Constrained Ulam Dynamic Mode Decomposition: Approximation of the Perron-Frobenius Operator for Deterministic and Stochastic Systems

Debdipta Goswami[1], Emma Thackray[2], and Derek A. Paley[3]

Abstract— Dynamical systems described by ordinary and stochastic differential equations can be analyzed through the eigen-decomposition of the Perron-Frobenius (PF) and Koopman transfer operators. While the Koopman operator may be approximated by data-driven techniques, e.g., Extended Dynamic Mode Decomposition (EDMD), the approximation of the PF operator uses a single-pass Monte Carlo approach in Ulam's method, which requires a sufficiently long time step. This paper proposes a finite-dimensional approximation technique for the PF operator that uses multi-pass Monte Carlo data to pose and solve a constrained EDMD-like least-squares problem to approximate the PF operator on a finite-dimensional basis. The basis functions used to project the PF operator are the characteristic functions of the state-space partitions. The results are analyzed theoretically and illustrated using deterministic and time-homogeneous stochastic systems.

Index Terms— Estimation, stochastic systems, optimization.

I. INTRODUCTION

The operator-theoretic approach to dynamical systems deals with the evolution of measurable maps under the system dynamics. The two main candidates of this approach are the Koopman operator and its dual, the Perron-Frobenius operator. While the Koopman operator is useful in studying observables, the Perron-Frobenius (PF) operator acts on the space of densities. Hence the PF operator is important when dealing with uncertainties in the system, especially when the likelihood of the state is given in the form of a probability density function under a suitable absolutely continuous probability measure. Formally, the PF operator is an infinite-dimensional operator operating on the space of L^1 functions.

The PF operator is used extensively to analyze the global behavior of dynamical systems, especially for fluid dynamics [1], and to estimate almost-invariant sets with efficient toolboxes like GAIO [2]. The PF operator, being able to transport density in a dynamical system, is also important in recursive estimation problems. The approximated PF operator can be used to transport density with less computational effort than solving a partial differential equation. Hence an accurate and efficient approximation technique of the PF operator is needed.

[1]Debdipta Goswami is a graduate student in the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742. `goswamid@umd.edu`

[2]Emma Thackray is an undergraduate student in the Department of Physics, Mount Holyoke College, South Hadley, MA 01075. `thack22e@mtholyoke.edu`

[3]Derek A. Paley is the Willis H. Young Jr. Professor of Aerospace Engineering Education in the Department of Aerospace Engineering and Institute for Systems Research, University of Maryland, College Park, MD 20742. `dpaley@umd.edu`

As the PF operator operates on infinite-dimensional spaces, it is customary to project it onto the span of suitable basis functions to approximate their dominant actions in finite dimensions. This projection is usually accomplished by Galerkin methods using a weak approximation of the operand functions [3], [4].

One technique to approximate the PF operator is Ulam's method [5], where the chosen basis functions are the characteristic functions of grids spanning the state space. The PF operator is reduced to a Markov state transition matrix and approximated by one-pass Monte Carlo simulation of a large number of initial conditions. However, this method requires that the time step of the simulation be sufficiently long [5] to avoid degeneracy. This limitation poses a difficulty when using the approximated PF operator for estimation, because precise density transport over small time intervals may be necessary to accommodate frequent observations. To enable PF approximation for a short time step, a multi-pass approach is necessary.

Multi-pass data-driven approaches, e.g., Dynamic Mode Decomposition (DMD), are ubiquitous for the approximation of the Koopman operator, the adjoint of the PF operator. ([6] describes the relationship between DMD and the Koopman operator; [7] demonstrates the use of DMD on experimental fluid flow data; and [8] extends mean ergodic theorem using Laplace averages to construct Koopman eigenfunctions.) The Extended DMD (EDMD) [9] uses a time series of the observable data to approximate the eigenvalues of the Koopman operator on a finite-dimensional basis set chosen from a dictionary of appropriate functions.

EDMD solves an unconstrained least-squares problem to approximate the Koopman operator in the form of a matrix. [10] proposes a structured version of EDMD that takes the positivity of the Koopman operator into account to generate a more accurate estimate of the Koopman eigenfunctions. To replicate this approach for the PF operator, two challenges need to be addressed. Unlike Koopman, the PF operator is Markov and, hence preserves measure. Moreover, the PF operator operates on densities, which implies it is not directly observable, so it must be approximated from a time series of Monte Carlo simulations.

The manuscript combines Ulam's method and constrained EDMD to remove the problem of long time steps in Ulam's method and to incorporate the accuracy of EDMD. The basis functions are chosen as the characteristic functions of the grids over the state space. It is shown analytically that the PF operator projected onto this basis constructs a Markov chain, and its eigenfunctions can be approximated from the

eigenvectors of the resultant stochastic matrix. The results are derived for both deterministic and time-homogeneous stochastic systems.

The contribution of this work is to develop the relationship between Ulam's method and EDMD for the approximation of the PF operator along dynamics governed by both ordinary and time-homogeneous stochastic differential equations. The new approximation technique for the PF operator solves a constrained optimization problem, unlike the unconstrained problem in EDMD. We demonstrate the performance of the proposed technique for several linear and nonlinear systems. Unlike Ulam's method, the proposed method utilizes multi-pass Monte Carlo data in order to support short time steps and hence can be effectively used for density transport in a recursive estimation algorithm.

The manuscript is organized as follows. Section II gives a brief description of the PF operator for deterministic and stochastic systems. Section III defines the weak approximation of the PF operator, summarizes Galerkin projection methods, and describes the proposed Constrained Ulam Dynamic Mode Decomposition (CU-DMD). It also gives the theoretical justification of the weak approximation and its role in approximating the PF eigenfunctions. Section IV illustrates the performance of CU-DMD on three test cases. Section V summarizes the paper and our ongoing work.

II. PERRON-FROBENIUS OPERATOR

Historically, transfer operators like the Perron-Frobenius (PF) operator are used in ergodic theory to study measure-theoretic characterization. The PF operator is described below in the context of both deterministic and stochastic systems.

A. Deterministic Systems

Let $\mathbb{X} \subset \mathbb{R}^d$ be a compact manifold and $f : \mathbb{X} \to \mathbb{X}$ be a smooth time-invariant vector field. Consider the autonomous time-invariant ODE

$$\dot{x} = f(x). \qquad (1)$$

Let $\phi_f : \mathbb{R} \times \mathbb{X} \to \mathbb{X}$ be the flow map of the ODE (1), i.e., $\phi_f(t, x_0)$ is a solution of the ODE (1) with the initial condition $x(0) = x_0$.

Definition 1: A semigroup of operator $\mathcal{P}^\tau : \tau \geq 0$ is said to be the *Perron-Frobenius* (PF) operator if $\mathcal{P}^\tau : L^1(\mathbb{X}) \to L^1(\mathbb{X})$ is defined by [5]

$$\mathcal{P}^\tau \rho(\cdot) = \rho \circ \phi_f(-\tau, \cdot)|\det(D_x \phi_f(-\tau, \cdot))|, \qquad (2)$$

where D_x denotes the Jacobian with respect to the space variable x. If $\rho(\cdot)$ is a probability density function (PDF) with respect to an absolutely continuous probability measure ν, then $\mathcal{P}^\tau \rho$ is another PDF with respect to the absolutely continuous probability measure $\nu \circ \phi(-\tau, \cdot)$. Specifically,

$$\int_B \mathcal{P}^\tau \rho \, d\nu = \int_{\phi_f(-\tau, B)} \rho \, d\nu, \qquad (3)$$

for any ν-measurable set B [11]. The PF operator translates a probability density function with time according to the flow

of the dynamics. It can be shown that the generator of this semigroup $\mathcal{P}^\tau \rho$ is given by $-\nabla \cdot (\rho f)$ [5].

Now suppose we define a time-varying PDF $\tilde{\rho}(t, x) \triangleq \mathcal{P}^t \rho(x)$. Then it can be shown that $\tilde{\rho}(t, x)$ satisfies the PDE [5]

$$\frac{\partial \tilde{\rho}}{\partial t} = -\nabla \cdot (\tilde{\rho} f) \qquad (4)$$

$$\tilde{\rho}(0, x) = \rho(x). \qquad (5)$$

B. Stochastic Systems

Let $\mathbb{X}$ and f be defined as before. Let $g : \mathbb{X} \to \mathbb{X}$ be another smooth time-invariant vector field. Let $(\mathbb{X}, \mathcal{B}(\mathbb{X}), \mathbb{P})$ be a probability triple with an absolutely continuous probability measure $\mathbb{P}$. Consider the following time-homogeneous Itô stochastic differential equation

$$dx_t = f(x_t)dt + g(x_t)dw_t, \, t > 0, \qquad (6)$$

where $x_t \in \mathbb{X}$ is a random process and w_t, $t > 0$, is a standard Wiener process. The stochastic dynamics (6) arise when there is process noise in the system defined by the ODE (1). The process noise enters the system as the diffusion term $g(x_t)$ in the Itô SDE. Here a flow map, being another random process, is difficult to define, but nonetheless the probability density may still be translated with a linear operator defined in terms of the transition density function.

Definition 2: The *transition density function* $p_\tau : \mathbb{X} \times \mathbb{X} \to \mathbb{R}_+$ is defined as [11]

$$\mathbb{P}[x_{t+\tau} \in \mathbb{B} | x_t = x] = \int_{\mathbb{B}} p_\tau(x, m)dm, \, \forall \mathbb{B} \in \mathcal{B}(\mathbb{X}). \quad (7)$$

The transition density function is the infinite-dimensional counterpart of the transition matrix entries for a Markov chain. Now we can define the stochastic PF operator.

Definition 3: Let $\rho \in L^1(\mathbb{X})$ be a probability density function. The *Perron-Frobenius* semigroup of operator $\mathcal{P}^\tau$, $\tau > 0$, is defined as $\mathcal{P}^\tau : L^1(\mathbb{X}) \to L^1(\mathbb{X})$ [11], such that

$$\mathcal{P}^\tau \rho(\cdot) = \int_{\mathbb{X}} p_\tau(m, \cdot)\rho(m)dm. \qquad (8)$$

The PF operator $\mathcal{P}^\tau$ as defined here is also a linear operator and $\mathcal{P}^\tau \rho$ has an infinitesimal generator $-\nabla \cdot (\rho f) + \frac{1}{2}\nabla^2(g^2\rho)$ [11].

III. NUMERICAL ESTIMATION OF THE PERRON-FROBENIUS OPERATOR

To obtain an explicit numerical estimate of the infinite-dimensional PF operator, we need to project it onto a suitable finite-dimensional basis of functions. As the PF operators for a specific dynamics also form a semigroup, we need to fix a time step size to extract one representative of the semigroup. The latter amounts to the discretization of the dynamics, generally performed using Galerkin methods employed by converting the operator equations (2) and (8) into a weakly approximated form, usually a Markov chain.

A. Weak Approximation

The weak approximation of the operator equation (2) or (8) is done by projecting it on a suitable set of basis functions. Here, the basis functions are chosen to be a family of characteristic functions of a partition of the state space, scaled accordingly by their Lebesgue measure. Let $D = \{B_i : i = 1, \ldots, M\}$ be a partition of the state space $\mathbb{X}$, usually a set of fine grids. The basis function ψ_i is taken to be $\psi_i \triangleq \chi_{B_i}$, where

$$\chi_{B_i}(x) = \begin{cases} 1, & \text{if } x \in B_i \\ 0, & \text{otherwise.} \end{cases}$$

The basis functions $\{\psi_1, \ldots, \psi_M\}$ are orthogonal. Now define a projection $\pi_M : L^1(\mathbb{X}) \to \text{sp}\{\psi_i, \ldots \psi_M\}$ to project ρ onto the span of these basis functions by

$$\pi_M \rho = \sum_{i=1}^M \left(\frac{1}{m(B_i)} \int_{B_i} \rho \, dm \right) \psi_i = \sum_{i=1}^M \frac{p^i}{m(B_i)} \psi_i, \quad (9)$$

where $m(\cdot)$ is the Lebesgue measure on $\mathbb{R}^d$ and $p^i \triangleq \int_{B_i} \rho \, dm$ are the weights of basis function ψ_i. Since the choice of basis is done before approximation, the projected density ρ is usually expressed as a vector $\mathbf{p} = (p^1, \ldots, p^n)$. This projection in turns restricts the infinite-dimensional operator $\mathcal{P}^\tau$ to a stochastic matrix P_τ given by

$$\begin{aligned} P_{\tau,ij} &= \mathbb{P}[\phi_f(t, x) \in B_j | x \in B_i] \\ &= \frac{\mathbb{P}[\phi_f(t, x) \in B_j, x \in B_i]}{\mathbb{P}[x \in B_i]} \\ &= \frac{m(B_i \cap \phi_f(-\tau, B_j))}{m(B_i)}, \end{aligned} \quad (10)$$

for deterministic system (2) where $\phi_f(-\tau, B_j) \triangleq \{\phi_f(-\tau, x) : x \in B_j\}$. In the last equality, the measure is changed to a standard Lebesgue measure, assuming we sample from a uniform distribution at $t = 0$ when computing $P_{\tau,ij}$.

For stochastic system (8), since the flow map is a random process, the P_τ matrix is given by

$$\begin{aligned} P_{\tau,ij} &= \mathbb{P}[x_{t+\tau} \in B_j | x_t \in B_i] \\ &= \mathbb{P}[x_\tau \in B_j | x_0 \in B_i], \text{ time homogeneity of (6)} \\ &= \frac{\mathbb{P}[x_\tau \in B_j, x_0 \in B_i]}{\mathbb{P}[x_0 \in B_i]} \\ &= \frac{k\mathbb{P}[x_\tau \in B_j, x_0 \in B_i]}{m(B_i)}, \end{aligned} \quad (11)$$

where $k > 0$ is the normalization factor such that the probability $\mathbb{P}[x_0 \in B_i] = \frac{m(B_i)}{k}$ corresponds to a uniform initial distribution.

With this formulation, the approximation of the PF operator is equivalent to the approximation of P_τ. The weak approximation, in effect, turns the PDF ρ into a Probability Mass Function (PMF) on each grid B_i, and the operator (2) becomes a Markov state transition equation [5]

$$\mathbf{p}_\tau = \mathbf{p}_0 P_\tau, \quad (12)$$

where $\mathbf{p}_\tau$ and $\mathbf{p}_0$ are the projection of $\mathcal{P}^\tau \rho$ and ρ, respec-

tively. Note that $\mathbf{p}_0$ and $\mathbf{p}_\tau$ are valid PMF since P_τ is a stochastic matrix. To see this, consider the projection of the transformed ρ, i.e.,

$$\begin{aligned} \pi_M \mathcal{P}^\tau \rho &= \pi_M \mathcal{P}^\tau \left(\sum_{i=1}^M \frac{p^i}{m(B_i)} \psi_i \right) \\ &= \sum_{i=1}^M \frac{p^i}{m(B_i)} \pi_M \left(\mathcal{P}^\tau \psi_i \right). \end{aligned} \quad (13)$$

Now, since $\{\psi_1, \ldots, \psi_M\}$ are orthogonal,

$$\pi_M \left(\mathcal{P}^\tau \psi_i \right) = \sum_{j=1}^M \frac{w_{ij}}{m(B_j)} \psi_j, \quad (14)$$

where the coefficients w_{ij} are given by

$$\begin{aligned} w_{ij} &= \int_{\mathbb{X}} \mathcal{P}^\tau(\psi_i(x)) \psi_j(x) dx \\ &= \int_{\mathbb{X}} \mathcal{P}^\tau(\chi_{B_i}(x)) \chi_{B_j}(x) dx \\ &= \int_{B_j} \mathcal{P}^\tau(\chi_{B_i}(x)) dx. \end{aligned} \quad (15)$$

For deterministic system (1), the coefficients are

$$\begin{aligned} w_{ij} &= \int_{B_j} \mathcal{P}^\tau(\chi_{B_i}(x)) dx \\ &= \int_{\phi_F(-\tau, B_j)} \chi_{B_i}(x) dx, \text{ from (3)} \\ &= m(B_i \cap \phi_f(-\tau, B_j)). \end{aligned}$$

Therefore,

$$\begin{aligned} \pi_M \mathcal{P}^\tau \rho &= \sum_{i=1}^M \frac{p^i}{m(B_i)} \sum_{j=1}^M \frac{m(B_i \cap \phi_f(-\tau, B_j))}{m(B_j)} \psi_j \\ &= \sum_{j=1}^M \frac{1}{m(B_j)} \sum_{i=1}^M p^i P_{\tau,ij} \psi_j, \end{aligned} \quad (16)$$

where $P_{\tau,ij} = \dfrac{m(B_i \cap \phi_f(-\tau, B_j))}{m(B_i)}$ from (10).

For stochastic system (6)

$$\begin{aligned} w_{ij} &= \int_{B_j} \mathcal{P}^\tau(\chi_{B_i}(x)) dx \\ &= \int_{B_j} \int_{\mathbb{X}} p_\tau(y, x) \chi_{B_i}(y) dy dx, \text{ from (8)} \\ &= \int_{\mathbb{X}} \int_{B_j} p_\tau(y, x) dx \chi_{B_i}(y) dy, \\ &\qquad\qquad \text{by Fubini's theorem} \\ &= \int_{\mathbb{X}} \mathbb{P}[x_{t+\tau} \in B_j | x_t = y] \chi_{B_i}(y) dy, \text{ from (7)} \\ &= \int_{\mathbb{X}} \mathbb{P}[x_\tau \in B_j | x_0 = y] \chi_{B_i}(y) dy, \\ &\qquad\qquad \text{by time-homogeneity of (6)} \\ &= \int_{B_i} \mathbb{P}[x_\tau \in B_j | x_0 = y] k d\mathbb{P}(y) \\ &= k\mathbb{P}[x_\tau \in B_j, x_0 \in B_i]. \end{aligned}$$

The uniform change of measure $dy = k d\mathbb{P}(y)$ with normalization factor $k > 0$ results from the fact that initial distribution (i.e., the distribution of x_0) is uniform. Hence, for the stochastic system,

$$\begin{aligned} \pi_M \mathcal{P}^\tau \rho &= \sum_{i=1}^M \frac{p^i}{m(B_i)} \sum_{j=1}^M \frac{k\mathbb{P}[x_\tau \in B_j, x_0 \in B_i]}{m(B_j)} \psi_j \\ &= \sum_{j=1}^M \frac{1}{m(B_j)} \sum_{i=1}^M p^i P_{\tau,ij} \psi_j, \end{aligned} \quad (17)$$

which is the same as (16) and differs only in the definition

of $P_{\tau,ij} = \dfrac{k\mathbb{P}[x_\tau \in B_j, x_0 \in B_i]}{m(B_i)}$ from (11).

Now define $\sum\limits_{i=1}^{M} p^i P_{\tau,ij} = p_\tau^j$, and $\mathbf{p}_\tau = (p_\tau^1, \ldots, p_\tau^M)$, which implies

$$\pi_M \mathcal{P}^\tau \rho = \sum_{j=1}^{M} \frac{p_\tau^j}{m(B_j)} \psi_j \qquad (18)$$

$$\mathbf{p}_\tau = \mathbf{p} P_\tau. \qquad (19)$$

Theorem 1: If (λ, φ) is an eigenvalue-eigenfunction pair of $\mathcal{P}^\tau$, λ is also an eigenvalue of P_τ with an left eigenvector $\mathbf{p}$ such that $\pi_M \varphi = \sum\limits_{j=1}^{M} \dfrac{p^j}{m(B_j)} \psi_j$

Proof: Since λ is an eigenvalue of $\mathcal{P}^\tau$ with the eigenfunction φ, then

$$\mathcal{P}^\tau \varphi = \lambda \varphi. \qquad (20)$$

Now let the projection of φ on $\mathrm{sp}\{\psi_1, \ldots, \psi_M\}$ be

$$\pi_M \varphi = \sum_{j=1}^{M} \frac{p^j}{m(B_j)} \psi_j, \qquad (21)$$

with appropriate coefficient vector $\mathbf{p}$. Then

$$\pi_M \mathcal{P}^\tau \varphi = \lambda \pi_M \varphi = \sum_{j=1}^{M} \frac{\lambda p^j}{m(B_j)} \psi_i \qquad (22)$$

since projection is a linear operation. But, from (16) or (17),

$$\pi_M \mathcal{P}^\tau \varphi = \sum_{j=1}^{M} \frac{\sum\limits_{i=1}^{M} p^i P_{\tau,ij}}{m(B_j)} \psi_j. \qquad (23)$$

Since $\{\psi_1, \ldots, \psi_M\}$ are linearly independent, from (22) and (23), $\lambda p^j = \sum\limits_{i=1}^{M} p^i P_{\tau,ij}$, $\forall\, j = 1, \ldots, M$, i.e., $\mathbf{p} P_\tau = \lambda \mathbf{p}$. Hence λ is an eigenvalue of P_τ with (left) eigenvector $\mathbf{p}$. ∎

The converse of Theorem 1 is, in general, not true. However, if we restrict the PF operator to the span of the basis functions, then the converse also holds.

Theorem 2: If $(\lambda, \mathbf{p})$ is an eigenvalue-(left) eigenvector pair of P_τ, then λ is also an eigenvalue of the restricted operator $\pi_M \mathcal{P}^\tau$ with eigenfunction $\varphi \triangleq \sum\limits_{i=1}^{M} \dfrac{p^i}{m(B_i)} \psi_i$.

Proof: Since $(\lambda, \mathbf{p})$ is an eigenvalue-(left) eigenvector pair of P_τ,

$$\mathbf{p} P_\tau = \lambda \mathbf{p}. \qquad (24)$$

Now from Eq. (18),

$$\pi_M \mathcal{P}^\tau \varphi = \sum_{j=1}^{M} \frac{p_\tau^j}{m(B_j)} \psi_j \text{ and } \mathbf{p}_\tau = \mathbf{p} P_\tau.$$

But since $\mathbf{p} P_\tau$, i.e., $\mathbf{p}_\tau = \lambda \mathbf{p}$, we get

$$\pi_M \mathcal{P}^\tau \varphi = \sum_{j=1}^{M} \frac{\lambda p^j}{m(B_j)} \psi_j = \lambda \sum_{j=1}^{M} \frac{p^j}{m(B_j)} \psi_j = \lambda \varphi. \quad (25)$$

Therefore λ is also an eigenvalue of the restricted operator $\pi_M \mathcal{P}^\tau$ with eigenfunction φ. ∎

B. Contstrained Ulam Dynamic Mode Decomposition

Ulam's method uses a Monte Carlo approach to numerically estimate the Markov state transition matrix P_τ. Within each B_i, a set of N test points $x_{i,1}, \ldots, x_{i,N}$ are defined and numerically integrated to obtain $\phi_f(\tau, x_{i,k})$, $k = 1, \ldots, n$, i.e., their final positions along the trajectories of the ODE (1) or the SDE (6). The estimated P_τ is given by [5]

$$P_{\tau,ij} \approx \frac{\#\{k : x_{i,k} \in B_i, \phi_f(\tau, x_{i,k}) \in B_j\}}{N}. \qquad (26)$$

The choice of τ is important and depends on the resolution of the partition D in this method. If the resolution is coarse, i.e., too few grid-cells (M is small), and τ is also small, then many of the test points will not leave their original grid cell B_i, and the estimated P_τ will be close to the identity matrix.

Extended Dynamic Mode Decomposition (EDMD) [9] is a method to extract the modes of a complex dynamical system by solving a least-squares problem. EDMD estimates the eigenvalues and eigenfunctions of the Koopman Operator [12], the dual of the PF operator, which operates on the space of L^∞ observables. The Koopman semigroup of operator $\mathcal{K}^t : L^\infty(\mathbb{X}) \to L^\infty(\mathbb{X})$ is defined as

$$(\mathcal{K}^t \varphi)(\cdot) = \varphi \circ \phi_f(t, \cdot). \qquad (27)$$

If we fix the time step $t = \tau$, then the ODE (1) becomes an iterative map $x((k+1)\tau) = \phi_f(\tau, x(k\tau))$, and we can drop τ and define $\phi_f(\tau, x) \triangleq F(x)$. The discrete-time dynamics become

$$x_{k+1} = F(x_k). \qquad (28)$$

The time-discretized version of Koopman operator is $\mathcal{K}^\tau \varphi(\cdot) = \varphi \circ F(\cdot)$. In EDMD, just like Ulam's method, the infinite-dimensional operator $\mathcal{K}^\tau$ is projected onto a finite-dimensional basis in $L^\infty(\mathbb{X})$ to represent it as a matrix K. Let $\{\psi_1, \ldots, \psi_M\}$ be the basis functions and, like in Ulam's method, we define $\pi_M : L^\infty(\mathbb{X}) \to \mathrm{sp}\{\psi_i, \ldots \psi_M\}$ to project φ onto the span of these basis functions. Then

$$\varphi(x) = \sum_{i=1}^{M} a_i \psi_i(x) \qquad (29)$$

$$\mathcal{K}^\tau \varphi(x) = \sum_{i=1}^{M} b_i \psi_i(x) + r, \qquad (30)$$

with residue r. Now, since φ is an observable, we can observe $\{\varphi(x_0), \ldots \varphi(x_{n+1})\}$ for any $n > 0$, where x_i are from the discretized dynamics (28). So we can estimate the matrix K by the least-squares formulation

$$K = \Psi_{x_0}^\dagger \Psi_{x_1}, \qquad (31)$$

where $\Psi_{x_0,ij} = \psi_i(x_j)$ and $\Psi_{X_1,ij} = \psi_i(x_{j+1})$, $i = 1, \ldots, M$, and $j = 0, \ldots, n$.

In the same light, the weak approximation of the PF operator can be thought of as projecting onto the basis function $\psi_i = \dfrac{1}{m(B_i)} \chi_{B_i}$. Since the basis functions are related to the

density of states, and cannot be readily observed, we need the help of Monte Carlo simulation.

Let $\mathbf{p}_0, \ldots, \mathbf{p}_{n+1}$ be $n+2$ subsequent PMF resulting from the operation of $\mathcal{P}^\tau$ on the initial density ρ projected on the $sp\{\psi_1, \ldots, \psi_M\}$. We know from (12) that $\mathbf{p}_{k+1} = \mathbf{p}_k P_\tau$. Define

$$\Psi_0 = [\mathbf{p}_0^T, \ldots, \mathbf{p}_n^T]^T, \quad \Psi_1 = [\mathbf{p}_1^T, \ldots, \mathbf{p}_{n+1}^T]^T. \quad (32)$$

Therefore, from the Markov relation, $\Psi_1 = P_\tau \Psi_0$, we have

$$P_\tau \approx \Psi_1 \Psi_0^\dagger. \quad (33)$$

But since we cannot observe directly the values of $\mathbf{p}_0, \ldots, \mathbf{p}_{n+1}$, we need to estimate them from Monte Carlo sampling. For this, we start from a uniform distribution of N particles in each of the M grids at $t = 0$. Then we successively integrate them forward for time interval τ to get their position at $t = \tau$. The number of particles in B_i will change from N to some value d_1^i, where the subscript 1 represents the first time step. Similarly we go on integrating for n more successive time steps to get the number of particles in each grid d_j^i for each time step j. Let $\mathbf{d_j} = (d_j^1, \ldots, d_j^M)$. Then empirically $\hat{\mathbf{p}}_j \triangleq \frac{1}{MN} \mathbf{d}_j \approx \mathbf{p}_j$ for $j = 0, \ldots, n+1$. Now, define empirical data matrices

$$\hat{\Psi}_0 = [\hat{\mathbf{p}}_0^T, \ldots, \hat{\mathbf{p}}_n^T]^T, \quad \hat{\Psi}_1 = [\hat{\mathbf{p}}_1^T, \ldots, \hat{\mathbf{p}}_{n+1}^T]^T. \quad (34)$$

Next find P_τ that minimizes the error between $\hat{\Psi}_1$ and $\hat{\Psi}_0 P_\tau$, using a constrained least-squares formulation:

$$\begin{aligned}
\underset{P_\tau}{\text{minimize}} \quad & \left\| \hat{\Psi}_1 - \hat{\Psi}_0 P_\tau \right\|_F \\
\text{subject to} \quad & P_{\tau,ij} \geq 0, \ i,j \in \{1, \ldots, M\} \\
& \sum_{j=1}^{M} P_{\tau,ij} = 1, \ i \in \{1, \ldots, M\}.
\end{aligned} \quad (35)$$

The problem (35) is a convex quadratic programming problem and yields a unique minimum that can be solved using gradient-descent or interior point methods. Unlike Ulam's method, this method is a multi-pass approach, which gives more accuracy with short time steps. It also has a distinct advantage over Ulam's method in stochastic systems, since there a very small increment in Wiener noise with a short time step. By solving for P_τ and exploiting Theorem 2, we can approximate the eigenvalues and eigenfunctions of the infinite-dimensional operator $\mathcal{P}^\tau$. Henceforth we refer to the problem (35) as Constrained Ulam DMD (CU-DMD).

IV. Numerical Simulations

CU-DMD is demonstrated on three different dynamical systems. First, consider a second-order linear system:

$$\dot{x} = Ax, \quad (36)$$

where $x \in \mathbb{R}^2$ and $A = \begin{bmatrix} 0 & 1 \\ -a & -b \end{bmatrix}$. The parameters $a = 0.5$ and $b = 0.1$ are chosen to produce damped oscillations. We have chosen time step $\tau = 0.1$ and a 40-by-40 grid, i.e., a total of 1600 grid-cells. For this system, if $\lambda_{1,2}$
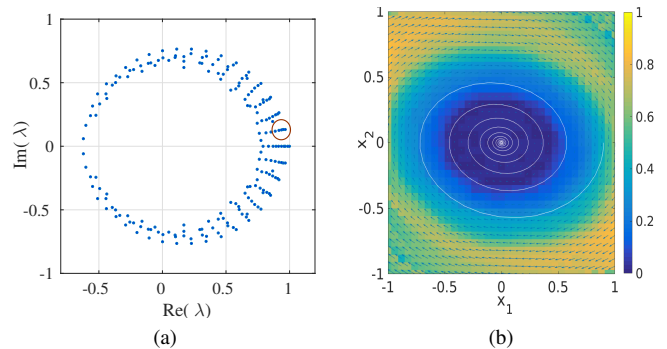

Fig. 1: Second-order linear system (36): (a) dominant spectra and (b) eigenfunction corresponding to $\lambda = 0.89 + i0.06$ (red circle).
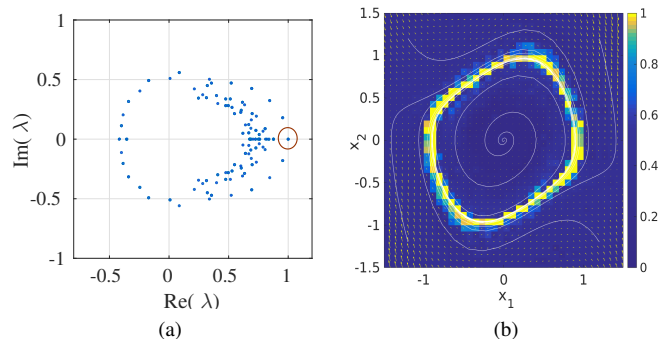

Fig. 2: Van der Pol system (37): (a) dominant spectra with unity eigenvalue circled and (b) the corresponding eigenfunction.

are the eigenvalues of $\exp(\tau A)$ with left eigenvectors $\mathbf{w}_{1,2}$, then $\lambda' = k\lambda_1 + l\lambda_2$ for $k, l \in \mathbb{R}$ is a Koopman eigenvalue with eigenfunction $\phi(x) = (\mathbf{w}_1^* x)^k (\mathbf{w}_2^* x)^l$ [6], [13]. Moreover $\phi(x)$ is also a PF eigenfunction with eigenvalue $\lambda = -\lambda' - \text{tr}(A)$ [14], if $|\lambda| \leq 1$. In this case $\lambda_2 = \bar{\lambda}_1$ and $\text{tr}(A) = -0.1$. Choosing $(k,l) = (-0.8, 0)$ and $(k,l) = (-0.4, -0.3)$ so that $|\lambda| < 1$ produces two PF eigenfunctions. Table I compares the exact eigenfunctions with those computed by CU-DMD using the error averaged over the state space. CU-DMD outperforms Ulam's method for time step $\tau = 0.1$. Moreover, the dominant 200 eigenvalues and the approximated eigenfunction corresponding to $\lambda = 0.89 - i0.06$ are shown in Fig. 1.

TABLE I: Error comparison for linear system (36)

PF Eigenvalue	Error	
	Ulam's method	CU-DMD
$0.89 \pm i0.06$	0.45	0.11
$0.79 \pm i0.05$	0.51	0.14

Next, consider the (scaled) Van der Pol oscillator system:

$$\begin{aligned}
\dot{x}_1 &= x_2 \\
\dot{x}_2 &= m(c - x_1^2)x_2 - x_1.
\end{aligned} \quad (37)$$

where $m = 2$ and $c = 0.2$. The time step is $\tau = 0.1$ and the grid is 40-by-40. The approximate spectra of the 100 dominant eigenvalues and the eigenfunction corresponding to $\lambda = 1$ are shown in Fig. 2. The eigenfunction traces out
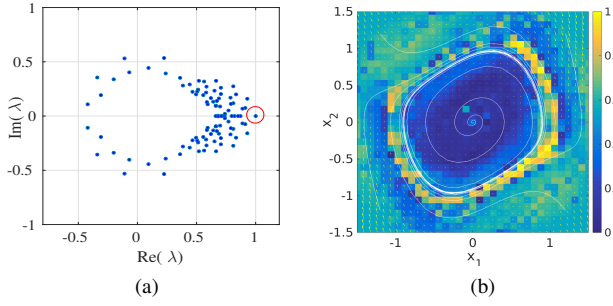
Fig. 3: The stochastic Van der Pol system (37) with noise (38): (a) dominant spectra with unity eigenvalue circled and (b) the corresponding eigenfunction.
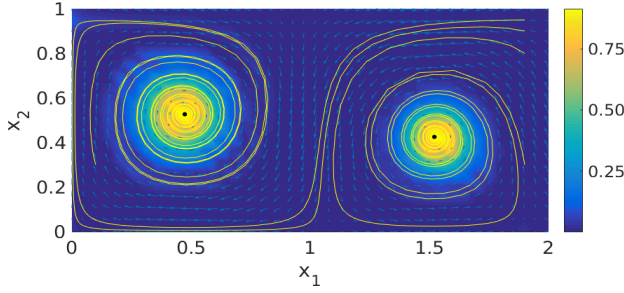

Fig. 4: The stochastic double-gyre system (39) approximated eigenfunction corresponding to unity eigenvalue.

a limit cycle, as expected.

To demonstrate CU-DMD for a stochastic system, we chose a time-homogeneous Itô stochastic differential equation of the form

$$dx_t = f(x_t)dt + \sqrt{2\sigma}dw_t, \ t > 0, \tag{38}$$

where dw_t is the standard Wiener increment and $\sigma = 0.1$ is the variance parameter. The drift field $f(x)$ is the deterministic Van der Pol system (37). The dominant spectra and the eigenfunction corresponding to the unity eigenvalue for the stochastic Van der Pol system are shown in the Fig. 3; the eigenfunction is not entirely concentrated on the limit-cycle due to the diffusion term. Diffusive behavior is also demonstrated using a double gyre system [5] corrupted with a Wiener noise. Here the drift field is

$$f(x) = \begin{bmatrix} -\pi a \sin\left(\frac{\pi x_1}{s}\right)\cos\left(\frac{\pi x_2}{s}\right) - \mu x_1 \\ -\pi a \cos\left(\frac{\pi x_1}{s}\right)\sin\left(\frac{\pi x_2}{s}\right) - \mu x_2 \end{bmatrix}, \tag{39}$$

where $a = 0.2$, $s = 1$, and $\mu = 0.1$. For the deterministic double-gyre system, the eigenfunction corresponding to unity is singular (concentrated on the stable equilibria) and not shown here due to space constraint. For the stochastic case, the eigenfunction spreads out from the stable equilibria as illustrated in Fig. 4.

V. CONCLUSION

This paper provides a new approximation method for computing eigenvalues and eigenfunctions of the Perron-Frobenius (PF) operator by combining the accuracy of Extended Dynamic Mode Decomposition and the Galerkin projection used in Ulam's method. The CU-DMD algorithm successfully approximates the PF operator and its eigenfunctions for smaller time steps than Ulam's Method. CU-DMD utilizes time-series data from Monte Carlo simulations and constrained quadratic programming to generate a Markov state-transition matrix to approximate the PF operator. The analytical justifications for the Galerkin projection and the eigenfunction approximation from the basis functions are provided. The algorithm is demonstrated on several non-linear systems with and without diffusive Wiener noise, and captures the modes of the system reasonably well. Potential future work includes the convergence proof for the approximation in the Hilbert space of the transfer operators.

REFERENCES

[1] G. Froyland, R. M. Stuart, and E. van Sebille, "How well-connected is the surface of the global ocean?" *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 24, no. 3, p. 033126, 2014.

[2] M. Dellnitz, G. Froyland, and O. Junge, "The algorithms behind GAIO — set oriented numerical methods for dynamical systems," in *Ergodic Theory, Analysis, and Efficient Simulation of Dynamical Systems*. Berlin, Heidelberg: Springer, 2001, pp. 145–174.

[3] S. Klus, P. Koltai, and C. Schtte, "On the numerical approximation of the Perron-Frobenius and Koopman operator," *Journal of Computational Dynamics*, vol. 3, no. 1, pp. 51–79, 2016.

[4] J. Ding, Q. Du, and T. Li, "High order approximation of the Frobenius-Perron operator," *Applied Mathematics and Computation*, vol. 53, no. 2, pp. 151–171, 1993.

[5] G. Froyland and K. Padberg, "Almost-invariant sets and invariant manifolds connecting probabilistic and geometric descriptions of coherent structures in flows," *Physica D: Nonlinear Phenomena*, vol. 238, no. 16, pp. 1507 – 1523, 2009.

[6] C. W. Rowley, I. Mezić, S. Bagheri, P. Schlatter, and D. S. Henningson, "Spectral analysis of nonlinear flows," *Journal of Fluid Mechanics*, vol. 641, pp. 115–127, 2009.

[7] P. J. Schmid, "Dynamic mode decomposition of numerical and experimental data," *Journal of Fluid Mechanics*, vol. 656, p. 528, 2010.

[8] R. Mohr and I. Mezić, "Construction of eigenfunctions for scalar-type operators via Laplace averages with connections to the Koopman operator," *arXiv:1403.6559 [math.SP]*, Mar. 2014. [Online]. Available: arxiv.org/abs/1403.6559

[9] Q. Li, F. Dietrich, E. M. Bollt, and I. G. Kevrekidis, "Extended dynamic mode decomposition with dictionary learning: A data-driven adaptive spectral decomposition of the Koopman operator," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 27, no. 10, p. 103111, 2017.

[10] B. Huang and U. Vaidya, "Data-driven approximation of transfer operators: Naturally Structured Dynamic Mode Decomposition," *arXiv:1709.06203 [math.DS]*, Sep. 2017. [Online]. Available: arxiv.org/abs/1709.06203

[11] S. Klus, F. Nüske, P. Koltai, H. Wu, I. Kevrekidis, C. Schütte, and F. Noé, "Data-driven model reduction and transfer operator approximation," *Journal of Nonlinear Science*, vol. 28, no. 3, pp. 985–1010, 2018.

[12] B. O. Koopman, "Hamiltonian systems and transformation in Hilbert space," *Proceedings of National Academy of Sciences*, vol. 17, pp. 315–318, 1931.

[13] A. Mauroy and I. Mezić, "Global stability analysis using the eigenfunctions of the Koopman operator," *IEEE Transactions on Automatic Control*, vol. 61, no. 11, pp. 3356–3369, Nov 2016.

[14] A. Lasota and M. Mackey, *Chaos, fractals, and noise: stochastic aspects of dynamics*, ser. Applied mathematical sciences. Springer-Verlag, 1994.