An MCMC Method for Uncertainty Set Generation via Operator-Theoretic Metrics

Anand Srinivasan and Naoya Takeishi

Abstract—Model uncertainty sets are required in many control problems, such as robust control and prediction with uncertainty, but there is no definite methodology to generate uncertainty sets for general nonlinear dynamical systems. In this paper, we propose a method for model uncertainty set generation via Markov chain Monte Carlo. The proposed method samples distributions over dynamical systems via metrics over transfer operators, and is applicable to general nonlinear dynamical systems. We adapt Hamiltonian Monte Carlo for sampling high-dimensional transfer operators in a computationally efficient manner. We present numerical examples to validate the proposed method for uncertainty set generation.

I. INTRODUCTION

Generating model uncertainty sets of dynamical systems is a universal problem in situations such as robust control, prediction with uncertainty, and scenario optimization. For example, in a min-max model predictive control (MPC) problem, one would like to solve

$$\min_{\{u\}} \max_{F \in \mathcal{F}} J(\{u\}, x_0, F),$$

where \(J\) is some loss function, \(u\) denotes control signals, \(x_0\) is an initial state, and \(F\) is a dynamics model. Here, \(\mathcal{F}\) is a model uncertainty set with which the worst-case performance is to be optimized and thus is key to good performance of the robust controller. However, the way to configure a good \(\mathcal{F}\) is not trivial in general.

A possible strategy for model uncertainty set generation is via Bayesian inference, with which we can create an uncertainty set from an inferred posterior. One of the common difficulties in Bayesian inference is that we need to prepare appropriate priors and observation models, which is sometimes challenging when the target dynamics are nonlinear. Moreover, the computational procedures of Bayesian inference may depend on the specific parametrization of dynamics models.

As for the treatment of nonlinear dynamics, the operator-theoretic view on dynamical systems [1] is attracting attention recently. Considering linear operators (called transfer operators) that represent the transition of observable functions or density functions, we can analyze, identify, and control nonlinear dynamical systems using linear techniques, and there is literature also in the control community including [2]–[9]. However, it has not been discussed well to deal with uncertainty within such an operator-theoretic view on dynamical systems.

The aim of this work lies in leveraging the powerful machinery of the operator-theoretic view on dynamical systems for model uncertainty set generation. To this end, we develop a sampling method using a metric over transfer operators [10] on nonlinear dynamical systems. The proposed method is based on a Hamiltonian Monte Carlo in the space defined by the metric, and we provide heuristics for efficiently sampling transfer operators with constraints. Moreover, it is agnostic of model parametrization, and thus we can use the proposed method also as a versatile Bayesian inference algorithm. We show the validity of the proposed method via numerical examples of model uncertainty set generation.

II. RELATED WORK

Uncertainty is inevitable as data are always finite and may include observation noise, but the intersection of uncertainty and the operator-theoretic view on dynamical systems has not yet been explored well. Takeishi et al. [11] discussed a probabilistic interpretation of a technique called dynamic mode decomposition (DMD), which is a widely used algorithm for computing transfer operators, but they only considered the uncertainty of spectral components of the operators. Morton et al. [12] considered the uncertainty of linear transition operators via the uncertainty of observable function values for a model based on neural networks. These methods are useful in certain contexts but not generally applicable to model uncertainty set generation.

Complementarily to model uncertainty, state uncertainty has been at the center of interests in the control community. Several researchers have investigated the computation of state uncertainty using operator-theoretic techniques (see, e.g., [13]–[15]), but in most cases they do not discuss the model uncertainty in details.

III. BACKGROUND

In this section we review technical backgrounds on the transfer operators and the kernel function based on them.

A. Transfer operator theory of nonlinear dynamics

A transfer operator is a linear map \(\mathcal{L}\) acting on functions \(\phi : \mathcal{M} \to \mathbb{C}\) on phase space \(\mathcal{M}\) of a dynamical system \(F : \mathcal{M} \to \mathcal{M}\). Due to linearity of composition, \(\mathcal{L}\) is a linear map for any \(F\). This follows the general motif of trading low-dimensional but nonlinear dynamics for high-dimensional but linear dynamics common in mathematical physics. In fluid mechanics, statistical physics, climate sciences, control theory, and many other fields, the use of transfer operators in describing dynamical systems is classically useful and has
seen a recent surge of popularity for their use data-driven system identification [1].

When \( \phi \) are phase-space densities \( p_M(x) \), \( L \) is referred to as the Perron–Frobenius operator \( P \) which acts as the pushforward operator for the Markov process \( p_M^+(x) = \mathcal{P}p_M(x) \). For example, in the special case of Hamiltonian dynamics, \( P \) is known as the Liouville operator \( L = -(\partial_p H)\partial_x + (\partial_x H)\partial_p \) and is indispensable for the study of both equilibrium and nonequilibrium behavior in statistical physics.

More generally, when \( \phi \) are observables \( \phi : \mathcal{M} \to \mathbb{C} \), \( L \) is known as the Koopman operator \( \mathcal{K} \), named for Koopman and von Neumann’s [16] pioneering study on the spectrum of transfer operators for characterizing dynamical systems. Seminal work by Mezić [1] showed the utility of transfer operator theory in data-driven system identification and model reduction. The idea of identifying Koopman eigenfunctions in order to obtain linear descriptions of nonlinear systems led to a natural analysis of nonlinear dynamical systems via the spectrum of the Koopman operator:

\[
\mathcal{K} \phi = \Lambda^t \phi,
\]

where \( \phi \) are known eigenfunctions of the Koopman operator and \( \Lambda \) its eigenvalues. This led to the SVD-based and least-squares solutions for \( \mathcal{K} \) by Schmid et al. [17], called DMD.

Since then, DMD has been extended with dictionaries of nonlinear basis functions (namely extended DMD or EDMD [18]), observables in reproducing kernel Hilbert spaces [19], [20], neural networks (e.g., [21]), and the incorporation of control [4]. In this paper, we build upon this prior work by incorporating the ideas of uncertainty quantification into transfer operator theory to construct an uncertainty set for governing dynamics when only observations are known.

**B. Transfer operators for stochastic processes**

To discuss uncertainty quantification in the context of system identification, let us move to a fundamentally probabilistic setting where observations \( X_t \in \mathcal{M} \) represent a stochastic process. Below, we provide a brief background on the computation of Koopman operators in a stochastic setting, as introduced by Klus et al. [20].

Let a dynamical system \( F : \mathcal{M} \to \mathcal{M} \) have invariant probability measure \( \mu \), compactly supported over a measurable subspace \( \Omega \subseteq \mathcal{M} \). Let the sequence \( \{X_t|t \geq 0\} \) be an ergodic Markov process with transition density given by \( p(y|x) = \mathbb{P}t(F(X) = y|X = x) \). Given an observable \( g \in L^2(\mathcal{M}, \mu) \), let us define the **stochastic Koopman operator** \( \mathcal{K} \) as the conditional expectation:

\[
\mathcal{K}g(x) = \mathbb{E}[g(F(X))|X = x] = \int_{\Omega} p(y|x)g(y)\,d\mu(y). \tag{1}
\]

Its right-adjoint, the Perron–Frobenius operator, directly maps marginal distributions:

\[
p^+(x) = \mathcal{P}p(x) = \int_{\Omega} p(y|x)p(x)\,d\mu(x).
\]

Let observables \( \phi \) lie in an inner product space \( \mathcal{H} \) (in [20], \( \mathcal{H} \) is taken to be an RKHS), but for the purposes of prediction and control, we may assume an explicit observable map \( \Phi \) (consisting of basis functions \( \{\phi_i\} \) in order to have easy access to its preimage, as is done in EDMD). Now, we may define the Gramian \( C_{XX} \) (also referred to as a cross-covariance operator [22]):

\[
C_{XY} := \mathbb{E}_{XY}[\Phi(X) \otimes \Phi(Y)] = \int_{\Omega \times \Omega} \Phi(x) \otimes \Phi(y)\,dp(x,y).
\]

Using the relation \( C_{XY} = \mathbb{E}[f(X)X]C_{XX} \) [22], we can express the stochastic Koopman and Perron–Frobenius operators (1) in terms of the cross-covariance operators as:

\[
\mathcal{K} = C_{XY}(C_{XX} + \epsilon I)^{-1}, \tag{2}
\]

\[
\mathcal{P} = C_{YX}(C_{XX} + \epsilon I)^{-1}, \tag{3}
\]

which push forward observables \( g \in \mathcal{H} \) and densities \( p(x) \) on \( \mathcal{M} \), respectively.

**C. Kernels over dynamical systems**

Ishikawa et al. [10] introduced a kernel over dynamical systems defined by their respective Perron–Frobenius operators. For two dynamical systems \( (D_1, D_2) \) specified by their initial values and maps \( (X_{1,0}, f_1), (X_{2,0}, f_2) \),

\[
k^{m,T}_{P_{1,2}}((X_{1,0}, f_1), (X_{2,0}, f_2)) \tag{4}
\]

\[
:= \text{Tr} \left( \bigwedge_{t=0}^{T-1} (L_{1,t}P_{1}^{t}I_{1} \ast (L_{2,t}P_{2}^{t}I_{2})) \right), \tag{5}
\]

where \( I : \mathbb{C}^n \to \mathcal{H} \) is an initial value operator, \( P^t : \mathcal{H} \to \mathcal{H} \) is the \( t \)th iterate of the Perron–Frobenius operator, \( L_{h} : \mathcal{H} \to \mathcal{H}_{ab} \) is an observable operator, and \( m \) is a hyperparameter. The details of these operators are in [10], where it is also shown that \( k^{m,T}_{P_{1,2}} \) is a positive-definite kernel.

**IV. PROPOSED METHOD**

In this section, we first define a kernel and a pseudo-metric over Koopman operators utilizing the previous studies on operator-theoretic kernels over dynamical systems [10], [23]. Then, we present sampling procedures for dynamical systems using the kernel. The procedures depend on the Hamiltonian Monte Carlo method [24], and we introduce some heuristic modifications.

**A. Defining a kernel over Koopman operators**

Whereas the kernel proposed by Ishikawa et al. [10] is defined for Perron–Frobenius operators acting on state-space densities, we adapt it for the Koopman operator as follows. Let the initial value and observable operators \( I, L_{h} \) be already applied, and the kernel (4) evaluated on a dynamical system \( \mathcal{K} \) which acts in this (observable) space. Then we simplify (4) to:

\[
k^{m,T}_{K_{1,2}}(f_1, f_2) := \text{Tr} \left( \bigwedge_{t=0}^{T-1} (K_{1}^{t}(K_{2}^{t})^T) \right). \tag{6}
\]

However, this kernel is convergent only for semi-stable \( \mathcal{K} \), that is, those with spectral radius \( \rho(\mathcal{K}) \leq 1 \) [10]. To ensure convergence of \( k^{m,T}_{K_{1,2}} \) everywhere, we introduce an
exponential discounting factor $\lambda \geq 0$ (and, replacing the operator $K$ with its finite-dimensional approximation $K$):

$$k_{K}^{m,T,\lambda}(K_1, K_2) = \text{Tr} \left( \sum_{t=0}^{m-1} e^{-\lambda t} (K_1^t)(K_2^t)^T \right).$$

(7)

Note that the discounting factor has been adopted also in previous studies on dynamical system kernels [23, 25].

Let us show the convergence of (7) as $T \rightarrow \infty$ informally. We first observe that $\sum_{t=0}^{T} A^t$ converges if $\lim_{T \rightarrow \infty} \|A^t\|_F = 0$ for any matrix $A$. Using the well-known result that the product of two convergent series is convergent, it suffices to show that $\lim_{T \rightarrow \infty} \|e^{-\lambda T/2} K^t\|_F = 0$ for all $K$. Using Gelfand’s formula,

$$\|e^{-\lambda T/2} K^t\|_F \leq e^{-\lambda T/2} \|K^t\|_F < e^{-\lambda T/2} \rho(K)^t.$$

When $\rho(K) > 1$, the corresponding series is convergent if $\lambda > 2 \log \rho(K)$. Thus $k_{K}^{m,T,\lambda}$ converges for all $K$ and appropriate $\lambda$.

B. Sampling from distributions over transfer operators

We propose a posterior inference procedure for dynamical systems models given an inner product defined over transfer operators. We define a (pseudo-)metric bounded in $[0,1]$ for operators $K_1, K_2$, using a cosine similarity, as

$$d_k(K_1, K_2) := \sqrt{1 - \frac{\langle K_1, K_2 \rangle_k}{\langle K_1, K_1 \rangle_k \langle K_2, K_2 \rangle_k}},$$

where $\langle \cdot, \cdot \rangle_k$ denotes the inner product induced by a positive-definite kernel $k$. As a baseline, we may also consider the standard linear kernel $k(A, B) = \text{Tr}(A^T B)$, which is not necessarily a good option for transfer operators.

Let $D \in [0,1]$ be a random variable with density $p_D$. Furthermore, let $K_0$ be a nominal transfer operator (such as ones estimated by DMD). Then we define the likelihood of any dynamical system $K$ as:

$$\mathcal{L}(K \mid X) := p_D(d_k(K, K_0))$$

(8)

For example, we may assume $p_D$ is a Beta distribution with $\beta \gg \alpha$, or an exponential distribution with vanishing density past 1. Using this we may easily construct an uncertainty set of radius $r$ as $\Delta = \{ K \mid \mathcal{L}(K \mid X) \mid d_k(K, K_0) \leq r \}$.

Sampling $K$ from $\mathcal{L}$ can be done in a number of ways, and is our main topic of discussion. In what follows, we assume a uniform prior over $K$ in the chosen kernel distance $d_k$, although any prior may be used. The conceptually simplest algorithm is rejection sampling: for any uniformly perturbed $K$, accept if $u < c \mathcal{L}(K \mid X)$ for $u \sim [0, 1]$ and convergence parameter $c$. Unfortunately, even generating the initial uniform perturbations may be computationally intractable due to the high dimensionality of the samples. This results in low acceptance rates for rejection sampling as well as random-walk MCMC methods such as Metropolis–Hastings. We therefore turn our focus to gradient-based MCMC methods which are able to generate distant proposals and achieve dimensionality-independent acceptance rates.

C. High-dimensional sampling via Hamiltonian Monte Carlo

In his seminal work [24], Neal introduced Hamiltonian Monte Carlo (HMC), which uses the gradient of the likelihood to simulate stochastic Hamiltonian dynamics whose stationary distribution is the posterior (8). It is well suited to our case since many kernels over dynamical systems are basically differentiable.

We adapt HMC for transfer operator sampling by introducing an auxiliary momentum variable $R$ which is of the same dimension as $K$. Let us define the potential and the Hamiltonian of an operator $K$ as:

$$U(K) = -\log \left[ p_D(d_k(K, K_0)) \right]$$

(9)

$$H(K) = U(K) + \frac{1}{2} \text{Tr}(RTR)$$

(10)

We note an important constraint on $U$, and by extension $d_k$: that it must be continuous everywhere. Indeed, whereas the Perron–Frobenius kernel (4) is convergent only for semi-stable dynamics, the discounted kernel (7) is continuous and convergent for all $K$. With this, we generate samples $\{K\}$ about a nominal $K_0$ via Hamiltonian dynamics in the potential defined by (9) using the leapfrog integrator (we refer the reader to [24] for details).

D. Heuristic: HMC with uniform prior

The mixing time of HMC is highly sensitive to the choice of discretization parameters (in particular, $n$-leapfrog, step size). In practice, Neal recommends $\epsilon \sim O(d^{1/4})$ [24], and there also exist adaptive step-setting algorithms such as the No-U-Turn Sampler [26]; however, we find a tradeoff between computational efficiency (samples/step) and sufficient exploration of the model space when using HMC for transfer operators. To compensate, we use a pre-run of HMC in a zero-potential to generate a uniform prior of samples, which are then used as initial conditions for parallel HMC sampling from the posterior (8). We find that this accelerates the mixing time and wall-clock time for sampling significantly, and makes good use of hardware parallelism where available.

E. Heuristic: HMC with spectral constraints

In the interest of producing meaningful samples, one may wish to impose constraints using prior knowledge of the underlying system. Such constraints can be readily expressed as boundary conditions on HMC without changing its stationary solution or reversibility (under some assumptions) – known as Reflective HMC [27]. As an example, suppose that we have knowledge that the system is structurally stable under any realistic perturbation; then, we can (roughly) encode this as a constraint on the spectral radius $\rho(K) - \epsilon \leq \rho(K) \leq \rho(K_0) + \epsilon$.

Extending the reflective HMC procedure from [27], we describe a leapfrog integrator (Algorithm 1) which ensures $f(K) \in [a,b]$ for any differentiable, scalar-valued $f$. In numerical experiments, we use $f(K) = \rho(K)$ and $[a,b] = [\rho(K_0) - .01, \rho(K_0) + .01]$.
Algorithm 1 HMC with scalar constraints

```
procedure BOUNDEDLEAPFROG(K, R, f, a, b)
  R ← 1/2K unfit K U
  for i ← 1 to L do
    while |e − e'| > δ do ▷ Reflect till exhaustion
      e' ← e
      K, R, e ← STEP(K, R, f, a, b)
    R ← 1/2K unfit K U
  return K, R

function STEP(K, R, f, a, b)
  if f(K + eR) > b then
    return K, R
  else if f(K + eR) < a then
    ▷ Defined similarly
  else
    return K + eR, R, 0 ▷ Otherwise, take full step
```

Algorithm 2 Robust Koopman MPC

```
procedure MPC(A_d, B_d, β, h)
  T ∝ h
  λ ← 2 log p(A_d)
  Δ ← SAMPLE(n, A_d, β, T, λ)
  u_0 ← 0
  while True do
    x_0 ← APPLYINPUT(u_0)
    u_* ← min_{u ∈ Δ} J_f(A_d, x_0, u) s.t. (12), (13)
    u_0 ← u_*(0)

  function SAMPLE(A_d, β, T, λ)
    d(x) ← 1 − k^{t+1}[x,K_0]^2
    U(x) ← − log (BETA1(1, β, d(x))
    for K_0 in {K_0} do
      A_d ← LEAPFROGHMC(K, U(x = 0, A_d)
    return {A_d}
```

G. Application: Robust control

The use of operator-theoretic formalism for model-predictive control has been explored extensively (see, e.g., [28], [29]), whereby the description of a controlled nonlinear system can be written as linear dynamics using Koopman eigenfunctions [4], [30]. A controlled nonlinear dynamical system of the form \( \dot{x} = f(x, u) \) can be written in discretized, Koopman-linear form as:

\[
\begin{align*}
  z_{t+1} &= A_d z_t + B_d u_t, \\
  z_0 &= \Psi(z_0), \\
  \hat{x}_t &= C z_t,
\end{align*}
\]

where \( A_d \) is a discrete-time predictor for the uncontrolled system, \( \Psi \) spans an invariant subspace of Koopman eigenfunctions, \( B_d \) is the control influence matrix, and \( C \) is a preimage matrix. \( A_d, B_d, \text{ and } C \) may be simultaneously computed from data snapshots \( X \) and \( Y \) for control inputs \( U \) as:

\[
\begin{bmatrix}
A_d & C \\
B_d & \cdots
\end{bmatrix} = \begin{bmatrix}
\Psi(Y) \\
\Psi(X) \\
U
\end{bmatrix}^T \begin{bmatrix}
(\Psi(X))^T \\
(U^n) \\
(U^n)
\end{bmatrix}^T.
\]

Subsequently, one solves the standard quadratic programming MPC problem to obtain a controller for the linear system \( z_t \).

In Algorithm 2, we present a robust optimization procedure for the Koopman MPC problem when there is uncertainty in the model \( A_d \). This may arise from intrinsic process noise in \( \dot{x} \), observation noise in \( X \), limited data, or even more interestingly, uncertainty in the true Koopman eigenfunctions \( \Psi \). We note that our method is different from robust computation of a single Koopman operator from noisy data [31], [32]; rather, we describe a robust control procedure via generation of an explicit uncertainty set for use when any of the above uncertainties exist.

F. Computation of samples in practical settings

We will give two formulations of HMC (9) for finite arguments, one explicitly over transfer operators, and one implicitly over observed trajectories. Either may be used.

a) Formulation 1: Assume a transfer operator \( K \) for a dynamical system is approximated as a matrix \( K \in \mathbb{R}^{d \times d} \). Then, (7) simplifies to:

\[
k^{k,T}_{K_0}(K, K_0) = \sum_{I \subseteq [d], |I| = m} \det \left( \sum_{t=0}^{T-1} e^{-\lambda t} K^t(K_0^t)^T \right)_{[I, I]},
\]

where \( A_{[I, I]} \) denotes the submatrix given by indices \( I \). This formulation can be used when we have a nominal Koopman operator estimation \( K_0 \) and want to generate perturbed \( K \)'s.

b) Formulation 2: Assume observations \( x_i \in \mathbb{R}^d \) of a dynamical system are given as a matrix \( X \). Then, we may define the Hamiltonian (9) for trajectories as:

\[
H(X) = \log \left[ p_D(d(x_i, X_0)) \right] + 1/2(R, R).
\]

The kernel (7) can be defined for trajectories of length \( T \) in a similar fashion to (11). However, an explicit discounting term is no longer needed. As an example, for \( m = 2 \):

\[
k^{2,T}_{K_0}(X, X_0) = \sum_{i,j \in [1, T]} \det \begin{bmatrix}
k(x_i, x_j) & k(x_i, x_0, j) \\
k(x_0, i, x_j) & k(x_0, i, x_0, j)
\end{bmatrix}
\]

for some feature kernel \( k(\cdot, \cdot) \). Here, \( x_i \) and \( x_0, j \) are the elements of \( X \) and \( X_0, j \), respectively. If we have explicit observables \( \phi \), simply let \( k(x, y) = \langle \phi(x), \phi(y) \rangle \). This formulation can be used when we have a nominal trajectory \( X_0 \). Then, while we learn a dynamics model \( F_\theta \) parameterized by \( \theta \), we generate trajectories from \( F_\theta \) and use them as \( X \), by which we can perform inference on \( \theta \) via \( X \).
V. NUMERICAL EXAMPLES

In numerical experiments, we show how our sampling procedure compares against baseline perturbation methods in generating meaningful perturbations of dynamical systems. For example, methods in robust optimization (RO) take the general form

$$\min_{\theta} \max_{A \in \Delta} J(A, \theta)$$

where \(A\) is a model and \(\Delta A\) is an uncertainty structure. \(\Delta A\) may have a particular form, e.g. block-diagonal, or unstructured, e.g. \(\Delta A \sim N(0_{d \times d}, \Sigma^2)\). The goodness of the RO minimizer depends solely on the choice of uncertainty set. Either of these uncertainty structures essentially induces a distribution over the norms of perturbations; we note that this is an assumption, and the subject of our testing is whether this assumption is valid when it is known the perturbed matrices represent dynamical systems.

A norm-bounded perturbation set essentially translates into our framework as sampling with a trace kernel:

$$k(A, A_0) = \text{Tr}(A^T A_0),$$

which is not necessarily suitable for dynamical systems. In the following experiments, we use this as a baseline for comparison. By contrast, the proposed \((11)\) induces a distribution over divergences between powers of operators.

We demonstrate that the latter better preserves key properties of dynamical systems such as structural stability and attractor basins, while effectively exploring dynamics space, on both linear systems of ODEs and nonlinear systems via the Koopman operator.

A. 2-dimensional LTI systems

We first consider linear systems of ODEs of the form \(\dot{x} = Ax\) via discretization as \(A_d = e^{A \Delta t}\). We use simple 2x2 systems in order to clearly characterize the dynamics in a trace-determinant plot. Using the discounted kernel \((11)\), we are able to generate perturbations of both source- and saddle-types in addition to the semistable regimes. We use spread parameter \(\beta = 5\), HMC step \(\epsilon = 10^{-4}\), HMC leapfrog \(L = 100\), and generate \(N = 1000\) samples with \(k = 50\) initial conditions for every test shown. We compare the following two kernels:

$$\text{Tr}(A^T A_0) \text{ in } (15) \text{ and } k_{m,T,\lambda}^K(A, A_0) \text{ in } (11),$$

which we call a trace kernel and a Koopman kernel, respectively.

The strength of the proposed method, using \(k_{m,T,\lambda}^K\), can be seen when the nominal system is within a region of structural instability (see the two center systems, Figure 1). The trace kernel perturbations venture easily into spiral sink or spiral sources, which are distant in dynamics terms but very close in absolute norm. In these cases, the proposed method using the Koopman kernel retains a much tighter spread in the trace-determinant plane. Furthermore, it can be seen from the posterior distributions that the proposed method is able to explore distant dynamics while staying bounded within structurally similar regions.

B. Unforced Duffing oscillator

Next, we consider perturbations of a nonlinear system via the Koopman operator. We use the unforced Duffing equation:

$$\dot{x} = y, \quad \dot{y} = -0.3y + x - x^3,$$

whose basins of attraction are like in Figure 2.

We use simulated trajectories of length \(t = 400\) seconds with 8000 samples per trajectory across 144 initial conditions in the range \([-2, -2] \times [2, 2]\). Using 15 polynomial observables with maximum degree 5, we compute the Koopman operator \(K\) as \(K = Y X^T (X X^T)^\dagger\). For all experiments we use spread \(\beta = 5\), HMC step \(\epsilon = 5 \times 10^{-5}\), HMC

---

\(1\) All source codes are available on https://github.com/ooblahman/koopman-robust-control.
leapfrog $L = 200$, $N = 2000$ samples, and $k = 200$ initial conditions. In the shown perturbations, trajectories in the left and right basins of attraction are highlighted in red and blue, respectively (Figures 3a and 3b).

We immediately observe some qualitative differences between the two perturbation sets. First, it is apparent that perturbations via the Koopman kernel preserve attractor structure in most samples, versus almost none in the baseline setting. In the case of the Duffing oscillator, this is a defining feature, and such preservations are important consideration for any robust prediction or control procedure over dynamics models. Second, a large proportion of samples produced by the baseline method are diverging; these would need to be manually filtered out if used in a robust optimization setting. We observe in experiments that this can be mitigated by restricting the norm of perturbations (i.e., increasing $\beta$), but this comes at the cost of decreased exploration of dynamics space, and changes the robustness of an RO solution using the perturbation set (moreover, manual filtering changes the posterior, altering the RO problem). We also find that a spectral radius constraint (Algorithm 1) alleviates many of these concerns with the baseline method, however, non-convex reflection is not a trivial procedure to implement in HMC and is not a typically used method in generating uncertainty sets.

Finally, while attractors are mostly preserved in our method, the attractor basins seem to undergo some geometry warping. This suggests an interpretation of our perturbation method as warping the underlying potential wells, which may have meaningful physical interpretations.

VI. CONCLUSIONS

In this work, we developed a method for sampling from distributions over dynamical systems defined via the operator-theoretic metrics. We suggested to use the method for model uncertainty sets generation, which is a universal problem in control problems such as robust control. The future direction of research includes to express constraints over sampled dynamical systems where we may have domain-specific knowledge.

REFERENCES

[1] I. Mezić, “Spectral properties of dynamical systems, model reduction and decompositions,” *Nonlinear Dynamics*, vol. 41, pp. 309–325, Aug 2005.

[2] A. Mauroy, F. Forni, and R. Sepulchre, “An operator-theoretic approach to differential positivity,” in *Proceedings of the 54th IEEE Conference on Decision and Control*, pp. 7028–7033, 2015.

[3] J. Annoni, P. Seiler, and M. R. Jovanović, “Sparsity-promoting dynamic mode decomposition for systems with inputs,” in *Proceedings of the 55th IEEE Conference on Decision and Control*, pp. 6506–6511, 2016.

[4] J. L. Proctor, S. L. Brunton, and J. N. Kutz, “Dynamic mode decomposition with control,” *SIAM Journal on Applied Dynamical Systems*, vol. 15, no. 1, pp. 142–161, 2016.

[5] A. Surana, “Koopman operator based observer synthesis for control-affine nonlinear systems,” in *Proceedings of the 55th IEEE Conference on Decision and Control*, pp. 6492–6499, 2016.

[6] A. Mauroy and J. Goncalves, “Linear identification of nonlinear systems: A lifting technique based on the Koopman operator,” in *Proceedings of the 55th IEEE Conference on Decision and Control*, pp. 6500–6505, 2016.

[7] A. Mauroy and I. Mezić, “Global stability analysis using the eigenfunctions of the Koopman operator,” *IEEE Transactions on Automatic Control*, vol. 51, no. 1, pp. 3356–3369, 2016.

[8] A. Surana, M. O. Williams, M. Morari, and A. Banaszuk, “Koopman operator framework for constrained state estimation,” in *Proceedings of the 56th IEEE Conference on Decision and Control*, pp. 94–101, 2017.

[9] N. Takeishi, T. Yairi, and Y. Kawahara, “Factorially switching dynamic mode decomposition for Koopman analysis of time-variant systems,” in *Proceedings of the 57th IEEE Conference on Decision and Control*, pp. 6402–6408, 2018.

[10] I. Ishikawa, K. Fuji, M. Ikeda, Y. Hashimoto, and Y. Kawahara, “Metric on nonlinear dynamical systems with Perron–Frobenius operators,” in *Advances in Neural Information Processing Systems 30*, pp. 2856–2866, 2018.

[11] N. Takeishi, Y. Kawahara, Y. Tabei, and T. Yairi, “Bayesian dynamic mode decomposition,” in *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, pp. 2814–2821, 2017.

[12] J. Morton, F. D. Witherden, and M. J. Kochenderfer, “Deep variational Koopman models: Inferring Koopman observations for uncertainty-aware dynamics modeling and control,” in *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, pp. 3173–3179, 2019.

[13] A. Surana and A. Banaszuk, “Linear observer synthesis for nonlinear systems using Koopman Operator framework,” IFAC-PapersOnLine, vol. 49, no. 18, pp. 716–723, 2016.

[14] T. Shnitzer and R. T. and Jean Jacques Slotine, “Diffusion maps Kalman filter for a class of systems with gradient flows,” arXiv:1711.09598, 2017.

[15] J. J. Meyers, A. M. Leonard, J. D. Rogers, and A. R. Gerlach, “Koopman operator approach to optimal control selection under uncertainty;
in *Proceedings of the 2019 American Control Conference*, pp. 2964–2971, 2019.

[16] B. O. Koopman and J. v. Neumann, “Dynamical systems of continuous spectra,” *Proceedings of the National Academy of Sciences*, vol. 18, no. 3, pp. 255–263, 1932.

[17] P. J. Schmid, “Dynamic mode decomposition of numerical and experimental data,” *Journal of Fluid Mechanics*, vol. 656, pp. 5–28, 2010.

[18] M. O. Williams, I. G. Kevrekidis, and C. W. Rowley, “A data-driven approximation of the Koopman operator: Extending dynamic mode decomposition,” *Journal of Nonlinear Science*, vol. 25, no. 6, pp. 1307–1346, 2015.

[19] Y. Kawahara, “Dynamic mode decomposition with reproducing kernels for Koopman spectral analysis,” in *Advances in Neural Information Processing Systems*, vol. 29, pp. 911–919, 2016.

[20] S. Klus, I. Schuster, and K. Muandet, “Eigendecompositions of transfer operators in reproducing kernel Hilbert spaces,” *Journal of Nonlinear Science*, vol. 30, p. 283315, Aug 2019.

[21] N. Takeishi, Y. Kawahara, and T. Yairi, “Learning Koopman invariant subspaces for dynamic mode decomposition,” in *Advances in Neural Information Processing Systems*, vol. 30, pp. 1130–1140, 2017.

[22] L. Song, J. Huang, A. Smola, and K. Fukumizu, “Hilbert space embeddings of conditional distributions with applications to dynamical systems,” in *Proceedings of the 26th Annual International Conference on Machine Learning*, pp. 961–968, 2009.

[23] K. Fujii, Y. Inaba, and Y. Kawahara, “Koopman spectral kernels for comparing complex dynamics: Application to multiagent sport plays,” in *Lecture Notes in Computer Science*, vol. 10536, pp. 127–139, 2017.

[24] R. M. Neal, *MCMC Using Hamiltonian Dynamics*, ch. 5. CRC Press, 2011.

[25] S. Vishwanathan, A. J. Smola, and R. Vidal, “Binet–Cauchy kernels on dynamical systems and its application to the analysis of dynamic scenes,” *Int. J. Comput. Vis.*, vol. 73, no. 1, p. 95119, 2007.

[26] M. D. Hoffman and A. Gelman, “The No-U-Turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo,” *Journal of Machine Learning Research*, vol. 15, no. 47, pp. 1593–1623, 2014.

[27] H. Mohasel Afshar and J. Domke, “Reflection, refraction, and Hamiltonian Monte Carlo,” in *Advances in Neural Information Processing Systems* 28, pp. 3007–3015, 2015.

[28] H. Arbabi, M. Korda, and I. Mezić, “A data-driven Koopman model predictive control framework for nonlinear partial differential equations,” in *Proceedings of the 57th IEEE Conference on Decision and Control*, pp. 6409–6414, 2018.

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

[30] J. L. Proctor, S. L. Brunton, and J. N. Kutz, “Generalizing Koopman theory to allow for inputs and control,” *SIAM Journal on Applied Dynamical Systems*, vol. 17, no. 1, pp. 909–930, 2018.

[31] M. S. Hemati, C. W. Rowley, E. A. Deem, and L. N. Cattafesta, “De-biasing the dynamic mode decomposition for applied Koopman spectral analysis of noisy datasets,” *Theoretical and Computational Fluid Dynamics*, vol. 31, no. 4, pp. 349–368, 2017.

[32] S. Sinha, B. Huang, and U. Vaidya, “On robust computation of Koopman operator and prediction in random dynamical systems,” *Journal of Nonlinear Science*, 2019.