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \omega & \omega^2 & \dots & \omega^{T-1} \\ 1 & \omega^2 & \omega^4 & \dots & \omega^{2(T-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{T-1} & \omega^{2(T-1)} & \dots & \omega^{(T-1)^2} \end{bmatrix}. \quad (10)$$

The coefficient matrix above is an instance of the Vandermonde matrix.

Definition 2. Let $a_1, \dots, a_m \in \mathbb{C}$ and $n \in \mathbb{N}$. The associated Vandermonde matrix is defined as

$$\mathcal{V}_{n|a_1, \dots, a_m} = \begin{bmatrix} 1 & a_1 & a_1^2 & \dots & a_1^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_m & a_m^2 & \dots & a_m^{n-1} \end{bmatrix}. \quad (11)$$

Basic properties. For pairwise distinct nodes a_1, \dots, a_m , the Vandermonde matrix satisfies:

- $\det \mathcal{V}_{m|a_1, \dots, a_m} = \prod_{1 \leq j < i \leq m} (a_i - a_j)$.
- $\text{rank } \mathcal{V}_{n|a_1, \dots, a_m} = \min\{m, n\}$.
- Viewing $\mathcal{V}_{n|a_1, \dots, a_m}$ as a linear map $F: \mathbb{C}^m \rightarrow \mathbb{C}^n$, F is injective if and only if $m \leq n$, surjective if and only if $m \geq n$, and bijective if and only if $m = n$.

In particular,

$$\det \mathcal{V}_{T|\omega^0, \omega^1, \dots, \omega^{T-1}} = \prod_{0 \leq m < n < T} (\omega^n - \omega^m) \neq 0,$$

so Equation (10) has a unique solution for (c_0, \dots, c_{T-1}) .

3.2. Formulation of DKMD

Assume that the Koopman mode decomposition of a function f is expressed as a countably infinite sum of eigenfunctions $\{\phi_n\}_{n \in \mathbb{N}}$ of the Koopman generator \mathcal{K} . For example, assume that the spectrum of \mathcal{K} consists only of a countably infinite set of isolated eigenvalues $\{\lambda_n\}_{n \in \mathbb{N}}$:

$$f = \sum_{n=1}^{\infty} \phi_n.$$

Then, without loss of generality, we may assume that the values $\{e^{\lambda_n}\}_{n \in \mathbb{N}}$ are pairwise distinct.

Given finitely many samples $x_t = f(t)$ for $t = 0, 1, \dots, T-1$, DKMD seeks to satisfy

$$x_t = \sum_{n=1}^{\infty} c_n e^{\lambda_n t}, \quad \text{where } c_n := \phi_n(0). \quad (12)$$

In the same way as in the DFT case, since the infinitely many unknowns $\{c_n\}_{n=1}^{\infty}$ and $\{\lambda_n\}_{n=1}^{\infty}$ make the system indefinite, we restrict their number to a finite $\ell < \infty$. Then the system can be written compactly as

$$(x_0, x_1, \dots, x_{T-1}) = (c_1, c_2, \dots, c_\ell) \mathcal{V}_{T|e^{\lambda_1}, \dots, e^{\lambda_\ell}}. \quad (13)$$

If $e^{\lambda_1}, \dots, e^{\lambda_\ell}$ are pairwise distinct and $\ell \geq T$, the Vandermonde matrix represents a surjective linear mapping from \mathbb{C}^ℓ onto \mathbb{C}^T . Then we have the following:

Proposition 2. For any $(x_0, \dots, x_{T-1}) \in \mathbb{C}^T$, pairwise distinct $e^{\lambda_1}, \dots, e^{\lambda_\ell}$, and $\ell \geq T$, there exists $(c_1, \dots, c_\ell) \in \mathbb{C}^\ell$ satisfying Equation (13).

Thus, in contrast to the DFT case, where the given values x_0, \dots, x_{T-1} are decomposed into exactly T Fourier components, performing DKMD requires determining the number ℓ of Koopman modes within the range $\ell < T$. Specifically, in the nonlinear system (13), the Koopman degree ℓ itself appears as an additional unknown, alongside $e^{\lambda_1}, \dots, e^{\lambda_\ell}$ and c_1, \dots, c_ℓ .

In this regard, the primary contribution of this paper is to establish criteria for determining ℓ , and to present computationally efficient methods for estimating it both exactly and approximately, particularly when the given observations are contaminated by noise.

In Section 5, we review an existing method from the literature for solving the equation in the special case where ℓ is known.

4. Definitions and Notations

The following definitions and descriptions are essential for the subsequent sections.

We extend the aforementioned formalization of DKMD from the decomposition of \mathbb{C} -valued functions to that of \mathbb{C}^m -valued functions for a dimension $m \geq 1$. This modification does not alter the fundamental nature of the problem but better aligns with real-world applications of DKMD, such as fluid dynamics.

Definition 3 (DKMD). Let $[x_0 \dots x_{T-1}]$ be a matrix of observables at times $t = 0, \dots, T-1$, where each x_t is an m -dimensional column vector. The discrete Koopman mode decomposition (DKMD) of this observable matrix is a matrix factorization given by

$$[x_0 \dots x_{T-1}] = [m_1 \dots m_\ell] \mathcal{V}_{T|\mu_1, \dots, \mu_\ell}, \quad (14)$$

where $\mu_1, \dots, \mu_\ell \in \mathbb{C}$ are pairwise distinct.

Definition 4 (Koopman Eigenvalues, Modes, and Degree). Given a DKMD as in Equation (14), we refer to ℓ as the Koopman degree, μ_1, \dots, μ_ℓ as the Koopman eigenvalues, and m_1, \dots, m_ℓ as the Koopman modes.

For convenience, let \mathbb{X} and \mathbb{M} denote the observable matrix and the matrix of modes, respectively:

$$\mathbb{X} = [x_0 \dots x_{T-1}], \quad \mathbb{M} = [m_1 \dots m_\ell].$$

Then the DKMD can be expressed compactly as

$$\mathbb{X} = \mathbb{M} \mathcal{V}_{T|\mu_1, \dots, \mu_\ell}. \quad (15)$$

Table 1 summarizes the principal notations used throughout this article.

Table 1. Notation summary.

Notation	Description
ℓ	Koopman degree.
$\mathbf{x}_0, \dots, \mathbf{x}_{T-1} \in \mathbb{C}^m$	Column vectors of observables.
$\mu_1, \dots, \mu_\ell \in \mathbb{C}$	Koopman eigenvalues of an ℓ -degree DKMD.
$\mathbf{m}_1, \dots, \mathbf{m}_\ell \in \mathbb{C}^m$	Koopman modes of an ℓ -degree DKMD.
$\mathbb{X} \in \mathbb{C}^{m \times T}$	Observable matrix $[\mathbf{x}_0 \dots \mathbf{x}_{T-1}]$.
$\mathbb{M} \in \mathbb{C}^{m \times \ell}$	Mode matrix $[\mathbf{m}_1 \dots \mathbf{m}_\ell]$.
$\mathbb{X}_i^j \in \mathbb{C}^{m \times (j-i+1)}$	Submatrix $[\mathbf{x}_i \dots \mathbf{x}_j]$ for $0 \leq i \leq j < T$.
$\mathbb{X}\langle i \rangle \in \mathbb{C}^T$	The i th row vector of \mathbb{X} .
$\mathcal{V}_{n a_1, \dots, a_k} \in \mathbb{C}^{k \times n}$	Vandermonde matrix (Definition 2).
$\mathcal{H}_{k \mathbb{X}} \in \mathbb{C}^{(k+1) \times m(T-k)}$	Hankel matrix (Definition 6).
$\dim_{\mathcal{H}}(k \mathbb{X})$	k th Hankel dimension of \mathbb{X} , defined as $\text{rank } \mathcal{H}_{k \mathbb{X}}$ (Definition 10).
$\text{codim}_{\mathcal{H}}(k \mathbb{X})$	k th Hankel codimension of \mathbb{X} , defined as $k + 1 - \dim_{\mathcal{H}}(k \mathbb{X})$ (Definition 10).
L	The smallest ℓ such that $\text{codim}_{\mathcal{H}}(\ell \mathbb{X}) > 0$.
A_{ij}	Entry of a matrix A at row i and column j .
$\ A\ _F$	Frobenius norm of A : $\ A\ _F = \sqrt{\sum_{i,j} A_{ij} ^2}$.
A^+	Moore–Penrose pseudoinverse of A ; $B = CA^+$ minimizes $\ C - BA\ _F$.
$\mathbb{X}^T \in \mathbb{C}^{n \times m}$	Transpose of $\mathbb{X} \in \mathbb{C}^{m \times n}$.
$\mathbb{X}^* \in \mathbb{C}^{n \times m}$	Conjugate transpose of $\mathbb{X} \in \mathbb{C}^{m \times n}$.
$\mathbb{V}(\mathbb{M})$	Subspace spanned by the column vectors of \mathbb{M} .
$[\mathbf{A} \ \mathbf{B}] \in \mathbb{C}^{k \times (m+n)}$	Matrix obtained by appending the n columns of $\mathbf{B} \in \mathbb{C}^{k \times n}$ to $\mathbf{A} \in \mathbb{C}^{k \times m}$.
$V^\perp \subset \mathbb{C}^n$	Orthogonal complement of a subspace $V \subseteq \mathbb{C}^n$.
v_i	The i th component of a vector v .
$\mathbf{0}^n \in \mathbb{C}^n$	n -dimensional zero row vector $(0 \dots 0)$.
$\mathbf{v}^{a,b} \in \mathbb{C}^{n+a+b}$	Column vector defined as $\mathbf{v}^{a,b} = (\mathbf{0}^a \ \mathbf{v}^T \ \mathbf{0}^b)^T$ for $\mathbf{v} \in \mathbb{C}^n$.
$\text{diag}(A_1, \dots, A_k)$	Block diagonal matrix with diagonal blocks A_1, \dots, A_k .

5. Computing DKMD for Known Degrees (Related Work)

In this section, we introduce the *vector Prony method* [16], which estimates the unknown Koopman eigenvalues μ_1, \dots, μ_ℓ and the Koopman mode matrix \mathbb{M} that satisfy Equation (15), given the observable matrix \mathbb{X} and the Koopman degree ℓ .

The procedure first computes the Koopman eigenvalues, and subsequently computes the Koopman modes based on the obtained eigenvalues.

5.1. Computing the Koopman Eigenvalues

We first introduce the characteristic polynomial associated with a DKMD, whose roots correspond to the Koopman eigenvalues.

Definition 5 (Characteristic Polynomial). For a DKMD $\mathbb{X} = \mathbb{M}\mathcal{V}_{T|\mu_1, \dots, \mu_\ell}$, the polynomial

$$f(X) = \prod_{i=1}^{\ell} (X - \mu_i)$$

is called the characteristic polynomial of the DKMD.

If the characteristic polynomial of a DKMD is expressed as

$$f(X) = X^\ell + a_{\ell-1}X^{\ell-1} + \dots + a_0,$$

then the following recurrence relation holds:

Proposition 3. For all integers j with $0 \leq j \leq T - \ell - 1$,

$$\mathbf{x}_{j+\ell} + a_{\ell-1}\mathbf{x}_{j+\ell-1} + \cdots + a_0\mathbf{x}_j = \mathbf{0}.$$

Proof. The statement follows from

$$\mathbf{x}_{j+\ell} + a_{\ell-1}\mathbf{x}_{j+\ell-1} + \cdots + a_0\mathbf{x}_j = \mathbb{M} \mathbf{diag}(\mu_1^j, \dots, \mu_\ell^j) \begin{bmatrix} f(\mu_1) & \dots & f(\mu_\ell) \end{bmatrix}^\top = \mathbf{0}.$$

□

Using Hankel matrices (Definition 6), the statement of Proposition 3 can be compactly expressed as

$$\mathcal{H}_{0|\mathbb{X}_\ell^{T-1}} = - \begin{bmatrix} a_0 & \dots & a_{\ell-1} \end{bmatrix} \mathcal{H}_{\ell-1|\mathbb{X}_0^{T-1}}. \quad (16)$$

Definition 6 (Hankel Matrix). Given a matrix $\mathbb{X} = [\mathbf{x}_0, \dots, \mathbf{x}_{T-1}]$, the k th Hankel matrix of \mathbb{X} is defined for $0 \leq k \leq T - 1$ as

$$\mathcal{H}_{k|\mathbb{X}} = \begin{bmatrix} (\mathbb{X}_0^k)^\top & (\mathbb{X}_1^{k+1})^\top & \dots & (\mathbb{X}_{T-k-1}^{T-1})^\top \end{bmatrix},$$

where $\mathcal{H}_{k|\mathbb{X}} \in \mathbb{C}^{(k+1)m \times (T-k)}$.

Although $(a_0, \dots, a_{\ell-1})$ that satisfy Equation (16) may not exist, and even if they exist they may not be unique, the following least-squares optimality always holds:

$$- \mathcal{H}_{0|\mathbb{X}_\ell^{T-1}} \mathcal{H}_{\ell-1|\mathbb{X}_0^{T-1}}^+ \in \arg \min_{a_0, \dots, a_{\ell-1}} \left\| \mathcal{H}_{0|\mathbb{X}_\ell^{T-1}} + \begin{bmatrix} a_0 & \dots & a_{\ell-1} \end{bmatrix} \mathcal{H}_{\ell-1|\mathbb{X}_0^{T-1}} \right\|_F.$$

This represents the least-squares solution for the characteristic coefficients when Equation (16) is inconsistent.

5.2. Computing the Koopman Modes

Once the eigenvalues μ_1, \dots, μ_ℓ are computed, the linear equation (14) determines \mathbb{M} . Since

$$\mathbb{X} \mathcal{V}_{T|\mu_1, \dots, \mu_\ell}^+ \in \arg \min_{\mathbb{M}} \left\| \mathbb{X} - \mathbb{M} \mathcal{V}_{T|\mu_1, \dots, \mu_\ell} \right\|_F$$

holds, if a solution exists, $\mathbb{X} \mathcal{V}_{T|\mu_1, \dots, \mu_\ell}^+$ provides one possible solution, which is not necessarily unique.

Equation (15) can be viewed as a factorization of linear mappings:

$$\mathbb{X} : \mathbb{C}^T \rightarrow \mathbb{C}^m, \quad \mathcal{V}_{T|\mu_1, \dots, \mu_\ell} : \mathbb{C}^T \rightarrow \mathbb{C}^\ell, \quad \mathbb{M} : \mathbb{C}^\ell \rightarrow \mathbb{C}^m.$$

The condition for the existence of such an \mathbb{M} is

$$\ker \mathbb{X} \supseteq \ker \mathcal{V}_{T|\mu_1, \dots, \mu_\ell}.$$

Under this condition, the solution is unique if and only if $\mathcal{V}_{T|\mu_1, \dots, \mu_\ell}$ is surjective, which holds if and only if $\ell \leq T$ and the nodes μ_1, \dots, μ_ℓ are pairwise distinct.

Proposition 4. For an observable matrix \mathbb{X} and a Koopman degree ℓ with $\ell \leq T$, a DKMD of \mathbb{X} with Koopman eigenvalues μ_1, \dots, μ_ℓ , if it exists, it is unique.

6. The Contributions of This Article

Given an observable matrix, the vector Prony method introduced in the previous section computes a DKMD for a specified Koopman degree ℓ . However, a DKMD does not always exist for the given degree, and even when it does, it may not be unique.

Definition 7 (Feasible Degree). *A Koopman degree ℓ is said to be feasible for the observables if a DKMD of degree ℓ exists for the given observable matrix.*

Although the Koopman degree ℓ must satisfy $\ell < T$ (Proposition 2), a DKMD may exist for multiple values of ℓ . Thus, selecting the *optimal* feasible degree is essential to obtain an optimal DKMD. For this selection, we consider two independent principles:

Minimality: The optimal degree should be the smallest among all feasible degrees. This principle is analogous to Occam’s razor, favoring the simplest representation that adequately explains the observations.

Uniqueness: The optimal degree should correspond to a unique DKMD. If multiple DKMDs exist for a given ℓ , as described in a later section, the set of such decompositions forms a continuum, where different DKMDs yield distinct eigenvalues and modes. Consequently, any particular DKMD extracted from this continuum—such as one obtained by the vector Prony method—may fail to reproduce the true dynamics precisely.

In this article, we demonstrate that if a degree satisfies the uniqueness criterion, it also satisfies the minimality criterion, but the converse does not necessarily hold. Thus, between these two principles, we adopt the uniqueness criterion as the standard for selecting the optimal Koopman degree.

Definition 8 (Uniquely Feasible Degree). *A Koopman degree is said to be uniquely feasible if a DKMD of that degree exists and is unique for the given observables.*

In summary, the objective of this paper is to establish a theoretical framework for uniquely feasible degrees. Specifically, we demonstrate and present the following:

- A uniquely feasible degree for a given observable matrix, if it exists, is the smallest among all feasible degrees.
- Several structural properties of uniquely feasible degrees lead to computationally efficient algorithms for determining them.
- These algorithms are further extended to handle noisy observables via least-squares formulations.

7. Finding Uniquely Feasible Degrees

7.1. Key Indices: Hankel Dimension and Codimension

We first summarize the results of the previous sections in a theorem that provides a necessary and sufficient condition for an observable matrix to admit a DKMD.

For notational convenience, we first introduce the following definition.

Definition 9 (Square-free coefficient vector). *A vector $(a_0 \dots a_{n-1} \ 1) \in \mathbb{C}^{n+1}$ is called a square-free coefficient vector if the algebraic equation*

$$X^n + a_{n-1}X^{n-1} + \dots + a_0 = 0$$

has no repeated roots.

Theorem 1 (Feasibility condition). *Let \mathbb{X} be an observable matrix and $\ell < T$. The following statements are equivalent:*

1. \mathbb{X} admits a DKMD with Koopman degree ℓ ; equivalently, ℓ is feasible for \mathbb{X} .
2. There exists a square-free coefficient vector $(a_0 \dots a_{\ell-1} \ 1)$ satisfying

$$(a_0 \dots a_{\ell-1} \ 1) \mathcal{H}_{\ell|\mathbb{X}} = 0. \quad (17)$$

Proof. (1 \Rightarrow 2) Suppose \mathbb{X} admits a DKMD with Koopman eigenvalues μ_1, \dots, μ_ℓ . The coefficient vector $\mathbf{a} = \begin{pmatrix} a_0 & \dots & a_{\ell-1} & 1 \end{pmatrix}$ of the characteristic polynomial of this DKMD is square-free by the definition of DKMD (Definition 3), and the assertion of Proposition 3 can be restated as

$$\mathbf{a} \left(\mathbb{X}_i^{i+\ell} \right)^\top = \mathbf{0} \quad \text{for } i = 0, \dots, T - \ell - 1,$$

which further implies Equation (17).

(2 \Rightarrow 1) Let μ_1, \dots, μ_ℓ denote the distinct roots of the square-free polynomial $X^\ell + a_{\ell-1}X^{\ell-1} + \dots + a_0 = 0$. We consider $\mathbb{X}, \mathcal{V}_{T|\mu_1, \dots, \mu_\ell}$, and \mathbb{M} as representing linear mappings $F : \mathbb{C}^T \rightarrow \mathbb{C}^m, G : \mathbb{C}^T \rightarrow \mathbb{C}^\ell$, and $H : \mathbb{C}^\ell \rightarrow \mathbb{C}^m$, respectively. Then, the rank-nullity theorem implies

$$\ker G = \bigoplus_{i=0}^{T-\ell-1} \mathbb{C} \left(\mathbf{0}^i \quad \mathbf{a} \quad \mathbf{0}^{T-\ell-1-i} \right)^\top,$$

and $\ker F \supseteq \ker G$ follows from

$$\mathbb{X} \left(\mathbf{0}^i \quad \mathbf{a} \quad \mathbf{0}^{T-\ell-1-i} \right)^\top = \mathbb{X}_i^{i+\ell} \mathbf{a}^\top = \mathbf{0}$$

for all $i = 0, \dots, T - \ell - 1$, which implies the existence of H . \square

The square-free coefficient vector in Equation (17) lies in the orthogonal complement $\mathbb{V}(\mathcal{H}_{\ell|\mathbb{X}})^\perp$, which is the subspace of $\mathbb{C}^{\ell+1}$ consisting of all vectors orthogonal to every column of $\mathcal{H}_{\ell|\mathbb{X}}$. Theorem 1 thus shows that the existence of a DKMD depends on the structure of $\mathbb{V}(\mathcal{H}_{k|\mathbb{X}})$ and its orthogonal complement. This motivates the following key indices.

Definition 10 (Hankel dimension and codimension). *Given an observable matrix \mathbb{X} , the Hankel dimension and codimension of order k ($k < T$) are defined as*

$$\dim_{\mathcal{H}}(k | \mathbb{X}) = \text{rank } \mathcal{H}_{k|\mathbb{X}}, \quad \text{codim}_{\mathcal{H}}(k | \mathbb{X}) = k + 1 - \dim_{\mathcal{H}}(k | \mathbb{X}).$$

Using these indices, we can restate Theorem 1 as follows:

Corollary 1. *For an observable matrix \mathbb{X} and a Koopman degree $\ell < T$, a necessary condition for the existence of an ℓ -degree DKMD of \mathbb{X} is*

$$\text{codim}_{\mathcal{H}}(\ell | \mathbb{X}) \geq 1.$$

However, the condition $\text{codim}_{\mathcal{H}}(\ell | \mathbb{X}) \geq 1$ is not a sufficient condition due to the following two reasons:

- Even if we can find $\begin{pmatrix} a_0 & \dots & a_{\ell-1} & a_\ell \end{pmatrix}$ with

$$\begin{pmatrix} a_0 & \dots & a_{\ell-1} & a_\ell \end{pmatrix} \mathcal{H}_{\ell|\mathbb{X}} = \mathbf{0},$$

it may happen that $a_\ell = 0$, meaning that the polynomial corresponding to $\begin{pmatrix} a_0 & \dots & a_{\ell-1} & a_\ell \end{pmatrix}$ is of degree lower than ℓ , which cannot induce a DKMD of degree ℓ .

- Even if

$$\begin{pmatrix} a_0 & \dots & a_{\ell-1} & 1 \end{pmatrix} \mathcal{H}_{\ell|\mathbb{X}} = \mathbf{0},$$

the resulting polynomial

$$X^\ell + a_{\ell-1}X^{\ell-1} + \dots + a_0 = 0$$

may have repeated roots, rendering it unsuitable as a characteristic polynomial.

Our purpose here is to identify a Koopman degree ℓ that admits a unique DKMD. We can represent the condition for a uniquely feasible degree using the Hankel codimension, assuming that ℓ is feasible, that is, an ℓ -degree DKMD exists.

Corollary 2. For a feasible degree ℓ of an observable matrix \mathbb{X} with $\ell < T$, the following hold:

1. The DKMD is unique if and only if $\text{codim}_{\mathcal{H}}(\ell | \mathbb{X}) = 1$.
2. The set of ℓ -degree DKMDs forms a continuum if and only if $\text{codim}_{\mathcal{H}}(\ell | \mathbb{X}) > 1$.

Proof. Since ℓ is feasible, $\mathbb{V}(\mathcal{H}_{\ell|\mathbb{X}})^{\perp}$ is non-empty, and there exists a coefficient vector

$$\begin{pmatrix} a_0 & \dots & a_{\ell-1} & 1 \end{pmatrix}^{\top} \in \mathbb{V}(\mathcal{H}_{\ell|\mathbb{X}})^{\perp}$$

which determines the characteristic polynomial of a DKMD of \mathbb{X} .

If $\text{codim}_{\mathcal{H}}(\ell | \mathbb{X}) = 1$, that is, if $\dim \mathbb{V}(\mathcal{H}_{\ell|\mathbb{X}})^{\perp} = 1$, then only $\begin{pmatrix} a_0 & \dots & a_{\ell-1} & 1 \end{pmatrix}$ can determine a characteristic polynomial, implying that the DKMD is unique by Proposition 4.

On the other hand, if $\text{codim}_{\mathcal{H}}(\ell | \mathbb{X}) > 1$, there exists a coefficient vector

$$\begin{pmatrix} b_0 & \dots & b_{\ell-1} & 1 \end{pmatrix} \in \mathbb{V}(\mathcal{H}_{\ell|\mathbb{X}})^{\perp} \setminus \mathbb{C} \begin{pmatrix} a_0 & \dots & a_{\ell-1} & 1 \end{pmatrix}.$$

When we define

$$f_{\alpha}(X) = (1 - \alpha)f(X) + \alpha(X^{\ell} + b_{\ell-1}X^{\ell-1} + \dots + b_0),$$

since $f_0(X)$ is square-free, there exists $\epsilon > 0$ such that $f_{\alpha}(X)$ remains square-free for any $\alpha \in (-\epsilon, \epsilon)$. In particular, each distinct value of α yields a distinct DKMD. \square

7.2. The Koopman Dimension and Codimension for $m = 1$

In this section, we investigate the Hankel dimension and codimension in the restricted case where $m = 1$, i.e., when \mathbb{X} consists of a single row with T components. These results will be extended to the general case $m \geq 1$ in Section 7.4.6.

Lemma 1. Let \mathbb{X} be an observable matrix with one row and T columns. Furthermore, suppose that \mathbb{X} admits a DKMD of the form

$$\mathbb{X} = \begin{bmatrix} m_1 & \dots & m_{\ell} \end{bmatrix} \mathcal{V}_{T|\mu_1, \dots, \mu_{\ell}}$$

such that $\ell \leq \frac{T+1}{2}$ and each m_i is nonzero. Then, the leftmost $\ell \times \ell$ block submatrix of $\mathcal{H}_{\ell-1|\mathbb{X}}$ has nonzero determinant.

Proof. The leftmost $\ell \times \ell$ block submatrix is expressed as:

$$\begin{bmatrix} \left(\mathbb{X}_0^{\ell-1}\right)^{\top} & \left(\mathbb{X}_1^{\ell}\right)^{\top} & \dots & \left(\mathbb{X}_{\ell-1}^{2\ell-2}\right)^{\top} \end{bmatrix} = \left(\mathcal{V}_{\ell|\mu_1, \dots, \mu_{\ell}}\right)^{\top} \mathbf{diag}(m_1, \dots, m_{\ell}) \mathcal{V}_{\ell|\mu_1, \dots, \mu_{\ell}}.$$

The determinant of this matrix is nonzero, because μ_1, \dots, μ_{ℓ} are mutually distinct, implying $\det \mathcal{V}_{\ell|\mu_1, \dots, \mu_{\ell}} \neq 0$, and m_1, \dots, m_{ℓ} are nonzero, implying $\det \mathbf{diag}(m_1, \dots, m_{\ell}) \neq 0$. \square

Theorem 2. Let \mathbb{X} , m_1, \dots, m_ℓ , and μ_1, \dots, μ_ℓ be as in Lemma 1. Assume further that $\ell \leq \frac{T+1}{2}$. Then, the Hankel dimension and codimension are given by:

$$\dim_{\mathcal{H}}(k | \mathbb{X}) = \begin{cases} k+1, & \text{for } k \in \{0, \dots, \ell-1\}; \\ \ell, & \text{for } k \in \{\ell, \dots, T-\ell\}; \\ T-k, & \text{for } k \in \{T-\ell+1, \dots, T-1\}; \end{cases}$$

$$\text{codim}_{\mathcal{H}}(k | \mathbb{X}) = \begin{cases} 0, & \text{for } k \in \{0, \dots, \ell-1\}; \\ k+1-\ell, & \text{for } k \in \{\ell, \dots, T-\ell\}; \\ 2k-T+1, & \text{for } k \in \{T-\ell+1, \dots, T-1\}. \end{cases}$$

Proof. Note that $\mathcal{H}_{k|\mathbb{X}}$ consists of $k+1$ rows and $T-k$ columns.

If $0 \leq k \leq \ell-1$, equivalently if $k+1 \leq \ell$ and $T-k \geq \ell$, the leftmost $(k+1) \times \ell$ submatrix of $\mathcal{H}_{k|\mathbb{X}}$ coincides with the top $(k+1) \times \ell$ submatrix of $\mathcal{H}_{\ell-1|\mathbb{X}}$, implying its rank is $k+1$ by Lemma 1. Thus, all $k+1$ rows of $\mathcal{H}_{k|\mathbb{X}}$ are linearly independent, and $\dim_{\mathcal{H}}(k | \mathbb{X}) = k+1$ follows.

If $T-\ell+1 \leq k \leq T-1$, equivalently if $k+1 > \ell$ and $T-k < \ell$, the entire $\mathcal{H}_{k|\mathbb{X}}$ has $T-k$ columns. Its top $\ell \times (T-k)$ submatrix coincides with the leftmost $\ell \times (T-k)$ submatrix of $\mathcal{H}_{\ell-1|\mathbb{X}}$, implying its rank is $T-k$ by Lemma 1. Thus, all $T-k$ columns of $\mathcal{H}_{k|\mathbb{X}}$ are linearly independent, and $\dim_{\mathcal{H}}(k | \mathbb{X}) = T-k$ follows.

If $\ell \leq k \leq T-\ell$, equivalently if $k+1 > \ell$ and $T-k \geq \ell$, the top-left $\ell \times \ell$ submatrix of $\mathcal{H}_{k|\mathbb{X}}$ coincides with that of $\mathcal{H}_{\ell-1|\mathbb{X}}$, and hence, Lemma 1 implies that the leftmost ℓ columns of $\mathcal{H}_{k|\mathbb{X}}$ are linearly independent.

On the other hand, we let

$$\prod_{i=1}^{\ell} (X - \mu_i) = X^{\ell} - a_{\ell-1}X^{\ell-1} - \dots - a_0$$

denote the characteristic polynomial. Then each eigenvalue μ_i satisfies $\mu_i^{\ell} = a_{\ell-1}\mu_i^{\ell-1} + \dots + a_0$. For the i -th element x_i of \mathbb{X} with $i \geq \ell$,

$$\begin{aligned} x_i &= \sum_{j=1}^{\ell} m_j \mu_j^i = \sum_{j=1}^{\ell} m_j \mu_j^{i-\ell} \cdot \mu_j^{\ell} = \sum_{j=1}^{\ell} m_j \mu_j^{i-\ell} \left(\sum_{n=0}^{\ell-1} a_n \mu_j^n \right) \\ &= \sum_{n=0}^{\ell-1} a_n \left(\sum_{j=1}^{\ell} m_j \mu_j^{i-\ell+n} \right) = \sum_{n=0}^{\ell-1} a_n x_{i-\ell+n}. \end{aligned}$$

This recurrence relation shows that for any $i \geq \ell$, the element x_i is determined by $x_{i-\ell}, \dots, x_{i-1}$. Consequently, any column of $\mathcal{H}_{k|\mathbb{X}}$ is a linear combination of the leftmost ℓ columns, which have been proven to be linearly independent. Therefore, $\dim_{\mathcal{H}}(k | \mathbb{X}) = \ell$.

The claims about the Hankel codimension follow directly from the definition $\text{codim}_{\mathcal{H}}(k | \mathbb{X}) = k+1 - \dim_{\mathcal{H}}(k | \mathbb{X})$. \square

Corollary 3. If $\ell \leq \frac{T}{2}$, then $k = \ell$ is the unique value satisfying $\text{codim}_{\mathcal{H}}(k | \mathbb{X}) = 1$.

If $\ell = \frac{T+1}{2}$, then \mathbb{X} admits no uniquely feasible degree.

7.3. Examples

In this section, we introduce three examples of an observable matrix \mathbb{X} , each demonstrating different properties regarding the existence of a uniquely feasible degree:

- No Koopman degree ℓ satisfies $\text{codim}_{\mathcal{H}}(\ell | \mathbb{X}) = 1$, meaning that no uniquely feasible degree exists (Example 2).

- A Koopman degree ℓ with $\text{codim}_{\mathcal{H}}(\ell | \mathbb{X}) = 1$ exists, but the corresponding characteristic polynomial is not square-free. As a result, a uniquely feasible degree does not exist (Example 3).
- A uniquely feasible degree exists, ensuring that a DKMD is uniquely determined for the degree (Example 4).

These examples illustrate the conditions under which a DKMD is uniquely determined and the role of Hankel codimension in establishing uniqueness.

Example 2. Consider the observable matrix

$$\mathbb{X} = \begin{bmatrix} 1 & 1 & 1 & 1 & 3 & 5 & 7 \end{bmatrix}.$$

The Hankel matrices for $\ell = 1, 2, 3$ are computed as:

$$\mathcal{H}_{1|\mathbb{X}} = \begin{bmatrix} 1 & 1 & 1 & 1 & 3 & 5 \\ 1 & 1 & 1 & 3 & 5 & 7 \end{bmatrix}, \quad \mathcal{H}_{2|\mathbb{X}} = \begin{bmatrix} 1 & 1 & 1 & 1 & 3 \\ 1 & 1 & 1 & 3 & 5 \\ 1 & 1 & 3 & 5 & 7 \end{bmatrix}, \quad \text{and} \quad \mathcal{H}_{3|\mathbb{X}} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 3 \\ 1 & 1 & 3 & 5 \\ 1 & 3 & 5 & 7 \end{bmatrix}.$$

These computations yield:

$$\text{codim}_{\mathcal{H}}(1 | \mathbb{X}) = \text{codim}_{\mathcal{H}}(2 | \mathbb{X}) = \text{codim}_{\mathcal{H}}(3 | \mathbb{X}) = 0,$$

implying that \mathbb{X} admits no DKMD for these degrees by Corollary 1.

On the other hand, for $\ell = 4, 5, 6$, we have

$$\mathcal{H}_{4|\mathbb{X}} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 3 \\ 1 & 3 & 5 \\ 3 & 5 & 7 \end{bmatrix}, \quad \mathcal{H}_{5|\mathbb{X}} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 3 \\ 3 & 5 \\ 5 & 7 \end{bmatrix}, \quad \text{and} \quad \mathcal{H}_{6|\mathbb{X}} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 3 \\ 5 \\ 7 \end{bmatrix}.$$

It follows that

$$\text{codim}_{\mathcal{H}}(4 | \mathbb{X}) = 2, \quad \text{codim}_{\mathcal{H}}(5 | \mathbb{X}) = 4, \quad \text{and} \quad \text{codim}_{\mathcal{H}}(6 | \mathbb{X}) = 6,$$

implying that DKMDs form a continuum by Corollary 2.

For example, for $\ell = 4$,

$$\mathbb{V}(\mathcal{H}_{4|\mathbb{X}})^{\perp} = \mathbb{C} \begin{pmatrix} -1 & -1 & 0 & -1 & 1 \end{pmatrix}^{\top} \oplus \mathbb{C} \begin{pmatrix} -1 & 1 & 0 & 0 & 0 \end{pmatrix}^{\top}.$$

Thus, the possible candidates for characteristic polynomials take the form:

$$f_a(x) = x^4 - x^3 - x - 1 + a(x - 1), \quad (18)$$

whose discriminant is computed as:

$$D = -27a^4 + 54a^3 - 783a^2 - 900a - 500.$$

Since $f_a(X) = 0$ has no repeated roots if and only if $D \neq 0$, it follows that 4-degree DKMDs form a continuum.

Example 3. When we let $\mathbb{X} = \begin{bmatrix} 1 & 1 & -3 & 5 & -7 \end{bmatrix}$, the Hankel matrices for $\ell = 1, 2, 3$ are computed as:

$$\mathcal{H}_{1|\mathbb{X}} = \begin{bmatrix} 1 & 1 & -3 & 5 \\ 1 & -3 & 5 & -7 \end{bmatrix}, \quad \mathcal{H}_{2|\mathbb{X}} = \begin{bmatrix} 1 & 1 & -3 \\ 1 & -3 & 5 \\ -3 & 5 & -7 \end{bmatrix}, \quad \text{and} \quad \mathcal{H}_{3|\mathbb{X}} = \begin{bmatrix} 1 & 1 \\ 1 & -3 \\ -3 & 5 \\ 5 & -7 \end{bmatrix},$$

implying

$$\text{codim}_{\mathcal{H}}(1 | \mathbb{X}) = 0, \quad \text{codim}_{\mathcal{H}}(2 | \mathbb{X}) = 1, \quad \text{and} \quad \text{codim}_{\mathcal{H}}(3 | \mathbb{X}) = 2.$$

In fact, we have $\mathbb{V}(\mathcal{H}_{1|\mathbb{X}})^\perp = (0)$, and furthermore,

$$\mathbb{V}(\mathcal{H}_{2|\mathbb{X}})^\perp = \mathbb{C}(1 \ 2 \ 1)^\top \quad \text{and} \quad \mathbb{V}(\mathcal{H}_{3|\mathbb{X}})^\perp = \mathbb{C}(1 \ 2 \ 1 \ 0)^\top \oplus \mathbb{C}(0 \ 1 \ 2 \ 1)^\top.$$

This implies that there exists no DKMD for any of these degrees because:

1. For $\ell = 1$, no characteristic polynomial exists.
2. For $\ell = 2$, if a DKMD existed, the corresponding characteristic polynomial would be

$$X^2 + 2X + 1 = (X + 1)^2,$$

which is not square-free.

3. For $\ell = 3$, if a DKMD existed, the corresponding characteristic polynomial would be of the form

$$X^3 + 2X^2 + X + a(X^2 + 2X + 1) = (X + a)(X + 1)^2,$$

which is also not square-free.

On the other hand, for $\ell = 4$, we have $\mathcal{H}_{4|\mathbb{X}} = \mathbb{X}^\top$ and

$$\begin{aligned} \mathbb{V}(\mathcal{H}_{4|\mathbb{X}})^\perp &= \mathbb{C}(7 \ 0 \ 0 \ 0 \ 1)^\top \oplus \mathbb{C}(1 \ 2 \ 1 \ 0 \ 0)^\top \\ &\oplus \mathbb{C}(0 \ 1 \ 2 \ 1 \ 0)^\top \oplus \mathbb{C}(0 \ 0 \ 1 \ 2 \ 1)^\top, \end{aligned}$$

implying that there exists a DKMD whose characteristic polynomial is

$$X^4 + 7 = \prod_{k=0}^3 \left(X - \sqrt[4]{7} e^{\frac{i\pi}{4}(2k+1)} \right).$$

Furthermore, all possible characteristic polynomials of DKMDs for these observables are of the form

$$f_{a,b,c}(X) = a(X^4 + 7) + (1 - a)(X^4 + 2X^3 + X^2) + b(X^3 + 2X^2 + X) + c(X^2 + 2X + 1).$$

Since $f_{a,b,c}(X)$ is square-free for $(a, b, c) = (1, 0, 0)$, and since the set of $(a, b, c) \in \mathbb{C}^3$ for which $f_{a,b,c}(X)$ is square-free is an open subset of \mathbb{C}^3 , we conclude that the set of possible four-degree DKMDs for \mathbb{X} forms a continuum.

Example 4. We expand \mathbb{X} of Example 2 by adding one more dimension to the observables. Let the observable matrix \mathbb{X} be given by

$$\mathbb{X} = \begin{bmatrix} 1 & 1 & 2 & 3 & 5 & 8 & 13 \\ 1 & 1 & 1 & 1 & 3 & 5 & 7 \end{bmatrix}.$$

For $\ell = 1, 2, 3$, the corresponding Hankel matrices are computed as:

$$\mathcal{H}_{1|\mathbb{X}} = \begin{bmatrix} 1 & 1 & 1 & 1 & 2 & 1 & 3 & 1 & 5 & 3 & 8 & 5 \\ 1 & 1 & 2 & 1 & 3 & 1 & 5 & 3 & 8 & 5 & 13 & 7 \end{bmatrix},$$

$$\mathcal{H}_{2|\mathbb{X}} = \begin{bmatrix} 1 & 1 & 1 & 1 & 2 & 1 & 3 & 1 & 5 & 3 \\ 1 & 1 & 2 & 1 & 3 & 1 & 5 & 3 & 8 & 5 \\ 2 & 1 & 3 & 1 & 5 & 3 & 8 & 5 & 13 & 7 \end{bmatrix}, \quad \text{and} \quad \mathcal{H}_{3|\mathbb{X}} = \begin{bmatrix} 1 & 1 & 1 & 1 & 2 & 1 & 3 & 1 \\ 1 & 1 & 2 & 1 & 3 & 1 & 5 & 3 \\ 2 & 1 & 3 & 1 & 5 & 3 & 8 & 5 \\ 3 & 1 & 5 & 3 & 8 & 5 & 13 & 7 \end{bmatrix},$$

implying

$$\text{codim}_{\mathcal{H}}(1 | \mathbb{X}) = \text{codim}_{\mathcal{H}}(2 | \mathbb{X}) = \text{codim}_{\mathcal{H}}(3 | \mathbb{X}) = 0.$$

On the other hand,

$$\text{codim}_{\mathcal{H}}(4 | \mathbb{X}) = 1, \quad \text{codim}_{\mathcal{H}}(5 | \mathbb{X}) = 2, \quad \text{and} \quad \text{codim}_{\mathcal{H}}(6 | \mathbb{X}) = 5$$

follows from

$$\mathcal{H}_{4|\mathbb{X}} = \begin{bmatrix} 1 & 1 & 1 & 1 & 2 & 1 \\ 1 & 1 & 2 & 1 & 3 & 1 \\ 2 & 1 & 3 & 1 & 5 & 3 \\ 3 & 1 & 5 & 3 & 8 & 5 \\ 5 & 3 & 8 & 5 & 13 & 7 \end{bmatrix}, \quad \mathcal{H}_{5|\mathbb{X}} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 2 & 1 & 3 & 1 \\ 3 & 1 & 5 & 3 \\ 5 & 3 & 8 & 5 \\ 8 & 5 & 13 & 7 \end{bmatrix}, \quad \text{and} \quad \mathcal{H}_{6|\mathbb{X}} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 5 & 3 \\ 8 & 5 \\ 13 & 7 \end{bmatrix}.$$

Therefore, if \mathbb{X} admits a uniquely feasible degree, it must be four. In fact,

$$\mathbb{V}(\mathcal{H}_{4|\mathbb{X}})^{\perp} = \mathbb{C}(-1 \quad -1 \quad 0 \quad -1 \quad 1)^{\top}$$

holds, and the corresponding characteristic polynomial is

$$f(x) = x^4 - x^3 - x - 1,$$

which is square-free, implying that $\ell = 4$ is uniquely feasible.

7.4. Important Properties of The Hankel Dimension and Codimension

In this subsection, we introduce important properties of the Hankel dimension and Hankel codimension, which play a crucial role in designing algorithms to determine uniquely feasible degrees, including the following. For the convenience of description, we denote the minimum feasible degree as:

$$L = \min\{\ell \in [0, T) \mid \text{codim}_{\mathcal{H}}(\ell | \mathbb{X}) > 0\}. \quad (19)$$

Best Possible Upper Bound of a Uniquely Feasible Degree.

Let $r = \text{rank } \mathbb{X}$. If $\ell > \frac{rT}{r+1}$, then $\text{codim}_{\mathcal{H}}(\ell | \mathbb{X}) > 1$. Hence, no uniquely feasible degree can exceed $\frac{rT}{r+1}$, which is also the sharpest possible upper bound.

Monotonic Increase of the Hankel Codimension.

The Hankel codimension $\text{codim}_{\mathcal{H}}(\ell | \mathbb{X})$ is strictly increasing with respect to ℓ over the interval $[L, T)$.

Equivalence Between Unique and Minimal Feasibility.

The monotonicity of the codimension implies that, if ℓ is uniquely feasible, then $\ell = L$. In particular, if $\text{codim}_{\mathcal{H}}(L | \mathbb{X}) > 1$, no uniquely feasible degree exists.

Saturation of the Hankel Dimension.

If $L \leq \frac{T}{2}$ and $\ell \in [L, T - L]$, then

$$\dim_{\mathcal{H}}(\ell | \mathbb{X}) = L, \quad \text{codim}_{\mathcal{H}}(\ell | \mathbb{X}) = \ell + 1 - L.$$

In particular, $\text{codim}_{\mathcal{H}}(L | \mathbb{X}) = 1$ holds, which implies L is the only candidate for a uniquely feasible degree.

7.4.1. Invariance under Basis Transformations

For $\mathbb{X} \in \mathbb{C}^{m \times T}$, $\mathbb{Y} \in \mathbb{C}^{n \times T}$, and $A \in \mathbb{C}^{n \times m}$, we assume

$$\mathbb{Y} = A\mathbb{X}.$$

Their corresponding Hankel matrices are computed as

$$\mathcal{H}_{k|\mathbb{Y}} = \mathcal{H}_{k|\mathbb{X}} \mathbf{diag}(\underbrace{A^T, \dots, A^T}_{T-k}),$$

where $\mathbf{diag}(\underbrace{A^T, \dots, A^T}_{T-k})$ is defined as the $m(T-k) \times n(T-k)$ matrix given by

$$\mathbf{diag}(\underbrace{A^T, \dots, A^T}_{T-k}) = \begin{bmatrix} A^T & & 0 \\ & \ddots & \\ 0 & & A^T \end{bmatrix}.$$

This implies, in particular,

$$\dim_{\mathcal{H}}(k | \mathbb{Y}) \leq \dim_{\mathcal{H}}(k | \mathbb{X}) \quad \text{and} \quad \text{codim}_{\mathcal{H}}(k | \mathbb{Y}) \geq \text{codim}_{\mathcal{H}}(k | \mathbb{X}).$$

Furthermore, if a square-free polynomial $f(x) = x^k + a_{k-1}x^{k-1} + \dots + a_0$ is a characteristic polynomial of a DKMD for \mathbb{X} , it is also a characteristic polynomial of a DKMD for \mathbb{Y} . In fact, we have

$$\begin{pmatrix} a_0 & \dots & a_{k-1} & 1 \end{pmatrix} \mathcal{H}_{k|\mathbb{Y}} = \begin{pmatrix} a_0 & \dots & a_{k-1} & 1 \end{pmatrix} \mathcal{H}_{k|\mathbb{X}} \mathbf{diag}(\underbrace{A^T, \dots, A^T}_{T-k}) = \mathbf{0}.$$

Furthermore, if in addition there exists $B \in \mathbb{C}^{m \times n}$ such that $\mathbb{X} = B\mathbb{Y}$, then

$$\dim_{\mathcal{H}}(k | \mathbb{Y}) = \dim_{\mathcal{H}}(k | \mathbb{X}) \quad \text{and} \quad \text{codim}_{\mathcal{H}}(k | \mathbb{Y}) = \text{codim}_{\mathcal{H}}(k | \mathbb{X})$$

holds, and the correspondence of characteristic polynomials is bijective.

The condition for $\mathbb{Y} = A\mathbb{X}$ is that the row space of \mathbb{Y} is contained in the row space of \mathbb{X} , and symmetrically, the condition for $\mathbb{X} = B\mathbb{Y}$ is that the row space of \mathbb{Y} contains the row space of \mathbb{X} . Hence, both $\mathbb{Y} = A\mathbb{X}$ and $\mathbb{X} = B\mathbb{Y}$ hold if and only if the row space of \mathbb{Y} and the row space of \mathbb{X} are identical. In other words, under the hypothesis $\mathbb{Y} = A\mathbb{X}$, the necessary and sufficient condition for the existence of $B \in \mathbb{C}^{m \times n}$ such that $\mathbb{X} = B\mathbb{Y}$ is $\text{rank } \mathbb{Y} = \text{rank } \mathbb{X}$.

If $\mathbb{Y} = A\mathbb{X}$ and $\text{rank } \mathbb{Y} = \text{rank } \mathbb{X}$ hold, the minimum $\min_B \|\mathbb{X} - B\mathbb{Y}\|_F$ is zero, and hence, $\|\mathbb{X} - \mathbb{X}\mathbb{Y}^+\mathbb{Y}\|_F = 0$ holds, implying

$$\mathbb{X} = \mathbb{X}\mathbb{Y}^+\mathbb{Y}.$$

Furthermore, this bijective correspondence between characteristic polynomials yields a bijective correspondence between DKMDs. In fact, $\mathbb{X} = \mathbb{M}\mathcal{V}_{T|\mu_1, \dots, \mu_k}$, which is a DKMD of \mathbb{X} , yields

$$\mathbb{Y} = (\mathbb{A}\mathbb{M})\mathcal{V}_{T|\mu_1, \dots, \mu_k},$$

which is a DKMD of \mathbb{Y} . In reverse, a DKMD of \mathbb{Y} $\mathbb{Y} = \mathbb{M}\mathcal{V}_{T|\mu_1, \dots, \mu_k}$ corresponds to a DKMD of \mathbb{X}

$$\mathbb{X} = (\mathbb{X}\mathbb{Y}^+\mathbb{M})\mathcal{V}_{T|\mu_1, \dots, \mu_k}.$$

The latter correspondence is the inverse of the former by Proposition 4.

Thus, we have:

Theorem 3. *If $\text{rank } \mathbb{Y} = \text{rank } \mathbb{X}$ holds for $\mathbb{Y} = \mathbb{A}\mathbb{X}$, then the following statements hold:*

1. $\dim_{\mathcal{H}}(k | \mathbb{X}) = \dim_{\mathcal{H}}(k | \mathbb{Y})$ for each $k \in \{0, 1, \dots, T-1\}$.
2. $\text{codim}_{\mathcal{H}}(k | \mathbb{X}) = \text{codim}_{\mathcal{H}}(k | \mathbb{Y})$ for each $k \in \{0, 1, \dots, T-1\}$.
3. The set of characteristic polynomials for \mathbb{X} is identical with that for \mathbb{Y} .
4. A DKMD $\mathbb{X} = \mathbb{M}\mathcal{V}_{T|\mu_1, \dots, \mu_k}$ of \mathbb{X} can be converted to a DKMD $\mathbb{Y} = (\mathbb{A}\mathbb{M})\mathcal{V}_{T|\mu_1, \dots, \mu_k}$ of \mathbb{Y} , while a DKMD $\mathbb{Y} = \mathbb{M}\mathcal{V}_{T|\mu_1, \dots, \mu_k}$ of \mathbb{Y} can be converted to a DKMD $\mathbb{X} = (\mathbb{X}\mathbb{Y}^+\mathbb{M})\mathcal{V}_{T|\mu_1, \dots, \mu_k}$ of \mathbb{X} . These conversions yield a bijective correspondence between the set of DKMDs for \mathbb{X} and that for \mathbb{Y} .

As an application of Theorem 4, the following two cases are particularly important:

- When \mathbb{A} is a nonsingular $m \times m$ matrix, we have $\text{rank } \mathbb{Y} = \text{rank } \mathbb{X}$ automatically, and thus Theorem 4 provides the invariance of the Hankel dimension, Hankel codimension, characteristic polynomials, and DKMDs under basis transformations in \mathbb{C}^m .
- When $\mathbb{A} \in \mathbb{C}^{r \times m}$ with $r = \text{rank } \mathbb{X}$ is selected to satisfy $\text{rank } (\mathbb{A}\mathbb{X}) = \text{rank } \mathbb{X}$, the matrix $\mathbb{Y} = \mathbb{A}\mathbb{X}$ has fewer rows than \mathbb{X} , and thus \mathbb{Y} requires less computation to obtain DKMDs than \mathbb{X} .

7.4.2. Invariance under Basis Transformations

For $\mathbb{X} \in \mathbb{C}^{m \times T}$, $\mathbb{Y} \in \mathbb{C}^{n \times T}$, and $\mathbb{A} \in \mathbb{C}^{n \times m}$, we assume

$$\mathbb{Y} = \mathbb{A}\mathbb{X}.$$

Their corresponding Hankel matrices are computed as

$$\mathcal{H}_{k|\mathbb{Y}} = \mathcal{H}_{k|\mathbb{X}} \mathbf{diag}(\underbrace{\mathbb{A}^T, \dots, \mathbb{A}^T}_{T-k}),$$

where $\mathbf{diag}(\underbrace{\mathbb{A}^T, \dots, \mathbb{A}^T}_{T-k})$ is defined as the $m(T-k) \times n(T-k)$ matrix given by

$$\mathbf{diag}(\underbrace{\mathbb{A}^T, \dots, \mathbb{A}^T}_{T-k}) = \begin{bmatrix} \mathbb{A}^T & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbb{A}^T \end{bmatrix}.$$

This implies, in particular,

$$\dim_{\mathcal{H}}(k | \mathbb{Y}) \leq \dim_{\mathcal{H}}(k | \mathbb{X}) \quad \text{and} \quad \text{codim}_{\mathcal{H}}(k | \mathbb{Y}) \geq \text{codim}_{\mathcal{H}}(k | \mathbb{X}).$$

Furthermore, if a square-free polynomial $f(x) = x^k + a_{k-1}x^{k-1} + \dots + a_0$ is a characteristic polynomial of a DKMD for \mathbb{X} , it is also a characteristic polynomial of a DKMD for \mathbb{Y} . In fact, we have

$$\begin{pmatrix} a_0 & \dots & a_{k-1} & 1 \end{pmatrix} \mathcal{H}_{k|\mathbb{Y}} = \begin{pmatrix} a_0 & \dots & a_{k-1} & 1 \end{pmatrix} \mathcal{H}_{k|\mathbb{X}} \underbrace{\text{diag}(\underbrace{A^\top, \dots, A^\top}_{T-k})}_{T-k} = \mathbf{0}.$$

Furthermore, if there exists $B \in \mathbb{C}^{m \times n}$ such that $\mathbb{X} = B\mathbb{Y}$,

$$\dim_{\mathcal{H}}(k | \mathbb{Y}) = \dim_{\mathcal{H}}(k | \mathbb{X}) \quad \text{and} \quad \text{codim}_{\mathcal{H}}(k | \mathbb{Y}) = \text{codim}_{\mathcal{H}}(k | \mathbb{X}),$$

holds, and the correspondence of characteristic polynomials is bijective.

The condition for $\mathbb{Y} = A\mathbb{X}$ is that the row space of \mathbb{Y} is contained in the row space of \mathbb{X} , and symmetrically, the condition for $\mathbb{X} = B\mathbb{Y}$ is that the row space of \mathbb{Y} contains the row space of \mathbb{X} . Hence, both $\mathbb{Y} = A\mathbb{X}$ and $\mathbb{X} = B\mathbb{Y}$ hold, if and only if the row space of \mathbb{Y} and the row space of \mathbb{X} are identical. In other words, under the hypothesis $\mathbb{Y} = A\mathbb{X}$, the necessary and sufficient condition for the existence of $B \in \mathbb{C}^{m \times n}$ such that $\mathbb{X} = B\mathbb{Y}$ is $\text{rank } \mathbb{Y} = \text{rank } \mathbb{X}$.

If $\mathbb{Y} = A\mathbb{X}$ and $\text{rank } \mathbb{Y} = \text{rank } \mathbb{X}$ hold, the minimum $\min_B \|\mathbb{X} - B\mathbb{Y}\|_F$ is zero, and hence, $\|\mathbb{X} - \mathbb{X}\mathbb{Y}^+\mathbb{Y}\|_F = 0$ holds, implying

$$\mathbb{X} = \mathbb{X}\mathbb{Y}^+\mathbb{Y}.$$

Furthermore, the aforementioned bijective correspondence between characteristic polynomials yields a bijective correspondence between DKMDs. In fact, $\mathbb{X} = \mathbb{M}\mathcal{V}_{T|\mu_1, \dots, \mu_k}$, which is a DKMD of \mathbb{X} , yields

$$\mathbb{Y} = (A\mathbb{M})\mathcal{V}_{T|\mu_1, \dots, \mu_k},$$

which is a DKMD of \mathbb{Y} . In reverse, a DKMD of \mathbb{Y} $\mathbb{Y} = \mathbb{M}\mathcal{V}_{T|\mu_1, \dots, \mu_k}$, corresponds to a DKMD of \mathbb{X}

$$\mathbb{X} = (\mathbb{X}\mathbb{Y}^+\mathbb{M})\mathcal{V}_{T|\mu_1, \dots, \mu_k}.$$

The latter correspondence is the inverse of the former by Proposition 4.

Thus, we have:

Theorem 4. *If $\text{rank } \mathbb{Y} = \text{rank } \mathbb{X}$ holds for $\mathbb{Y} = A\mathbb{X}$, then the following statements hold:*

1. $\dim_{\mathcal{H}}(k | \mathbb{X}) = \dim_{\mathcal{H}}(k | \mathbb{Y})$ for each $k \in \{0, 1, \dots, T-1\}$.
2. $\text{codim}_{\mathcal{H}}(k | \mathbb{X}) = \text{codim}_{\mathcal{H}}(k | \mathbb{Y})$ for each $k \in \{0, 1, \dots, T-1\}$.
3. The set of characteristic polynomials for \mathbb{X} is identical with that for \mathbb{Y} .
4. A DKMD $\mathbb{X} = \mathbb{M}\mathcal{V}_{T|\mu_1, \dots, \mu_k}$ of \mathbb{X} can be converted to a DKMD $\mathbb{Y} = (A\mathbb{M})\mathcal{V}_{T|\mu_1, \dots, \mu_k}$ of \mathbb{Y} , while a DKMD $\mathbb{Y} = \mathbb{M}\mathcal{V}_{T|\mu_1, \dots, \mu_k}$ of \mathbb{Y} can be converted to a DKMD $\mathbb{X} = (\mathbb{X}\mathbb{Y}^+\mathbb{M})\mathcal{V}_{T|\mu_1, \dots, \mu_k}$ of \mathbb{X} . These conversions yields a bijective correspondence between the set of DKMDs for \mathbb{X} and that for \mathbb{Y} .

As application of Theorem 4, the following two cases are particularly important.

- When A is a nonsingular $m \times m$ matrix, we have $\text{rank } \mathbb{Y} = \text{rank } \mathbb{X}$ automatically, and thus Theorem 4 provides the invariance of the Hankel dimension, Hankel codimension, characteristic polynomials and DKMDs under basis transformation in \mathbb{C}^m .
- If $A \in \mathbb{C}^{r \times m}$ can be selected to satisfy $\text{rank } \mathbb{X} = \text{rank } A\mathbb{X}$. since $\mathbb{Y} = A\mathbb{X}$ has fewer rows than \mathbb{X} , \mathbb{Y} requires less computation to obtain DKMDs than \mathbb{X} .

7.4.3. The Best Possible Upper Bound for A Uniquely Feasible Degree

Although we have seen that a uniquely feasible degree is always less than T , we can determine the best possible upper bound. First, we see:

Proposition 5. For an observable matrix \mathbb{X} with T columns, we have $r = \text{rank } \mathbb{X} < T$ if \mathbb{X} admits a uniquely feasible degree.

Proof. If ℓ is a uniquely feasible degree, Proposition 2 implies $\ell < T$. On the other hand, $\mathbb{X} = \mathbb{M}\mathcal{V}_{T|\mu_1, \dots, \mu_\ell}$ implies $r \leq \ell$. The assertion follows. \square

Based on this Proposition, we now establish the best possible upper bound for a Koopman degree that admits a unique DKMD.

Theorem 5. If $\text{codim}_{\mathcal{H}}(\ell | \mathbb{X}) = 1$, then $\ell \leq \frac{rT}{r+1}$. Furthermore, $\lfloor \frac{rT}{r+1} \rfloor$ is the best possible upper bound for a uniquely feasible degree.

Proof. When we take $\mathbb{Y} = \mathbb{A}\mathbb{X} \in \mathbb{C}^{r \times T}$ with $\text{rank } \mathbb{Y} = r$, Theorem 4 implies

$$\dim_{\mathcal{H}}(\ell | \mathbb{X}) = \dim_{\mathcal{H}}(\ell | \mathbb{Y}) = \text{rank } \mathcal{H}_{\ell|\mathbb{Y}} \leq r(T - \ell),$$

equivalently,

$$\text{codim}_{\mathcal{H}}(\ell | \mathbb{X}) \geq \ell + 1 - r(T - \ell).$$

Since $\text{codim}_{\mathcal{H}}(\ell | \mathbb{X}) = 1$ by hypothesis, $\ell + 1 - r(T - \ell) \leq 1$, which yields

$$\ell \leq \frac{rT}{r+1}.$$

To show that $\lfloor \frac{rT}{r+1} \rfloor$ is the best possible upper bound for a uniquely feasible degree, we construct an $r \times T$ observable matrix \mathbb{X} with $\text{rank } \mathbb{X} = r$ and $\dim_{\mathcal{H}}(\ell | \mathbb{X}) = \ell$.

First, we determine a sufficiently long series \mathbb{Y} with mutually distinct eigenvalues (μ_1, \dots, μ_ℓ) and nonzero modes (m_1, \dots, m_ℓ) :

$$\mathbb{Y} = \begin{bmatrix} x_0 & \dots & x_{\hat{T}-1} \end{bmatrix} = \begin{bmatrix} m_1 & \dots & m_\ell \end{bmatrix} \mathcal{V}_{\hat{T}|\mu_1, \dots, \mu_\ell},$$

where \hat{T} is chosen sufficiently large so that we can construct an observable matrix

$$\mathbb{X} = \begin{bmatrix} x_{a_1} & \dots & x_{a_1+T-1} \\ x_{a_2} & \dots & x_{a_2+T-1} \\ \vdots & \ddots & \vdots \\ x_{a_r} & \dots & x_{a_r+T-1} \end{bmatrix}.$$

Without loss of generality, we let $0 = a_1 < a_2 < \dots < a_r$ and further suppose $a_{i+1} - a_i \leq T - \ell$ for $i = 1, \dots, r-1$.

We now verify that the columns of $\mathcal{H}_{\ell|\mathbb{X}}$ correspond exactly to the first $a_r + T - \ell$ columns of $\mathcal{H}_{\ell|\mathbb{Y}}$. The i -th row of \mathbb{X} contributes $T - \ell$ consecutive columns to $\mathcal{H}_{\ell|\mathbb{X}}$, where the last such column is $(x_{a_i+T-\ell-1} \dots x_{a_i+T-1})^T$. On the other hand, the first column contributed by the $(i+1)$ -th row is $(x_{a_{i+1}} \dots x_{a_{i+1}+\ell})^T$. Since $a_{i+1} - a_i \leq T - \ell$ implies $a_{i+1} \leq a_i + T - \ell$, we have $a_{i+1} \leq a_i + T - \ell - 1 + 1$, which means the columns contributed by the i -th and $(i+1)$ -th rows are contiguous (or overlapping) in $\mathcal{H}_{\ell|\mathbb{Y}}$, with no gaps. Consequently, the set of columns of $\mathcal{H}_{\ell|\mathbb{X}}$ coincides exactly with the first $a_r + T - \ell$ columns of $\mathcal{H}_{\ell|\mathbb{Y}}$.

Thus, $\text{rank } \mathcal{H}_{\ell|\mathbb{X}} = \ell$ holds if and only if

$$a_r + T - \ell \geq \ell$$

by Theorem 2. Since a_r satisfies $r - 1 \leq a_r \leq (r - 1)(T - \ell)$, the condition for such an a_r to exist is that

$$2\ell - T \leq (r - 1)(T - \ell), \quad \text{i.e.,} \quad \ell \leq \frac{rT}{r + 1}.$$

Since ℓ must be an integer, the largest feasible value is $\ell = \lfloor \frac{rT}{r+1} \rfloor$, completing the proof. \square

7.4.4. Monotonic increase of the Hankel codimension

In this section, we establish that the Hankel codimension increases strictly monotonically beyond a certain threshold. This property plays a crucial role in identifying uniquely feasible degrees.

Theorem 6. *In the domain $\ell \in \{L, L + 1, \dots, T - 1\}$, $\text{codim}_{\mathcal{H}}(\ell \mid \mathbb{X})$ is strictly monotonically increasing.*

Proof. Let $L \leq \ell < k < T$. First, for a nonzero vector $w \in \mathbb{V}(\mathcal{H}_{L|\mathbb{X}})^\perp$, define

$$w^{0,\ell-L} = (w^\top \underbrace{0 \dots 0}_{\ell-L})^\top \in \mathbb{C}^{\ell+1}.$$

This is a nonzero vector in $\mathbb{V}(\mathcal{H}_{\ell|\mathbb{X}})^\perp$, so we have $\text{codim}_{\mathcal{H}}(\ell \mid \mathbb{X}) > 0$.

Next, let v_1, \dots, v_d be a basis of $\mathbb{V}(\mathcal{H}_{\ell|\mathbb{X}})^\perp$, where $d = \text{codim}_{\mathcal{H}}(\ell \mid \mathbb{X})$. By reordering if necessary, we may assume that

$$j_1 := \max\{j : (v_1)_j \neq 0\} \geq \max\{j : (v_i)_j \neq 0\}$$

holds for all $i \in \{1, \dots, d\}$.

We claim that the following $d + (k - \ell)$ vectors are linearly independent:

$$v_1^{1,k-\ell-1}, v_1^{2,k-\ell-2}, \dots, v_1^{k-\ell,0}, v_1^{0,k-\ell}, v_2^{0,k-\ell}, \dots, v_d^{0,k-\ell},$$

which all belong to $\mathbb{V}(\mathcal{H}_{k|\mathbb{X}})^\perp \subset \mathbb{C}^{k+1}$. To prove the claim, let us consider

$$\sum_{s=1}^{k-\ell} c_s v_1^{s,k-\ell-s} + \sum_{i=1}^d c'_i v_i^{0,k-\ell} = \mathbf{0}.$$

The vector $v_1^{s,k-\ell-s}$ has the form $(\mathbf{0}^s v_1^\top \mathbf{0}^{k-\ell-s})^\top \in \mathbb{C}^{k+1}$, so its $(j_1 + s)$ -th component equals $(v_1)_{j_1} \neq 0$.

First, we prove $c_s = 0$ for $s = k - \ell, k - \ell - 1, \dots, 1$ by backward induction on s .

For $s = k - \ell$, consider the $(j_1 + k - \ell)$ -th component of the left-hand side. Among all vectors in the sum, only $v_1^{k-\ell,0}$ has a nonzero value at this position, namely $(v_1)_{j_1}$. The vectors $v_i^{0,k-\ell}$ for $i = 1, \dots, d$ have the form $(v_i^\top \mathbf{0}^{k-\ell})^\top$, and since $k - \ell \geq 1$, we have $j_1 + k - \ell > j_1$, which implies that their $(j_1 + k - \ell)$ -th component is zero. Similarly, for $t < k - \ell$, the vector $v_1^{t,k-\ell-t}$ has nonzero components only up to position $j_1 + t < j_1 + k - \ell$, so its $(j_1 + k - \ell)$ -th component is also zero. Therefore, $c_{k-\ell}(v_1)_{j_1} = 0$, which gives $c_{k-\ell} = 0$.

For $s < k - \ell$, assuming $c_{s+1} = \dots = c_{k-\ell} = 0$ by the induction hypothesis, we consider the $(j_1 + s)$ -th component of the left-hand side. By the same argument as above, only $v_1^{s,k-\ell-s}$ has a nonzero value at position $j_1 + s$, namely $(v_1)_{j_1}$. Therefore, $c_s(v_1)_{j_1} = 0$, which gives $c_s = 0$.

This reduces the equation to $\sum_{i=1}^d c'_i v_i^{0,k-\ell} = \mathbf{0}$, which implies $c'_i = 0$ for $i = 1, \dots, d$ by the linear independence of v_1, \dots, v_d .

Therefore,

$$\text{codim}_{\mathcal{H}}(k \mid \mathbb{X}) \geq d + (k - \ell) = \text{codim}_{\mathcal{H}}(\ell \mid \mathbb{X}) + (k - \ell) > \text{codim}_{\mathcal{H}}(\ell \mid \mathbb{X}),$$

which completes the proof. \square

7.4.5. Equivalence Between Unique and Minimal Feasibility

If a uniquely feasible degree exists, it must coincide with the minimum feasible degree L . This conclusion follows directly from Theorem 6 as an immediate corollary.

Corollary 4. *If \mathbb{X} admits a uniquely feasible degree, it must be L .*

Proof. While Corollary 2 states that a uniquely feasible degree ℓ must satisfy $\text{codim}_{\mathcal{H}}(\ell | \mathbb{X}) = 1$, Theorem 6 ensures that this condition is met only when $\ell = L$. \square

7.4.6. Saturation of $\text{dim}_{\mathcal{H}}(\ell | \mathbb{X})$

In this section, we present a theorem that extends Theorem 2. This result plays a crucial role in the development of algorithms for determining uniquely feasible degrees, particularly in the case where $L \leq \frac{T+1}{2}$ for the minimum feasible degree L .

Theorem 7. *If $L \leq \frac{T+1}{2}$ and an L -degree DKMD exists, then:*

$$\begin{aligned} \text{dim}_{\mathcal{H}}(\ell | \mathbb{X}) &= \begin{cases} \ell + 1, & \text{for } 0 \leq \ell \leq L - 1, \\ L, & \text{for } L \leq \ell \leq T - L; \end{cases} \\ \text{codim}_{\mathcal{H}}(\ell | \mathbb{X}) &= \begin{cases} 0, & \text{for } 0 \leq \ell \leq L - 1, \\ \ell - L + 1, & \text{for } L \leq \ell \leq T - L. \end{cases} \end{aligned}$$

Proof. We express the L -degree DKMD given by hypothesis in the form

$$\mathbb{X} = \mathbb{M}\mathcal{V}_{T|\mu_1, \dots, \mu_L}.$$

No column of \mathbb{M} is a zero vector, since L is the minimum feasible degree. Indeed, if

$$\mathbb{X} = \mathbb{M}'\mathcal{V}_{T|\mu_1, \dots, \mu_{L-1}}$$

held for some $\mathbb{M}' \in \mathbb{C}^{m \times (L-1)}$, then the coefficient vector

$$\begin{pmatrix} b_0 & b_1 & \dots & b_{L-2} & 1 \end{pmatrix}$$

of the polynomial

$$(X - \mu_1) \cdots (X - \mu_{L-1}) = X^{L-1} + b_{L-2}X^{L-2} + \cdots + b_0$$

would belong to $\mathbb{V}(\mathcal{H}_{L-1|\mathbb{X}})^\perp$, contradicting the definition of L .

By Theorem 4, we can transform \mathbb{X} by multiplying a matrix A such that $\text{rank } \mathbb{X} = \text{rank}(A\mathbb{X})$ without changing the assertion. In particular, we may assume the following:

- $\mathbb{X} \in \mathbb{C}^{r \times T}$ with $\text{rank } \mathbb{X} = r$. To achieve this, take i_1, \dots, i_r such that the rows $\mathbb{X}\langle i_1 \rangle, \dots, \mathbb{X}\langle i_r \rangle$ are linearly independent. Then, define $A \in \mathbb{C}^{r \times m}$ so that the j -th row of A has 1 as the i_j -th component and 0 for the other components.
- The first row of \mathbb{M} has no zero component: $m_{1i} \neq 0$ for $i = 1, 2, \dots, L$. Since each column of \mathbb{M} is nonzero and $\mathbb{X} \in \mathbb{C}^{r \times T}$, we can find a nonsingular matrix $A \in \mathbb{C}^{r \times r}$ such that the first row of $A\mathbb{M}$ has no zero component.

After this transformation, we apply Theorem 2 to the first row of \mathbb{X} :

$$\mathbb{X}\langle 1 \rangle = \begin{pmatrix} m_{11} & \dots & m_{1L} \end{pmatrix} \mathcal{V}_{T|\mu_1, \dots, \mu_L},$$

and obtain

$$\dim_{\mathcal{H}}(\ell | \mathbb{X}) \geq \dim_{\mathcal{H}}(\ell | \mathbb{X}\langle 1 \rangle) = \begin{cases} \ell + 1, & \text{for } 0 \leq \ell \leq L - 1, \\ L, & \text{for } L \leq \ell \leq T - L, \end{cases}$$

since the columns of $\mathcal{H}_{\ell|\mathbb{X}\langle 1 \rangle}$ are among those of $\mathcal{H}_{\ell|\mathbb{X}}$. For $0 \leq \ell \leq L - 1$, in particular, the equality $\dim_{\mathcal{H}}(\ell | \mathbb{X}) = \ell + 1$ holds because $\dim_{\mathcal{H}}(\ell | \mathbb{X}) \leq \ell + 1$ by definition.

Leveraging this evaluation of lower bound, the claim $\dim_{\mathcal{H}}(\ell | \mathbb{X}) = L$ for $L \leq \ell \leq T - L$ can be proven by induction on ℓ .

For the base case $\ell = L$, the definition of L gives $\text{codim}_{\mathcal{H}}(L | \mathbb{X}) \geq 1$, which implies $\dim_{\mathcal{H}}(L | \mathbb{X}) = L + 1 - \text{codim}_{\mathcal{H}}(L | \mathbb{X}) \leq L$. Combined with $\dim_{\mathcal{H}}(L | \mathbb{X}) \geq L$ shown above, we have $\dim_{\mathcal{H}}(L | \mathbb{X}) = L$.

For the inductive step, assume $\ell > L$ and $\dim_{\mathcal{H}}(\ell - 1 | \mathbb{X}) = L$, i.e., $\text{codim}_{\mathcal{H}}(\ell - 1 | \mathbb{X}) = \ell - L$. By Theorem 6, $\text{codim}_{\mathcal{H}}(\ell | \mathbb{X}) > \text{codim}_{\mathcal{H}}(\ell - 1 | \mathbb{X}) = \ell - L$, which gives

$$\dim_{\mathcal{H}}(\ell | \mathbb{X}) = \ell + 1 - \text{codim}_{\mathcal{H}}(\ell | \mathbb{X}) < \ell + 1 - (\ell - L) = L + 1,$$

and consequently, $\dim_{\mathcal{H}}(\ell | \mathbb{X}) \leq L$. Combined with $\dim_{\mathcal{H}}(\ell | \mathbb{X}) \geq L$, we conclude $\dim_{\mathcal{H}}(\ell | \mathbb{X}) = L$.

Finally, the expressions for $\text{codim}_{\mathcal{H}}(\ell | \mathbb{X})$ follow directly from the definition of the Hankel codimension. \square

Although Theorem 7 requires $L \leq \frac{T+1}{2}$, if $L = \frac{T+1}{2}$ (which occurs only when T is odd), then $L = \frac{T+1}{2} > \frac{T-1}{2} = T - L$ holds, implying that the range $L \leq \ell \leq T - L$ in Theorem 7 is empty. In this boundary case, the theorem provides information only for $\ell = 0, 1, \dots, L - 1 = \frac{T-1}{2}$, namely that $\dim_{\mathcal{H}}(\ell | \mathbb{X}) = \ell + 1$ and $\text{codim}_{\mathcal{H}}(\ell | \mathbb{X}) = 0$ in this range.

For $L < \frac{T+1}{2}$, which for integer L is equivalent to $L \leq \frac{T}{2}$, we have:

Corollary 5. *If L is feasible and satisfies $L \leq \frac{T}{2}$, then L is a uniquely feasible degree.*

Proof. Since L is feasible, an L -degree DKMD exists. The condition $L \leq \frac{T}{2}$ implies $L < \frac{T+1}{2}$, so Theorem 7 applies and gives $\text{codim}_{\mathcal{H}}(L | \mathbb{X}) = 1$. By Corollary 2, this means that L is uniquely feasible. \square

8. Algorithms

Leveraging the five properties mentioned in Section 7.4, we develop efficient algorithms to search for a uniquely feasible degree L and to determine an L -degree characteristic polynomial, which provides eigenvalues of a DKMD. Once mutually distinct eigenvalues are obtained, the associated Koopman modes are calculated as described in Section 5.2. Our algorithms are categorized into:

1. One that applies to the case where $L \leq \frac{T}{2}$ and determines L by $\dim_{\mathcal{H}}\left(\lfloor \frac{T}{2} \rfloor | \mathbb{X}\right)$;
2. Another that performs a binary search to determine L in the case when $L > \frac{T}{2}$.

To introduce the algorithms, we start by addressing a theoretical scenario where the observables exactly consist of a finite number of wave components. In this case, an *exact DKMD* is obtained. Subsequently, we consider a practical scenario where the observables comprise a finite number of dominant wave components, an infinite number of minor wave components, and noise. Here, our algorithms focus on extracting only the dominant wave components, effectively filtering out the minor components and the noise, resulting in an *approximate DKMD*.

8.1. A Theoretical Scenario

We first present Algorithm 1, which reduces the dimension of each observable so that decomposing the reduced observable matrix is equivalent to decomposing the original one. This reduction is practically useful for more efficient computation and clearer understanding of the underlying structure.

We then present algorithms that decompose an observable matrix for two cases: $L \leq \frac{T}{2}$ (Algorithms 2 and 3) and $\frac{T}{2} < L \leq \frac{rT}{r+1}$ (Algorithm 4).

8.1.1. Dimension Reduction

Although each observable vector, *i.e.*, a column vector of \mathbb{X} , is of dimension m , the rank r of \mathbb{X} can be smaller than m . In this case, Algorithm 1 determines a dimensionally reduced observable matrix $\mathbb{Y} \in \mathbb{C}^{r \times T}$ and a conversion matrix $A \in \mathbb{C}^{r \times m}$ such that $\mathbb{Y} = A\mathbb{X}$ and $\text{rank } \mathbb{Y} = \text{rank } \mathbb{X} = r$. By Theorem 4, the Hankel dimensions and codimensions, as well as the set of characteristic polynomials, are invariant under this conversion, and DKMDs for \mathbb{X} and those for \mathbb{Y} are mutually converted by matrix multiplication by A and $\mathbb{X}\mathbb{Y}^+$.

Such A and \mathbb{Y} can be constructed by selecting r linearly independent rows of \mathbb{X} and defining A so that these rows become the rows of $\mathbb{Y} = A\mathbb{X}$. Since these r rows of \mathbb{X} are linearly independent and form all the rows of \mathbb{Y} , we have $\text{rank } \mathbb{Y} = r = \text{rank } \mathbb{X}$.

Algorithm 1 Dimension reduction of the observable matrix.

Require: $\mathbb{X} \in \mathbb{C}^{m \times T}$

Ensure: Matrices $A \in \mathbb{C}^{r \times m}$ and $\mathbb{Y} \in \mathbb{C}^{r \times T}$ for $r = \text{rank } \mathbb{X}$ with $\mathbb{Y} = A\mathbb{X}$ and $\text{rank } \mathbb{Y} = r$

- 1: Find $i_1, \dots, i_r \in \{1, 2, \dots, m\}$ such that the row vectors $\mathbb{X}\langle i_1 \rangle, \dots, \mathbb{X}\langle i_r \rangle$ are linearly independent;
 - 2: Determine a matrix $A \in \mathbb{C}^{r \times m}$ such that the (j, i_j) component is 1 and all other components are 0 for $j = 1, \dots, r$;
 - 3: **return** A and $\mathbb{Y} = A\mathbb{X}$.
-

By applying Algorithm 1, we can reduce the problem of decomposing $\mathbb{X} \in \mathbb{C}^{m \times T}$ to the problem of decomposing $\mathbb{Y} \in \mathbb{C}^{r \times T}$ with $r = \text{rank } \mathbb{X} \leq m$. This reduction provides benefits in terms of computational efficiency and clearer understanding of the data structure when executing the algorithms presented below. However, these algorithms are formulated in general terms and do not require that such a dimension reduction has been performed.

8.1.2. Case $L \leq \frac{T}{2}$

Algorithm 2 first investigates whether $L \leq \frac{T}{2}$ by leveraging the equivalence between $L \leq \left\lfloor \frac{T}{2} \right\rfloor$ and $\text{codim}_{\mathcal{H}}\left(\left\lfloor \frac{T}{2} \right\rfloor \mid \mathbb{X}\right) > 0$. Indeed, if $L \leq \left\lfloor \frac{T}{2} \right\rfloor$, then $\text{codim}_{\mathcal{H}}\left(\left\lfloor \frac{T}{2} \right\rfloor \mid \mathbb{X}\right) \geq \text{codim}_{\mathcal{H}}(L \mid \mathbb{X}) > 0$ holds by Theorem 6. Conversely, if $\text{codim}_{\mathcal{H}}\left(\left\lfloor \frac{T}{2} \right\rfloor \mid \mathbb{X}\right) > 0$, then $L \leq \left\lfloor \frac{T}{2} \right\rfloor$ by the definition of L .

If this investigation reveals $L \leq \frac{T}{2}$, the algorithm identifies L as $\dim_{\mathcal{H}}\left(\left\lfloor \frac{T}{2} \right\rfloor \mid \mathbb{X}\right)$ by Theorem 7, and then executes Algorithm 3 to determine whether L is uniquely feasible. If $L > \frac{T}{2}$, the algorithm returns the value `continue`, indicating that Algorithm 4 should be used.

When invoked, Algorithm 3 verifies the following:

1. A vector $(a_0, a_1, \dots, a_L, 1)^T$ exists in $\mathbb{V}(\mathcal{H}_{L|\mathbb{X}})^\perp$. This can be efficiently verified by performing a QR decomposition of $\mathcal{H}_{L|\mathbb{X}}$.
2. If such a vector exists, verify that the polynomial

$$x^{L+1} + a_L x^L + \dots + a_1 x + a_0 = 0$$

has no repeated roots.

If both conditions are satisfied, this confirms that L is feasible, and we can then apply Theorem 7. As a result, we have $\text{codim}_{\mathcal{H}}(L \mid \mathbb{X}) = 1$, meaning that L is uniquely feasible. Thus, the polynomial obtained in the verification is square-free and serves as the characteristic polynomial of the unique L -degree DKMD of \mathbb{X} . In this case, the algorithm returns the obtained characteristic polynomial. Otherwise, it returns the value `no_solution`, indicating that no uniquely feasible degree exists.

Algorithm 2 Search for an L -degree characteristic polynomial when $L \leq \frac{T}{2}$.

Require: $\mathbb{X} \in \mathbb{C}^{m \times T}$

Ensure: The signal continue if $L > \frac{T}{2}$; the characteristic polynomial if $L \leq \frac{T}{2}$ is uniquely feasible; no_solution if $L \leq \frac{T}{2}$ is not uniquely feasible.

- 1: **if** $\text{codim}_{\mathcal{H}}\left(\left\lfloor \frac{T}{2} \right\rfloor \mid \mathbb{X}\right) > 0$ **then**
 - 2: Let $L = \text{dim}_{\mathcal{H}}\left(\left\lfloor \frac{T}{2} \right\rfloor \mid \mathbb{X}\right)$;
 - 3: Execute Algorithm 3;
 - 4: **return** the return value of Algorithm 3;
 - 5: **else**
 - 6: **return** continue
 - 7: **end if**
-

Algorithm 3 Determine the characteristic polynomial.

Require: $\mathbb{X} \in \mathbb{C}^{m \times T}$ and L

Ensure: An L -degree characteristic polynomial or no_solution

- 1: **if** $\exists (a_0 \ a_1 \ \dots \ a_L \ 1)^T \in \mathbb{V}(\mathcal{H}_{L|\mathbb{X}})^\perp$ **then**
 - 2: Let $f(x) = x^{L+1} + a_L x^L + \dots + a_1 x + a_0$;
 - 3: **if** $f(x) = 0$ has no repeated roots **then**
 - 4: **return** $f(x)$
 - 5: **end if**
 - 6: **end if**
 - 7: **return** no_solution
-

8.1.3. Case $\frac{T}{2} < L \leq \frac{rT}{r+1}$

Algorithm 4 details the procedure for cases when $L > \frac{T}{2}$. Note that if L is uniquely feasible, then $L \leq \frac{rT}{r+1}$ holds by Theorem 5, and this gives the best possible upper bound.

The algorithm first verifies whether $\text{codim}_{\mathcal{H}}\left(\left\lfloor \frac{rT}{r+1} \right\rfloor \mid \mathbb{X}\right) > 0$. If this condition does not hold, no uniquely feasible degree exists, and the algorithm returns the signal no_solution.

If $\text{codim}_{\mathcal{H}}\left(\left\lfloor \frac{rT}{r+1} \right\rfloor \mid \mathbb{X}\right) > 0$, then L lies in the range $\left\lfloor \frac{T}{2} \right\rfloor < L \leq \left\lfloor \frac{rT}{r+1} \right\rfloor$. The algorithm utilizes a binary search to find L , leveraging the fact that $\text{codim}_{\mathcal{H}}(\ell \mid \mathbb{X})$ is a strictly increasing function by Theorem 6.

Since the identified L does not necessarily satisfy $\text{codim}_{\mathcal{H}}(L \mid \mathbb{X}) = 1$, the algorithm must verify $\text{codim}_{\mathcal{H}}(L \mid \mathbb{X}) = 1$ before executing Algorithm 3 to confirm that L is uniquely

8.2. A Practical Scenario

In practice, even if the Koopman operator has only discrete eigenvalues, the number of eigenvalues can be infinite, and in addition, observables can contain error signals. In such situations, the purpose of DKMD is to find a finite number of major wave components that most significantly affect the observables. From the viewpoint of executing our algorithms, the presence of minor components and errors makes direct computation of Hankel dimensions via QR decomposition impractical. In fact, $\text{rank } \mathcal{H}_{\ell|\mathbb{X}} = \ell + 1$ may always hold, which makes it impossible to identify the Hankel dimensions. In this section, we present a method to estimate Hankel dimensions for the major components via singular value decomposition (SVD) rather than QR decomposition.

We assume that the observables are represented as

$$x_t = \sum_{n=1}^{\ell} m_n \lambda_n^t + \sum_{n=1}^{\infty} m'_n \lambda_n'^t + \varepsilon_t,$$

Algorithm 4 Search for an L -degree characteristic polynomial when $L \in \left(\frac{T}{2}, \frac{rT}{r+1}\right]$.

Require: $\mathbb{X} \in \mathbb{C}^{m \times T}$ with $\text{codim}_{\mathcal{H}}\left(\left\lfloor \frac{T}{2} \right\rfloor \mid \mathbb{X}\right) = 0$

Ensure: Either the characteristic polynomial of the DKMD of \mathbb{X} for the uniquely feasible degree $L > \frac{T}{2}$, if present, or no_solution, otherwise.

```

1: if  $\text{codim}_{\mathcal{H}}\left(\left\lfloor \frac{rT}{r+1} \right\rfloor \mid \mathbb{X}\right) = 0$  then
2:   return no_solution
3: end if
4: Let  $l = \left\lfloor \frac{T}{2} \right\rfloor$ ;
5: Let  $h = \left\lfloor \frac{rT}{r+1} \right\rfloor$ ;
6: while  $h - l > 1$  do
7:   Let  $k = \left\lfloor \frac{l+h}{2} \right\rfloor$ ;
8:   if  $\text{codim}_{\mathcal{H}}(k \mid \mathbb{X}) = 0$  then
9:     Let  $l = k$ ;
10:  else if  $\text{codim}_{\mathcal{H}}(k \mid \mathbb{X}) = 1$  then
11:    Let  $L = k$ ;
12:    Execute Algorithm 3;
13:    return the return value of Algorithm 3;
14:  else
15:    Let  $h = k$ ;
16:  end if
17: end while
18: if  $\text{codim}_{\mathcal{H}}(h \mid \mathbb{X}) = 1$  then
19:   Let  $L = h$ ;
20:   Execute Algorithm 3;
21:   return the return value of Algorithm 3;
22: else
23:   return no_solution;
24: end if

```

where λ_n^t and $\lambda_n'^t$ represent major and minor wave components, respectively, and ε_t is noise. We define:

$$\hat{\mathbf{x}}_t = \sum_{n=1}^{\ell} \hat{m}_n \hat{\lambda}_n^t, \quad \hat{\mathbb{X}} = [\hat{\mathbf{x}}_0 \cdots \hat{\mathbf{x}}_{T-1}];$$

$$\mathbf{x}'_t = \sum_{n=1}^{\infty} m'_n \lambda_n'^t + \varepsilon_t, \quad \mathbb{X}' = [\mathbf{x}'_0 \cdots \mathbf{x}'_{T-1}],$$

and our basic assumption is that \mathbf{x}'_t is a small perturbation. Our aim is to estimate $\text{rank } \mathcal{H}_{\ell|\hat{\mathbb{X}}}$ from $\mathcal{H}_{\ell|\mathbb{X}}$, taking advantage of the fact that $\text{rank } \mathcal{H}_{\ell|\hat{\mathbb{X}}}$ equals the number of positive singular values of $\mathcal{H}_{\ell|\hat{\mathbb{X}}}$.

We consider $A = \hat{A} + A' \in \mathbb{C}^{m \times n}$ with $m \leq n$. Let $\sigma_1 \geq \cdots \geq \sigma_m \geq 0$, $\hat{\sigma}_1 \geq \cdots \geq \hat{\sigma}_m \geq 0$, and $\sigma'_1 \geq \cdots \geq \sigma'_m \geq 0$ be the singular values of A , \hat{A} , and A' , respectively. Then,

$$|\sigma_i - \hat{\sigma}_i| \leq \sigma'_1$$

holds for all $i \in \{1, \dots, m\}$. This can be proven as follows. By the Courant–Fischer min-max theorem [17], the i -th singular value of A satisfies

$$\sigma_i = \min_{\substack{V \subseteq \mathbb{C}^n \\ \dim V = n-i+1}} \max_{\substack{v \in V \\ \|v\|=1}} \|Av\|.$$

Since $A = \widehat{A} + A'$, we have $\sigma_i - \widehat{\sigma}_i \leq \sigma'_1$ as follows:

$$\begin{aligned} \sigma_i &= \min_{\substack{V \subseteq \mathbb{C}^n \\ \dim V = n-i+1}} \max_{\substack{v \in V \\ \|v\|=1}} \|Av\| \\ &\leq \min_{\substack{V \subseteq \mathbb{C}^n \\ \dim V = n-i+1}} \max_{\substack{v \in V \\ \|v\|=1}} \left(\|\widehat{A}v\| + \|A'v\| \right) \\ &\leq \min_{\substack{V \subseteq \mathbb{C}^n \\ \dim V = n-i+1}} \max_{\substack{v \in V \\ \|v\|=1}} \|\widehat{A}v\| + \sigma'_1 = \widehat{\sigma}_i + \sigma'_1. \end{aligned}$$

By the same reasoning, we also obtain $\widehat{\sigma}_i - \sigma_i \leq \sigma'_1$.

Furthermore, if $\text{rank } \widehat{A} = r < m$, that is,

$$\widehat{\sigma}_1 \geq \dots \geq \widehat{\sigma}_r > 0 = \widehat{\sigma}_{r+1} = \dots = \widehat{\sigma}_m,$$

then we have

$$\sigma_i \begin{cases} \geq \widehat{\sigma}_i - \sigma'_1 \geq \widehat{\sigma}_r - \sigma'_1, & \text{if } i \leq r, \\ \leq \sigma'_1, & \text{if } i > r. \end{cases}$$

Therefore, if $\widehat{\sigma}_r$ is sufficiently larger than σ'_1 , there exists a large gap between σ_i for $i \leq r$ and σ_i for $i > r$, and hence, we can estimate r from $\sigma_1, \dots, \sigma_m$. Thus, if we can assume that the smallest positive singular value of $\mathcal{H}_{\ell|\widehat{\mathbb{X}}}$ is sufficiently larger than the largest singular value of $\mathcal{H}_{\ell|\mathbb{X}'}$, we can apply this method to estimate $\dim_{\mathcal{H}}(\ell | \widehat{\mathbb{X}})$.

8.3. Time Complexities

The computationally intensive operations in these algorithms primarily involve executing QR decomposition (QRD), singular value decomposition (SVD), and solving equations (EqS). Table 2 demonstrates that the algorithms execute these computations only a small number of times, and consequently, prove to be highly efficient.

Table 2. Time complexity of the algorithms in the theoretical and practical scenarios.

Algorithms	Theoretical		Practical	
	QRD	EqS	SVD	EqS
Algorithm 1	1	0	1	0
Algorithm 2	1	0	1	0
Algorithm 3	1	1	1	1
Algorithm 4	$< \log_2 T$	0	$< \log_2 T$	0

9. Simulations

Through simulations, we investigate the accuracy of Algorithms 2 to 4 in estimating Koopman eigenvalues and making predictions.

9.1. Synthetic Datasets Used in the Simulations

The synthetic datasets of observables are constructed with $m = 2$ and $T = 50$, which makes \mathbb{X} a 2×50 matrix, and are randomly generated by performing the following steps:

1. Sample as many distinct Koopman eigenvalues, each classified as either major or minor, as specified in Table 3. Let $\lambda_1, \dots, \lambda_N$ denote these eigenvalues. For each λ_i , its complex conjugate $\bar{\lambda}_i$ must also be in the set. Furthermore, every conjugate pair $(\lambda, \bar{\lambda})$ is sampled independently as follows:

- $|\lambda| = |\bar{\lambda}|$ is sampled according to a log-normal distribution with parameters $\mu = 0$ and $\sigma = 0.01$, whose probability density function is $\frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln x)^2}{2\sigma^2}\right)$. The median, mean, and variance of this distribution are $e^0 = 1$, $e^{\frac{\sigma^2}{2}}$, and $e^{2\sigma^2} - e^{\sigma^2}$, respectively.
- $\arg \lambda = -\arg \bar{\lambda}$ is sampled uniformly from the interval $(0, \pi)$.

The distribution of $|\lambda|$ is designed to restrict the occurrence of samples too far from 1, because $|\lambda|$ much larger than 1 causes the observables to diverge, while a component with $|\lambda|$ much smaller than 1 decays rapidly.

- Determine the Koopman mode $\mathbf{m}_i = (m_{1i}, m_{2i})^\top$ corresponding to λ_i with the following constraints:
 - $m_{1j} = \bar{m}_{1i}$ and $m_{2j} = \bar{m}_{2i}$ hold whenever $\lambda_j = \bar{\lambda}_i$;
 - The modes associated with the major eigenvalues must have significant magnitudes, while those associated with the minor eigenvalues must have smaller magnitudes.

To satisfy the second requirement, we use a function $\zeta : \mathbb{C} \rightarrow [0, 1]$ defined below, which has sharp peaks only at the sampled major eigenvalues $\lambda_{i_1}, \dots, \lambda_{i_k}$:

$$\zeta(\lambda \mid \lambda_{i_1}, \dots, \lambda_{i_k}) = \max_{j=1, \dots, k} \frac{2}{1 + e^{100[(|\lambda| - |\lambda_{i_j}|)^2 + (\arg \lambda - \arg \lambda_{i_j})^2]}}$$

For each Koopman eigenvalue λ_i , the mode \mathbf{m}_i is determined by sampling the argument of each component uniformly at random, while setting the magnitude to $\zeta(\lambda_i \mid \lambda_{i_1}, \dots, \lambda_{i_k})$, i.e., $|m_{1i}| = |m_{2i}| = \zeta(\lambda_i \mid \lambda_{i_1}, \dots, \lambda_{i_k})$.

- Construct \mathbb{X} as $\mathbb{X} = [\mathbf{m}_1, \dots, \mathbf{m}_N] \mathcal{V}_T |_{\lambda_1, \dots, \lambda_N}$. If the inclusion of noise is required, add to \mathbb{X} a noise matrix $[\varepsilon_{it}]$ with $i \in \{1, 2\}$ and $t \in \{0, 1, \dots, T-1\}$, where each ε_{it} is independently sampled from a normal distribution $\mathcal{N}(0, 0.01^2)$.

In addition, to evaluate the predictive accuracy of our algorithms, we compute observable values for $t \in \{50, 51, \dots, 79\}$ using the Koopman eigenvalues and Koopman modes determined above.

Table 3. Simulation scenarios.

Scenario number	Type	# Major eigenvalues	# Minor eigenvalues	Noise inclusion
1	$L \leq \frac{T}{2}$	10	0	No
2	$L > \frac{T}{2}$	30	0	No
3	$L \leq \frac{T}{2}$	10	90	Yes
4	$L > \frac{T}{2}$	30	70	Yes

9.2. Simulation Scenarios

We conduct simulations under the following four distinct scenarios:

- Scenarios 1 and 2 investigate the case where an exact DKMD is obtained via QR decomposition (Section 8.1).
- Scenarios 3 and 4 investigate the case where an approximated DKMD is obtained via singular value decomposition (Section 8.2).
- Scenarios 1 and 3 are used to investigate Algorithm 2.
- Scenarios 2 and 4 are used to investigate Algorithm 4.

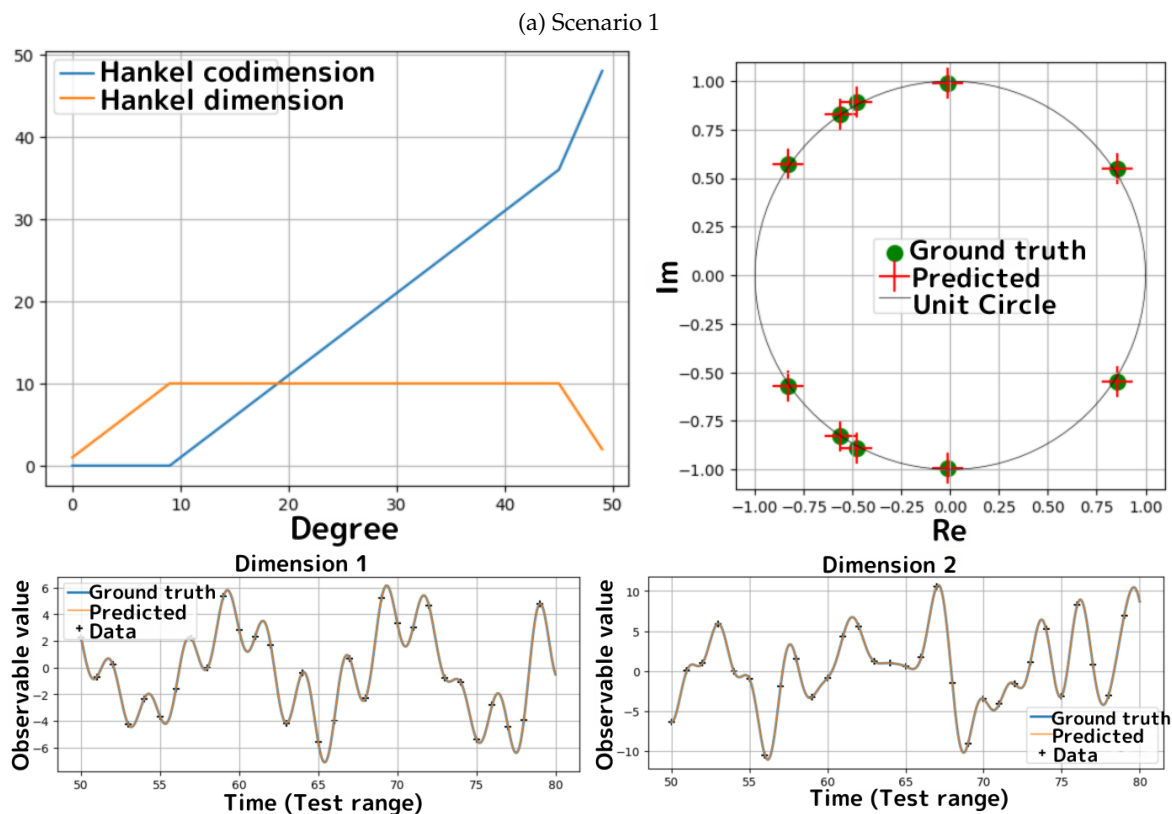
9.3. Results of the Simulations

We have obtained excellent results in the simulations for all scenarios. In the case of estimating an exact DKMD, the estimated Koopman eigenvalues and the predictions for $t \in \{50, 51, \dots, 79\}$ are identical to the ground truth within numerical precision. In the case of estimating an approximated DKMD, the estimated Koopman eigenvalues and the predictions for $t \in \{50, 51, \dots, 79\}$ show close agreement with the ground truth.

Scenarios 1 and 2: For both scenarios, Algorithms 2 and 4 correctly identify the ground truth uniquely feasible degrees. Furthermore, we observe that the estimated eigenvalues (the upper-right panels of Figure 3(a,b)) and the predictions for $t \in \{50, 51, \dots, 79\}$ (the bottom panels of Figure 3(a,b)) are identical to the ground truth within numerical precision.

Scenario 3: The upper-left panel in Figure 4(a) depicts the logarithm of the singular values of $\mathcal{H}_{25|\mathbb{X}}$. Evidently, we observe a significant gap between the top ten singular values and those that follow, leading to the conclusion that $\dim_{\mathcal{H}}(25 | \mathbb{X}) = 10$, indicating that $\ell = 10$ is the uniquely feasible degree. Furthermore, the estimated eigenvalues (the upper-right panel) and the predictions (the bottom panels) show excellent agreement with the ground truth.

Scenario 4: The upper-left panel in Figure 4(b) depicts the result of a singular value decomposition of $\mathcal{H}_{\ell|\mathbb{X}}$ when the binary search of Algorithm 4 visits $\ell = 30$. Since the top 30 singular values are significantly greater than those that follow, we can conclude that $\text{codim}_{\mathcal{H}}(30 | \mathbb{X}) = 1$, meaning that $\ell = 30$ is the uniquely feasible degree. The estimated eigenvalues and the predictions also show excellent agreement with the ground truth.



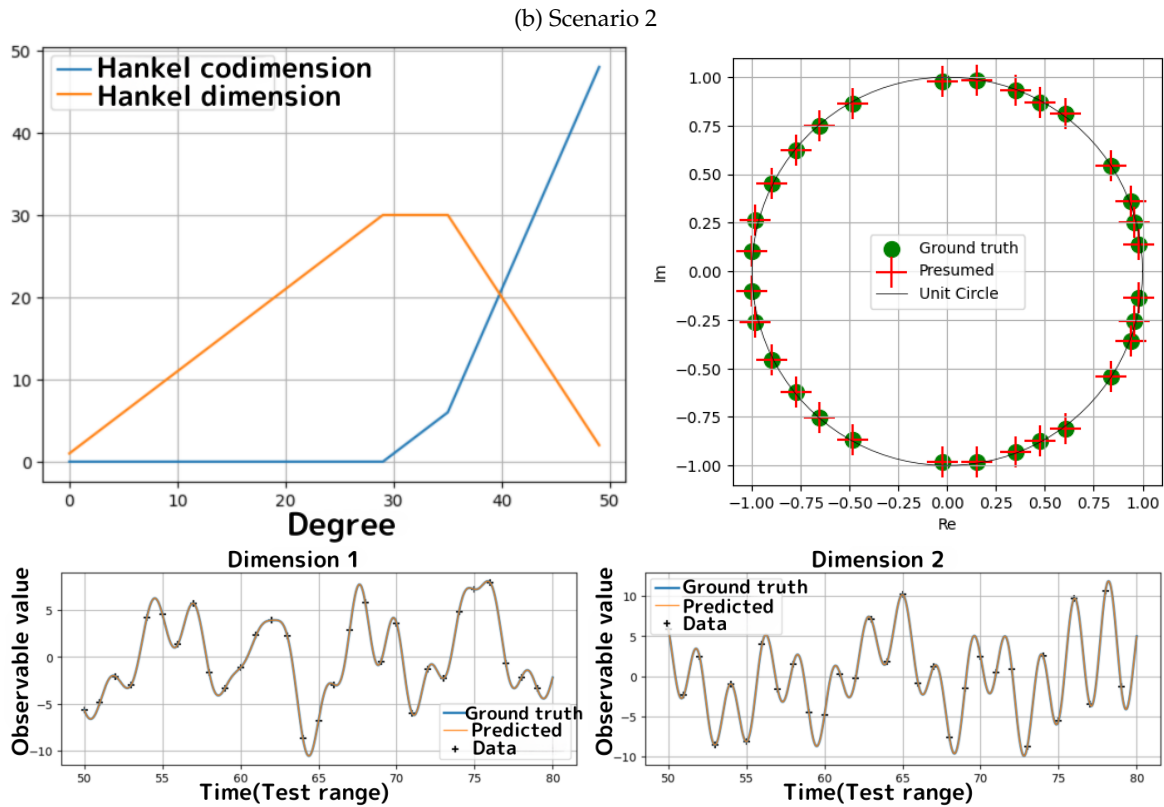
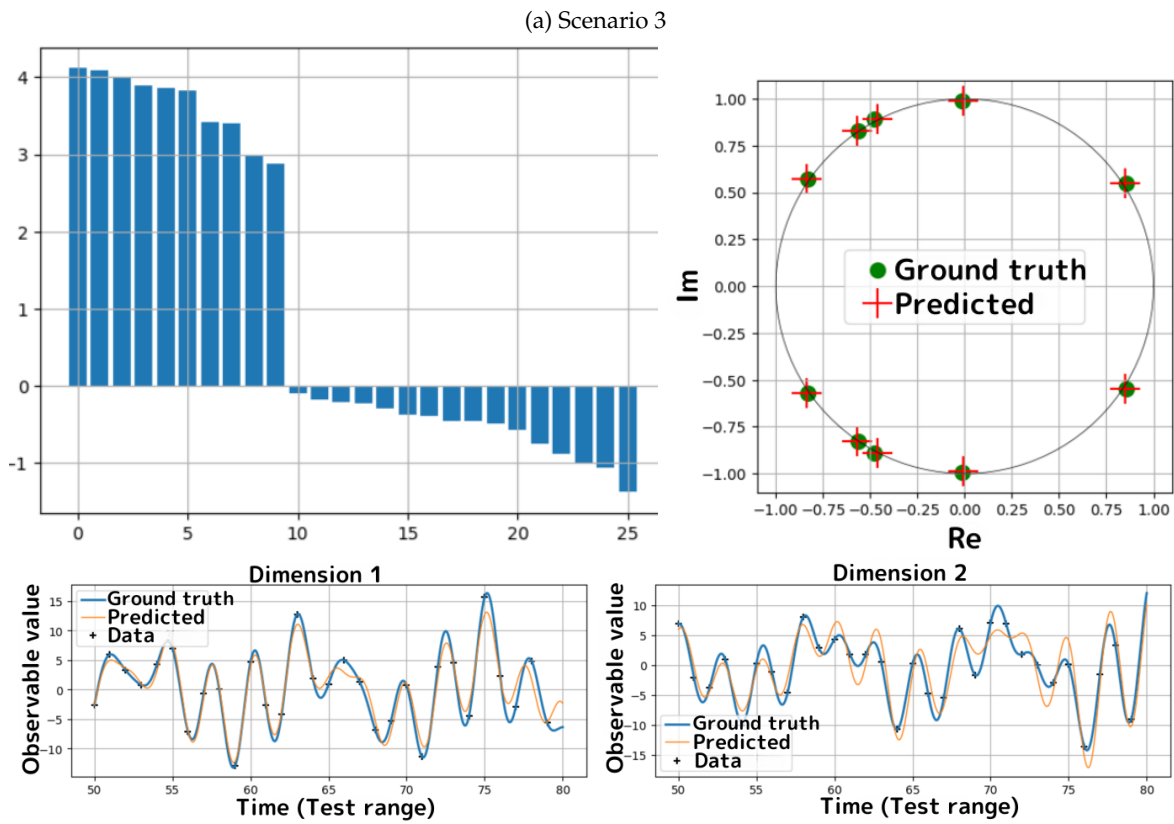


Figure 3. Results of simulations for Scenarios 1 and 2.



(b) Scenario 4

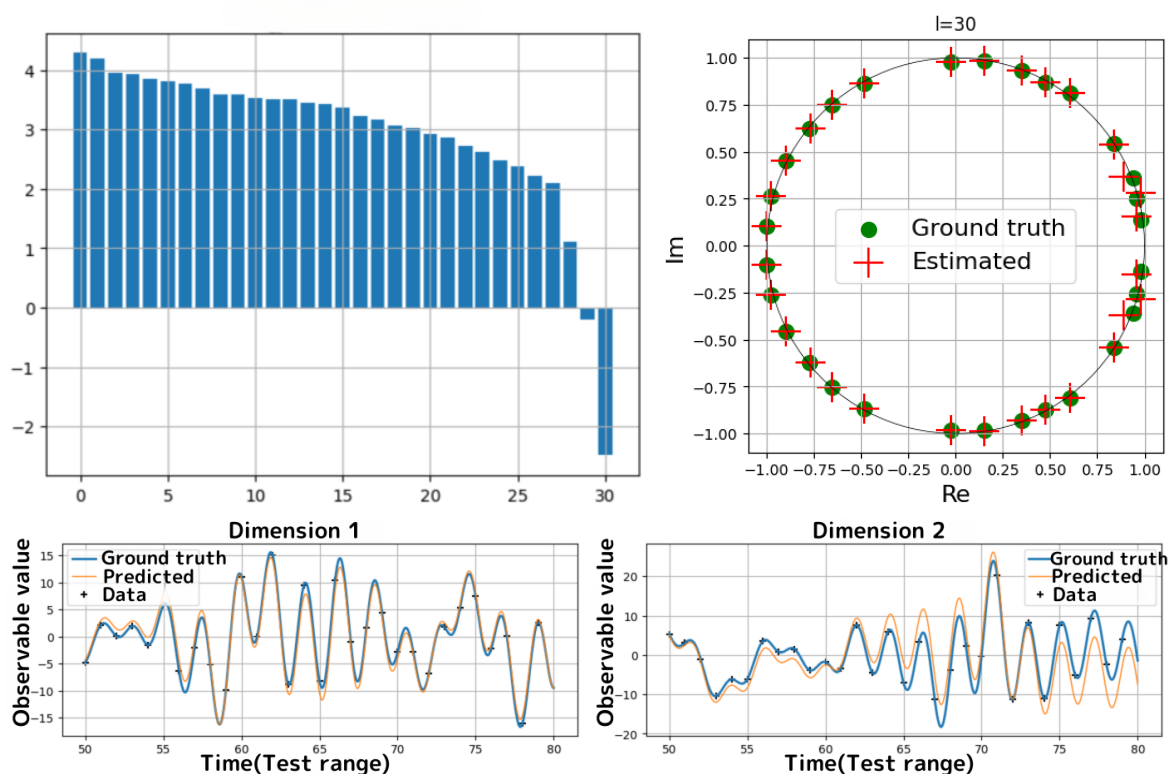


Figure 4. Results of simulations for Scenarios 3 and 4.

10. Conclusions

We have developed a theoretical framework to estimate the correct degree (uniquely feasible degree) of Discrete Koopman Mode Decomposition (DKMD) from a given dataset of observables. The degree of DKMD corresponds to the number of Koopman eigenvalues involved. We demonstrate that unless the correct degree is used, infinitely many DKMDs with different Koopman eigenvalues and modes can fit the training data. However, these variations lead to divergent predictions for future time steps, resulting in unreliable forecasts. Furthermore, the theory provides efficient algorithms to identify uniquely feasible degrees.

References

1. Rudin, W. *Real and Complex Analysis*, 3rd ed.; McGraw-Hill: New York, 1987.
2. Koopman, B.O. Hamiltonian systems and transformation in Hilbert space. *Proceedings of the national academy of sciences of the united states of america* **1931**, *17*, 315.
3. Schmid, P.J. Dynamic mode decomposition of numerical and experimental data. *Journal of fluid mechanics* **2010**, *656*, 5–28.
4. Rowley, C.W.; Mezić, I.; Bagheri, S.; Schlatter, P.; Henningson, D.S. Spectral analysis of nonlinear flows. *Journal of Fluid Mechanics* **2009**, *641*, 115–127. <https://doi.org/10.1017/S0022112009992059>.
5. Tu, J.H.; Rowley, C.W.; Luchtenburg, D.M.; Brunton, S.L.; Kutz, J.N. On dynamic mode decomposition: Theory and applications. *Journal of Computational Dynamics* **2014**, *1*, 391.
6. Taira, K.; Brunton, S.L.; Dawson, S.T.; Rowley, C.W.; Colonius, T.; McKeon, B.J.; Schmidt, O.T.; Gordeyev, S.; Theofilis, V.; Ukeiley, L.S. Modal analysis of fluid flows: An overview. *Aiaa Journal* **2017**, *55*, 4013–4041.
7. Brunton, S.L.; Brunton, B.W.; Proctor, J.L.; Kaiser, E.; Kutz, J.N. Chaos as an intermittently forced linear system. *Nature communications* **2017**, *8*, 1–9.
8. Brunton, B.W.; Johnson, L.A.; Ojemann, J.G.; Kutz, J.N. Extracting spatial–temporal coherent patterns in large-scale neural recordings using dynamic mode decomposition. *Journal of neuroscience methods* **2016**, *258*, 1–15.
9. Taylor, R.; Kutz, J.N.; Morgan, K.; Nelson, B.A. Dynamic mode decomposition for plasma diagnostics and validation. *Review of Scientific Instruments* **2018**, *89*, 053501.

10. Kaptanoglu, A.A.; Morgan, K.D.; Hansen, C.J.; Brunton, S.L. Characterizing magnetized plasmas with dynamic mode decomposition. *Physics of Plasmas* **2020**, *27*, 032108.
11. Kusaba, A.; Shin, K.; Shepard, D.; Kuboyama, T. Predictive Nonlinear Modeling by Koopman Mode Decomposition. In Proceedings of the 2020 International Conference on Data Mining Workshops (ICDMW). IEEE, 2020, pp. 811–819.
12. Fujii, K.; Takeishi, N.; Kibushi, B.; Kouzaki, M.; Kawahara, Y. Data-driven spectral analysis for coordinative structures in periodic human locomotion. *Scientific reports* **2019**, *9*, 1–14.
13. Berger, E.; Sastuba, M.; Vogt, D.; Jung, B.; Ben Amor, H. Estimation of perturbations in robotic behavior using dynamic mode decomposition. *Advanced Robotics* **2015**, *29*, 331–343.
14. Takeishi, N.; Kawahara, Y.; Yairi, T. Sparse nonnegative dynamic mode decomposition. In Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP). IEEE, 2017, pp. 2682–2686.
15. Rudin, W. *Functional Analysis*, second ed.; McGraw-Hill Series in Higher Mathematics, McGraw-Hill, Inc.: New York, St. Louis, San Francisco, 1991; pp. xvi + 424. Second edition.
16. Susuki, Y.; Mezic, I. A Prony approximation of Koopman mode decomposition. In Proceedings of the 2015 54th IEEE Conference on Decision and Control (CDC). IEEE, 2015, pp. 7022–7027. <https://doi.org/10.1109/CDC.2015.7403220>.
17. Horn, R.A.; Johnson, C.R. *Matrix Analysis*, 2nd ed.; Cambridge University Press, 2012.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.