



# Koopman Operator Spectrum for Random Dynamical Systems

Nelida Črnjarić-Žic<sup>1</sup> · Senka Maćešić<sup>1</sup> · Igor Mezić<sup>2</sup>

Received: 23 November 2017 / Accepted: 14 August 2019 / Published online: 13 September 2019  
© Springer Science+Business Media, LLC, part of Springer Nature 2019

## Abstract

In this paper, we consider the Koopman operator associated with the discrete and the continuous-time random dynamical system (RDS). We provide results that characterize the spectrum and the eigenfunctions of the stochastic Koopman operator associated with different types of linear RDS. Then we consider the RDS for which the associated Koopman operator family is a semigroup, especially those for which the generator can be determined. We define a stochastic Hankel–DMD algorithm for numerical approximations of the spectral objects (eigenvalues, eigenfunctions) of the stochastic Koopman operator and prove its convergence. We apply the methodology to a variety of examples, revealing objects in spectral expansions of the stochastic Koopman operator and enabling model reduction.

**Keywords** Stochastic Koopman operator · Random dynamical systems · Stochastic differential equations · Dynamic mode decomposition

**Mathematics Subject Classification** 37H10 · 47B33 · 37M99 · 65P99

---

Communicated by Alain Goriely.

---

✉ Nelida Črnjarić-Žic  
nelida@riteh.hr

Senka Maćešić  
senka.macesic@riteh.hr

Igor Mezić  
mezić@engr.ucsb.edu

<sup>1</sup> Faculty of Engineering, University of Rijeka, Vukovarska 58, 51000 Rijeka, Croatia

<sup>2</sup> Faculty of Mechanical Engineering and Mathematics, University of California, Santa Barbara, Santa Barbara, CA 93106, USA

## 1 Introduction

Prediction and control of the evolution of large complex dynamical systems is a modern-day science and engineering challenge. Some dynamical systems can be modeled well enough by using the standard mathematical tools, such as differential or integral calculus. In these cases the simplifications of the system are typically introduced by neglecting some of the phenomena with a small impact on the behavior of the system. However, there are many dynamical systems for which the mathematical model is too complex or even does not exist, but data can be obtained by monitoring some observables of the system. In this context, data-driven analysis methods need to be developed, including techniques to extract simpler, representative components of the process that can be used for modeling and prediction.

One approach for decomposing the complex systems into simpler structures is via the spectral decomposition of the associated Koopman operator. Koopman operator was introduced in Koopman (1931) in the measure-preserving setting, as a composition operator acting on the Hilbert space of square-integrable functions. The increased interest in the spectral operator-theoretic approach to dynamical systems in last decade starts with the works Mezić and Banaszuk (2004) and Mezić (2005) (see also the earlier Mezić and Banaszuk 2000), where the problem of decomposing the evolution of an ergodic dissipative dynamical system from the perspective of operator theory was studied, and data-driven methods for computation of eigenvalues and eigenfunctions were provided based on rigorous harmonic analysis methods. The application of the theory to complex systems are numerous, for example, in fluid flows (Bagheri 2013; Sharma et al. 2016), infectious disease dynamics (Proctor and Eckhoff 2015), power systems (Susuki and Mezić 2012; Susuki et al. 2016), etc. The advantage of using the Koopman operator framework lies in the fact that it can be applied within the data-driven environment, even if the underlying mathematical model is not known. In addition to the analysis of the system properties the Koopman decomposition of complex dynamical structures can provide approximations to the evolution of a possibly high-dimensional system in lower dimensions, enabling model reduction. An overview of the spectral properties of the Koopman operator and its applications prior to 2012 is given in Mezić (2013), Budišić et al. (2012).

Of particular importance in Koopman operator theory in the deterministic case is the interplay between the geometry of the underlying dynamics and the level sets of the eigenfunctions of the Koopman operator. Namely, invariant sets can be determined as level sets of Koopman eigenfunctions at eigenvalue 0 (continuous time) or 1 (discrete time), (Mezić 1994; Mezić and Wiggins 1999), periodic sets can be determined as level sets of Koopman operator eigenfunctions at the imaginary axis (continuous time) or the unit circle (discrete time), the concept of isostables—level sets of eigenfunctions associated with eigenvalues with negative real part (continuous time) or inside the unit circle (discrete time) generalize the geometric concept of Fenichel fibers (Mauroy et al. 2013). Also, zero level sets of certain families of eigenfunctions determine stable, unstable, and center manifolds to invariant sets (Mezić 2015, 2017a). In this paper, we enable generalization of such concepts by developing the theory of stochastic Koopman operators associated with RDS.

A variety of methods for determining the numerical approximation of the Koopman spectral decomposition, known under the name Koopman mode decomposition (KMD) have been developed. KMD consists of the triple of Koopman eigenvalues, eigenfunctions, and modes, which are the building blocks for the evolution of observables under the dynamics of a dynamical system. A general method for computing the Koopman modes, based on the rigorous theoretical results for the generalized Laplace transform, is known under the name generalized Laplace analysis (GLA) (Mezić 2013; Budišić et al. 2012), whose first version was advanced in Mezić and Banaszuk (2000). Another method is the dynamic mode decomposition (DMD) method. DMD was firstly introduced in Schmid (2008) for the study of the fluid flows, without any reference to the Koopman operator. The connection between the KMD and DMD was first pointed out in Rowley et al. (2009). Like GLA, DMD is a data-driven technique. Due to the fact it could be numerically implemented relatively easily, since it relies on standard linear algebra concepts, this method become extremely popular in the data-driven systems research community. Starting with (Schmid 2010) many versions of the DMD algorithm have been introduced to efficiently compute the spectral objects of the Koopman operator under various assumptions on the data. In Tu et al. (2014) the exact DMD algorithm was introduced, while in Williams et al. (2015) the extension of the DMD algorithm, under the name extended dynamic mode decomposition (EDMD) was proposed. These two methods rely on the more standard approaches to representing a linear operator with respect to a specific basis (finite section method), rather than a sampling method—as in the companion matrix-based DMD (see Drmač et al. 2018, section 2.1). Susuki and Mezić (2015) introduced a further extension of the DMD algorithm, which combines the Prony method with the DMD algorithm, so that the Hankel matrix is used instead of the companion matrix to compute the Koopman spectrum on the single observable as well as on the vector of observable functions. Arbabi and Mezić (2017) referred to this algorithm as to the Hankel–DMD and proved that under certain assumptions the obtained eigenvalues and eigenfunctions converge to the exact eigenvalues and eigenfunctions of the Koopman operator. The assumptions were removed, and further connections between the spectra of EDMD matrices and eigenvalues of the Koopman operator were proven in Korda and Mezić (2018).

The majority of works analyzing or practically using the spectral properties of the Koopman operator assume that the dynamical system under consideration is autonomous. Similarly, the proposed numerical algorithms for evaluating KMD were almost exclusively applied to autonomous dynamical systems. The generalization of the Koopman operator framework to nonautonomous system was introduced in Mezić and Surana (2016), where the definitions of the nonautonomous Koopman eigenvalues, eigenfunction and modes, as building blocks of the dynamical system, are given. KMD for nonautonomous systems was studied in Maćešić et al. (2018), where the possibility of using the Arnoldi-like DMD algorithms for evaluating the Koopman eigenvalues and eigenfunctions in the nonautonomous case was carefully explored and the appropriate extensions of the algorithm were proposed.

Another possible generalization of the Koopman operator framework is its extension to random dynamical systems (RDS) and stochastic systems. There is a long history of analyzing the dynamics of the stochastic Markov processes through the study of the spectral properties of the associated transfer operators (Dynkin 1965; Yosida 1980;

Lasota and Mackey 1994; Arnold 1974). In Williams et al. (2015), it was realized that if the data provided to the EDMD algorithm are generated by a Markov stochastic process instead of a deterministic dynamical system, the algorithm approximates the eigenfunctions of the Kolmogorov backward equation. Klus and coworkers used the Ulam's method and EDMD algorithm in the series of papers (Klus et al. 2016; Klus and Schütte 2016) to approximate the spectral objects of the Koopman operator and its adjoint Perron–Frobenius operator. There is a variety of numerical techniques used to approximate different transfer operators associated with the stochastic systems, as well as their spectral objects. A review can be found in Klus et al. (2018).

In this paper we consider the stochastic Koopman operator associated with the discrete- and continuous-time RDS, using the definition and classification of RDS from Arnold (1998). The Koopman operator for discrete RDS was first introduced in Mezić and Banaszuk (2000), where the action of the stochastic Koopman operator on an observable was defined by taking an expectation of values of the observable at the next time step. The eigenspace at 1 was shown to relate to invariant sets of stochastic dynamics, followed by Mezić and Banaszuk (2004), Mezić (2005) in which analysis of the geometry of eigenspaces associated with any additional eigenvalues on the unit circle was pursued and shown to relate to periodic sets of the underlying RDS. Here we extend the definition of the stochastic Koopman operator to the continuous-time RDS and explore some properties of the associated stochastic Koopman operator family, its generators, eigenvalues and eigenfunctions.

The characterization of the dynamics of such systems by using the eigenvalues and the eigenfunctions of the related generators was studied in recent papers (Shnitzer et al. 2017; Tantet et al. 2017; Giannakis 2019). Giannakis (2019) develops a framework for the KMD based on the representation of the Koopman operator in a smooth orthonormal bases determined from the time-ordered noisy data through the diffusion map algorithm. Using this representation, the Koopman eigenfunctions are approximated as the eigenfunctions of the related advection–diffusion operator. A similar approach by using the manifold learning technique via diffusion maps was used in Shnitzer et al. (2017) to capture the inherent coordinates for building an intrinsic representation of the dynamics generated by the Langevin stochastic differential equation. The linear operator is then used to describe the evolution of the constructed coordinates on the state space of the dynamical system. The obtained coordinates are approximations of the eigenfunctions of the stochastic Koopman generator, so that the described approach is closely connected with the Koopman operator techniques for building the representation of the system dynamics.

In order to numerically approximate the spectral objects of the stochastic Koopman operator, we explore the possibility of application of DMD algorithms that were originally developed for evaluating KMD in deterministic systems. As already mentioned, algorithms that are typically used to extract relevant spectral information are, for example, the Schmid DMD algorithm (Schmid 2010; Schmid et al. 2011), the exact DMD algorithm (Tu et al. 2014), and the Arnoldi-like algorithm (Rowley et al. 2009). The application of DMD algorithm to noisy data is studied in Hemati et al. (2017) and Takeishi et al. (2017). In order to remove the bias errors produced by using the standard DMD algorithms on data with the observation noise that can arise, for example, as a consequence of imprecise measurements, Hemati et al. (2017) devel-

oped the total least-squares DMD. Takeishi et al. (2017) considered the numerical approximations of spectral objects of the stochastic Koopman operator for the RDS with observation noise by using the DMD algorithm, and proved its convergence to the stochastic Koopman operator under certain assumptions, following the work on deterministic systems in Arbabi and Mezić (2017). Due to the systematic error produced by the standard DMD algorithm they developed the version of the algorithm that also takes into account the observation noise and refer to it as the subspace DMD algorithm. Here we provide the convergence proof for the sHankel–DMD algorithm, an extension of the Hankel–DMD algorithm.

There exist some numerical issues that can arise in DMD algorithms applied to deterministic systems, leading to poor approximations of eigenvalues and eigenvectors. In order to overcome these difficulties, Drmač et al. (2018) proposed a data-driven algorithm for computing DMD, called DMD refined Rayleigh Ritz (DMD RRR) algorithm, which enables selection of Ritz pairs based on the data-driven computation of the residual, and substantially improves the quality of the retained spectral objects. We use it in the current context of RDS. In most considered RDS examples, we determine first the Koopman eigenvalues and eigenfunctions of the related deterministic dynamical system and then explore the algorithm behavior on the considered RDS. Despite the fact that in some cases we noticed high sensitivity of the numerical algorithm to the noise introduced into the system, in most cases very satisfactory approximations are obtained.

The paper is organized as follows. In Sect. 2, the definition of the stochastic Koopman operator family for the discrete and the continuous-time RDS is given. The classes of RDS that we consider are: the discrete-time RDS, the continuous-time RDS generated by the random differential equations (RDE) and the continuous-time RDS generated by the stochastic differential equations (SDE). In accordance with this classification of RDS, we provide some results that characterize the spectral objects of the Koopman operator associated with the linear RDS. In Sect. 3 we limit our considerations to the RDS for which the associated stochastic Koopman operator family is a semigroup. These are the RDS which can be identified with the time-homogeneous Markov stochastic processes. We also provide a characterization of the generators of the Koopman operator family. In Sect. 4, a brief description of the DMD RRR algorithm is provided and different approaches for its application to the RDS are described. We define the stochastic Hankel–DMD (sHankel–DMD) algorithm for approximating the eigenvalues and eigenfunctions of the stochastic Koopman operator and prove, under the assumption of ergodicity, its convergence to the true eigenvalues and eigenfunctions of the stochastic Koopman operator. Finally, the computations of Koopman eigenvalues and eigenfunctions by using the described DMD algorithms are illustrated on variety of numerical examples in Sect. 5.

## 2 Stochastic Koopman Operator and Linear RDS

First, we give the definition and a brief description of RDS, which follows terminology and results in Arnold (1998), Arnold and Crauel (1991).

Let  $(\Omega, \mathcal{F}, P)$  be a probability space and  $\mathbb{T}$  a semigroup (we can think of it as time). Suppose that  $\theta := (\theta(t))_{t \in \mathbb{T}}$  is a group or semigroup of measurable transformations of  $(\Omega, \mathcal{F}, P)$  preserving a measure  $P$  (i.e.,  $\theta(t)P = P$ ), such that a map  $(t, \omega) \rightarrow \theta(t)\omega$  is measurable. The quadruple  $(\Omega, \mathcal{F}, P, (\theta(t))_{t \in \mathbb{T}})$  is a metric dynamical system.

A random dynamical system (RDS) on the measurable space  $(M, \mathcal{B})$  over a dynamical system  $(\Omega, \mathcal{F}, P, (\theta(t))_{t \in \mathbb{T}})$  is a measurable map  $\varphi : \mathbb{T} \times \Omega \times M \rightarrow M$  satisfying a cocycle property with respect to  $\theta(\cdot)$ , which means that the mappings  $\varphi(t, \omega) := \varphi(t, \omega, \cdot) : M \rightarrow M$  form the cocycle over  $\theta(\cdot)$ , i.e.:

$$\varphi(0, \omega) = id_M, \quad \varphi(t + s, \omega) = \varphi(t, \theta(s)\omega) \circ \varphi(s, \omega), \quad \text{for all } s, t \in \mathbb{T}, \omega \in \Omega. \quad (1)$$

We call  $\theta$  a driving flow and  $(\Omega, \mathcal{F}, P, (\theta(t))_{t \in \mathbb{T}})$  a driving dynamical system.

If  $\mathbb{T}$  is discrete (for example,  $\mathbb{T} = \mathbb{Z}$  or  $\mathbb{T} = \mathbb{Z}^+$ ), we speak about the discrete-time RDS, if  $\mathbb{T}$  is continuous (for example,  $\mathbb{T} = \mathbb{R}$  or  $\mathbb{T} = \mathbb{R}^+$ ), we have the continuous-time RDS.

In the next definition, we introduce the Koopman operator family associated with the RDS.

**Definition 1** The stochastic Koopman operator  $\mathcal{K}^t$  associated with the RDS  $\varphi$  is defined on a space of functions (observables)  $f : M \rightarrow \mathbb{C}$  for which the functional

$$\mathcal{K}^t f(\mathbf{x}) = \mathbb{E}[f(\varphi(t, \omega)\mathbf{x})], \quad \mathbf{x} \in M \quad (2)$$

exists. We refer to the family of operators  $(\mathcal{K}^t)_{t \in \mathbb{T}}$  as to the stochastic Koopman operator family.

For example, if we work with continuous on a compact metric space and if the measure is finite, the functional above will exist. This follows from the fact that the continuous function on compact spaces is bounded, and its composition with measurable function results with bounded measurable function, which is integrable (Yosida 1980).

The material exposed in this work refers to the cases where  $M \subseteq \mathbb{R}^d$  and  $\mathcal{B}$  is the Borel  $\sigma$ -algebra of Borel sets in  $M$ , so that  $(M; \mathcal{B})$  is a standard space. For a fixed  $\mathbf{x} \in M$ ,  $(\varphi(t, \omega)\mathbf{x})_{t \in \mathbb{T}, \omega \in \Omega}$  is the  $M$ -valued stochastic process, which implies that an RDS defines a family of random processes with values in  $M$ . For each  $\mathbf{x} \in M$ , the corresponding probability measure  $\mathbb{P}_{\mathbf{x}}$  on the canonical space  $\bar{\Omega} \subseteq M^{\mathbb{T}}$  is induced by the probability measure  $P$  associated with the driving dynamical system and a process  $(\varphi(t, \omega)\mathbf{x})_{t \in \mathbb{T}, \omega \in \Omega}$  (see Dynkin 1965; Arnold 1998, Appendix A.2). We suppose then that the observable functions are measurable and integrable with respect to the induced probability measures.

In the case when spectral expansions of the Koopman operator are required, the specification of the space of the observable functions might become more detailed. For example, some Hilbert space will be required for the appropriate application of the numerical algorithms and in some examples (see Example 2).

**Definition 2** The observables  $\phi^t : M \rightarrow \mathbb{C}$  that satisfy equation

$$\mathcal{K}^t \phi^t(\mathbf{x}) = \lambda^S(t)\phi^t(\mathbf{x}) \quad (3)$$

we call the eigenfunctions of the stochastic Koopman operator, while the associated values  $\lambda^S(t)$  we call the stochastic Koopman eigenvalues.

The above definition can be viewed geometrically: the level sets of a stochastic Koopman operator eigenfunction are unique in the sense that the expectation of the value of the level set at time  $t$  is precisely  $\lambda^S(t)\phi^t(\mathbf{x})$ , i.e., they depend only on the level set that the dynamics started from at  $t = 0$ . An even stronger statement is available for eigenvalues on the unit circle: in that case, the precise location of the state of the system after time  $t$  is within a level set of the associated eigenfunction. For example, if  $\phi(\mathbf{x})$  is an eigenfunction associated with eigenvalue 1, then its level sets are invariant sets for the underlying RDS [see the proof in Mezić and Banaszuk (2004) for the discrete-time RDS].

There are three main types of RDS (Arnold 1998) that are particularly important from the point of view of practical applications. These are: the discrete-time RDS, the continuous-time RDS generated by RDE and the continuous-time RDS generated by SDE. In what follows we describe the general settings for each of these types and provide the results that characterize the eigenvalues and eigenfunctions of the stochastic Koopman operators associated with the linear RDS.

### 2.1 Stochastic Koopman Operator for Discrete-Time RDS

For a discrete-time RDS, we have  $\mathbb{T} = \mathbb{Z}^+ \cup \{0\}$ . Let  $(\Omega, \mathcal{F}, P, (\theta(t))_{t \in \mathbb{T}})$  be a given metric dynamical system and let denote  $\psi = \theta(1)$ . The discrete RDS  $\varphi(n, \omega)$  over  $\theta$  can be defined by the one step maps  $T(\omega, \cdot) : M \rightarrow M$

$$T(\omega, \cdot) := \varphi(1, \omega),$$

since by applying the cocycle property one gets

$$\varphi(n, \omega) = T(\psi^{n-1}(\omega), \cdot) \circ \dots \circ T(\psi(\omega), \cdot) \circ T(\omega, \cdot), \quad n \geq 1. \tag{4}$$

From the fact that maps  $\psi^i = \theta(i)$  preserve the measure  $P$ , it follows that the maps  $T(\psi^i(\omega), \cdot)$  are identically distributed, thus  $(T(\psi^i(\omega), \cdot))_{i \in \mathbb{T}}$  is a stationary sequence of random maps on  $M$  (Arnold 1998, Section 2.1). According to (1),  $\varphi(0, \omega) = id_M$ . The action of the discrete RDS  $\varphi(n, \omega)$  on  $\mathbf{x} \in M$  gives its value at  $n$ th step and we denote it as

$$T^n(\omega, \mathbf{x}) = \varphi(n, \omega)\mathbf{x}. \tag{5}$$

In the proposition that follows, we consider the eigenvalues and the eigenfunctions of the stochastic Koopman operator related to the discrete RDS induced by a linear map  $T$ .

**Proposition 1** *Suppose that  $\mathbf{A} : \Omega \rightarrow \mathbb{R}^{d \times d}$  is measurable and that the one step map  $T : \Omega \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  of a discrete RDS is defined by*

$$T(\omega, \mathbf{x}) = \mathbf{A}(\omega)\mathbf{x}. \tag{6}$$

Denote by  $\Phi(n, \omega)$  the linear RDS satisfying  $T^n(\omega, \mathbf{x}) = \Phi(n, \omega)\mathbf{x}$ , i.e.,

$$\Phi(n, \omega) = \mathbf{A}(\psi^{n-1}(\omega)) \cdots \mathbf{A}(\psi(\omega))\mathbf{A}(\omega).$$

Assume that  $\hat{\Phi}(n) = \mathbb{E}[\Phi(n, \omega)]$  are diagonalizable, with simple eigenvalues  $\hat{\lambda}_j(n)$  and left and right eigenvectors  $\hat{\mathbf{w}}_j^n, \hat{\mathbf{v}}_j^n, j = 1, \dots, d$ . Then the eigenfunctions of the stochastic Koopman operator  $\mathcal{K}^n$  are

$$\phi_j^n(\mathbf{x}) = \langle \mathbf{x}, \hat{\mathbf{w}}_j^n \rangle, \quad j = 1, \dots, d, \tag{7}$$

with the corresponding eigenvalues  $\lambda_j^S(n) = \hat{\lambda}_j(n)$ .

Moreover, if matrices  $\mathbf{A}(\omega), \omega \in \Omega$  commute and are diagonalizable with the simple eigenvalues  $\lambda_j(\omega)$  and corresponding left eigenvectors  $\mathbf{w}_j, j = 1, \dots, d$ , then

$$\hat{\mathbf{w}}_j^n = \mathbf{w}_j \quad \text{and} \quad \lambda_j^S(n) = \mathbb{E} \left[ \prod_{i=1}^n \lambda_j(\psi^{i-1}(\omega)) \right].$$

Furthermore,  $\hat{\mathbf{v}}_j^n, j = 1, \dots, d$  are the Koopman modes of the full-state observable and the following expansion is valid

$$\mathcal{K}^n \mathbf{x} = \sum_{j=1}^d \lambda_j^S(n) \langle \mathbf{x}, \hat{\mathbf{w}}_j^n \rangle \hat{\mathbf{v}}_j^n. \tag{8}$$

**Proof** The action of the stochastic Koopman operator on functions defined by (7) is equal to

$$\begin{aligned} \mathcal{K}^n \phi_j^n(\mathbf{x}) &= \mathbb{E}[\langle \Phi(n, \omega)\mathbf{x}, \hat{\mathbf{w}}_j^n \rangle] = \mathbb{E}[\langle \mathbf{x}, \Phi(n, \omega)^* \hat{\mathbf{w}}_j^n \rangle] = \langle \mathbf{x}, \mathbb{E}[\Phi(n, \omega)^*] \hat{\mathbf{w}}_j^n \rangle \\ &= \langle \mathbf{x}, \hat{\Phi}(n)^* \hat{\mathbf{w}}_j^n \rangle = \hat{\lambda}_j(n) \langle \mathbf{x}, \hat{\mathbf{w}}_j(n) \rangle = \hat{\lambda}_j(n) \phi_j^n(\mathbf{x}), \end{aligned} \tag{9}$$

where we have used that

$$\mathbb{E}[\Phi(n, \omega)^*] = (\mathbb{E}[\Phi(n, \omega)])^* = \hat{\Phi}(n)^*.$$

In the case when matrices  $\mathbf{A}(\omega), \omega \in \Omega$  commute and are diagonalizable, they are simultaneously diagonalizable (see Horn and Johnson 2013, Theorem 1.3.12), i.e., there exists a single invertible matrix  $\mathbf{V}$ , so that  $\mathbf{A}(\omega) = \mathbf{V}\Lambda(\omega)\mathbf{V}^{-1}$ , where  $\Lambda(\omega) = \text{diag}(\lambda_1(\omega), \dots, \lambda_d(\omega))$ . It is clear that the columns of the matrix  $\mathbf{V}$  are the common right eigenvectors of the matrices  $\mathbf{A}(\omega), \omega \in \Omega$ , and that  $\mathbf{W}^* = \mathbf{V}^{-1}$ , where  $\mathbf{W}$  is the matrix of the left eigenvectors. It is straightforward that

$$\Phi(n, \omega) = \mathbf{V} \prod_{i=1}^n \Lambda(\psi^{i-1}(\omega)) \mathbf{W}^*.$$



Therefore

$$\hat{\Phi}(n) = \mathbf{V} \mathbb{E} \left[ \prod_{i=1}^n \Lambda(\psi^{i-1}(\omega)) \right] \mathbf{W}^*,$$

and we easily conclude that  $\hat{\mathbf{w}}_j = \mathbf{w}_j$  and  $\lambda_j^S(n) = \mathbb{E} \left[ \prod_{i=1}^n \lambda_j(\psi^{i-1}(\omega)) \right]$ .

Equation (8) can be easily proved by using the decomposition of the state  $\mathbf{x}$  in the base  $\hat{\mathbf{w}}_j^n, \hat{\mathbf{v}}_j^n, j = 1, \dots, d$ , the linearity of the Koopman operator, and its action on the eigenfunctions given with (7), i.e.,

$$\mathcal{K}^n \mathbf{x} = \mathcal{K}^n \sum_{j=1}^d \langle \mathbf{x}, \hat{\mathbf{w}}_j^n \rangle \hat{\mathbf{v}}_j^n = \sum_{j=1}^d \mathcal{K}^n \langle \mathbf{x}, \hat{\mathbf{w}}_j^n \rangle \hat{\mathbf{v}}_j^n = \sum_{j=1}^d \lambda_j^S(n) \langle \mathbf{x}, \hat{\mathbf{w}}_j^n \rangle \hat{\mathbf{v}}_j^n. \tag{10}$$

□

**Remark 1** We will use the term principal eigenfunctions for the eigenfunctions of the form  $\phi_j(\mathbf{x}) = \langle \mathbf{x}, \mathbf{w}_j \rangle, j = 1, \dots, d$  that appear in Proposition 1 and in some other propositions in the rest of the paper. Also, the eigenvalues corresponding to them we call the principal eigenvalues. It follows from (8) that in the linear case, the action of the Koopman operators on the full-state observable can be derived by using just the principal eigenfunctions, eigenvalues and modes.

### 2.2 Stochastic Koopman Operator for RDS Generated by the Random Differential Equations

Suppose that  $\mathbb{T} = \mathbb{R}$  or  $\mathbb{T} = \mathbb{R}^+ \cup \{0\}$ . Let  $(\Omega, \mathcal{F}, P, (\theta(t))_{t \in \mathbb{T}})$  be a metric dynamical system. We consider a continuous-time RDS generated by the random differential equation (RDE) of the following form

$$\dot{\mathbf{x}} = F(\theta(t)\omega, \mathbf{x}), \tag{11}$$

defined on the manifold  $M \subseteq \mathbb{R}^d$ , where  $\theta(t)\omega$  is associated with the random dynamics. In this type of equations, the randomness refers just to the random parameters, which do not depend on the state of the system. RDE (11) generates an RDS  $\varphi$  over  $\theta$ , whose action is defined by

$$\varphi(t, \omega)\mathbf{x} = \mathbf{x} + \int_0^t F(\theta(s)\omega, \varphi(s, \omega)\mathbf{x}) ds. \tag{12}$$

The properties of this RDS under different regularity properties of the function  $F$  can be found in Arnold (1998, Section 2.2). A set of trajectories starting at  $\mathbf{x}$  that are generated by (11) is given by  $\varphi(t, \omega)\mathbf{x}$  and it defines the family of random variables. We say that this is a solution of the RDE with the initial condition  $\varphi(0, \omega)\mathbf{x} = \mathbf{x}$ . Since the solutions of the RDE are defined pathwise, for each fixed  $\omega$  the trajectory

can be determined as a solution of deterministic ordinary differential equation, so that the RDE (11) can be seen as a family of ordinary differential equations.

In the next proposition, we derive the eigenvalues and the eigenfunctions of the stochastic Koopman operators associated with the linear RDS.

**Proposition 2** *If  $\mathbf{A} : \Omega \rightarrow \mathbb{R}^{d \times d}$  and  $\mathbf{A} \in L^1(\Omega, \mathcal{F}, P)$ , then RDE*

$$\dot{\mathbf{x}} = \mathbf{A}(\theta(t)\omega)\mathbf{x}, \tag{13}$$

generates a linear RDS  $\Phi$  satisfying

$$\Phi(t, \omega) = \mathbf{I} + \int_0^t \mathbf{A}(\theta(s)\omega)\Phi(s, \omega)ds. \tag{14}$$

Assume that  $\hat{\Phi}(t) = \mathbb{E}[\Phi(t, \omega)]$  is diagonalizable, with simple eigenvalues  $\hat{\mu}_j^t$  and left eigenvectors  $\hat{\mathbf{w}}_j^t$ ,  $j = 1, \dots, d$ . Then

$$\phi_j^t(\mathbf{x}) = \langle \mathbf{x}, \hat{\mathbf{w}}_j^t \rangle, \quad j = 1, \dots, d, \tag{15}$$

are the principal eigenfunctions of the stochastic Koopman operator  $\mathcal{K}^t$  with corresponding principal eigenvalues  $\lambda_j^S(t) = \hat{\mu}_j^t$ ,  $j = 1, \dots, d$ .

Moreover, if matrices  $\mathbf{A}(\omega)$  commute and are diagonalizable with the simple eigenvalues  $\lambda_j(\omega)$  and corresponding left eigenvectors  $\mathbf{w}_j$ ,  $j = 1, \dots, d$ , then

$$\hat{\mathbf{w}}_j^t = \mathbf{w}_j \quad \text{and} \quad \lambda_j^S(t) = \mathbb{E} \left[ e^{\int_0^t \lambda_j(\theta(s)\omega)ds} \right].$$

Furthermore,  $\hat{\mathbf{v}}_j^t$ ,  $j = 1, \dots, d$  are the Koopman modes of the full-state observable and the following expansion is valid

$$\mathcal{K}^t \mathbf{x} = \sum_{j=1}^d \lambda_j^S(t) \langle \mathbf{x}, \hat{\mathbf{w}}_j^t \rangle \hat{\mathbf{v}}_j^t. \tag{16}$$

**Proof** The first part of the proposition follows from (Arnold 1998, Example 2.2.8),

Furthermore, the action of the stochastic Koopman operator on functions defined by (15) is equal to

$$\begin{aligned} \mathcal{K}^t \phi_j^t(\mathbf{x}) &= \mathbb{E}[\langle \Phi(t, \omega)\mathbf{x}, \hat{\mathbf{w}}_j^t \rangle] = \mathbb{E}[\langle \mathbf{x}, \Phi^*(t, \omega)\hat{\mathbf{w}}_j^t \rangle] = \langle \mathbf{x}, \mathbb{E}[\Phi^*(t, \omega)]\hat{\mathbf{w}}_j^t \rangle \\ &= \langle \mathbf{x}, \hat{\Phi}^*(t)\hat{\mathbf{w}}_j^t \rangle = \hat{\mu}_j^t \langle \mathbf{x}, \hat{\mathbf{w}}_j^t \rangle = \hat{\mu}_j^t \phi_j^t(\mathbf{x}) \end{aligned} \tag{17}$$

With the same argument as in the proof of Proposition 1, we have  $\mathbf{A}(\omega) = \mathbf{V}\mathbf{\Lambda}(\omega)\mathbf{W}^*$ , where  $\mathbf{W}$  and  $\mathbf{V}$  are matrices of common left and right eigenvectors, and  $\mathbf{\Lambda}(\omega) = \text{diag}(\lambda_1(\omega), \dots, \lambda_d(\omega))$ . It is straightforward that

$$\Phi(t, \omega) = e^{\int_0^t \mathbf{A}(\theta(s)\omega)ds} = \mathbf{V}e^{\int_0^t \mathbf{\Lambda}(\theta(s)\omega)ds}\mathbf{W}^*,$$

and

$$\hat{\Phi}(t) = \mathbf{V} \mathbb{E} \left[ e^{\int_0^t \Lambda(\theta(s)\omega) ds} \right] \mathbf{W}^*.$$

We easily conclude that  $\hat{\mathbf{w}}_j = \mathbf{w}_j$  and  $\lambda_j^S(t) = \mathbb{E} \left[ e^{\int_0^t \lambda_j(\theta(s)\omega) ds} \right]$ . The proof of (16) is same as the proof of (8) in Proposition 2.  $\square$

**Example 1** (Linear scalar RDS) Suppose that linear scalar RDE is given by

$$\dot{x} = a(\omega)x, \tag{18}$$

where  $a : \Omega \rightarrow \mathbb{R}$  is random variable with finite moments. Observe that  $\theta(t) = id$ , which means that the probability space does not change with time. If the moment generating function defined by  $\mathcal{M}_a(t) = \mathbb{E}[e^{ta(\omega)}]$  is analytic for  $|t| < R$ , then for the initial condition  $\varphi(0, \omega)x = x$  and  $t < R$  there exists a unique solution of (18), which can be expressed as (Strand 1970)

$$\varphi(t, \omega)x = x e^{\int_0^t a(\omega) ds}. \tag{19}$$

The action of the stochastic Koopman operator on the full-state observable function  $\phi(x) = x$  is then

$$\mathcal{K}^t x = \mathbb{E} [\varphi(t, \omega)x] = \mathbb{E} \left[ x e^{\int_0^t a(\omega) ds} \right] = \mathbb{E} \left[ e^{a(\omega)t} \right] x. \tag{20}$$

Thus,  $\phi(x) = x$  is the eigenfunction of the stochastic Koopman operator and the corresponding eigenvalue satisfies

$$\lambda^S(t) = \mathbb{E} \left[ e^{a(\omega)t} \right] = \mathcal{M}_a(t).$$

### 2.3 Stochastic Koopman Operator for RDS Generated by the Stochastic Differential Equations

Let  $\mathbb{T} = \mathbb{R}^+$ ,  $M = \mathbb{R}^d$  and  $T \in \mathbb{T}$ . Suppose that the stochastic process  $X_t(\omega)$ ,  $t \in [0, T]$ ,  $\omega \in \Omega$  is obtained as a solution of the nonautonomous stochastic differential equation (SDE)

$$dX_t = G(t, X_t)dt + \sigma(t, X_t)dW_t, \tag{21}$$

where  $G : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  and  $\sigma : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^{d \times r}$  are  $L^2$  measurable. Here  $W_t = (W_t^1, \dots, W_t^r)$  denotes the  $r$ -dimensional Wiener process with independent components and standard properties, i.e.,  $\mathbb{E}(W_t^i) = 0$ ,  $i = 1, \dots, r$ ,  $\mathbb{E}(W_t^i W_s^j) = \min\{t, s\} \delta_{ij}$ ,  $i, j = 1, \dots, r$  ( $\delta_{ij}$  is the Kronecker delta symbol). The solution  $X_t(\omega)$  of (21) with the initial condition  $X_{t_0}(\omega)$  is formally defined in terms of Itô integral as

(see Arnold 1974, Chapter 6)

$$X_t(\omega) = X_{t_0}(\omega) + \int_{t_0}^t G(s, X_s(\omega))ds + \int_{t_0}^t \sigma(s, X_s(\omega))dW_s. \tag{22}$$

The probability space on which the process is considered can be identified with  $\Omega = C_0(\mathbb{R}^+, \mathbb{R}^r)$  (space of continuous functions satisfying  $\omega(t_0) = 0$ ). Then,  $\omega \in \Omega$  is identified with the canonical realization of the Wiener process such that  $\omega(t) = W_t(\omega)$ . If  $\mathcal{F}$  is the Borel  $\sigma$ -algebra,  $P$  the measure generated by the Wiener process, and  $\theta(t)$  defined by the ‘‘Wiener shifts’’

$$\theta(t)\omega(\cdot) = \omega(t + \cdot) - \omega(t), \tag{23}$$

$(\Omega, \mathcal{F}, P, (\theta(t))_{t \in \mathbb{T}})$  becomes a metric dynamical system (see Arnold 1998, Appendix A). It is a driving dynamical system for the two-parameter family of RDS  $\varphi(t, t_0, \omega)$  that is given by

$$\varphi(t, t_0, \omega)\mathbf{x} = X_t(\omega),$$

where  $X_t(\omega)$  is given with (22) for the initial condition  $X_{t_0}(\omega) = \mathbf{x}$ . The basic existence and uniqueness results for the SDE of the form (21) that are valid under different regularity and boundedness properties of the functions  $G$  and  $\sigma$  can be found in Arnold (1974, Section 6.3), Pavliotis (2014, Section 3.3).

The Koopman operator family  $\mathcal{K}^{t,t_0}$  related to this RDS is defined by

$$\mathcal{K}^{t,t_0} f(\mathbf{x}) = \mathbb{E}[f(\varphi(t, t_0, \omega)\mathbf{x})]. \tag{24}$$

In this more general setting with the two-parameter family of Koopman operators (24), the eigenfunctions  $\phi^{t,t_0} : M \rightarrow \mathbb{C}$  and eigenvalues  $\lambda^S(t, t_0)$  of the Koopman operator  $\mathcal{K}^{t,t_0}$  defined on a finite-time interval satisfy

$$\mathcal{K}^{t,t_0} \phi^{t,t_0}(\mathbf{x}) = \lambda^S(t, t_0) \phi^{t,t_0}(\mathbf{x}). \tag{25}$$

The following two propositions treat two classes of linear SDE. In the first one, the random part of equations models the additive noise, and in the second one, it models the multiplicative noise.

**Proposition 3** *Let the linear SDE with additive noise be defined by*

$$dX_t = \mathbf{A}(t)X_t dt + \sum_{i=1}^r b^i(t) dW_t^i, \tag{26}$$

where  $\mathbf{A}(t)$  is  $d \times d$  matrix of functions and  $b^i(t)$ ,  $i = 1, \dots, r$  are  $d$ -dimensional vector functions. Assume that the fundamental matrix  $\Phi(t, t_0)$  satisfying the matrix differential equation

$$\dot{\Phi} = \mathbf{A}(t)\Phi, \quad \Phi(t_0) = \mathbf{I} \tag{27}$$

is diagonalizable, with simple eigenvalues  $\hat{\mu}_j^{t,t_0}$  and left eigenvectors  $\hat{\mathbf{w}}_j^{t,t_0}$ ,  $j = 1, \dots, d$ . Then

$$\phi_j^{t,t_0}(\mathbf{x}) = \langle \mathbf{x}, \hat{\mathbf{w}}_j^{t,t_0} \rangle, \quad j = 1, \dots, d, \tag{28}$$

are the eigenfunctions of the stochastic Koopman operator  $\mathcal{K}^{t,t_0}$ , with corresponding eigenvalues

$$\lambda_j^S(t, t_0) = \hat{\mu}_j^{t,t_0}.$$

If matrices  $\mathbf{A}(t)$  commute and are diagonalizable with the simple eigenvalues  $\lambda_j(t)$  and corresponding left eigenvectors  $\mathbf{w}_j$ ,  $j = 1, \dots, d$ , then

$$\hat{\mathbf{w}}_j^{t,t_0} = \mathbf{w}_j \quad \text{and} \quad \lambda_j^S(t, t_0) = e^{\int_{t_0}^t \lambda_j(s) ds}. \tag{29}$$

**Proof** Since the solution of (26) with the initial condition  $X_{t_0}(\omega) = \mathbf{x}$  is given by, (see Arnold 1974, Section 8.2)

$$X_t(\omega) = \Phi(t, t_0) \left( \mathbf{x} + \sum_{i=1}^r \int_{t_0}^t \Phi^{-1}(s, t_0) b^i(s) dW_s^i \right), \tag{30}$$

we have

$$\begin{aligned} \mathcal{K}^{t,t_0} \phi_j^{t,t_0}(\mathbf{x}) &= \mathbb{E} \left[ \phi_j^{t,t_0}(X_t(\omega)) \right] \\ &= \mathbb{E} \left[ \langle \Phi(t, t_0) \mathbf{x}, \hat{\mathbf{w}}_j^{t,t_0} \rangle + \left\langle \sum_{i=1}^r \int_{t_0}^t \Phi(t, t_0) \Phi^{-1}(s, t_0) b^i(s) dW_s^i, \hat{\mathbf{w}}_j^{t,t_0} \right\rangle \right] \\ &= \langle \Phi(t, t_0) \mathbf{x}, \hat{\mathbf{w}}_j^{t,t_0} \rangle + \sum_{i=1}^r \left\langle \mathbb{E} \left[ \int_{t_0}^t \Phi(t, t_0) \Phi^{-1}(s, t_0) b^i(s) dW_s^i \right], \hat{\mathbf{w}}_j^{t,t_0} \right\rangle \\ &= \hat{\mu}_j^{t,t_0} \langle \mathbf{x}, \hat{\mathbf{w}}_j^{t,t_0} \rangle = \hat{\mu}_j^{t,t_0} \phi_j^{S,t,t_0}(\mathbf{x}), \end{aligned} \tag{31}$$

where we used that  $\mathbb{E} \left[ \int_{t_0}^t \mathbf{F}(s) dW_s \right] = 0$  (see Arnold 1974, Theorem 4.4.14) applied to  $\mathbf{F}(s) = \Phi(t, t_0) \Phi^{-1}(s, t_0) b^i(s)$ . With this we proved the first statement.

Since, in the commutative case, the fundamental matrix can be expressed in the form  $\Phi(t, t_0) = e^{\int_{t_0}^t \mathbf{A}(s) ds}$ , its eigenvectors coincide with the eigenvectors of the matrix  $\mathbf{A}(t)$  and eigenvalues are given by (29).  $\square$

**Proposition 4** Let the linear SDE with multiplicative noise be defined by

$$dX_t = \mathbf{A}(t) X_t dt + \sum_{i=1}^r \mathbf{B}^i(t) X_t dW_t^i, \tag{32}$$

where  $\mathbf{A}(t)$ ,  $\mathbf{B}^i(t)$ ,  $i = 1, \dots, r$  are  $d \times d$  matrices of functions. Denote with  $\Phi(t, t_0)$  the fundamental matrix satisfying the matrix SDE

$$d\Phi = \mathbf{A}\Phi dt + \sum_{i=1}^r \mathbf{B}^i(t)\Phi dW_t^i, \quad \Phi(t_0) = \mathbf{I} \quad (33)$$

and assume that  $\hat{\Phi}(t, t_0) = \mathbb{E}[\Phi(t, t_0)]$  is diagonalizable, with simple eigenvalues  $\hat{\mu}_j^{t,t_0}$  and left eigenvectors  $\hat{\mathbf{w}}_j^{t,t_0}$ ,  $j = 1, \dots, d$ . Then

$$\phi_j^{t,t_0}(\mathbf{x}) = \langle \mathbf{x}, \hat{\mathbf{w}}_j^{t,t_0} \rangle, \quad j = 1, \dots, d, \quad (34)$$

are the eigenfunctions of the stochastic Koopman operator  $\mathcal{K}^{t,t_0}$ , with corresponding eigenvalues

$$\lambda_j^S(t, t_0) = \hat{\mu}_j^{t,t_0}.$$

If the matrices  $\mathbf{A}(t)$ ,  $\mathbf{B}^i(t)$ ,  $i = 1, \dots, r$  commute, i.e., if  $\mathbf{A}(t)\mathbf{A}(s) = \mathbf{A}(s)\mathbf{A}(t)$ ,  $\mathbf{A}(t)\mathbf{B}^i(s) = \mathbf{B}^i(s)\mathbf{A}(t)$ ,  $\mathbf{B}^i(t)\mathbf{B}^j(s) = \mathbf{B}^j(s)\mathbf{B}^i(t)$  for  $i, j = 1, \dots, r$  and all  $s, t$ , and if the matrices  $\mathbf{A}(t)$  are diagonalizable with the simple eigenvalues  $\lambda_j(t)$  and corresponding left eigenvectors  $\mathbf{w}_j$ ,  $j = 1, \dots, d$ , then

$$\hat{\mathbf{w}}_j^{t,t_0} = \mathbf{w}_j \quad \text{and} \quad \lambda_j^S(t, t_0) = e^{\int_{t_0}^t \lambda_j(s) ds}. \quad (35)$$

**Proof** For the fundamental matrix  $\Phi(t, t_0)$  and the initial condition  $X_{t_0}(\omega) = \mathbf{x}$ , the solution of (32) is equal to, (see Arnold 1974, Section 8.5)

$$X_t(\omega) = \Phi(t, t_0)\mathbf{x}, \quad (36)$$

thus

$$\begin{aligned} \mathcal{K}^{t,t_0} \phi_j^{t,t_0}(\mathbf{x}) &= \mathbb{E} \left[ \phi_j^{t,t_0}(X_t(\omega)) \right] = \mathbb{E} \left[ \langle \Phi(t, t_0)\mathbf{x}, \hat{\mathbf{w}}_j^{t,t_0} \rangle \right] \\ &= \langle \hat{\Phi}(t, t_0)\mathbf{x}, \hat{\mathbf{w}}_j^{t,t_0} \rangle = \mu_j^{t,t_0} \langle \mathbf{x}, \hat{\mathbf{w}}_j^{t,t_0} \rangle = e^{\hat{\lambda}_j(t,t_0)} \phi_j^{t,t_0}(\mathbf{x}), \end{aligned} \quad (37)$$

For the case with commutative matrices  $\mathbf{A}(t)$  and  $\mathbf{B}^i(t)$ ,  $i = 1, \dots, r$ , the fundamental matrix  $\Phi(t, t_0)$  can be expressed in an explicit form as, (see Arnold 1974, Sections 8.4, 8.5)

$$\Phi(t, t_0) = e^{\int_{t_0}^t \left( \mathbf{A}(s) - \frac{1}{2} \sum_{i=1}^r \mathbf{B}^i(s)\mathbf{B}^i(s)^T \right) ds + \int_{t_0}^t \sum_{i=1}^r \mathbf{B}^i(s) dW_s^i}. \quad (38)$$

Since

$$\begin{aligned} \hat{\Phi}(t, t_0) &= \mathbb{E}[\Phi(t, t_0)] = e^{\int_{t_0}^t \left( \mathbf{A}(s) - \frac{1}{2} \sum_{i=1}^r \mathbf{B}^i(s)\mathbf{B}^i(s)^T \right) ds} \mathbb{E} \left[ e^{\int_{t_0}^t \sum_{i=1}^r \mathbf{B}^i(s) dW_s^i} \right] \\ &= e^{\int_{t_0}^t \left( \mathbf{A}(s) - \frac{1}{2} \sum_{i=1}^r \mathbf{B}^i(s)\mathbf{B}^i(s)^T \right) ds} e^{\int_{t_0}^t \frac{1}{2} \sum_{i=1}^r \mathbf{B}^i(s)\mathbf{B}^i(s)^T ds} = e^{\int_{t_0}^t \mathbf{A}(s) ds}, \end{aligned} \quad (39)$$

the eigenvectors of  $\hat{\Phi}(t, t_0)$  coincide with the eigenvectors of the matrix  $\mathbf{A}(t)$  and eigenvalues are given by (35), which proves the statement. Here we used the fact that  $\mathbb{E} \left[ e^{\int_{t_0}^t \mathbf{B}(s) dW_s} \right] = e^{\frac{1}{2} \int_{t_0}^t \mathbf{B}(s) \mathbf{B}(s)^T ds}$  (see Arnold 1974, Theorem 8.4.5).  $\square$

### 3 Semigroups of Koopman Operators and Their Generators

In this section, we are interested in RDS for which the associated stochastic Koopman operator family satisfies the semigroup property. Strong results for the properties of the stochastic Koopman operator family can be obtained in more specific settings such as linear setting we analyzed before and Markov setting. Therefore, we limit here to the particular types of RDS and consider only those that are Markovian (Crauel 1991), or even more precisely, to those RDS for which the family of processes  $\{(\varphi(t, \omega)\mathbf{x}, \mathbf{x} \in M)\}$  is a time-homogeneous Markov family. Denote by  $\mathcal{F}_t^{\mathbf{x}}$  the  $\sigma$ -algebra generated by the “past” of the stochastic process, i.e.,  $\mathcal{F}_t^{\mathbf{x}} = \sigma\{\varphi(s, \omega)\mathbf{x}, s \leq t\}$ . The Markov property implies that for  $s \leq t$  and every random variable  $Y$ , measurable with respect to  $\mathcal{F}_t^{\mathbf{x}}$ , the following relation holds

$$\mathbb{E}[Y | \mathcal{F}_s^{\mathbf{x}}] = \mathbb{E}[Y | \varphi(s, \omega)\mathbf{x}]. \tag{40}$$

Moreover, for such  $Y$ , the following equality, known as the Chapman–Kolmogorov equation, is valid

$$\mathbb{E}[Y | \varphi(s, \omega)\mathbf{x}] = \mathbb{E}[\mathbb{E}[Y | \varphi(t, \omega)\mathbf{x}] | \varphi(s, \omega)\mathbf{x}]. \tag{41}$$

The Chapman–Kolmogorov equation implies the semigroup property of the stochastic Koopman operators, i.e.,

$$\mathcal{K}^{t+s} = \mathcal{K}^s \circ \mathcal{K}^t. \tag{42}$$

For the settings of Markov process, the Koopman operator is known under the name Markov propagator or transition operator and its properties have been studied for a long time (Dynkin 1965; Yosida 1980).

There are two important and well-known classes of RDS, which can be identified with the Markov family. These are the discrete-time RDS generated by an independent identically distributed process and the continuous-time RDS generated by the stochastic differential equations where the noise is modeled using Wiener process. In each of these cases the probability space associated with the stochastic process modeling the noise can be identified with the canonical measure-preserving dynamical system (Arnold 1998, Appendix A.2).

In what follows we describe briefly the canonical dynamical system  $(\theta(t))_{t \in \mathbb{T}}$  that is induced by some given stochastic process  $\tilde{\xi}$ . We will use this concept in the next subsection. Let suppose that  $\tilde{\xi} = (\tilde{\xi}_t)_{t \in \mathbb{T}}, \tilde{\xi}_t : \tilde{\Omega} \rightarrow B$  is a  $B$ -valued stochastic process on a probability space  $\tilde{\Omega}$ , where  $(B, \mathcal{B})$  is a measurable state space. The given process and the probability measure on  $\tilde{\Omega}$  induce a probability measure  $P$  on  $B^{\mathbb{T}}$ , so that one can define a new probability space  $(\Omega, \mathcal{F}, P) = (B^{\mathbb{T}}, \mathcal{B}^{\mathbb{T}}, P)$  where  $\mathcal{B}^{\mathbb{T}}$  is the

$\sigma$ -algebra generated by the collection of the cylinder sets. The canonical realization  $\xi$  of the stochastic process  $\tilde{\xi}$  is defined on  $(\Omega, \mathcal{F}, P)$  by the coordinate functions  $\xi_t(\omega) = \omega(t)$ ,  $\omega \in \Omega$ . The shift transformations  $\theta(t) : B^{\mathbb{T}} \rightarrow B^{\mathbb{T}}$  given by

$$\theta(t)\omega(\cdot) := \omega(t + \cdot), \quad t \in \mathbb{T}, \quad (43)$$

constitute the semigroup or group of measurable transformations. Notice that the canonical realization  $\xi_t(\omega)$  can be viewed as the composition of the shift transformation (43) and of the canonical projection  $\pi : B^{\mathbb{T}} \rightarrow B$  defined by

$$\pi(\omega(\cdot)) = \omega(0), \quad (44)$$

i.e.,  $\xi_t(\omega) = \omega(t) = \pi(\theta(t)\omega)$ .

If time  $\mathbb{T}$  is discrete ( $\mathbb{T} = \mathbb{Z}$  or  $\mathbb{T} = \mathbb{Z}^+$ ) a map  $(t, \omega) \rightarrow \theta(t)\omega$  is measurable and  $(\theta(t))_{t \in \mathbb{T}}$  is a measurable dynamical system. In the continuous-time case ( $\mathbb{T} = \mathbb{R}$  or  $\mathbb{T} = \mathbb{R}^+$ ), when  $B = \mathbb{R}^m$  or when  $B$  is a Polish space,  $(\theta(t))_{t \in \mathbb{T}}$  could become a measurable dynamical system after some redefinition of the probability measure  $P$  and of the  $\sigma$ -algebra set (see Arnold 1998, Appendix A.2).

### 3.1 Discrete-Time RDS

Let now assume that for  $\mathbb{T} = \mathbb{Z}^+ \cup \{0\}$ ,  $\omega = (\omega_i)_{i \in \mathbb{T}}$  is a canonical realization of the stochastic process with the associated driving system composed by the shift transformation maps (43) as described in the previous paragraph. If we assume that the discrete RDS  $\varphi(n, \omega)$  is defined by the one step map  $T(\omega, \cdot) : M \rightarrow M$  of the form

$$T(\omega, \cdot) = T_0(\pi(\omega), \cdot), \quad (45)$$

where  $\pi$  denotes the canonical projection (44), by taking into account (45) in (4), we get

$$\varphi(n, \omega) = T_0(\pi(\psi^{n-1}(\omega)), \cdot) \circ \cdots \circ T_0(\pi(\psi(\omega)), \cdot) \circ T_0(\pi(\omega), \cdot), \quad n \geq 1. \quad (46)$$

If  $\omega$  is an i.i.d. stochastic process, the sequence  $\{T(\psi^i(\omega), \cdot) = T_0(\pi(\psi^{i-1}(\omega)), \cdot), i \geq 1\}$  is an i.i.d. sequence of random maps, so that RDS (46) generates the time-homogeneous Markov family  $\{\varphi(n, \omega)\mathbf{x}, \mathbf{x} \in M\}$  (Arnold 1998, Chapter 2.1). In this case the stochastic Koopman operator family is a semigroup, so that  $\mathcal{K}^n = \mathcal{K}^1 \circ \cdots \circ \mathcal{K}^1 = (\mathcal{K}^1)^n$ ,  $n > 0$ . Therefore one can think about  $\mathcal{K}^1$  as the generator of the stochastic Koopman semigroup, and we denote it by  $\mathcal{K}^S$ . According to (2),  $\mathcal{K}^1$  is determined by using the one step map  $T(\omega, \cdot)$  as

$$\mathcal{K}^S f(\mathbf{x}) = \mathcal{K}^1 f(\mathbf{x}) = \mathbb{E}[f(T(\omega, \mathbf{x}))]. \quad (47)$$

It follows from the semigroup property that if  $\lambda^S$  is the eigenvalue of the stochastic Koopman generator with the associated eigenfunction  $\phi(x)$ , then  $(\lambda^S)^n$  and  $\phi(x)$  are the eigenvalue and the eigenfunction of the operator  $\mathcal{K}^n$ .



**Remark 2** Note that in the more general case presented in Sect. 2.1, the random maps were identically distributed but not necessarily independent, so that the future state of the process obtained by the action of their composition (4) could depend on the past behavior of the system and not only on the present state.

**Example 2** (Noisy rotation on the circle) We describe here the example considered in Junge et al. (2004). A deterministic dynamical system representing the rotation on the unit circle  $S^1$  is defined by

$$T(x) = x + \vartheta, \tag{48}$$

where  $\vartheta \in S^1$  is constant number. We consider here its stochastic perturbation, i.e., a discrete RDS over the dynamical system  $\theta = (\theta(t))_{t \in \mathbb{T}}$ , where  $\theta(t)$  are shift transformations (43), defined by the one step map  $T : \Omega \times S^1 \rightarrow S^1$  of the form (45):

$$T(\omega, x) = x + \vartheta + \pi(\omega). \tag{49}$$

Here  $\omega \in [-\delta/2, \delta/2]^{\mathbb{Z}}$  is a canonical realization of a stochastic process and  $\pi(\cdot)$  is the canonical projection defined by (44). We suppose that the coordinates  $\omega_i$  are i.i.d. with uniform distribution on the interval  $[-\delta/2, \delta/2]$  for some  $\delta > 0$ . According to (47), the action of the associated stochastic Koopman generator on an observable function  $f : S^1 \rightarrow \mathbb{C}$  is given by

$$\mathcal{K}^S f(x) = \mathbb{E}[f(T(\omega, x))] = \frac{1}{\delta} \int_{-\delta/2}^{\delta/2} f(x + \vartheta + \omega_0) d\omega_0. \tag{50}$$

For the functions

$$\phi_j(x) = \exp(i2\pi jx), \quad j \in \mathbb{Z}, \tag{51}$$

the following equality holds

$$\begin{aligned} \mathcal{K}^S \phi_j(x) &= \frac{1}{\delta} \int_{-\delta/2}^{\delta/2} \exp(i2\pi j(x + \vartheta + \omega_0)) d\omega_0 \\ &= \frac{\sin(j\pi\delta)}{j\pi\delta} \exp(i2\pi j\vartheta) \exp(i2\pi jx) \\ &= \frac{\sin(j\pi\delta)}{j\pi\delta} \exp(i2\pi j\vartheta) \phi_j(x). \end{aligned} \tag{52}$$

We easily conclude that (51) are the eigenfunctions of the stochastic Koopman generator with corresponding eigenvalues

$$\lambda_j^S = \frac{\sin(j\pi\delta)}{j\pi\delta} \exp(i2\pi j\vartheta), \quad j \in \mathbb{Z} \tag{53}$$

For any function  $f : L^2(S^1) \rightarrow \mathbb{C}$ , we have the spectral expansion

$$\mathcal{K}^n f(x) = \sum_{j \in \mathbb{Z}} c_j \left( \frac{\sin(j\pi\delta)}{j\pi\delta} \right)^n \exp(i2\pi jn\vartheta) \exp(i2\pi jx). \tag{54}$$

where  $c_j$  are the Fourier coefficients of  $f$ . Clearly,  $\mathcal{K}^n f(x) \rightarrow c_0$  as  $t \rightarrow \infty$ . It is known that the eigenvalues of the Koopman generator related to the deterministic dynamical system (48) lie on the unit circle and are equal to  $\lambda_j = \exp(i2\pi j\vartheta)$ , while the eigenfunctions are the same as in the random case. Moreover, it is interesting to observe that for rational  $\vartheta$  the eigenspaces in deterministic case, i.e., for  $\delta = 0$ , are infinite dimensional, while they are finite dimensional in the stochastic case ( $\delta > 0$ ) due to the compactness of the operator. To be more precise, in this example, the eigenspaces become one dimensional.

**Remark 3** Consider the case when randomness is additive, i.e., when the one-step map  $T : \Omega \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  is given by

$$T(\omega, \mathbf{x}) = \mathbf{A}\mathbf{x} + \pi(\theta(t)\omega),$$

where  $\theta(t)$  is shift transformation and  $\omega : \Omega \rightarrow \mathbb{R}^d$  is the canonical realization of a process with i.i.d. components. Suppose that  $\mathbb{E}[\omega] = 0$  and that the matrix  $\mathbf{A}$  is diagonalizable with simple eigenvalues  $\lambda_j, j = 1, \dots, d$ . Then the eigenfunctions of the stochastic Koopman generator  $\mathcal{K}^S$  are principal eigenfunctions of the form  $\phi_j(\mathbf{x}) = \langle \mathbf{x}, \mathbf{w}_j \rangle, j = 1, \dots, d$ , where  $\mathbf{w}_j, j = 1, \dots, d$  are left eigenvectors of  $\mathbf{A}$ , while its eigenvalues coincide with the eigenvalues of the matrix  $\mathbf{A}$ .

### 3.1.1 Continuous-Time RDS

Let suppose that the stochastic Koopman operator family satisfies semigroup property. Define the generator of the stochastic Koopman family  $(\mathcal{K}^t)_{t \in \mathbb{T}}$  acting on the observable functions  $f \in C_b^1(\mathbb{R}^d)$  ( $C_b^1(\mathbb{R}^d)$  is the space of continuously differentiable functions on  $\mathbb{R}^d$  with bounded and continuous first derivative) by the limit

$$\mathcal{K}^S f(\mathbf{x}) = \lim_{t \rightarrow 0^+} \frac{\mathcal{K}^t f(\mathbf{x}) - f(\mathbf{x})}{t}, \tag{55}$$

if it exists. For the Koopman operators associated with the RDS generated by RDE, we have the following proposition.

**Proposition 5** *If the solution of RDE (11) is differentiable with respect to  $t$  and the stochastic Koopman family  $(\mathcal{K}^t)_{t \in \mathbb{T}}$  is a semigroup, then the action of the generator  $\mathcal{K}^S$  on  $f \in C_b^1(\mathbb{R}^d)$  is equal to*

$$\mathcal{K}^S f(\mathbf{x}) = \mathbb{E}[F(\omega, \mathbf{x})] \cdot \nabla f(\mathbf{x}). \tag{56}$$

**Proof**

$$\begin{aligned}
 \mathcal{K}^S f(\mathbf{x}) &= \lim_{t \rightarrow 0^+} \frac{\mathcal{K}^t f(\mathbf{x}) - f(\mathbf{x})}{t} = \lim_{t \rightarrow 0^+} \frac{\mathbb{E}[f(X(t, \omega, \mathbf{x}))] - f(\mathbf{x})}{t} \\
 &= \lim_{t \rightarrow 0^+} \frac{\mathbb{E}[f(X(t, \omega, \mathbf{x})) - f(\mathbf{x})]}{t} = \mathbb{E} \left[ \lim_{t \rightarrow 0^+} \frac{f(X(t, \omega, \mathbf{x})) - f(X(0, \omega, \mathbf{x}))}{t} \right] \\
 &= \mathbb{E} \left[ \frac{d}{dt} f(X(t, \omega, \mathbf{x})) \Big|_{t=0} \right] = \mathbb{E} \left[ \nabla f(\mathbf{x}) \cdot \frac{d}{dt} X(t, \omega, \mathbf{x}) \Big|_{t=0} \right] \\
 &= \mathbb{E}[F(\omega, \mathbf{x})] \cdot \nabla f(\mathbf{x}). \tag{57}
 \end{aligned}$$

The swapping of the order of limit and expectation in the second line is justified by the dominated convergence theorem and the fact that the convergence  $\frac{f(X(t, \omega, \mathbf{x})) - f(\mathbf{x})}{t}$  is uniform for all  $\omega$  and  $x$ , since the derivative of  $f$  is bounded and the solution is differentiable (see Lasota and Mackey 1994, Section 7.6 for the proof in the deterministic case).  $\square$

**Corollary 1** *Suppose that a stochastic Koopman generator  $\mathcal{K}^S$  associated with RDE (11) exists. If  $\phi_1$  and  $\phi_2$  are the eigenfunctions of  $\mathcal{K}^S$  with the associated eigenvalues  $\lambda_1$  and  $\lambda_2$ , then  $\phi_1\phi_2$  is also an eigenfunction with the associated eigenvalue  $\lambda_1 + \lambda_2$ .*

**Proof** Since  $\mathcal{K}^S \phi_i(\mathbf{x}) = \lambda_i \phi_i(\mathbf{x}) = \mathbb{E}[F(\omega, \mathbf{x})] \cdot \nabla \phi_i(\mathbf{x})$ , for  $i = 1, 2$ , we have

$$\begin{aligned}
 \mathcal{K}^S(\phi_1\phi_2)(\mathbf{x}) &= \mathbb{E}[F(\omega, \mathbf{x})] \cdot \nabla(\phi_1\phi_2)(\mathbf{x}) \\
 &= \mathbb{E}[F(\omega, \mathbf{x})] \cdot \nabla \phi_1(\mathbf{x}) \cdot \phi_2(\mathbf{x}) + \mathbb{E}[F(\omega, \mathbf{x})] \cdot \nabla \phi_2(\mathbf{x}) \cdot \phi_1(\mathbf{x}) \\
 &= \lambda_1 \phi_1(\mathbf{x})\phi_2(\mathbf{x}) + \lambda_2 \phi_2(\mathbf{x})\phi_1(\mathbf{x}) = (\lambda_1 + \lambda_2)(\phi_1\phi_2)(\mathbf{x})
 \end{aligned}$$

$\square$

**Corollary 2** *Let  $\phi \in C_b^1(\mathbb{R}^d)$  be an eigenfunction associated with eigenvalue  $\lambda$  of the stochastic Koopman generator  $\mathcal{K}^S$  associated with an RDE (11). Then*

$$\frac{d}{dt} \phi(\varphi(t, \omega)\mathbf{x}) = \lambda \phi(\varphi(t, \omega)\mathbf{x}) + \tilde{F}(\omega, \mathbf{x}) \cdot \nabla \phi(\varphi(t, \omega)\mathbf{x}), \tag{58}$$

where

$$\tilde{F}(\omega, \mathbf{x}) = F(\omega, \mathbf{x}) - \mathbb{E}[F(\omega, \mathbf{x})].$$

**Corollary 3** *Suppose that a stochastic Koopman generator  $\mathcal{K}^S$  associated with linear RDE (13) exists. Also, assume that  $\hat{\mathbf{A}} = \mathbb{E}[\mathbf{A}(\omega)]$  is diagonalizable, with simple eigenvalues  $\hat{\lambda}_j$  and left eigenvectors  $\hat{\mathbf{w}}_j$ ,  $j = 1, \dots, d$ . Then*

$$\phi_j(\mathbf{x}) = \langle \mathbf{x}, \hat{\mathbf{w}}_j \rangle, \quad j = 1, \dots, d, \tag{59}$$

are the principal eigenfunctions of the generator  $\mathcal{K}^S$ , while  $\lambda_j^S = \hat{\lambda}_j$  are the associated principal eigenvalues.

**Proof** According to (57), the action of the generator  $\mathcal{K}^S$  of the Koopman operator family associated with the linear RDS generated by (13) on  $f \in C_b^1(\mathbb{R}^d)$  is equal to

$$\mathcal{K}^S f(\mathbf{x}) = \mathbb{E}[\mathbf{A}(\omega)\mathbf{x}] \cdot \nabla f(\mathbf{x}) = \hat{\mathbf{A}}\mathbf{x} \cdot \nabla f(\mathbf{x}). \quad (60)$$

Thus

$$\mathcal{K}^S \phi_j(\mathbf{x}) = \langle \hat{\mathbf{A}}\mathbf{x}, \mathbf{w}_j \rangle = \langle \mathbf{x}, \hat{\mathbf{A}}^* \mathbf{w}_j \rangle = \hat{\lambda}_j \phi_j(\mathbf{x}),$$

which proves the statement.  $\square$

**Remark 4** Provided that the assumptions of Corollary 3 are valid, the principal eigenfunctions  $\phi_j(\mathbf{x})$  given by (59) are the eigenfunctions of each Koopman operator  $\mathcal{K}^t$  also, with the corresponding principal eigenvalues  $\lambda_j^S(t) = e^{\hat{\lambda}_j t}$ . The set of principal eigenfunctions does not cover all the eigenfunctions of the Koopman operator as we discuss next. According to Corollary 1, over the space of real analytic functions,

$$\phi(\mathbf{x}) = \phi_1^{n_1}(\mathbf{x}) \cdots \phi_d^{n_d}(\mathbf{x}), \quad \lambda = \sum_{j=1}^d n_j \hat{\lambda}_j,$$

with  $n_j \in \mathbb{N}^+ \cup \{0\}$ ,  $j = 1, \dots, d$ , are the eigenvalues and eigenfunctions of the Koopman generator. Thus, like in deterministic case, any analytic observable function  $f$  can be represented as linear combination of powers of the principal eigenfunctions (Mezić 2017b) and its evolution under the RDS can be obtained using spectral expansion formula.

Another type of RDS which could be identified with the time-homogeneous Markov family are the RDS generated by the SDE of the form (21) with autonomous functions  $G$  and  $\sigma$ , i.e.,

$$dX_t = G(X_t)dt + \sigma(X_t)dW_t. \quad (61)$$

In this case the stochastic differential equation generates the one-parameter family of RDS  $\varphi(t, \omega) := \varphi(t, 0, \omega) = \varphi(t + t_0, t_0, \omega)$ , so that the corresponding stochastic Koopman operator family and the associated stochastic eigenvalues and eigenfunctions depend on parameter  $t$ , only. In this autonomous setting, we denote by  $X_t(\mathbf{x})$  the solution of Eq. (61) for the initial condition  $X_0(\omega) = \mathbf{x}$ . In accordance with (22) it is equal to

$$X_t(\mathbf{x}) = \mathbf{x} + \int_0^t G(X_s)ds + \int_0^t \sigma(X_s)dW_s. \quad (62)$$

There are numerous books dealing with the properties of the solutions of such SDE, for example (Cohen and Elliott 2015, Chapter 16; Da Prato 2014; Pavliotis 2014). Here we reference some of the books containing proofs that the generated solutions form Markov family and moreover, that the associated transfer operators form semigroup: (Arnold 1974, 1998, Section 2.3; Lasota and Mackey 1994, Chapter 11; Cohen and Elliott 2015, Chapter 17; Da Prato 2014, Section 9).

In most of the stated references the generator of the stochastic solutions  $X_t$  is studied. This generator is in the stochastic literature known under the name backward Kolmogorov operator and is a second-order elliptic type differential operator. It appears as the right-hand side of the backward Kolmogorov equation, whose solution, with the initial condition prescribed as the value of the observable function  $f(\mathbf{x})$ , gives the time evolution of the average value of the observable function, i.e., the value  $\mathcal{K}^t f(\mathbf{x})$  (Pavliotis 2014, Sections 2.5, 3; Cohen and Elliott 2015, Section 17.4; Da Prato 2014, Section 9.3). Thus, the Kolmogorov operator is also the generator of the stochastic Koopman semigroup, and its action over the space of real bounded, continuous, and twice differentiable functions on  $\mathbb{R}^d$  with bounded and continuous first and second derivatives, which we denote by  $C_b^2(\mathbb{R}^d)$ , is given in the following proposition.

**Proposition 6** *The action of the generator of the stochastic Koopman family  $\mathcal{K}^S$  on  $f \in C_b^2(\mathbb{R}^d)$  is given by*

$$\mathcal{K}^S f(\mathbf{x}) = G(\mathbf{x})\nabla f(\mathbf{x}) + \frac{1}{2}\text{Tr}\left(\sigma(\mathbf{x})(\nabla^2 f(\mathbf{x}))\sigma(\mathbf{x})^T\right), \tag{63}$$

where  $Tr$  denotes the trace of the matrix.

The following proposition follows from the proof of the Feynman–Kac formula (Pavliotis 2014, Section 3.4; Cohen and Elliott 2015, Section 17.4).

**Proposition 7** *Let  $\phi \in C_b^2(\mathbb{R}^d)$  be an eigenfunction of the stochastic Koopman generator  $\mathcal{K}^S$  associated with RDS generated by SDE (61) with the corresponding eigenvalue  $\lambda$ . Then*

$$d\phi(X_t) = \lambda\phi(X_t)dt + \nabla\phi(X_t)\sigma(X_t)dW_t. \tag{64}$$

$\phi$  is the eigenfunction of the stochastic Koopman operator  $\mathcal{K}^t$  also, i.e.,

$$\mathcal{K}^t\phi(\mathbf{x}) = e^{\lambda t}\phi(\mathbf{x}). \tag{65}$$

**Proof** Suppose  $d = 1$ . It follows from Itô’s formula

$$dY_t = f'(X_t)dX_t + \frac{1}{2}f''(X_t)\sigma(X_t)^2dt \tag{66}$$

that the eigenfunction  $\phi(X_t)$  evolves according to

$$\begin{aligned} d\phi(X_t) &= \phi'(X_t)G(X_t)dt + \frac{1}{2}\phi''(X_t)\sigma(X_t)^2dt + \phi'(X_t)\sigma(X_t)dW_t \\ &= \mathcal{K}^S\phi(X_t)dt + \phi'(X_t)\sigma(X_t)dW_t \\ &= \lambda\phi(X_t)dt + \phi'(X_t)\sigma(X_t)dW_t, \end{aligned} \tag{67}$$

where in the last equality we used that  $\mathcal{K}^S\phi(x) = \lambda\phi(x)$ . By using the similar procedure, the Eq. (64) valid in multidimensional case can be easily derived.

The fact that  $\phi$  is an eigenfunction of each Koopman semigroup member  $\mathcal{K}^t$  follows from the spectral mapping theorem (Engel and Nagel 2001, Chapter IV.3).  $\square$

The fact that the eigenfunctions related to the generator of the dynamical system evolve according to (67), is used in Shnitzer et al. (2017). In that paper the approximations of eigenfunctions computed via diffusion maps are used to parametrize the state of the RDS using the linear operator similarly as in Koopman framework.

**Remark 5** Unlike in the case of RDS generated by the RDE, the product of eigenfunctions of the stochastic Koopman generator associated with RDS generated by SDE is not necessarily an eigenfunction. This easily follows from Proposition 6. However, the eigenfunctions in many cases satisfy a recurrence relationship [e.g., Hermite polynomials, which are Koopman eigenfunctions of Ornstein–Uhlenbeck processes obtained as a solution of the Ornstein–Uhlenbeck SDE (Pavliotis 2014, Section 4.4)] and thus can be deduced from the principal eigenfunctions. This reduces the problem to the analysis of principal eigenfunctions, and thus  $d$  objects in an  $d$ -dimensional space, remarkable for a nominally infinite-dimensional representation.

## 4 Numerical Approximations of the Stochastic Koopman Operator

One of the goals of this work is to compute the numerical approximations of the spectral objects of the stochastic Koopman operators, i.e., their eigenvalues and eigenfunctions. This is performed by extending the DMD algorithms, that are originally developed for approximating the spectral object of the Koopman operators in deterministic settings, to the stochastic framework. DMD algorithms and the spectral analysis of the Koopman operator are connected in the following way. Suppose that the restriction of the infinite-dimensional stochastic Koopman operator  $\mathcal{K}^t$  to an appropriate  $n$ -dimensional subspace of functions is closed under the action of the Koopman operator (see Drmač et al. 2018). Let  $\mathbb{K} \in \mathbb{C}^{n \times n}$  denote its representation by an  $n \times n$ -dimensional matrix. The goal of the DMD numerical algorithms is to determine the spectrum and associated eigenvectors of the finite-dimensional linear operator  $\mathbb{K}$ , and those are in turn the eigenvalues and coordinates of eigenfunctions of the associated Koopman operator in the chosen finite-dimensional basis. If the considered subspace is not closed under the action of the Koopman operator, one can expect that under certain assumptions, the operator  $\mathbb{K}$  at least approximates the underlying dynamics (Arbabi and Mezić 2017; Korda and Mezić 2018).

The DMD algorithm provides us with a decomposition of the pair  $(\mathbf{X}_m, \mathbf{Y}_m)$  given by the eigenvalues and the eigenvectors of the operator  $\mathbb{K}$ , which is in the data-driven settings not known explicitly, while it is known that its action on the range  $\mathbf{X}_m$  should be equal to  $\mathbf{Y}_m$  (Tu et al. 2014). For  $n \gg m$  and full column range of  $\mathbf{X}_m$ , there exists an exact solution of the equation  $\mathbf{Y}_m = \mathbb{K}\mathbf{X}_m$ , while for  $m \gg n$  the equation is satisfied in a least-squares sense (Tu et al. 2014).

Here we use the enhanced DMD algorithm, proposed in Drmač et al. (2018), which we denote hereafter by DMD RRR. We briefly describe two types of DMD algorithm and propose their application in the stochastic framework. The first type of DMD algorithm is the modification of the standard DMD approach that uses snapshot pairs in deterministic framework, and the second one is obtained from the DMD applied to the Hankel matrix. In the stochastic framework, in both considered approaches, instead

of the application of the algorithm to the individual snapshots along a trajectory, it is applied to the expected values of the observable functions at the appropriate time moment.

### 4.1 The DMD Algorithms for RDS

Let  $\mathbf{f} = (f_1, \dots, f_n)^T : M \rightarrow \mathbb{C}^n$  be a vector-valued observable on the state space. For an  $\mathbf{x} \in M$ , let  $\mathbf{f}^k(\omega, \mathbf{x}) = \mathbf{f}(T^k(\omega, \mathbf{x}))$ ,  $k = 0, 1, \dots$  be a vector-valued observable series on the trajectory of the considered discrete-time RDS. Denote its expectation by  $\mathbf{f}^k(\mathbf{x}) = \mathbb{E}[\mathbf{f}^k(\omega, \mathbf{x})] = \mathcal{K}^k \mathbf{f}(\mathbf{x})$ . For the continuous-time RDS  $\varphi$ , we choose the time step  $\Delta t$  and define the values of the observables in the same way as in the discrete-time case by taking  $T^k(\omega, \mathbf{x}) = \varphi(k\Delta t, \omega)\mathbf{x}$ . The expected value of the observable at the time moment  $t_k = k\Delta t$ , i.e., the value of the action of the stochastic Koopman operator  $\mathcal{K}^k$  on  $\mathbf{f}$ , is evaluated as  $\mathbf{f}^k(\mathbf{x}) = \mathcal{K}_{\Delta t}^k \mathbf{f}(\mathbf{x}) = \mathbb{E}[\mathbf{f}(\varphi(k\Delta t, \omega)\mathbf{x})]$ .

In this stochastic framework, the DMD algorithm is applied to the matrices  $\mathbf{X}_m, \mathbf{Y}_m \in \mathbb{C}^{n \times m}$ , which are for the chosen initial states  $\mathbf{x}_1, \dots, \mathbf{x}_m \in M$  defined by

$$\mathbf{X}_m = (\mathbf{f}^0(\mathbf{x}_1) \ \mathbf{f}^0(\mathbf{x}_2) \ \dots \ \mathbf{f}^0(\mathbf{x}_m)) \quad \text{and} \quad \mathbf{Y}_m = (\mathbf{f}^k(\mathbf{x}_1) \ \mathbf{f}^k(\mathbf{x}_2) \ \dots \ \mathbf{f}^k(\mathbf{x}_m)). \quad (68)$$

In this case we expect that the DMD algorithm provides us with the eigenvalues and eigenfunctions of the finite-dimensional approximation operator associated with the Koopman operator  $\mathcal{K}^{tk}$ . In the case when the Koopman operator family is a semigroup, we have  $\mathcal{K}_{\Delta t}^k = (\mathcal{K}_{\Delta t}^1)^k$ , and thus it makes sense to apply the DMD algorithm to the time-delayed expected values of the observable function, which means that for the chosen initial condition  $\mathbf{x}_0$ , we define the matrices  $\mathbf{X}_m$  and  $\mathbf{Y}_m$  as

$$\mathbf{X}_m = (\mathbf{f}^0(\mathbf{x}_0) \ \mathbf{f}^1(\mathbf{x}_0) \ \dots \ \mathbf{f}^{m-1}(\mathbf{x}_0)) \quad \text{and} \quad \mathbf{Y}_m = (\mathbf{f}^1(\mathbf{x}_0) \ \mathbf{f}^2(\mathbf{x}_0) \ \dots \ \mathbf{f}^m(\mathbf{x}_0)). \quad (69)$$

In this case, we expect that the algorithm provide us with the eigenvalues and eigenfunctions of the approximation operator associated with  $\mathcal{K}_{\Delta t}^1$ . The proof of the convergence of the DMD type algorithm with input matrices (69) where  $f_1, \dots, f_n$  span the finite-dimensional invariant subspace of the stochastic Koopman operator, to its eigenvalues and eigenfunctions is given in Takeishi et al. (2017) for the RDS in which the noise is modeled by i.i.d. random variables and under the assumption of ergodicity.

We describe now the crucial steps of the DMD RRR algorithm. The DMD RRR algorithm starts in the same way as other SVD-based DMD algorithms, i.e., with the SVD decomposition for low-dimensional approximation of data:  $\mathbf{X}_m = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^* \approx \mathbf{U}_r \mathbf{\Sigma}_r \mathbf{V}_r^*$ , where  $\mathbf{\Sigma} = \text{diag}((\sigma_i)_{i=1}^{\min(m,n)})$ ,  $\sigma_i$  are singular values arranged in the descending order, i.e.,  $\sigma_1 \geq \dots \geq \sigma_{\min(m,n)} \geq 0$ , and  $r$  is the dimension of the approximation space.  $\mathbf{U}_r$  and  $\mathbf{U}_r$  are composed from first  $r$  columns of matrices  $\mathbf{U}$  and  $\mathbf{U}$  and  $\mathbf{\Sigma}_r = \text{diag}((\sigma_i)_i = 1^r)$ . Then  $\mathbf{Y}_m = \mathbb{K}\mathbf{X}_m$  can be approximated as

$$\mathbf{Y}_m \approx \mathbb{K}\mathbf{U}_r \mathbf{\Sigma}_r \mathbf{V}_r^*. \quad (70)$$

Since  $\mathbb{K}\mathbf{U}_r = \mathbf{Y}_m \mathbf{V}_r \boldsymbol{\Sigma}_r^{-1}$ , the Rayleigh quotient matrix  $\mathbf{S}_r = \mathbf{U}_r^* \mathbb{K} \mathbf{U}_r$  with respect to the range  $\mathbf{U}_r$  can be computed in this data-driven setting as

$$\mathbf{S}_r = \mathbf{U}_r^* \mathbf{Y}_m \mathbf{V}_r \boldsymbol{\Sigma}_r^{-1}. \quad (71)$$

Each eigenpair  $(\lambda, \mathbf{w})$  of  $\mathbf{S}_r$  generates the corresponding Ritz pair  $(\lambda, \mathbf{U}_r \mathbf{w})$ , that is a candidate for the approximation of the eigenvalue and eigenvector of the Koopman operator.

Here we emphasize few crucial points at which the DMD RRR algorithm is improved in comparison with the standard DMD algorithms (Schmid 2010; Tu et al. 2014). The first point of difference refers to the dimension  $r$  of the reduction space. Instead of defining the dimension of the space a priori or to take into account the spectral gap in singular values, it is proposed in Drmač et al. (2018) to take into account user supplied tolerance  $\varepsilon$ , which is then used for defining  $r$  as the largest index satisfying  $\sigma_r \geq \sigma_1 \varepsilon$ . The algorithm is, according to Drmač et al. (2018), further enhanced with the residual computation for each Ritz vector pair  $(\lambda, \mathbf{U}_r \mathbf{w})$

$$\eta = \|\mathbb{K}(\mathbf{U}_r \mathbf{w}) - \lambda(\mathbf{U}_r \mathbf{w})\|_2 = \|(\mathbf{Y}_m \mathbf{V}_r \boldsymbol{\Sigma}_r^{-1}) \mathbf{w} - \lambda(\mathbf{U}_r \mathbf{w})\|_2, \quad (72)$$

where  $\|\cdot\|_2$  stands for  $L_2$  norm. Then the vectors at which the required accuracy is not attained are not taken into account. The final improvement compared to the standard algorithms refers to scaling of the initial data, i.e., if the matrix  $D_x$  is defined with  $D_x = \text{diag}(\|\mathbf{X}_m(:, i)\|_2)_{i=0}^{m-1}$ , we set  $\mathbf{X}_m = \mathbf{X}_m^{(1)} D_x$  and  $\mathbf{Y}_m = \mathbf{Y}_m^{(1)} D_x$  and proceed with  $\mathbf{X}_m^{(1)}$  and  $\mathbf{Y}_m^{(1)}$  as data matrices.

## 4.2 Stochastic Hankel–DMD (sHankel–DMD) Algorithm

Here we define the stochastic Hankel matrix and describe the application of the DMD type algorithm on it. We use the Hankel–DMD method described in Arbabi and Mezić (2017) for deterministic dynamical system and expand the ideas to the stochastic settings. We first limit to the discrete-time RDS and assume that the associated stochastic Koopman operator family satisfies the semigroup property, since only in such case the results of the DMD algorithm applied to the Hankel matrix are meaningful. Otherwise, the situation similar to the nonautonomous deterministic case arise and the application of the DMD algorithm to the time-delayed expected values of the observables becomes an attempt to approximate different operators, which results with significant errors (Maćešić et al. 2018).

The Hankel matrix in the stochastic framework is defined as follows. For a scalar observable  $f : M \rightarrow \mathbb{C}$ , we define the vector of  $n$  observations along the trajectory that starts at  $\mathbf{x} \in M$  and is generated by the one step discrete random map  $T$ :

$$\mathbf{f}_n(\omega, \mathbf{x}) = \left( f(\mathbf{x}), f(T(\omega, \mathbf{x})), \dots, f(T^{n-1}(\omega, \mathbf{x})) \right)^T. \quad (73)$$



Let  $\mathbf{f}_n^k(\mathbf{x})$ ,  $k = 0, 1, \dots$  denote the expectation of  $\mathbf{f}_n(\theta(k)\omega, T^k(\omega, \mathbf{x}))$  over the trajectories of length  $k$ , i.e.,

$$\begin{aligned} \mathbf{f}_n^k(\mathbf{x}) &= \mathbb{E} \left[ \mathbf{f}_n(\theta(k)\omega, T^k(\omega, \mathbf{x})) \right] \\ &= \left( \mathcal{K}^k f(\mathbf{x}), \mathcal{K}^k f(T(\omega, \mathbf{x})), \dots, \mathcal{K}^k f(T^{n-1}(\omega, \mathbf{x})) \right)^T. \end{aligned} \tag{74}$$

Observe that the components of  $\mathbf{f}_n^k(\mathbf{x})$  are the values of the function  $\mathcal{K}^k f$  along the trajectory of length  $n$  starting at  $\mathbf{x} \in M$ . Numerically, the approximations of  $\mathcal{K}^k f(T^j(\omega, \mathbf{x}))$ ,  $j = 0, \dots, n - 1$  are obtained as expectations of the observable function values over the multiple trajectories of length  $k$  starting from  $T^j(\omega, \mathbf{x})$ ,  $j = 0, \dots, n - 1$ .

The stochastic Hankel matrix of dimension  $n \times m$ , associated with the trajectories starting at  $\mathbf{x} \in M$  and generated by the map  $T$  is defined by

$$\begin{aligned} \mathbf{H}_{n \times m}(\omega, \mathbf{x}) &= \left( \mathbf{f}_n^0(\mathbf{x}) \ \mathbf{f}_n^1(\mathbf{x}) \ \dots \ \mathbf{f}_n^{m-1}(\mathbf{x}) \right) \\ &= \begin{pmatrix} f(\mathbf{x}) & \mathcal{K}^1 f(\mathbf{x}) & \dots & \mathcal{K}^{m-1} f(\mathbf{x}) \\ f(T(\omega, \mathbf{x})) & \mathcal{K}^1 f(T(\omega, \mathbf{x})) & \dots & \mathcal{K}^{m-1} f(T(\omega, \mathbf{x})) \\ \vdots & \vdots & \ddots & \vdots \\ f(T^{n-1}(\omega, \mathbf{x})) & \mathcal{K}^1 f(T^{n-1}(\omega, \mathbf{x})) & \dots & \mathcal{K}^{m-1} f(T^{n-1}(\omega, \mathbf{x})) \end{pmatrix}. \end{aligned} \tag{75}$$

The columns of  $\mathbf{H}_{n \times m}$  are approximations of functions in the Krylov subspace

$$\mathbb{K}_m(\mathcal{K}, f) = \text{span} \left( f, \mathcal{K}^1 f, \dots, \mathcal{K}^{m-1} f \right), \tag{76}$$

obtained by sampling the values of functions  $\mathcal{K}^k f$ ,  $k = 0, \dots, m - 1$  along the trajectory of length  $n$  starting at  $\mathbf{x} \in M$ .

When the DMD algorithm is applied to the stochastic Hankel matrix, the input data matrix  $\mathbf{X}_m$  is defined by taking the first  $m$  columns of the Hankel matrix  $\mathbf{H}_{n \times (m+1)}(\omega, \mathbf{x})$ , while the data matrix  $\mathbf{Y}_m$  is formed from the last  $m$  columns of the same matrix. We refer to this methodology as to the stochastic Hankel–DMD (sHankel–DMD) algorithm.

As already mentioned, when we consider the continuous-time RDS, we could for the chosen  $\Delta t$  associate with it a discrete RDS by defining the one-step map as  $T(\omega, \mathbf{x}) = \varphi(\Delta t, \omega)\mathbf{x}$ . Then the Koopman operator  $\mathcal{K}^k$  should be replaced with the operator  $\mathcal{K}_{\Delta t}^k$ .

### 4.3 Convergence of the sHankel–DMD Algorithm

It was proved in Arbabi and Mezić (2017) that for ergodic systems, under the assumption that observables are in an invariant subspace of the Koopman operator, the eigenvalues and eigenfunctions obtained by the extended DMD algorithm applied to Hankel matrix, converge to the true Koopman eigenvalues and eigenfunctions of the considered system. The convergence of the DMD algorithm with input matrices

$\mathbf{X}_m$  and  $\mathbf{Y}_m$  defined by (69) to the eigenvalues and the eigenfunctions of the Koopman operator is proved in Takeishi et al. (2017) for the class of RDS in which the noise is modeled by i.i.d. random variables, under the assumption of ergodicity, and of the existence of the finite-dimensional invariant subspace. Here we prove that under the same assumptions, the convergence is accomplished for the sHankel–DMD algorithm. Our proof is based on the fact that the eigenvalues and the eigenfunctions obtained by DMD algorithm correspond to the matrix that is similar to the companion matrix, which represents the finite-dimensional approximation of the Koopman operator in the Krylov basis. We limit the considerations to the discrete-time RDS. Suppose that the dynamics on the compact invariant set  $A \subseteq M$  is generated by the measure-preserving map  $T(\omega, \cdot) : A \rightarrow A$  for each  $\omega \in \Omega$ . We recall from Arnold (1998, Section 1.4) that a probability measure  $\nu$  is invariant for RDS  $\varphi(n, \omega) = T^n(\omega, \cdot)$  if it is invariant for the skew product flow  $\Theta(n)(\omega, x) = (\theta(n)\omega, T^n(\omega, x))$  generated by  $T$  and  $\theta(t)$ , i.e., if  $\Theta(n)\nu = \nu$  and if  $\pi_\Omega \nu = \mathbb{P}$  where  $\pi_\Omega$  denotes the canonical projection  $\Omega \times A \rightarrow \Omega$ . Invariant measures always exist for a continuous RDS on compact space  $A$  (see Arnold 1998, Theorem 1.5.10). If  $A$  is a Polish space, the measure  $\nu$  on  $\Omega \times A$  could be written as product of measures, i.e.,  $d\nu(\omega, x) = d\mu_\omega(x)d\mathbb{P}(\omega)$ , i.e., for  $f \in L^1(\nu)$

$$\int_{\Omega \times A} f d\nu = \int_{\Omega} \int_A f(\omega, x) d\mu_\omega(x) d\mathbb{P}(\omega).$$

If the skew product dynamical system  $\Theta$  is ergodic on  $\Omega \times A$  with respect to some invariant measure  $\nu$  and if  $\theta(n)$  is ergodic with respect to  $\mathbb{P}$ , we say that  $\varphi$  is ergodic with respect to the invariant measure  $\nu$ . Under the assumption of ergodicity of  $\Theta$  with respect to the measure  $\nu$ , the Birkhoff's ergodic theorem states that the time average of observable  $f \in L^2(\Omega \times A; \nu)$  under  $\Theta$  is given by

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} f(\theta(k)\omega, T^k(\omega, \mathbf{x})) = \int_{\Omega \times A} f(\omega, x) d\nu, \quad \text{a. e. on } \Omega \times A. \quad (77)$$

The following proposition shows that under the assumption of ergodicity and Markovian property of RDS, the sHankel–DMD algorithm provide us with approximations of eigenvalues and eigenfunctions that converge to the true Koopman eigenvalues and eigenfunctions.

**Proposition 8** *Suppose that the dynamics on the compact invariant set  $A \subseteq M$  is given by the one step map  $T(\omega, \cdot) : A \rightarrow A$  for each  $\omega \in \Omega$  and that the associated discrete-time RDS  $\varphi$  is ergodic with respect to some invariant measure  $\nu$ . Assume additionally that the processes  $\{\varphi(n, \omega)\mathbf{x}, \mathbf{x} \in A\}$  form a Markov family. Denote by  $\mu$  the marginal measure  $\mu = \mathbb{E}_\mathbb{P}(\nu)$  on  $A$ .*

*Let the Krylov subspace  $\mathbb{K}_m(\mathcal{K}, f)$  span an  $r$ -dimensional subspace of the Hilbert space  $\mathcal{H} = L^2(A, \mu)$ , with  $r < m$ , invariant under the action of the stochastic Koopman operator. Then for almost every  $\mathbf{x} \in A$ , as  $n \rightarrow \infty$ , the eigenvalues and eigenfunctions obtained by applying DMD algorithm to the first  $r + 1$  columns of the*

$n \times (m + 1)$ -dimensional stochastic Hankel matrix, converge to the true eigenvalues and eigenfunctions of the stochastic Koopman operator.

**Proof** Consider the observables  $f : A \rightarrow \mathbb{R}$  belonging to the Hilbert space  $\mathcal{H} = L^2(A, \mu)$ . Due to the ergodicity of  $\Theta$ , in accordance with the Birkhoff’s ergodic theorem, (77) is valid, thus we get

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k(\omega, \mathbf{x})) = \int_{\Omega \times A} f(x) d\nu = \int_{\Omega} \int_A f(x) d\mu_{\omega}(x) d\mathbb{P}(\omega) = \int_A f d\mu, \tag{78}$$

where the last equality follows from the fact that  $\mu = \mathbb{E}_{\mathbb{P}}(\nu) = \mathbb{E}_{\mathbb{P}}(\mu_{\omega})$ . For observables  $f, g \in \mathcal{H}$  let the vectors of  $n$  observations along the trajectory starting at  $\mathbf{x} \in A$  of the RDS generated by the map  $T$  be denoted by  $\mathbf{f}_n(\omega, \mathbf{x})$  and  $\mathbf{g}_n(\omega, \mathbf{x})$  and defined by (73). If we denote the data-driven inner product by  $\langle \mathbf{f}_n(\omega, \mathbf{x}), \mathbf{g}_n(\omega, \mathbf{x}) \rangle$ , we have

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} [\langle \mathbf{f}_n(\omega, \mathbf{x}), \mathbf{g}_n(\omega, \mathbf{x}) \rangle] &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k(\omega, \mathbf{x}))g^*(T^k(\omega, \mathbf{x})) \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} fg^*(T^k(\omega, \mathbf{x})) = \int_A fg^* d\mu = \langle f, g \rangle_{\mathcal{H}} \text{ for a.e. } \mathbf{x} \end{aligned} \tag{79}$$

with respect to the measure  $\mu$ . Using the assumption that  $\mathbb{K}_m(\mathcal{K}, f)$  spans  $r$ -dimensional subspace of  $\mathcal{H}$ , which is invariant for the stochastic Koopman operator, the restriction of the Koopman operator to this subspace is finite dimensional and can be realized by an  $r \times r$  matrix. The representation of this matrix in the basis formed by the functions  $(f, \mathcal{K}f, \dots, \mathcal{K}^{m-1}f)$  is given with the companion matrix

$$\mathbf{C} = \begin{pmatrix} 0 & 0 & \dots & 0 & c_0 \\ 1 & 0 & \dots & 0 & c_1 \\ 0 & 1 & \dots & 0 & c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & c_{r-1} \end{pmatrix}, \tag{80}$$

where the vector  $\mathbf{c} = (c_0, c_1, \dots, c_{r-1})^T$ , obtained by using least-squares approximation is equal to

$$\mathbf{c} = \mathbf{G}^{-1} (\langle f, \mathcal{K}^r f \rangle_{\mathcal{H}}, \langle \mathcal{K}^1 f, \mathcal{K}^r f \rangle_{\mathcal{H}}, \dots, \langle \mathcal{K}^{r-1} f, \mathcal{K}^r f \rangle_{\mathcal{H}})^T. \tag{81}$$

Here  $\mathbf{G}$  denotes the Gramian matrix with elements  $G_{ij} = \langle \mathcal{K}^{i-1} f, \mathcal{K}^{j-1} f \rangle_{\mathcal{H}}$ ,  $i, j = 1, \dots, r$  (see Arbabi and Mezić 2017; Drmač et al. 2018).

Consider now the stochastic Hankel matrix  $\mathbf{H}_{n \times (r+1)}(\omega, \mathbf{x})$  of dimension  $n \times (r + 1)$  along a trajectory starting at  $\mathbf{x}$  and the DMD algorithm applied to the matrices  $\mathbf{X}_r = (\mathbf{f}_n^0(\mathbf{x}) \mathbf{f}_n^1(\mathbf{x}) \dots \mathbf{f}_n^{r-1}(\mathbf{x}))$  and  $\mathbf{Y}_r = (\mathbf{f}_n^1(\mathbf{x}) \mathbf{f}_n^2(\mathbf{x}) \dots \mathbf{f}_n^r(\mathbf{x}))$ . In what follows

we will prove that the eigenvalues and eigenvectors provided by DMD algorithm converge to the eigenvalues and eigenvectors of the stochastic Koopman operator. We will use the fact that these eigenvalues and eigenvectors are computed from the eigenvalues and eigenvectors of the matrix  $\mathbf{S}_r$  (71), which is, according to Drmač et al. (2018, Proposition 3.1), in the case considered here, similar to the numerical companion matrix  $\tilde{\mathbf{C}}$  that is defined hereafter. We denote by  $\tilde{\mathbf{C}}$  the numerical companion matrix computed as a best approximation of the least-squares problem  $\tilde{\mathbf{C}} = \arg \min_{\mathbf{B} \in \mathbb{C}^{r \times r}} \|\mathbf{Y}_r - \mathbf{X}_r \mathbf{B}\|$  (Drmač et al. 2018) by applying the companion matrix algorithm (Arbabi and Mezić 2017; Drmač et al. 2018) to matrices  $\mathbf{X}_r$  and  $\mathbf{Y}_r$ . Using the assumption that the Krylov subspace spans an  $r$ -dimensional subspace, which means that the matrix  $\mathbf{X}_r$  has a full column rank, the pseudoinverse is of the form  $\mathbf{X}_r^\dagger = (\mathbf{X}_r^* \mathbf{X}_r)^{-1} \mathbf{X}_r^*$  and the matrix  $\tilde{\mathbf{C}}$  is

$$\begin{aligned} \tilde{\mathbf{C}} &= \mathbf{X}_r^\dagger \mathbf{Y}_r = (\mathbf{X}_r^* \mathbf{X}_r)^{-1} \mathbf{X}_r^* \mathbf{Y}_r \\ &= \left( \frac{1}{n} \mathbf{X}_r^* \mathbf{X}_r \right)^{-1} \left( \frac{1}{n} \mathbf{X}_r^* \mathbf{Y}_r \right) = \tilde{\mathbf{G}}^{-1} \left( \frac{1}{n} \mathbf{Y}_r \mathbf{X}_r^* \right). \end{aligned} \tag{82}$$

Here  $\tilde{\mathbf{G}}$  denotes the numerical Grammian matrix whose elements are equal to

$$\begin{aligned} \tilde{G}_{ij}(\omega, \mathbf{x}) &= \frac{1}{n} \langle \mathbf{f}_n^{i-1}(\mathbf{x}), \mathbf{f}_n^{j-1}(\mathbf{x}) \rangle \\ &= \frac{1}{n} \sum_{k=0}^{n-1} \mathcal{K}^{i-1} f(T^k(\omega, \mathbf{x})) \mathcal{K}^{j-1} f^*(T^k(\omega, \mathbf{x})) \\ &= \frac{1}{n} \sum_{k=0}^{n-1} (\mathcal{K}^{i-1} f)(\mathcal{K}^{j-1} f^*)(T^k(\omega, \mathbf{x})), \quad i, j = 1, \dots, r. \end{aligned} \tag{83}$$

Now, by using (79), we conclude that

$$\lim_{n \rightarrow \infty} \tilde{G}_{ij}(\omega, \mathbf{x}) = \langle \mathcal{K}^{i-1} f, \mathcal{K}^{j-1} f \rangle_{\mathcal{H}}, \quad i, j = 1, \dots, r \quad \text{for a.e. } \mathbf{x} \tag{84}$$

with respect to the measure  $\mu$ . From (82) we get that the last column  $\tilde{\mathbf{c}} = (\tilde{c}_0(\omega, \mathbf{x}), \tilde{c}_1(\omega, \mathbf{x}), \dots, \tilde{c}_{r-1}(\omega, \mathbf{x}))^T$  of  $\tilde{\mathbf{C}}$  is equal to

$$\tilde{\mathbf{c}} = \tilde{\mathbf{G}}^{-1} \frac{1}{n} \left( \langle \mathbf{f}_n^0(\mathbf{x}), \mathbf{f}_n^r(\mathbf{x}) \rangle, \langle \mathbf{f}_n^1(\mathbf{x}), \mathbf{f}_n^r(\mathbf{x}) \rangle, \dots, \langle \mathbf{f}_n^{r-1}(\mathbf{x}), \mathbf{f}_n^r(\mathbf{x}) \rangle \right)^T. \tag{85}$$

Since

$$\lim_{n \rightarrow \infty} \langle \mathbf{f}_n^{j-1}(\mathbf{x}), \mathbf{f}_n^r(\mathbf{x}) \rangle = \langle \mathcal{K}^{j-1} f, \mathcal{K}^r f \rangle_{\mathcal{H}} \quad j = 1, \dots, r, \quad \text{for a.e. } \mathbf{x} \tag{86}$$

with respect to the measure  $\mu$  and the matrix  $\tilde{\mathbf{G}}^{-1}$  converges to the inverse of the true Grammian matrix  $\mathbf{G}^{-1}$ , it follows that the components of  $\tilde{\mathbf{c}}$  converge to the components of  $\mathbf{c}$  (81). As in Arbabi and Mezić (2017, Proposition 3.1), we use now the fact that

the eigenvalues of the matrix  $\tilde{\mathbf{C}}$  are the roots of a polynomial having coefficients  $\tilde{c}_i$ . Therefore the eigenvalues of  $\tilde{\mathbf{C}}$  are continuously dependent on these coefficients, which implies that the eigenvalues of the companion matrix  $\tilde{\mathbf{C}}$  converge to the eigenvalues of the exact companion matrix, i.e., to the exact eigenvalues of the stochastic Koopman operator. As we already mentioned before, the eigenvalues and eigenvectors provided by DMD algorithm are computed from the eigenvalues and eigenvectors of the matrix  $\mathbf{S}_r$  (71), which is for the full column rank matrix  $\mathbf{X}_r$ , similar to the companion matrix  $\tilde{\mathbf{C}}$  (see Drmač et al. 2018, Proposition 3.1), thus the conclusion follows.  $\square$

## 5 Numerical Examples

In this section, we illustrate the computation of the approximations of spectral objects of the stochastic Koopman operator on a variety of numerical examples. In all the examples we use the proposed methodologies in combination with the DMD RRR algorithm. First, we consider two discrete RDS examples: Example 2 and the linear RDS. Due to the fact that in the first case the semigroup property for the Koopman operator family is valid, the input matrices of DMD algorithm are defined by (69). We show that good approximations of the eigenvalues and eigenfunctions are obtained. In the second linear case, the DMD algorithm is applied on input matrices (68) and the differences between the exact and computed eigenvalues are determined. In Sect. 5.2 the approximation of time-dependent principal eigenvalues for the continuous-time RDS generated by the linear RDE are determined by applying DMD RRR algorithm on data matrices defined by (68) where  $\mathbf{Y}_m$  changes in time. In these two linear test examples, discrete and continuous one, the solutions do not form Markov family, so that the associated stochastic Koopman operator family is not a semigroup. In Sect. 5.3, we apply DMD RRR algorithm on input matrices defined by (68) and (69) to the RDS generated by scalar linear and scalar nonlinear SDE and approximate its eigenvalues and eigenfunctions. Finally, we apply the sHankel–DMD algorithm to two RDS generated by SDE systems (stochastic Stuart–Landau and noisy Van der Pol system) whose solutions converge to the asymptotic limit cycles and to the RDS generated by noisy Lotka–Volterra system with an asymptotically stable equilibrium point. In most considered examples, we first apply the DMD RRR algorithm to the related deterministic dynamical system (without noise) and then explore its behavior on the RDS.

### 5.1 Discrete RDS Examples

#### 5.1.1 Noisy Rotation on the Circle

Consider the dynamical system defined in Example 2. We use the following set of observables:

$$f_j(x) = \cos(j2\pi x), g_j(x) = \sin(j2\pi x), j = 1, \dots, n_1,$$

arranged in the vector-valued observable as

$$\mathbf{f} = (f_1, \dots, f_{n_1}, g_1, \dots, g_{n_1})^T.$$

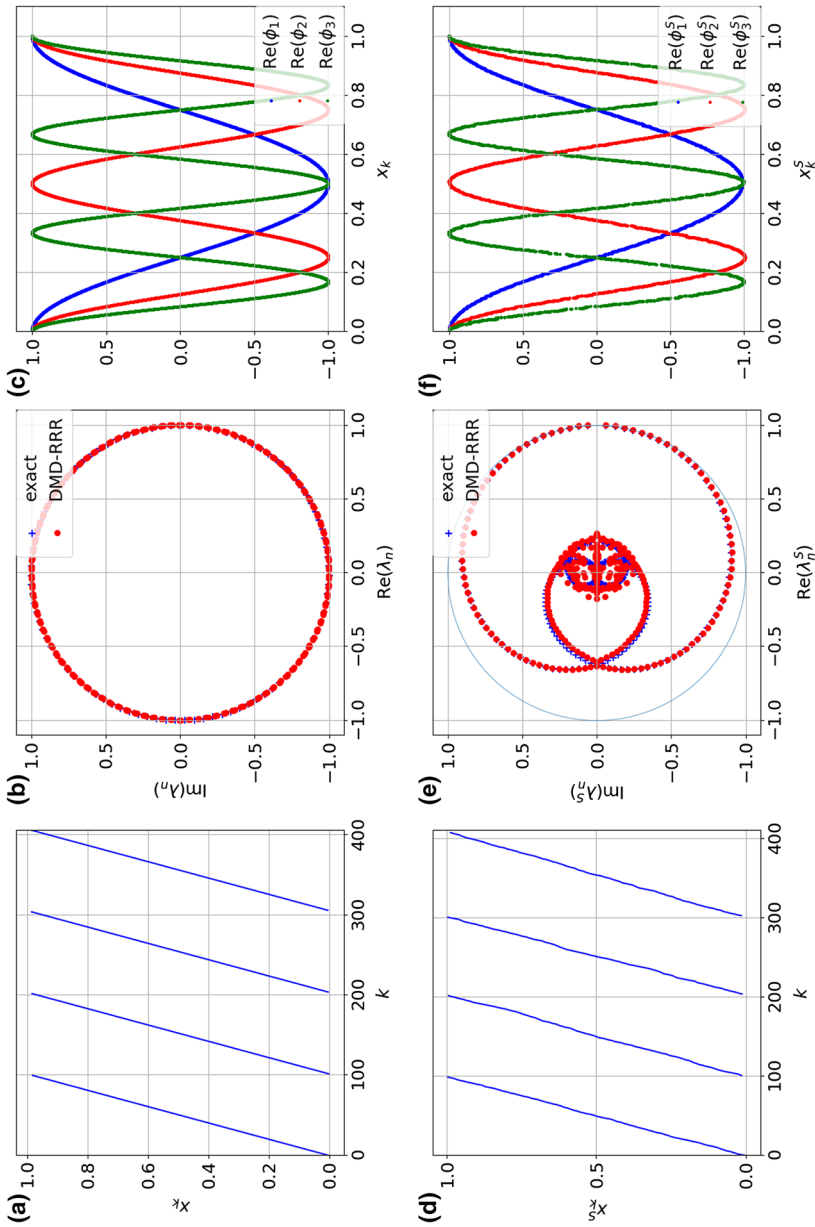
The input matrices  $\mathbf{X}_m$  and  $\mathbf{Y}_m$  of the DMD RRR algorithm are defined by (69). We consider both cases: deterministic and stochastic. In deterministic case, it follows from Tu et al. (2014), that for  $m \gg n$ , the matrix obtained by DMD algorithm is the approximation of the associated Koopman operator in a least-square sense. Therefore, in this example, due to the ergodicity and as a consequence of the law of large numbers, as  $m \rightarrow \infty$ , the obtained values provided by DMD algorithm converge to the eigenvalues and eigenvectors of the finite-dimensional approximating matrix of the Koopman operator. According to Takeishi et al. (2017), the algorithm converge in the stochastic case also to the eigenvalues and eigenfunctions of the stochastic Koopman operator. It is interesting to mention that to recover the spectrum of the stochastic Koopman operator in the considered ergodic system, it was enough to take in the matrices  $\mathbf{X}_m$  and  $\mathbf{Y}_m$  only the values of the observable vector on one long enough trajectory, i.e., to take  $\mathbf{X}_m = (\mathbf{f}(\mathbf{x}_0) \mathbf{f}(T(\omega, \mathbf{x}_0)) \dots \mathbf{f}(T^{m-1}(\omega, \mathbf{x}_0)))$  and  $\mathbf{Y}_m = (\mathbf{f}(T(\omega, \mathbf{x}_0)) \mathbf{f}(T^2(\omega, \mathbf{x}_0)) \dots \mathbf{f}(T^m(\omega, \mathbf{x}_0)))$ , where  $\mathbf{x}_0$  denotes the chosen initial state.

We take  $n_1 = 150$ , which gives  $n = 300$  observable functions, and use  $m = 5000$  sequential expectations of the observable functions to determine the matrices  $\mathbf{X}_m$  and  $\mathbf{Y}_m$ . The numerical results for the flow with the parameters  $\vartheta = \pi/320$  and  $\delta = 0.01$  obtained by using the DMD RRR algorithm are presented in Fig. 1. We note that the eigenvalues obtained in the deterministic case lie, as expected, on unit circle. The real parts of eigenfunctions  $\phi_i(x)$ ,  $i = 1, 2, 3$ , which we recover numerically, are presented in subfigure (c) of Fig. 1. They closely coincide with the theoretically established eigenfunctions given by (51). The results presented in subfigure (e) of Fig. 1 show that numerically captured eigenvalues of the stochastic Koopman operator coincide very well with theoretical eigenvalues (53). In the stochastic case, the real parts of the eigenfunctions belonging to the first three eigenvalues  $\lambda_i^S$ ,  $i = 1, 2, 3$  are presented in subfigure (f) of Fig. 1. As shown in Example 2, the eigenfunctions are given by (51). Two remarks are in order: (1) The deterministic and stochastic Koopman operators commute in this case, and thus the eigenfunctions are the same and the addition of noise does not perturb them (cf. Giannakis 2019 where the numerically computed eigenfunctions in a similar commuting situation were not sensitive to the noise introduced in the system). (2) The computed eigenvalues are the approximation of the eigenvalues associated with the restriction of the stochastic Koopman to the subspace spanned by the components of the vector observable. Since that subspace is invariant under the action of the Koopman operator, the eigenvalues are, as expected, computed with high accuracy.

Results very similar to the presented ones were obtained when the observable vector is constructed with the observable functions:  $f_j(x) = e^{j2\pi i x}$  and  $g_j(x) = e^{-j2\pi i x}$ ,  $j = 1, \dots, n_1$ .

### 5.1.2 Discrete Linear RDS

Here we consider the discrete linear RDS generated by a map (6) with



**Fig. 1** Rotation on circle defined by (48) and (49). Deterministic case  $\theta = \pi/320$ ; **a** solution; **b** Koopman eigenvalues; **c** real part of eigenfunctions. Stochastic case  $\theta = \pi/320$ ,  $\delta = 0.01$ ; **d** solution; **e** stochastic Koopman eigenvalues; **f** real part of eigenfunctions

$$\mathbf{A}(\omega) = \begin{pmatrix} 0 & \pi(\omega) \\ -\pi(\omega) & 0 \end{pmatrix}, \quad (87)$$

where  $\pi$  is given by (44). We suppose that the dynamical system  $\theta = (\theta(t))_{t \in \mathbb{Z}^+}$  is defined by the shift transformations (43). The coordinates  $\omega_i$ ,  $i \in \mathbb{Z}^+$  of  $\omega \in \Omega$  are induced by i.i.d. random variables with the following distribution

$$P(\omega_i = 1) = p_1, \quad P(\omega_i = 2) = 1 - p_1, \quad 0 < p_1 < 1.$$

As proved in Proposition 1, the principal eigenvalues of the stochastic Koopman operator are equal to the eigenvalues of the matrix  $\hat{\mathbf{A}} = \mathbb{E}[\mathbf{A}(\omega)]$ , while the associated eigenfunctions are given by (7). We take  $p_1 = 0.75$  and select  $N = 10^4$  initial points uniformly distributed over  $[0, 1] \times [0, 1]$ . For every chosen initial point  $\mathbf{x}_{j,0}$ ,  $j = 1, \dots, N$ , we determine the random trajectory  $\mathbf{x}_{j,k}$ ,  $k = 1, 2, \dots$ , where  $\mathbf{x}_{j,k}$  denotes state value at  $k$ th step. The DMD RRR algorithm is applied to the full-state observables by taking states on each of the trajectory separately, i.e., we take as the algorithm input the matrices  $\mathbf{X}_{m,j}, \mathbf{Y}_{m,j} \in \mathbb{R}^{2,m}$  defined by  $\mathbf{X}_{m,j} = (\mathbf{x}_{j,0}, \mathbf{x}_{j,1}, \dots, \mathbf{x}_{j,m-1})$  and  $\mathbf{Y}_{m,j} = (\mathbf{x}_{j,1}, \mathbf{x}_{j,2}, \dots, \mathbf{x}_{j,m})$ ,  $j = 1, \dots, N$ . In each computation associated with initial value  $\mathbf{x}_{j,0}$ , we obtain the eigenvalue pair  $\lambda_{1,2}^{(j)}$ , which should approximate the principal Koopman eigenvalues  $\hat{\lambda}_{1,2}$ . For fixed  $m$  and for the chosen set of initial conditions, we get samples of approximating eigenvalues  $\lambda_{1,2}^{(j)}$ ,  $j = 1, \dots, N$ . To estimate the error, for the obtained sets of approximating eigenvalues, we evaluate the  $L_1$ ,  $L_2$  and  $L_\infty$  norms of the difference between the exact eigenvalues and the computed eigenvalues. The norms determined for different values of parameter  $m$  are presented in Fig. 2. One could notice that the accuracy of the obtained eigenvalues increases monotonically with the number  $m$  used in the computations, and that the error is  $\mathcal{O}(\frac{1}{\sqrt{m}})$ , as would be expected for a random process of this type. It is worth noticing that the norm of the state values  $\|\mathbf{x}_{j,m}\|$  increases with  $m$  and could become quite large for large values of  $m$ . Thus, the condition number of the input matrices of the DMD RRR algorithm could be very large. However, the scaling that is applied in this enhanced DMD RRR algorithm prevent the instabilities that otherwise could be introduced if the standard DMD algorithm without scaling is applied.

## 5.2 Continuous-Time Linear RDS

We consider now the continuous-time linear RDS that evolves according to (13) where

$$\mathbf{A}(\omega) = \begin{pmatrix} \pi(\omega) & 1 \\ -b^2 & \pi(\omega) \end{pmatrix}, \quad (88)$$

$\theta(t)$  are shift transformations given by (43), and  $\pi$  is like in the previous test example given by (44). It is supposed that the coordinates  $\omega(t)$ ,  $t \in \mathbb{R}^+$  of  $\omega \in \Omega$  are piecewise constant of the form  $\omega(t) = \sum_{i=1}^{\infty} \omega_i \mathbf{1}_{((i-1)\Delta t^s, i\Delta t^s]}(t)$ , where  $\Delta t^s$  is the chosen fixed time interval between switches and  $\omega_i$ ,  $i \in \mathbb{Z}^+$  are i.i.d. random variables with the



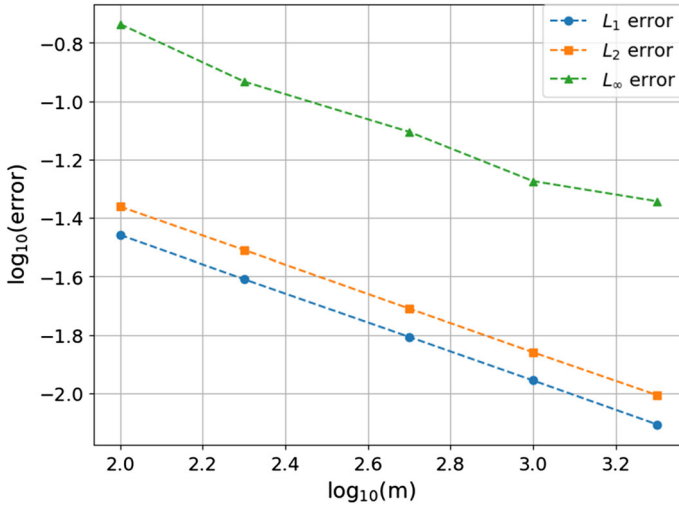


Fig. 2 Discrete linear RDS defined by (87).  $L_1$ ,  $L_2$ , and  $L_\infty$  errors of approximated Koopman eigenvalues

following distribution

$$P(\omega_i = a_1) = p_1, \quad P(\omega_i = a_2) = 1 - p_1, \quad 0 < p_1 < 1. \quad (89)$$

In the considered case, the obtained RDS is not Markovian and the stochastic Koopman operators  $(\mathcal{K}^t)_{t \in \mathbb{T}}$  do not satisfy semigroup property.

We are interested in the numerical approximations of the eigenvalues of the associated stochastic Koopman operators. Therefore we apply the DMD RRR algorithm to the matrices  $\mathbf{X}_m$  and  $\mathbf{Y}_m$  defined using (68) and the full-state observables. Using such approach, we expect to obtain the numerical approximations of the principal eigenvalues of the Koopman operators  $\mathcal{K}^{t_k} = \mathcal{K}_{\Delta t}^k, k = 0, 1, 2, \dots$ . Due to the switches of the matrices at time moments  $i \Delta t_s$ , we choose the time step  $\Delta t$  so that  $K \Delta t = \Delta t_s$  for some  $K \in \mathbb{N}$ .

Since the matrices  $A(\omega)$  commute and have simple eigenvalues, they are simultaneously diagonalizable, so that according to Proposition 2, the stochastic Koopman eigenvalues of the operator  $\mathcal{K}^t$  are equal  $\lambda_j^S(t) = \mathbb{E} \left( e^{\int_0^t \lambda_j(\theta(s)\omega) ds} \right), j = 1, 2$ . For large enough values of  $t$ , for example  $t = N \Delta t^s$ , after applying central limit theorem we get that the distribution of  $\int_0^t \lambda_j(\theta(s)\omega) ds = \sum_{i=1}^N \lambda_j(\omega_i) \Delta t^s$  is approximately normal with mean value  $\hat{\lambda}_j(t) = \mathbb{E}(\lambda_j(\omega_i))t$  and the variance  $\hat{\sigma}_j^2(t) = V(\lambda_j(\omega_i))\Delta t^s t$ , where  $V$  denotes variance. Here  $\lambda_j(\omega_i), j = 1, 2$  are the eigenvalues of the matrix  $\mathbf{A}(\omega_i)$ , thus we get

$$\hat{\lambda}_{1,2}(t) = (\hat{a} \pm bi)t, \quad \text{where } \hat{a} = p_1 a_1 + (1 - p_1)a_2,$$

and

$$\hat{\sigma}_{1,2}^2(t) = p_1(1 - p_1)(a_1 - a_2)^2 \Delta t^s t.$$

It follows that  $e^{\int_0^t \lambda_j(\theta(s)\omega) ds}$  is approximately lognormally distributed with the mean value

$$\lambda_j^S(t) = \mathbb{E} \left( e^{\int_0^t \lambda_j(\theta(s)\omega) ds} \right) = e^{\hat{\lambda}_j(t) + \frac{1}{2} \hat{\sigma}_j^2(t)}. \quad (90)$$

In the numerical computations, we take  $a_1 = -0.1$ ,  $a_2 = 0.1$ ,  $b = 2$ , and  $\Delta t^S = \frac{\pi}{30}$ . The numerical approximations of the expected values of the solution based on  $N = 100$  sampled trajectories and for the values of  $p_1 = 0.75$ ,  $p_1 = 0.5$ , and  $p_1 = 0.25$  are presented in Fig. 3. By taking into account  $m = 100$  initial points, we approximate the eigenvalues of the Koopman operators  $\mathcal{K}^{t_k}$ ,  $t_k = k\Delta t$ , where  $\Delta t = \frac{\pi}{60}$ . The relative error between the eigenvalues (90) and the eigenvalues computed using DMD RRR algorithm is presented in Fig. 4. One can conclude that the computations are stable since the relative errors in all three considered cases have similar magnitudes despite the fact that for  $p_1 = 0.25$  the norm of the solution significantly increases. As in the previous example, this is a consequence of the scaling of data used in the DMD RRR algorithm. As expected, the variance of errors increase with time.

### 5.3 Stochastic Differential Equations Examples

In this section we consider the flow induced by the autonomous SDE of the form (61) with random force accounting for the effect of the noise and disturbances of the system. In the considered examples the numerical solutions are determined by using the Euler–Mayurama method and the Runge–Kutta method for SDE developed in Rößler (2010). We use the fact that the convergence of Euler–Maruyama method is, for the examples we consider here, proved in Hutzenhaler and Jentzen (2015), so that with very small time step we obtain the reference solutions and then check the numerical solutions obtained with Runge–Kutta method. Nevertheless, the Runge–Kutta method performs significantly better and needs less computational effort to attain the same accuracy since much larger time step can be used.

#### 5.3.1 Linear Scalar SDE

We consider the SDE

$$dX = \mu X dt + \sigma dW_t, \quad (91)$$

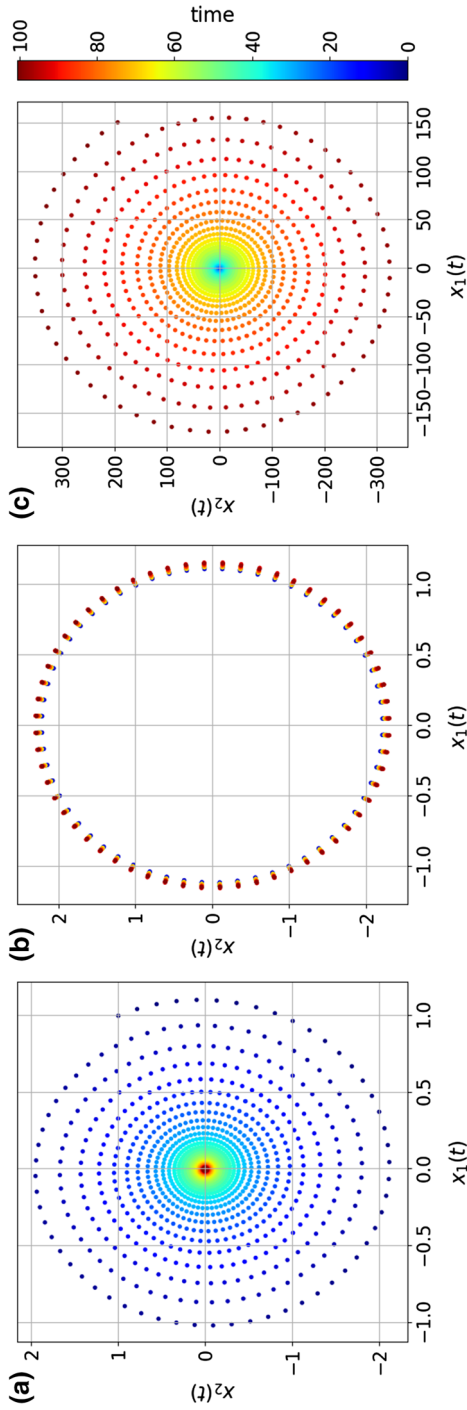
with  $\mu < 0$  and  $\sigma > 0$ , that generates the one-dimensional Ornstein–Uhlenbeck process (Pavliotis 2014). It is known that in the deterministic case, i.e., for  $\sigma = 0$ , the Koopman eigenvalues are equal to

$$\lambda_n = n\mu, \quad n = 0, 1, 2, \dots$$

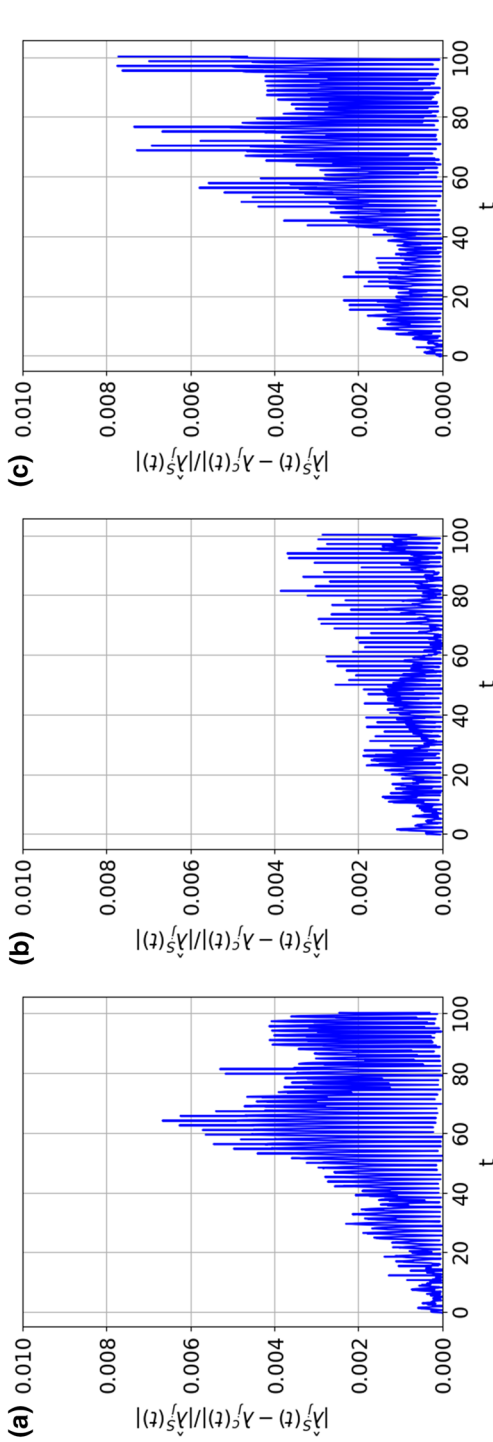
and the related Koopman eigenfunctions are

$$\phi_n(x) = x^n.$$

In the stochastic case, the spectral properties of the stochastic Koopman generator associated with the given equation were studied, for example in Gaspard et al. (1995),



**Fig. 3** Numerical solutions  $x_1^S(t)$ ,  $x_2^S(t)$  of linear RDS defined by (88) for different values of parameter  $p_1$ : **a**  $p_1 = 0.75$ ; **b**  $p_1 = 0.5$ ; **c**  $p_1 = 0.25$ . The color coding refers to time  $t$  (Color figure online)



**Fig. 4** Linear RDS defined by (88). The relative error between the eigenvalues  $\lambda_j^c(t)$  and the computed eigenvalues  $\lambda_j^s(t)$  using DMD RRR algorithm for different values of parameter  $p_1$ : **a**  $p_1 = 0.25$ ; **b**  $p_1 = 0.5$ ; **c**  $p_1 = 0.75$

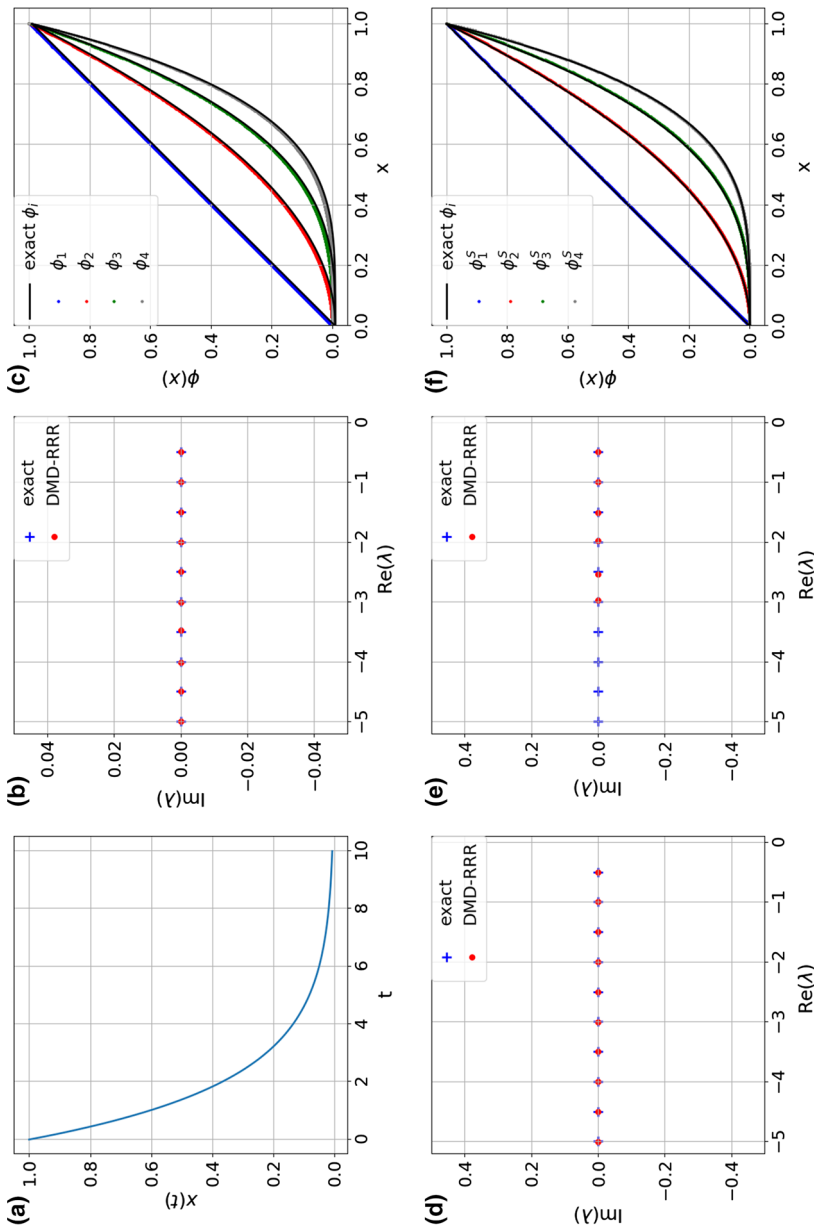
Pavliotis (2014) (under the name backward Kolmogorov equation or Fokker–Planck equation). Its eigenvalues are the same as in deterministic case, while the eigenfunctions are (see Gaspard et al. 1995; Pavliotis 2014)

$$\phi_n(x) = a_n H_n(\alpha x), \quad \alpha = \sqrt{\frac{|\mu|}{\sigma}}, \quad n = 0, 1, 2, \dots$$

Here  $H_n$  denotes Hermite polynomials and  $a_n$  are normalizing parameters.

It is known that the accuracy of the eigenvalues and eigenfunctions obtained by DMD algorithms is closely related to the choice of observable functions. In the deterministic case, when we take monomials  $f_j(x) = x^j$ ,  $j = 1, \dots, n$  as observable functions, satisfactory results for the first  $n$  eigenvalues were obtained for moderate values of  $n$ , for example  $n = 10$ , while for larger values of  $n$  significant numerical errors arise and only few eigenvalues were captured correctly. That is a consequence of highly ill-conditioned system when monomials of higher order  $x^n$  are used, since their evolution, given by  $e^{n\mu t} x^n$ , tends quickly to zero. Therefore the columns of the matrix  $\mathbf{X}_m$  become highly linearly dependent. In the case with  $n = 10$  and  $m = 2000$  snapshots, the ten leading eigenvalues are determined with the accuracy greater than 0.01 (see Fig. 5b). It is worth mentioning that the standard DMD algorithm, which does not include scaling of data, was more sensitive and when it was applied on the same series of snapshots, at most three eigenvalues with the accuracy greater than 0.01 were computed.

In the stochastic case, the DMD RRR algorithm was applied on the same set of observable functions as in the deterministic case. We use both approaches described in the Sect. 4.1 for defining input matrices  $\mathbf{X}_m$  and  $\mathbf{Y}_m$ . In the first approach, for chosen  $m = 100$  initial points from the interval  $[-1, 1]$  and  $n$ -dimensional observable vector, we define  $\mathbf{X}_m, \mathbf{Y}_m \in \mathbb{R}^{n \times m}$  by (68), where  $\mathbf{Y}_m$  is determined for  $k = 100$  and by using  $N = 1000$  trajectories for each initial condition. In the second approach, for one chosen initial point, we generate  $N = 1000$  trajectories to determine the approximations of the expected values of the observable functions, which are then used in (69) to form input matrices  $\mathbf{X}_m$  and  $\mathbf{Y}_m$  for  $m = 2000$ . The eigenvalues computed by using the first approach lead to very accurately computed eigenvalues of the Koopman operator (see Fig. 5d) with the accuracy similar as in deterministic case. The accuracy depends on the number of computations  $N$ , as well as on the number and on the distribution of the initial points. In the second approach, the number of eigenvalues captured with satisfactory accuracy decreases and in the presented case only four of them are captured with the accuracy greater than  $10^{-2}$  (see Fig. 5e). The eigenfunctions associated with this four eigenvalues are captured with satisfactory accuracy (Fig. 5f). The reason for the lower accuracy could be a consequence of the errors that are introduced into the matrices  $\mathbf{X}_m$  and  $\mathbf{Y}_m$  when the expected values of the observable functions are approximated by averaging its values along the trajectories. When multiple initial conditions are used such type of error is introduced only in the matrix  $\mathbf{Y}_m$  and not in the matrix  $\mathbf{X}_m$ . The accuracy of the algorithm could be increased by increasing the number of computed trajectories since then, due to the law of large numbers, the more accurate approximations of expected values of the observables should be obtained.



**Fig. 5** Linear scalar equation (91). Deterministic case  $\mu = -0.5$ ; **a** solution; **b** Koopman eigenvalues; **c** Koopman eigenfunctions; Stochastic case  $\mu = -0.5$   $\sigma = 0.001$ ; **d** stochastic Koopman eigenvalues—1st approach; DMD RRR with (68); **e** stochastic Koopman eigenvalues—2nd approach; DMD RRR with (69); **f** stochastic Koopman eigenfunctions

### 5.3.2 Nonlinear Scalar SDE

One of the simplest nonlinear stochastic dynamical systems is the SDE of the form

$$dX = (\mu X - X^3)dt + \sigma dW_t. \quad (92)$$

Its solutions are of different nature for  $\mu > 0$ ,  $\mu = 0$  and  $\mu < 0$  (Gaspard et al. 1995) and therefore the related deterministic equation (for  $\sigma = 0$ ) usually appears in the literature when studying the bifurcation phenomena. We consider here the equation for fixed parameter  $\mu < 0$ . The eigenvalues and the eigenfunctions of the Liouville operator associated with this equation and its adjoint were derived in Gaspard et al. (1995). Since (for  $\sigma = 0$ ) the adjoint of the Liouville operator is actually the Koopman operator of the deterministic equation, we have from (Gaspard et al. 1995) that the eigenvalues are equal to

$$\lambda_n = n\mu, n = 0, 1, 2, \dots,$$

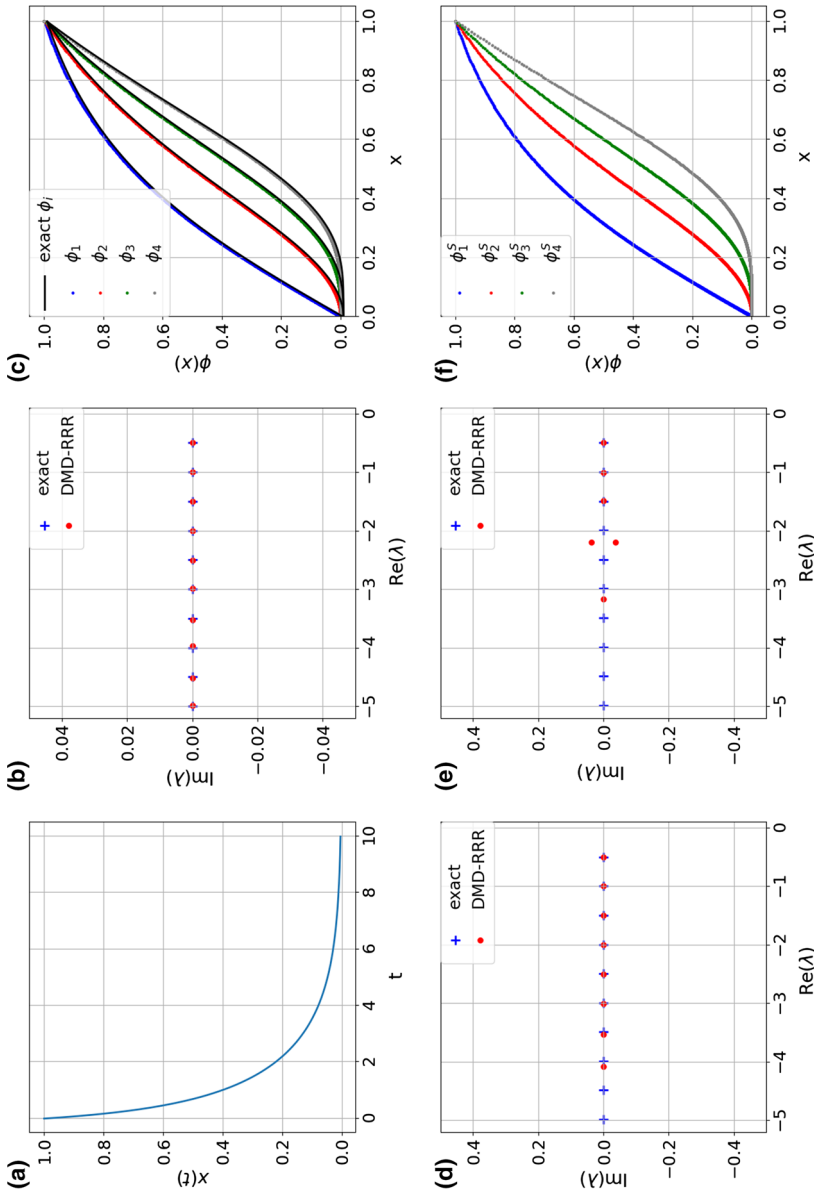
while the associated eigenfunctions are

$$\phi_n(x) = \left( \frac{x}{\sqrt{x^2 + |\mu|}} \right)^n.$$

In the stochastic case, when  $\sigma > 0$ , the eigenvalues and the eigenfunctions of the generator of the stochastic Koopman operator family can be evaluated by solving the associated backward Kolmogorov equation, which can be transformed to the Schrödinger equation (see Gaspard et al. 1995). We determine the approximations of eigenvalues numerically by solving Schrödinger equation using finite difference method. For small values of  $\sigma$  the eigenvalues are very similar to the deterministic case.

As in the linear scalar case, we first apply the DMD RRR algorithm to compute the eigenvalues and the eigenfunctions of the Koopman operator in the deterministic case. The specific set of observables used in the computation is of importance for the accuracy of the results. When we use the analytically known eigenfunctions or their linear combinations as the observable functions, the subspace spanned by these functions is closed under the action of the Koopman operator, so that its restriction on that subspace can be exactly described by a finite-dimensional matrix, which is the key matrix used for obtaining outputs of DMD RRR algorithm. As a consequence, the eigenvalues and the eigenfunctions are computed with high accuracy (see Fig. 6). On the other hand, when the set of observables was composed of monomials, the results were less accurate and only a few leading eigenvalues and eigenfunctions were computed with satisfactory accuracy.

In the stochastic case, similarly as in the previous linear example, we use both DMD RRR approaches and the same set of observable functions as in deterministic case. The results obtained by using  $m = 100$  initial conditions, and the expected value of the observable functions determined over  $N = 1000$  trajectories and after  $k = 100$  time



**Fig. 6** Nonlinear scalar equation (92). Deterministic case  $\mu = -0.5$ ; **a** solution; **b** Koopman eigenvalues; **c** Koopman eigenfunctions. Stochastic case  $\mu = -0.5, \sigma = 0.001$ ; **d** stochastic Koopman eigenvalues—1st approach: DMD RRR with (68); **e** stochastic Koopman eigenvalues—second approach: DMD RRR with (69); **f** stochastic Koopman eigenfunctions



steps give the same number of accurately computed eigenvalues as in deterministic case (see Fig. 6d). On the other hand, when the second approach associated with one initial point is used, only first three eigenvalues are computed with satisfactory accuracy despite the fact that the same number of computed trajectories  $N = 1000$  and a large number of columns  $m = 2000$  was used. The reason for lower accuracy could be in the error introduced into the matrices  $\mathbf{X}_m$  and  $\mathbf{Y}_m$  as was described in the linear scalar case.

### 5.3.3 Stuart–Landau Equations

The Stuart–Landau equations describe the behavior of a nonlinear oscillating system near the Hopf bifurcation. In polar coordinates  $(r, \theta)$ , the system reads

$$\begin{aligned} dr &= (\delta r - r^3)dt \\ d\theta &= (\gamma - \beta r^2)dt, \end{aligned} \tag{93}$$

where  $\beta, \gamma,$  and  $\delta$  are parameters of the system. The behavior of the system depends on these parameters. The parameter  $\delta$  associated with the Hopf bifurcation controls the stability of the fixed point  $x^*$  ( $\delta < 0$ ) and of the limit cycle  $\Gamma$  ( $\delta > 0$ ). We consider here the system with the limit cycle, i.e., for  $\delta > 0$ . According to Tantet et al. (2017), the base frequency of the limit cycle  $\Gamma : r = \sqrt{\delta}$  is equal to  $\omega_0 = \gamma - \beta\delta$  and the eigenvalues of the system are  $\lambda_{ln} = -2l\delta + in\omega_0, l \in \mathbb{N}, n \in \mathbb{Z}$ . The stochastic Stuart–Landau equations are defined by

$$\begin{aligned} dr &= \left( \delta r - r^3 + \frac{\epsilon^2}{r} \right) dt + \epsilon dW_r \\ d\theta &= (\gamma - \beta r^2)dt + \frac{\epsilon}{r} dW_\theta, \end{aligned} \tag{94}$$

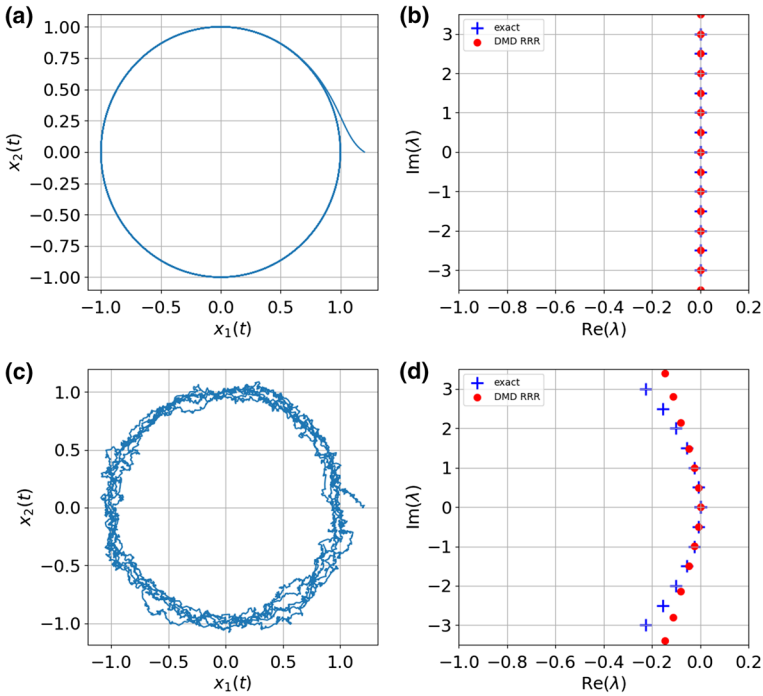
where  $W_r$  and  $W_\theta$  satisfy SDE system

$$\begin{aligned} dW_r &= \cos \theta dW_x + \sin \theta dW_y \\ dW_\theta &= -\sin \theta dW_x + \cos \theta dW_y, \end{aligned}$$

and  $dW_x$  and  $dW_y$  are independent Wiener processes. A detailed analysis of that system is provided in Tantet et al. (2017). For small noise, controlled by the parameter  $\epsilon$ , the approximation of Koopman eigenvalues are derived in Tantet et al. (2017), so that for the system with a stable limit cycle  $\Gamma$  ( $\delta > 0$ ), the eigenvalues are given by

$$\lambda_{ln} = \begin{cases} -\frac{n^2\epsilon^2(1+\beta^2)}{2\delta} + in\omega_0 + \mathcal{O}(\epsilon^4), l = 0 \\ -2l\delta + in\omega_0 + \mathcal{O}(\epsilon^2), l > 0 \end{cases} . \tag{95}$$

In order to determine the Koopman eigenvalues, we apply the sHankel–DMD algorithm to the scalar observable function of the form  $f(r, \theta) = \sum_{k=1}^K e^{\pm ik\theta}$ ,



**Fig. 7** Stuart–Landau system with parameters  $\delta = 0.5, \beta = 1, \gamma = 1$ . Deterministic case—Eq. (93): **a** solution; **b** Koopman eigenvalues. Stochastic case—Eq. (94): **c** solution; **d** stochastic Koopman eigenvalues. The threshold for the residuals is set to 0.001

$k = 1, \dots, K$  in the deterministic case, and to the function

$$f(r, \theta) = \sum_{k=1}^K e^{\pm ik(\theta - \beta \log(r/\delta))}, k = 1, \dots, K$$

in the stochastic case to calculate the eigenvalues  $\lambda_{0,n}, n = 1, \dots, K$ . The results for both cases, deterministic and stochastic, are presented in Fig. 7. For moderate values of  $n$ , quite good approximations of eigenvalues  $\lambda_{0,n}$  are obtained. We mention that to obtain the base rotating frequency and corresponding multiples, the time span of values or expected values of the observable functions used in the DMD RRR algorithm must include the basic period.

### 5.3.4 Noisy Van der Pol Oscillator

We analyze the two-dimensional test example in which the deterministic part is the system of two equations modeling standard Van der Pol oscillator. The stochastic part is modeled by the one-dimensional Wiener process so that the following system is

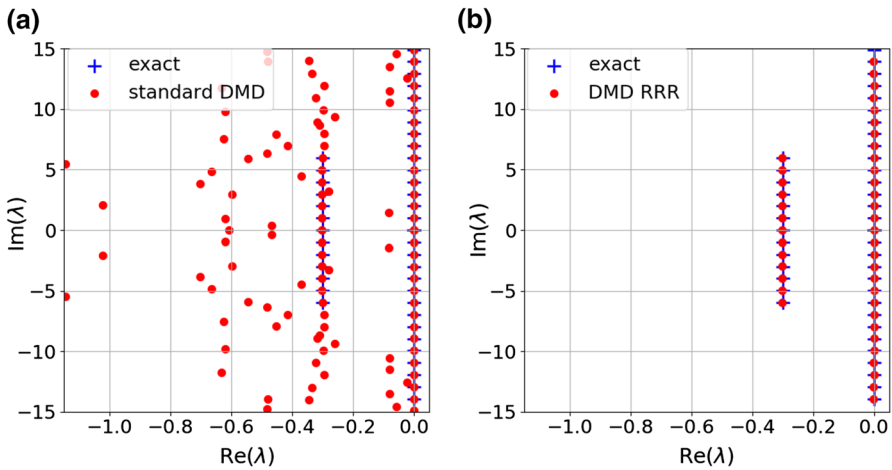
considered

$$\begin{aligned} dX_1 &= X_2 dt \\ dX_2 &= \left( \mu(1 - X_1^2)X_2 - X_1 \right) dt + \sqrt{2\epsilon} dW_t. \end{aligned} \quad (96)$$

It was proved in Arnold (1998) that the solution of the considered SDE system exists. Furthermore, it was proved in Hutzenhaler and Jentzen (2015) that the numerical solutions computed with the Euler–Maruyama method converge to the exact solution of the considered SDE, which is also true for the Runge–Kutta method for SDE (Rößler 2010) used here.

As in the previous examples, we consider first the deterministic system. It is well known that the dynamics converge to an asymptotic limit cycle whose basin of attraction is the whole  $\mathbb{R}^2$ . According to Strogatz (1994), for the base frequency of the limit cycle the expression  $\omega_0 = 1 - \frac{1}{16}\mu^2 + \mathcal{O}(\mu^3)$  is valid, thus for the parameter value  $\mu = 0.3$  the basic frequency of the limit cycle is  $\omega_0 \approx 0.995$ . For the chosen initial state, the base frequency computed by applying the Hankel–DMD RRR algorithm using the scalar observable function  $f(X_1, X_2) = X_1 + X_2 + \sqrt{X_1^2 + X_2^2}$ , is equal to 0.9944151. In Fig. 8 we present the eigenvalues computed with the standard DMD algorithm and the eigenvalues obtained by using DMD RRR algorithm for which the residuals of eigenvectors determined using (72) are smaller than the chosen threshold  $\eta_0 = 10^{-3}$ . The number of eigenvalues and eigenvectors obtained by using both DMD algorithms is 250 for the chosen Hankel matrix with  $n = 250$  columns and a larger number of rows, but for only few of them the residuals of the eigenvectors do not exceed the chosen threshold. The eigenvalues obtained using DMD RRR algorithm form a lattice structure containing the eigenvalues of the form  $\{k\omega_0, -\mu + k\omega_0 : k \in \mathbb{Z}\}$ , which is consistent with the theoretical results given in Mezić (2017b). As evident from Fig. 8, the standard DMD algorithm without residue evaluation computes many spurious eigenvalues. In Fig. 9 we present the solution and the eigenvalues for which the residuals of eigenvectors do not exceed the chosen threshold in DMD RRR. The time evolution of the real part of the first three eigenfunctions of the Van der Pol system presented in the same figure coincide with the results presented in Arbabi and Mezić (2017).

The same approach as in the deterministic case is applied on the solution obtained from the stochastic system (96). From Fig. 9 we can conclude that some sort of stochastic asymptotic limit cycle appears, i.e., for large  $t$  the solution is “smeared out” around the deterministic limit cycle. Such statistical equilibrium can be determined by solving the associated Fokker–Planck equation. In Fantuzzi et al. (2016) the energy bounds of the noisy van der Pol oscillator were determined. For small and moderate noise, the base frequency of the averaged limit cycle remains similar to the deterministic case (Leung 1995). Here we take the noise parameter  $\epsilon = 0.005$ . The sHankel–DMD algorithm is applied to the same observable function as in the deterministic case and by using the same matrix dimension. The eigenvalues and eigenfunctions of the related stochastic Koopman operator are not known analytically. They can be approximated by the numerical discretization of the associated generator, however their evaluation is out of the scope of this paper. Again, as in the deterministic case, only the eigen-



**Fig. 8** Van der Pol oscillator (96). Deterministic case: **a** eigenvalues obtained by using standard DMD algorithm; **b** eigenvalues obtained by using DMD RRR algorithm with the threshold for the residuals equal to 0.001

values for which the residuals of the corresponding eigenvectors do not exceed the threshold  $\eta_0 = 10^{-3}$  are selected and presented in Fig. 9. By comparing the results shown in Fig. 9, we conclude that the numerically determined time evolution of eigenfunctions in the stochastic case is very similar to the evolution of eigenfunctions in the deterministic case.

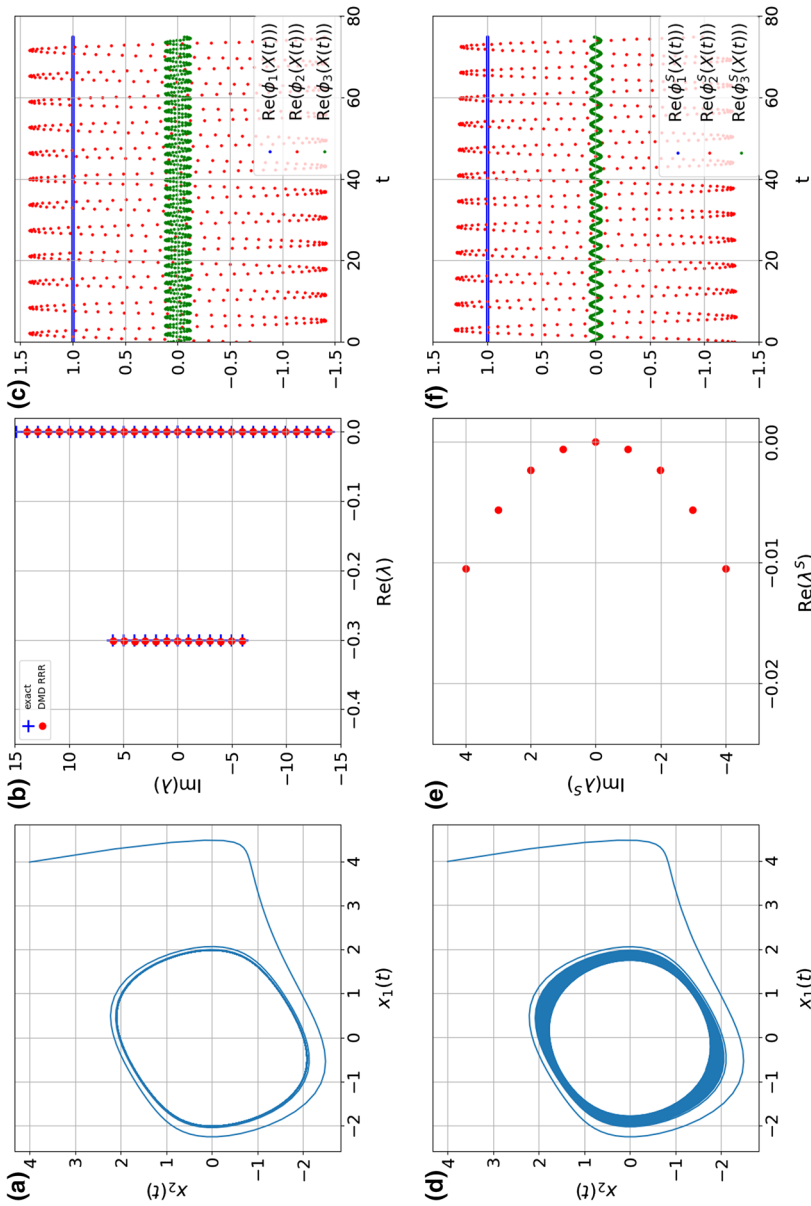
### 5.3.5 Noisy Lotka–Volterra (Predator–Prey) System

Here we consider the Lotka–Volterra system with competition, describing the predator prey system with two species. Such type of a model is typically used by biologists and ecologists for modeling the population dynamics. The corresponding noisy system is given by:

$$\begin{aligned} dX_1 &= (a_1 - b_1X_2 - c_1X_1) X_1dt + \sigma_1X_1dW_t^1 \\ dX_2 &= (-a_2 + b_2X_1 - c_2X_2) X_2dt + \sigma_2X_2dW_t^2. \end{aligned} \tag{97}$$

The model parameters  $a_1, b_1, c_1, a_2, b_2, c_2 > 0$  depend on the particular species under consideration. The intensity of the noise is modeled by nonnegative parameters  $\sigma_1$  and  $\sigma_2$ .

We first consider the properties of the deterministic system. The first quadrant  $\mathbb{R}_+^2 = \{(X_1, X_2) | X_1 \geq 0, X_2 \geq 0\}$  is the domain of evolution of the system. There is an equilibrium point on the  $x_1$ -axis  $(\frac{a_1}{c_1}, 0)$ . If the nullclines  $a_1 - b_1X_2 - c_1X_1 = 0$  and  $-a_2 + b_2X_1 + c_2X_2 = 0$  do not intersect, the equilibrium point on the positive  $x_1$  axis is unique. On the other hand, if these lines do intersect at the point  $(x_1^*, x_2^*)$ , it can be shown that this is an asymptotically stable equilibrium point whose basin of attraction is  $\mathbb{R}_+^2$  with the exception of the axes. The considered system can be linearized around the equilibrium point. As known from the Hartman–Grobman theorem, if all



**Fig. 9** Van der Pol oscillator (96). Deterministic case: **a** solution; **b** Koopman eigenvalues; **c** the time evolution of real part of Koopman eigenfunctions along trajectories. Stochastic case  $\epsilon = 0.005$ : **d** solution; **e** stochastic Koopman eigenvalues; **f** the time evolution of real part of stochastic Koopman eigenfunctions along trajectories. The threshold for the residuals is set to 0.001

the eigenvalues of the Jacobian matrix at the fixed point have negative (or positive) real part, then there exist a  $C^1$ -diffeomorphism  $\mathbf{h}$  defining the conjugacy map that locally transform the system to the linear one. Even more, as proved in Lan and Mezić (2013), if all the eigenvalues have negative real part (i.e., if the fixed point is exponentially stable), the  $C^1$ -diffeomorphism  $\mathbf{h}$  is defined on the whole basin of attraction and the system is conjugate to the linear one globally on the basin of attraction. This implies that the eigenvalues of the Koopman operator are the same as the eigenvalues of the Koopman operator associated with the linear system.

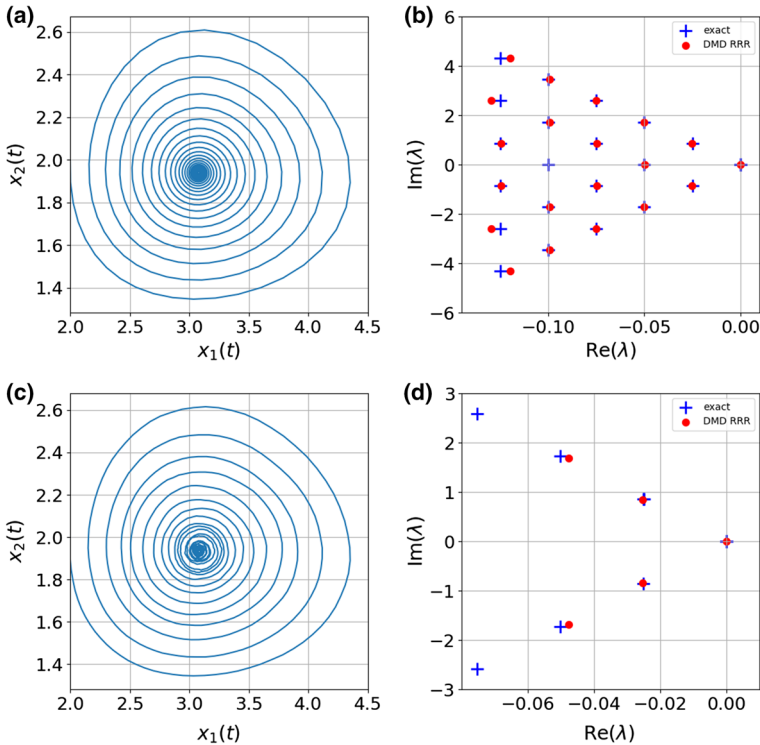
We select  $a_1 = 1.0$ ,  $b_1 = 0.5$ ,  $c_1 = 0.01$ ,  $a_2 = 0.75$ ,  $b_2 = 0.25$ ,  $c_2 = 0.01$ . The off-axes equilibrium point is equal to  $(x_1^*, x_2^*) = (3.07754, 1.93845)$ . Since the eigenvalues of the Jacobian matrix at this point are equal to  $\lambda_{1,2} = -0.02500799 \pm 0.863524i$ , we conclude that this is an exponentially stable fixed point and that the system is conjugate to the linear one on its basin of attraction.

To numerically evaluate the eigenvalues and eigenfunctions of the Koopman operator we use, as in the Van der Pol oscillator example, the Hankel–DMD RRR algorithm, where the Hankel matrix is defined for the scalar observable function  $f(X_1, X_2) = X_1 + X_2$ . We use the matrix with  $n = 250$  rows and  $m = 100$  columns. By taking into account only the eigenvalues related to the eigenvectors with the residuals smaller than  $\eta_0 = 10^{-3}$ , several leading eigenvalues of the Koopman operator are captured with high accuracy (see Fig. 10). The principal eigenvalues  $\lambda_{1,2}$  are calculated with the accuracy greater than  $10^{-6}$ .

The solution and the asymptotic properties of noisy Lotka–Volterra equation (97) are considered in Arató (2003). In the stochastic case, for small enough parameters  $\sigma_1$  and  $\sigma_2$ , one can determine the coordinates of the fixed point  $(\bar{x}_1^*, \bar{x}_2^*)$  as the solution of equations:  $a_1 - \frac{\sigma_1^2}{2} - b_1 X_2 - c_1 X_1 = 0$  and  $a_2 - \frac{\sigma_2^2}{2} - b_2 X_2 - c_2 X_1 = 0$ . According to Arató (2003), one can expect heuristically that  $\xi_1 = X_1 - \bar{x}_1^*$  and  $\xi_2 = X_2 - \bar{x}_2^*$  are components of the Ornstein–Uhlenbeck process in two dimensions, which is a solution of the corresponding linear SDE obtained by the linearization around the point  $(\bar{x}_1^*, \bar{x}_2^*)$ . For small enough  $\sigma_1$  and  $\sigma_2$ , the fixed point  $(\bar{x}_1^*, \bar{x}_2^*)$  remains asymptotically stable.

In the stochastic example that we consider here with the same parameters as in the deterministic case and for the variance parameters  $\sigma_1 = \sigma_2 = 0.05$ , the fixed point is equal to  $(\bar{x}_1^*, \bar{x}_2^*) = (3.08243, 1.93585)$ . After linearizing (97) around this point, we get that the eigenvalues of the corresponding Jacobian matrix are equal to  $\lambda_{1,2}^S = -0.02509 \pm 0.86363i$ . These eigenvalues coincide with the principal eigenvalues of the stochastic Koopman generator associated with the linear system.

Like in the deterministic case, we use the sHankel–DMD algorithm, applied to the same observable function as before. The expectations are determined as the average values over  $N = 1000$  trajectories (see Fig. 10). The principal eigenvalues of the Koopman operator in the considered stochastic case are evaluated by using the larger Hankel matrix than in deterministic case, with  $n = 750$  and  $m = 250$ . Using the residual threshold  $\eta_0 = 10^{-3}$ , we chose only the eigenvalues related to the eigenvectors having small enough residual. The remaining set of eigenvalues is presented in Fig. 10d. One can see that the principal eigenvalues are determined quite accurately, however, the higher order ones are missing.



**Fig. 10** Lotka–Volterra system (97). Deterministic case: **a** solution; **b** Koopman eigenvalues. Stochastic case: **c** solution; **d** stochastic Koopman eigenvalues—the exact eigenvalues refer to the determined eigenvalues  $\lambda_{1,2}^S$  that we heuristically expect to be valid in the stochastic case. The threshold for the residuals is set to 0.001

### 6 Conclusions

In this work, we consider the generators and spectral objects of stochastic Koopman operators associated with random dynamical systems. We explore linear RDS, without invoking Markovian property, and (generally nonlinear) RDS with the Markov property that implies the associated stochastic Koopman family is a semigroup.

To numerically compute the eigenvalues and the eigenfunctions of the stochastic Koopman operator family, we propose approaches that enable use of DMD algorithms in the stochastic framework for both Markovian and non-Markovian cases. We define the stochastic Hankel DMD algorithm and prove that under the assumption of ergodicity of RDS and the existence of finite-dimensional invariant Krylov subspace, it converges to the true eigenvalues and eigenfunctions of the stochastic Koopman operator.

The results presented for a variety of numerical examples show stability of the proposed numerical algorithms that reveal the spectral objects of the stochastic Koopman operators. Specifically, we use the DMD RRR algorithm (Drmač et al. 2018) that enables precise determination of spectral objects via control of the residuals.

There are many interesting questions remaining in the study of stochastic Koopman operators. Some of the most intriguing ones relate to relationship of the geometry of the level sets of stochastic Koopman operator eigenfunctions to the dynamics of the underlying RDS, initiated in Mezić and Banaszuk (2004) for the case of asymptotic dynamics, and eigenvalues on the unit circle. We are currently pursuing such work for the more general class of eigenvalues off the unit circle and their associated eigenfunctions. Furthermore, the application of the proposed algorithms on some more complex dynamical systems in the stochastic framework could be very interesting, thus we plan such investigations in our future work.

**Acknowledgements** This research has been supported by the DARPA Contract HR0011-16-C-0116 and HR0011-18-9-0033, AFOSR Grants FA9550-08-1-0217 and FA9550-17-C-0012, and support of scientific research of the University of Rijeka, Project No. 18-118-1257. We are thankful to Milan Korda, Allan Avila and anonymous referees for useful comments that helped to substantially improve the manuscript.

## References

- Arató, M.: A famous nonlinear stochastic equation: Lotka–Volterra model with diffusion. *Math. Comput. Model.* **38**, 709–726 (2003)
- Arbabi, H., Mezić, I.: Ergodic theory, dynamic mode decomposition and computation of spectral properties of the Koopman operator. *SIAM J. Appl. Dyn. Syst.* **16**(4), 2096–2126 (2017)
- Arnold, L.: *Stochastic Differential Equations*. Wiley, New York (1974)
- Arnold, L.: *Random Dynamical Systems*. Springer, New York (1998)
- Arnold, L., Crauel, H.: Random dynamical systems. Lypunov exponents. In: *Proceedings Oberwolfach 1990*. Springer Lecture notes in Mathematics, vol. 1486, pp. 1–22 (1991)
- Bagheri, S.: Koopman-mode decomposition of the cylinder wake. *J. Fluid Mech.* **726**, 596–623 (2013)
- Budišić, M., Mohr, R., Mezić, I.: Applied Koopmanism. *Chaos* **22**(4), 596–623 (2012)
- Cohen, S.N., Elliott, R.J.: *Stochastic Calculus and Applications*, 2nd edn. Birkhäuser, New York (2015)
- Crauel, H.: Markov measures for random dynamical systems. *Stoch. Stoch. Rep.* **37**(3), 153–173 (1991)
- Da Prato, G.J.Z.: *Stochastic Equations in Infinite Dimensions*, 2nd edn. Cambridge University Press, Cambridge (2014)
- Drmač, Z., Mezić, I., Mohr, R.: Data driven modal decompositions: analysis and enhancements. *SIAM J. Sci. Comput.* **40**(4), A2253–A2285 (2018)
- Dynkin, E.B.: *Markov Processes*, 1, 2, Grundlehren der Mathematischen Wissenschaften, vol. 121, 122. Springer, New York (1965)
- Engel, K., Nagel, R.: *One-Parameter Semigroups for Linear Evolution Operators*. Springer, New York (2001)
- Fantuzzi, G., Goluskin, D., Huang, D., Chernyshenko, S.: Bounds for deterministic and stochastic dynamical systems using sum-of-squares optimization. *SIAM J. Appl. Dyn. Syst.* **15**(4), 1962–1988 (2016)
- Gaspard, P., Nicolis, G., Provata, A., Tasaki, S.: Spectral signature of the pitchfork bifurcation: Liouville equation approach. *Phys. Rev. E* **51**(1), 74 (1995)
- Giannakis, D.: Data-driven spectral decomposition and forecasting of ergodic dynamical systems. *Appl. Comput. Harmon. Anal.* **47**(2), 338–396 (2019)
- Hemati, M.S., Rowley, C.W., Deem, E.A., Cattafesta, L.N.: De-biasing the dynamic mode decomposition for applied koopman spectral analysis. *Theor. Comput. Fluid. Dyn.* **31**(4), 349–368 (2017)
- Horn, R.A., Johnson, C.R.: *Matrix Analysis*, 2nd edn. Cambridge University Press, Cambridge (2013)
- Hutzenthaler, M., Jentzen, A.: *Numerical Approximations of Stochastic Differential Equations with Non-globally Lipschitz Continuous Coefficients*. *Memoirs of the American Mathematical Society*, vol. 236, no. 1112 (2015)
- Junge, O., Marsden, J.E., Mezić, I.: Uncertainty in the dynamics of conservative maps. In: *43rd IEEE Conference on Decision and Control* (2004)
- Klus, S., Koltai, P., Schütte, C.: On the numerical approximation of the Perron–Frobenius and Koopman operator. *J. Comput. Dyn.* **3**(1), 51–79 (2016)



- Klus, S., Nüske, F., Koltai, P., Wu, H., Kevrekidis, I., Schütte, C., Noé, F.: Data-driven model reduction and transfer operator approximation. *J. Nonlinear Sci.* **28**(3), 985–1010 (2018)
- Klus, S., Schütte, C.: Towards tensor-based methods for the numerical approximation of the Perron–Frobenius and Koopman operator. *J. Comput. Dyn.* **3**(2), 139–161 (2016)
- Koopman, B.: Hamiltonian systems and transformation in Hilbert space. *Proc. Natl. Acad. Sci. USA* **17**(5), 315 (1931)
- Korda, M., Mezić, I.: On convergence of extended dynamic mode decomposition to the Koopman operator. *J. Nonlinear Sci.* **28**(2), 687–710 (2018)
- Lan, Y., Mezić, I.: Linearization in the large of nonlinear systems and Koopman operator spectrum. *Physica D* **242**(1), 42–53 (2013). <https://doi.org/10.1016/j.physd.2012.08.017>
- Lasota, A., Mackey, M.C.: *Chaos, Fractals, and Noise, Stochastic Aspects of Dynamics*, 2nd edn. Springer, New York (1994)
- Leung, H.: Stochastic transient of a noisy van der Pol oscillator. *Physica A* **221**, 340–347 (1995)
- Maćešić, S., Črnjarić Žic, N., Mezić, I.: Koopman operator family spectrum for nonautonomous systems. *SIAM J. Appl. Dyn. Syst.* **17**(4), 2478–2515 (2018)
- Mauroy, A., Mezić, I., Moehlis, J.: Isostables, isochrons, and koopman spectrum for the action-angle representation of stable fixed point dynamics. *Physica D Nonlinear Phenom.* **261**, 19–30 (2013)
- Mezić, I.: On the geometrical and statistical properties of dynamical systems: theory and applications. Ph.D. thesis, California Institute of Technology (1994)
- Mezić, I.: Spectral properties of dynamical systems, model reduction and decompositions. *Nonlinear Dyn.* **41**(1), 309–325 (2005)
- Mezić, I.: Analysis of fluid flows via spectral properties of the Koopman operator. *Annu. Rev. Fluid Mech.* **45**, 357–378 (2013)
- Mezić, I.: On applications of the spectral theory of the Koopman operator in dynamical systems and control theory. In: 2015 IEEE 54th Annual Conference on Decision and Control (CDC), pp. 7034–7041. IEEE (2015)
- Mezić, I.: Koopman operator spectrum and data analysis (2017a). arXiv preprint [arXiv:1702.07597](https://arxiv.org/abs/1702.07597)
- Mezić, I.: Koopman Operator Spectrum and Data Analysis. ArXiv e-prints (2017b)
- Mezić, I., Banaszuk, A.: Comparison of systems with complex behavior: Spectral methods. In: Proceedings of the 39th IEEE Conference on Decision and Control, 2000, vol. 2, pp. 1224–1231. IEEE (2000)
- Mezić, I., Banaszuk, A.: Comparison of systems with complex behavior. *Physica D* **197**(1), 101–133 (2004)
- Mezić, I., Surana, A.: Koopman mode decomposition for periodic/quasi-periodic time dependence. *IFAC-PapersOnLine* **49**(18), 690–697 (2016). <https://doi.org/10.1016/j.ifacol.2016.10.246>
- Mezić, I., Wiggins, S.: A method for visualization of invariant sets of dynamical systems based on the ergodic partition. *Chaos Interdiscip. J. Nonlinear Sci.* **9**(1), 213–218 (1999)
- Pavliotis, G.A.: *Stochastic Processes and Applications: Diffusion Processes, the Fokker–Planck and Langevin Equations*, Volume 60 of Texts in Applied Mathematics. Springer, New York (2014)
- Proctor, J., Eckhoff, P.A.: Discovering dynamic patterns from infectious disease data using dynamic mode decomposition. *Int. Health* **7**, 139–145 (2015)
- Rößler, A.: Runge–Kutta methods for the strong approximation of solutions of stochastic differential equations. *SIAM J. Numer. Anal.* **48**(3), 922–952 (2010)
- Rowley, C., Mezić, I., Bagheri, S., Schlatter, P., Henningson, D.: Spectral analysis of nonlinear flows. *J. Fluid Mech.* **641**(1), 115–127 (2009)
- Schmid, P.: Dynamic mode decomposition of numerical and experimental data. In: Sixty-First Annual Meeting of the APS Division of Fluid Dynamics (2008)
- Schmid, P.: Dynamic mode decomposition of numerical and experimental data. *J. Fluid Mech.* **656**(1), 5–28 (2010)
- Schmid, P., Li, L., Juniper, M., Pust, O.: Applications of the dynamic mode decomposition. *Theor. Comput. Fluid Dyn.* **25**(1), 249–259 (2011)
- Sharma, A.S., Mezić, I., McKeon, B.J.: Correspondence between koopman mode decomposition, resolvent mode decomposition, and invariant solutions of the navier-stokes equations. *Phys. Rev. Fluids* **1**(3), 032,402 (2016)
- Shnitzer, T., Talmon, R., Slotine, J.: Manifold learning with contracting observers for data-driven time-series analysis. *IEEE Trans. Signal Proces.* **65**, 904–918 (2017)
- Strand, J.: Random ordinary differential equations. *J. Differ. Equ.* **7**, 538–553 (1970)
- Strogatz, S.H.: *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry and Engineering*. Perseus Books Publishing, New York (1994)

- Susuki, Y., Mezić, I.: Nonlinear Koopman modes and a precursor to power system swing instabilities. *IEEE Trans. Power Syst.* **27**(3), 1182–1191 (2012)
- Susuki, Y., Mezić, I.: A Prony approximation of Koopman mode decomposition. In: *Decision and Control (CDC), 2015 IEEE 54th Annual Conference on Decision and Control (2015)*. <https://doi.org/10.1109/CDC.2015.7403326>
- Susuki, Y., Mezić, I., Raak, F.T.H.: Applied Koopman operator theory for power system technology. *Nonlinear Theory Appl.* **7**(4), 430–459 (2016)
- Takeishi, N., Kawahara, Y., Yairi, T.: Subspace dynamic mode decomposition for stochastic Koopman analysis. *Phys. Rev. E* **96**, 033,310 (2017)
- Tantet, A., Chekroun, M.D., Dijkstra, H.A., Neelin, J.D.: Mixing Spectrum in Reduced Phase Spaces of Stochastic Differential Equations. Part II: Stochastic Hopf Bifurcation. *ArXiv e-prints* (2017)
- Tu, J.H., Rowley, C.W., Luchtenburg, D.M., Brunton, S.L., Kutz, J.N.: On dynamic mode decomposition: theory and applications. *J. Comput. Dyn.* **1**(2), 391–421 (2014)
- Williams, M., Kevrekidis, I., Rowley, C.: A data-driven approximation of the Koopman operator: extending dynamic mode decomposition. *J. Nonlinear Sci.* **25**(6), 1307–1346 (2015)
- Yosida, K.: *Functional Analysis*, 6th edn. Springer, New York (1980)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.