

Approximating equivariant evolution operators with recurrent reservoir computers

Fredy Vides^{1, a)}

Scientific Computing Innovation Center, School of Mathematics and Computer Science,
Universidad Nacional Autónoma de Honduras, Tegucigalpa, Honduras

(Dated: October 18, 2022)

In this document, some general results in structured matrix approximation theory with applications to autoregressive representation of recurrent reservoir computers are presented. Firstly, a generic nonlinear time delay embedding is considered for the time series data sampled from a system under study. Secondly, sparse least-squares and subspace rotation methods are applied to identify approximate representations of the output coupling matrices, that determine the autoregressive representations of the recurrent reservoir computers corresponding to some given system under consideration. Prototypical algorithms based on the aforementioned techniques, together with some applications to approximate identification and predictive simulation of equivariant nonlinear systems, that may or may not exhibit chaotic behavior are presented.

Time series model identification is an important task for data-driven dynamical system modeling and dynamic behavior forecasting. These model identification tasks can be approached as solutions to evolution operator identification problems. In this study some theoretical and computational methods for structure preserving identification of evolution operators, are presented.

I. INTRODUCTION

Reservoir computers provide interesting computational tools for system identification and simulation. Recently, some novel architectures named next generation reservoir computers have been devised. In this document, some elements of the theory and algorithmics corresponding to the computation of some particular types of recurrent reservoir computers are presented. The study reported in this document is focused on reservoir computers whose architecture can be approximately represented by either linear or nonlinear autoregressive vector models.

The main contribution of the work reported in this document is the application of *collaborative schemes* involving structured matrix approximation methods, together with linear and nonlinear autoregressive models. Some theoretical aspects of the aforementioned methods are described in §III. As a byproduct of the work reported in this document, a toolset of Python programs for semilinear sparse signal model computation based on the ideas presented in §III and §IV has been developed, and is available in²⁰.

The applications of the structure preserving operator identification technology developed as part of the work reported in this document, range from numerical simulation cyber-physical systems²¹, to climate simulation¹⁰.

A prototypical algorithm for the computation of recurrent reservoir computers based on the ideas presented in §III, is presented in §IV. Some illustrative computational implemen-

tations of the prototypical algorithm presented in §IV are documented in §V.

II. PRELIMINARIES AND NOTATION

The symbols \mathbb{R}^+ and \mathbb{Z}^+ will be used to denote the positive real numbers and positive integers, respectively. For any pair $p, n \in \mathbb{Z}^+$ the expression $d_p(n)$ will denote the positive integer $d_p(n) = n(p - 1)/(n - 1) + 1$. For any finite group G , the expression $|G|$ will be used to denote the order of G , that is, the number of elements of G . Given $\delta > 0$, let us consider the function defined by the expression

$$H_\delta(x) = \begin{cases} 1, & x > \delta \\ 0, & x \leq \delta \end{cases}.$$

Given a matrix $A \in \mathbb{C}^{m \times n}$ with singular values⁹ (§2.5.3) denoted by the expressions $s_j(A)$ for $j = 1, \dots, \min\{m, n\}$. We will write $\text{rk}_\delta(A)$ to denote the number

$$\text{rk}_\delta(A) = \sum_{j=1}^{\min\{m,n\}} H_\delta(s_j(A)).$$

For a nonzero matrix $A \in \mathbb{R}^{m \times n}$, the symbol A^+ will be used to denote the pseudoinverse⁹ (§5.5.4) of A .

Given a scalar time series $\Sigma = \{x_t\}_{t \geq 1} \subset \mathbb{R}^\infty$, a positive integer L and any $t \geq L$, we will write $\mathbf{x}_L(t)$ to denote the vector

$$\mathbf{x}_L(t) = \left[\mathbf{x}_L^{(1)}(t)^\top \ \mathbf{x}_L^{(2)}(t)^\top \ \dots \ \mathbf{x}_L^{(n)}(t)^\top \right]^\top \in \mathbb{R}^{nL},$$

with

$$\mathbf{x}_L^{(j)}(t) = \left[x_{t-L+1}^{(j)} \ x_{t-L+2}^{(j)} \ \dots \ x_{t-1}^{(j)} \ x_t^{(j)} \right]^\top \in \mathbb{R}^L.$$

for $1 \leq j \leq n$, where $x_{j,s}$ denotes the scalar j -component of each element x_s in the vector time series Σ , for $s \geq 1$.

The identity matrix in $\mathbb{R}^{n \times n}$ will be denoted by I_n , and we will write $\hat{e}_{j,n}$ to denote the matrices in $\mathbb{R}^{n \times 1}$ representing the canonical basis of \mathbb{R}^n (each $\hat{e}_{j,n}$ corresponds to the j -column of I_n). For any vector $x \in \mathbb{R}^n$, we will write $\|x\|$ to denote the

^{a)}Electronic mail: fredy.vides@unah.edu.hn

Euclidean norm of x . Given a matrix $X \in \mathbb{R}^{m \times n}$, the expression $\|X\|_F$ will denote the Frobenius norm of X .

For any integer $n > 0$, in this article we will identify the vectors in \mathbb{R}^n with column matrices in $\mathbb{R}^{n \times 1}$.

Given two matrices $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{p \times q}$, the tensor Kronecker tensor product $A \otimes B \in \mathbb{R}^{mp \times nq}$ is determined by the following operation.

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix}$$

For any integer $p > 0$ and any matrix $X \in \mathbb{R}^{m \times n}$, we will write $X^{\otimes p}$ to denote the operation determined by the following expression.

$$X^{\otimes p} = \begin{cases} X & , p = 1 \\ X \otimes X^{\otimes(p-1)} & , p \geq 2 \end{cases}$$

We will also use the symbol Π_p to denote the operator $\Pi_p : \mathbb{R}^n \rightarrow \mathbb{R}^{n^p}$ that is determined by the expression $\Pi_p(x) := x^{\otimes p}$, for each $x \in \mathbb{R}^n$.

For any matrix $A \in \mathbb{R}^{m \times n}$, we will denote by $\text{colsp}(A)$ the columns space of the matrix A . Given a list A_1, A_2, \dots, A_m such that for $1 \leq j \leq m$, $A_j \in \mathbb{R}^{n_j \times n_j}$ for some integer $n_j > 0$. The expression $A_1 \oplus A_2 \oplus \cdots \oplus A_m$ will denote the block diagonal matrix

$$A_1 \oplus A_2 \oplus \cdots \oplus A_m = \begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_m \end{bmatrix},$$

where the zero matrix blocks have been omitted.

We will write \mathbf{S}^1 to denote the set $\{z \in \mathbb{C} : |z| = 1\}$. Given any matrix $X \in \mathbb{R}^{m \times n}$, we will write X^\top to denote the transpose $X^\top \in \mathbb{R}^{n \times m}$ of X . A matrix $P \in \mathbb{C}^{n \times n}$ will be called an orthogonal projector whenever $P^2 = P = P^\top$. The group of orthogonal matrices in $\mathbb{R}^{n \times n}$ will be denoted by $\mathbb{O}(n)$ in this study. Given any matrix $A \in \mathbb{R}^{n \times n}$, we will write $\Lambda(A)$ to denote the spectrum of A , that is, the set of eigenvalues of A .

III. APPROXIMATE EVOLUTION OPERATOR REPRESENTATION

Let us consider discrete-time dynamical systems determined by the pair $(\hat{\Sigma}, \mathcal{T})$ with $\hat{\Sigma} \subset \mathbb{R}^n$ and where $\mathcal{T} : \hat{\Sigma} \rightarrow \hat{\Sigma}$ is an evolution operator. Given a finite group G and some matrix representation¹⁶ (Definition 3.1.1) $\boldsymbol{\pi} : G \rightarrow G_\boldsymbol{\pi} \mathbb{O}(n), g \mapsto g_\boldsymbol{\pi}$, the system $(\hat{\Sigma}, \mathcal{T})$ is said to be G -equivariant with respect to $\boldsymbol{\pi}$ if:

$$g_\boldsymbol{\pi} x = \mathcal{T}(g_\boldsymbol{\pi} x) \quad (\text{III.1})$$

for each $x \in \hat{\Sigma}$ and each $g_\boldsymbol{\pi} \in \boldsymbol{\pi}(G)$. When it is clear from the context, an explicit reference to $\boldsymbol{\pi}$ may be omitted.

Given $\varepsilon > 0$, some matrix representation $G_\boldsymbol{\pi} \subset \mathbb{O}(n)$ of a finite group G , and an orbit of a G -equivariant discrete-time

system to be identified, that can be represented by a vector times series $\Sigma = \{x_t\}_{t \geq 1} \subset \mathbb{R}^n$. We will study the problem of identifying a (generally nonlinear) operator $\hat{\mathcal{T}} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ such that

$$\|gx_{t+1} - \hat{\mathcal{T}}(gx_t)\| \leq \varepsilon, \quad (\text{III.2})$$

for each $1 \leq t \leq \tau$, each $g \in G_\boldsymbol{\pi}$ and some prescribed $\tau > 0$, with $x_{t+1} = \hat{\mathcal{T}}(x_t)$ for each $t \geq 1$.

A. Equivariant reservoir computers for approximate evolution operator representation

When for a matrix representation $\boldsymbol{\pi} : G \rightarrow \mathbb{O}(n)$ of some given finite group G , we consider the time series data $\Sigma \subset \mathbb{R}^n$ corresponding to an orbit determined by the difference equation

$$x_{t+1} = \mathcal{T}(x_t), \quad (\text{III.3})$$

for some G -equivariant discrete-time system $(\hat{\Sigma}, \mathcal{T})$ with respect to $\boldsymbol{\pi}$ to be identified. One may need to be preprocess the time series data before proceeding with the approximate representation of a suitable evolution operator. For this purpose, given some prescribed integer $L > 0$, one can consider the time series $\mathcal{D}_L(\Sigma)$ determined by the expression.

$$\mathcal{D}_L(\Sigma) = \{\mathbf{x}_L(t)\}_{t \geq L}$$

For the dilated time series $\mathcal{D}_L(\Sigma)$, the previously considered recurrence relation $x_{t+1} = \mathcal{T}(x_t), t \geq 1$, induces the following difference equations

$$\mathbf{x}_L(t+1) = \tilde{\mathcal{T}}(\mathbf{x}_L(t)), \quad (\text{III.4})$$

for $t \geq L$. Where $\tilde{\mathcal{T}}$ is some evolution operator to be identified such that

$$(\boldsymbol{\pi}(g) \otimes I_L) \mathbf{x}_L(t+1) = \tilde{\mathcal{T}}((\boldsymbol{\pi}(g) \otimes I_L) \mathbf{x}_L(t)), \quad (\text{III.5})$$

for each $g \in G$.

For any $p \geq 1$, let us consider the map $\bar{\partial}_p : \mathbb{R}^n \rightarrow \mathbb{R}^{d_p(n)}$ for $d_p(n) = n(n^p - 1)/(n - 1) + 1$, that is determined by the expression.

$$\bar{\partial}_p(x) := \begin{bmatrix} \Pi_1(x) \\ \Pi_2(x) \\ \vdots \\ \Pi_p(x) \\ 1 \end{bmatrix} = \begin{bmatrix} x^{\otimes 1} \\ x^{\otimes 2} \\ \vdots \\ x^{\otimes p} \\ 1 \end{bmatrix}$$

Given integers $p, L > 0$, an orbit $\Sigma = \{x_t\}_{t \geq 1} \subset \mathbb{R}^n$ of an equivariant system with finite symmetry group represented by a set of orthogonal matrices $G \subset \mathbb{R}^{n \times n}$. For a finite sample $\Sigma_N = \{x_t\}_{t=1}^T \subset \Sigma$, let us consider the matrices:

$$\mathbf{H}_{L,G}^{(0,p)}(\Sigma_T) = \left[\mathbf{h}_{L,g_1}^{(0,p)}(\Sigma_T) \cdots \mathbf{h}_{L,g_N}^{(0,p)}(\Sigma_T) \right] \quad (\text{III.6})$$

$$\mathbf{H}_{L,G}^{(1)}(\Sigma_T) = \left[\mathbf{h}_{L,g_1}^{(1)}(\Sigma_T) \cdots \mathbf{h}_{L,g_N}^{(1)}(\Sigma_T) \right]$$

where $\mathbf{h}_{L,g}^{(0,p)}(\Sigma_T)$ and $\mathbf{h}_{L,g}^{(1)}(\Sigma_T)$ are determined for any $g \in G$ by the following expressions.

$$\mathbf{h}_{L,g}^{(0,p)}(\Sigma_T) = [\bar{\partial}_p((g \otimes I_L)\mathbf{x}_L(L)) \ \cdots \ \bar{\partial}_p((g \otimes I_L)\mathbf{x}_L(T-1))] \quad (\text{III.7})$$

$$\mathbf{h}_{L,g}^{(1)}(\Sigma_T) = [(g \otimes I_L)\mathbf{x}_L(L+1) \ \cdots \ (g \otimes I_L)\mathbf{x}_L(T)]$$

The operator identification mechanism used in this study for dilated systems of the form (III.4), will be described by the expression:

$$\hat{\mathcal{T}}(\mathbf{x}_L(t)) = W\bar{\partial}_p(\mathbf{x}_L(t)), \quad t \geq L, \quad (\text{III.8})$$

for some matrix $W \in \mathbb{R}^{n \times d_p(n)}$ to be determined, with $d_p(n) = n(n^p - 1)/(n - 1) + 1$. Applying by the operator theoretic techniques and ideas presented in¹¹ and¹⁸, the matrix W in (III.8) can be estimated by approximately solving the matrix equation

$$W\mathbf{H}_{L,G}^{(0,p)}(\Sigma_T) = \mathbf{H}_{L,G}^{(1)}(\Sigma_T). \quad (\text{III.9})$$

The devices described by (III.8) are called equivariant recurrent reservoir computers (ERRC) in this paper.

For any given integers $L, n, p > 0$. Taking advantage of the maps $\bar{\partial}_p$, one can find a integer $0 < r_p(n) < d_p(n)$ together with a sparse matrix $R_{p,L}(n) \in \mathbb{R}^{r_p(n) \times d_p(n)}$, such that $R_{p,L}(n)^+ R_{p,L}(n) \bar{\partial}_p(x) \approx \bar{\partial}_p(x)$ for $x \in \mathbb{R}^{nL}$. The existence of the pair $r_p(n), R_{p,L}(n)$ is determined by the following theorem.

Theorem III.1. *Given $\varepsilon \in \mathbb{R}^+$ and $L, n, p \in \mathbb{Z}^+$. There are an integer $0 < r_p(n) < d_p(n)$ and a sparse matrix $R_{p,L}(n) \in \mathbb{R}^{r_p(n) \times d_p(n)}$ with $d_p(n)$ nonzero entries, such that $\|R_{p,L}(n)^+ R_{p,L}(n) \bar{\partial}_p(x) - \bar{\partial}_p(x)\| \leq \sqrt{d_p(n)L}\varepsilon$ for each $x \in \mathbb{R}^{nL}$.*

Proof. Let us consider the symmetric group \mathfrak{S}_{nL-1} on $nL-1$ letters, and let us consider any finite set of distinct points $\{\hat{x}_1, \dots, \hat{x}_{nL}, \hat{x}_{nL+1}\} \subset \mathbb{R}$ such that for each $1 \leq m \leq p$ and every pair on index sets $\{i_1, \dots, i_m\}, \{j_1, \dots, j_m\} \subset \{1, \dots, nL+1\}$:

$$\prod_{k=1}^m \hat{x}_{j_k} = \prod_{k=1}^m \hat{x}_{i_k}, \Leftrightarrow \exists \sigma \in \mathfrak{S}_{nL-1} : i_k = \sigma(j_k), \forall 1 \leq k \leq m. \quad (\text{III.10})$$

That is, the previously considered products coincide if and only if, the index set $\{i_1, \dots, i_m\}$ is a permutation of the set $\{j_1, \dots, j_m\}$, for each $1 \leq m \leq p$.

Let us now set $\mathbf{y} = v[\hat{x}_1 \ \hat{x}_2 \ \cdots \ \hat{x}_{nL}]^\top$, $d = d_p(n)$ and

$$\tilde{\mathbf{x}} = [\tilde{x}_1 \ \cdots \ \tilde{x}_d]^\top := \bar{\partial}_p(\mathbf{y}).$$

If in addition we consider the assignments:

$$\tilde{x}_d := \hat{x}_{nL+1},$$

$$R = e_{1,d}^\top.$$

Then, for each $j = 2, \dots, d$ one can find $1 \leq k_1(j), \dots, k_{n_j}(j) \leq d$ such that $|\tilde{x}_j - \tilde{x}_{k_m(j)}| \leq \varepsilon$, for every

$1 \leq m \leq n_j$. Consequently, if we set $R_0 := (1/n_j) \sum_{l=1}^{n_j} \hat{e}_{l,d}^\top$ and

$$R := \begin{bmatrix} R \\ R_0 \end{bmatrix},$$

whenever $k_1(j) = j$. After iterating on the previously described procedure for $2 \leq j \leq d$, we can define $R_{p,L}(n) := R$ and we can set the value of $r_p(n)$ as the number of rows of $R_{p,L}(n)$.

Given $x \in \mathbb{R}^{nL}$. Based on the structure of $R_{p,L}(n)$ determined by the constructive procedure used for its computation, it can be verified that $R_{p,L}(n)^+$ is well defined and in addition:

$$\|R_{p,L}(n)^+ R_{p,L}(n) \bar{\partial}_p(x) - \bar{\partial}_p(x)\| \leq \sqrt{d_p(n)}\varepsilon.$$

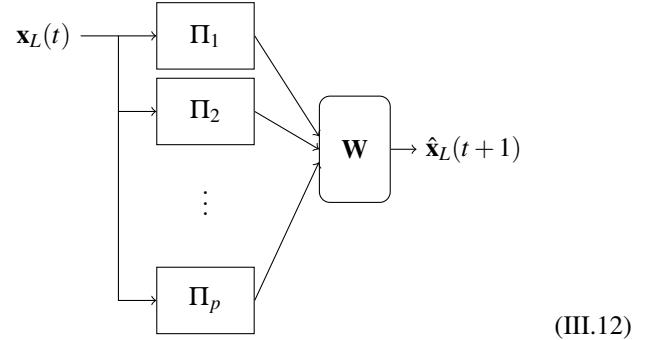
This completes the proof. \square

In order to reduce to computational effort corresponding to the solution of (III.9), using the matrix $R_{p,L}(n)$ described by Theorem III.1, one can obtain an approximate reduced representation of (III.9) determined by the expression.

$$\bar{W}R_{p,L}(n)\mathbf{H}_{L,G}^{(0,p)}(\Sigma_T) = \mathbf{H}_{L,G}^{(1)}(\Sigma_T) \quad (\text{III.11})$$

The architecture of the recurrent reservoir computers considered in this study was inspired by next generation reservoir computers⁸. Equivariant neural network models^{1,7} have found interesting applications to the computational simulation of physical systems.

Schematically, the recurrent models considered in this study can be described by a block diagram of the form,

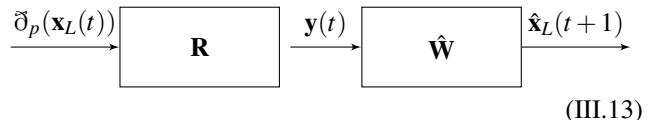


where for each $t \geq L$, the block \mathbf{W} is determined by the expression

$$\begin{aligned} \mathbf{W}(\Pi_1(\mathbf{x}_L(t)), \dots, \Pi_p(\mathbf{x}_L(t))) &:= \tilde{W} \begin{bmatrix} \Pi_1(\mathbf{x}_L(t)) \\ \vdots \\ \Pi_p(\mathbf{x}_L(t)) \end{bmatrix} + c_W \\ &= [\tilde{W} \ c_W] \bar{\partial}_p(\mathbf{x}_L(t)) \end{aligned}$$

and where the matrix $W = [\tilde{W} \ c_W]$ determined by (III.9).

The structure of the generic block \mathbf{W} ub (III.12) can be factored in the form



The layers \mathbf{R} and $\hat{\mathbf{W}}$ of the device (III.13) are determined by the expressions

$$\begin{aligned}\mathbf{R}(\mathbf{x}) &= \hat{R}\mathbf{x}, \\ \hat{\mathbf{W}}(\mathbf{y}) &= W\hat{R}^+\mathbf{y}\end{aligned}$$

for any pair of suitable vectors \mathbf{x}, \mathbf{y} . Where W is a sparse representation of an approximate solution (III.9) and \hat{R} is determined by Theorem III.1.

Using some suitable training subset Σ_S of a given data sample Σ_T from an arbitrary signal data $\Sigma = \{x_t\}_{t \geq 1} \subset \mathbb{R}^n$ under consideration, the parameters of the block \mathbf{W} of (III.12) are fitted using Σ_S . Simultaneously, a sparse representation of the matrix parameters corresponding to the block \mathbf{W} of the resulting model is computed.

Using the reservoir computer models described by (III.8), (III.12) and (III.13), we can compute approximate representations of evolution operators that satisfy (III.4) using the expression

$$\hat{\mathcal{T}}(\mathbf{x}_L(t)) := \hat{K}(\hat{\mathbf{W}} \circ \mathbf{R} \circ \mathfrak{D}_p(\mathbf{x}_L(t))) = \hat{K}W\mathfrak{D}_p(\mathbf{x}_L(t)), \quad (\text{III.14})$$

for each $t \geq L$. Furthermore, we can use the identified RRC model $\hat{\mathcal{T}}$ to simulate the behavior $x_{t+1} = \mathcal{T}(x_t)$ of the system described by (III.3) for $L \leq t \leq \tau$, by performing the operation:

$$\mathbf{T}(\mathbf{x}_L(t)) := \hat{K}\hat{\mathcal{T}}(\mathbf{x}_L(t)) = \hat{K}W\mathfrak{D}_p(\mathbf{x}_L(t)), \quad (\text{III.15})$$

for some $\tau > 0$ with

$$\hat{K} = \begin{bmatrix} \hat{e}_{1,nL}^\top \\ \hat{e}_{L+1,nL}^\top \\ \vdots \\ \hat{e}_{(n-1)L+1,nL}^\top \end{bmatrix}.$$

In this article, we will use the following notion of sparse representation. Given $\delta > 0$ and two matrices $A \in \mathbb{R}^{m \times n}$ and $X \in \mathbb{R}^{n \times p}$, a matrix $\hat{X} \in \mathbb{R}^{m \times p}$ is an approximate sparse representation of X , or a sparse representation of X for short, if $\|AX - A\hat{X}\|_F \leq C\delta$ for some $C > 0$ that does not depend on δ , and \hat{X} has fewer nonzero entries than X .

Given $p, L \in \mathbb{Z}^+$ and a matrix representation $\boldsymbol{\pi} : G \rightarrow G_{\boldsymbol{\pi}} \subset \mathbb{O}(n)$ of a finite group G , let us write $\boldsymbol{\pi}_{l,L}$ and $\boldsymbol{\pi}_{r,L}$ to denote the mappings defined for each $g \in G$ by the operations:

$$\boldsymbol{\pi}_{l,L}(g) := \boldsymbol{\pi}(g) \otimes I_L, \quad (\text{III.16})$$

$$\boldsymbol{\pi}_{r,L,p}(g) := \boldsymbol{\pi}_{l,L}(g) \oplus \boldsymbol{\pi}_{l,L}(g)^{\otimes 2} \oplus \cdots \oplus \boldsymbol{\pi}_{l,L}(g)^{\otimes p}.$$

For the previously considered integers, group and matrix mappings. Given $n, p, L \in \mathbb{Z}^+$. Let us consider the set

$\Phi_{G,\boldsymbol{\pi}}^{p,L}(n) \subset \mathbb{R}^{nL \times d_p(nL)}$ corresponding to the fixed point matrix set determined by the following expression.

$$\Phi_{G,\boldsymbol{\pi}}^{p,L}(n) = \{X \in \mathbb{R}^{nL \times d_p(nL)} : \boldsymbol{\pi}_{l,L}(g)^\top X \boldsymbol{\pi}_{r,L,p}(g) = X, g \in G\}.$$

Notation III.2. Given a $p, L \in \mathbb{Z}^+$ and a finite group G , a matrix representation $\boldsymbol{\pi} : G \rightarrow \mathbb{O}(n)$ is said to determine the equivariance condition of an ERRC, corresponding to a matrix solvent W of (III.9) and a map \mathfrak{D}_p , according to (III.8), if the conditions (III.1) are (approximately) satisfied for each $g\boldsymbol{\pi} \in \boldsymbol{\pi}(G)$.

Theorem III.3. Given $\delta > 0$, two integers $p, L > 0$, a finite group G , a sample $\Sigma_N = \{x_t\}_{t=1}^T$ from a G -equivariant system's orbit $\Sigma = \{x_t\}_{t \geq 1} \subset \mathbb{R}^n$ with $T > L$, and a matrix solvent $\bar{W} \in \mathbb{R}^{nL \times r_p(nL)}$ of (III.11) with $R_{p,L}(n)$ and $r_p(n)$ determined by Theorem III.1. If $r = \text{rk}_\delta(R_{p,L}(n)\mathbf{H}_{L,G}^{(0,p)}(\Sigma_T)) > 0$, then there is a sparse matrix $\hat{W} \in \mathbb{R}^{nL \times d_p(nL)}$ with at most $rd_p(nL)$ nonzero entries such that

$$\|\hat{W}R_{p,L}(n)\mathbf{H}_{L,G}^{(0,p)}(\Sigma_T) - \bar{W}R_{p,L}(n)\mathbf{H}_{L,G}^{(0,p)}(\Sigma_T)\|_F \leq K\delta, \quad (\text{III.17})$$

for $K = \sqrt{M(\min\{L, N-L+1\} - r)}(\sqrt{r}\|\hat{W}\|_F + \|\bar{W}\|_F)$. Furthermore, if $\boldsymbol{\pi} : G \rightarrow \mathbb{O}(n)$ is a matrix representations of G that determines the equivariance restrictions of the corresponding ERRC model, and if we set

$$\tilde{W} := \frac{1}{|G|} \sum_{g \in G} \boldsymbol{\pi}_{l,L}(g)^\top \hat{W}R_{p,L}(n)\boldsymbol{\pi}_{r,L,p}(g),$$

then $\tilde{W} \in \Phi_{G,\boldsymbol{\pi}}^{p,L}(n)$ and

$$\|\tilde{W}\mathfrak{D}_p(\mathbf{x}_L(t)) - \bar{W}R_{p,L}(n)\mathfrak{D}_p(\mathbf{x}_L(t))\| \leq K\delta, \quad (\text{III.18})$$

for each $L \leq t \leq T$.

Proof. By¹⁸ (Lemma 3.2) we will have that $\text{rk}_\delta((R_{p,L}(n)\mathbf{H}_{L,G}^{(0,p)}(\Sigma_T))^\top) = \text{rk}_\delta(R_{p,L}(n)\mathbf{H}_{L,G}^{(0,p)}(\Sigma_T)) > 0$. The estimate (III.17) follows by applying¹⁹ (Theorem 1) to the right hand side of the equality $\|\hat{W}R_{p,L}(n)\mathbf{H}_{L,G}^{(0,p)}(\Sigma_T) - \bar{W}R_{p,L}(n)\mathbf{H}_{L,G}^{(0,p)}(\Sigma_T)\|_F = \|(R_{p,L}(n)\mathbf{H}_{L,G}^{(0,p)}(\Sigma_T))^\top \hat{W}^\top - (R_{p,L}(n)\mathbf{H}_{L,G}^{(0,p)}(\Sigma_T))^\top \bar{W}^\top\|_F$, using a similar procedure to the one implemented in the proof of¹⁹ (Theorem 1).

Since G is finite, one can use direct computations to verify that $\tilde{W} \in \Phi_{G,\boldsymbol{\pi}}^{p,L}(n)$. And if we consider a presentation $G = \{g_1, \dots, g_{|G|}\}$, it can be seen that for each $L \leq t \leq T$.

$$\begin{aligned}
\|\tilde{W}\mathfrak{D}_p(\mathbf{x}_L(t)) - WR_{p,L}(n)\mathfrak{D}_p(\mathbf{x}_L(t))\| &\leq \frac{1}{|G|} \sum_{j=1}^{|G|} \left\| \boldsymbol{\pi}_{l,L}(g_j)^\top \hat{W}R_{p,L}(n)\boldsymbol{\pi}_{r,L,p}(g_j)\mathfrak{D}_p(\mathbf{x}_L(t)) - WR_{p,L}(n)\mathfrak{D}_p(\mathbf{x}_L(t)) \right\| \\
&= \frac{1}{|G|} \sum_{j=1}^{|G|} \left\| \hat{W}R_{p,L}(n)\boldsymbol{\pi}_{r,L,p}(g_j)\mathfrak{D}_p(\mathbf{x}_L(t)) - \boldsymbol{\pi}_{l,L}(g_j)WR_{p,L}(n)\mathfrak{D}_p(\mathbf{x}_L(t)) \right\| \\
&\leq \frac{1}{|G|} \sum_{j=1}^{|G|} \|\hat{W}R_{p,L}(n)\mathbf{H}_{L,G}^{(0,p)}(\Sigma_T) - WR_{p,L}(n)\mathbf{H}_{L,G}^{(0,p)}(\Sigma_T)\|_F \\
&\leq \|\hat{W}R_{p,L}(n)\mathbf{H}_{L,G}^{(0,p)}(\Sigma_T) - WR_{p,L}(n)\mathbf{H}_{L,G}^{(0,p)}(\Sigma_T)\|_F \leq K\delta.
\end{aligned}$$

This completes the proof. \square

B. Orhtognal Procrustes Problems and Linear Approximants of Recurrent Reservoir Computers

Given a finite group $G \subset \mathbb{O}(n)$ and some signal $\Sigma = \{\mathbf{x}_t\}_{t \geq 1} \subset \mathbb{R}^n$ from an orbit of a G -equivariant discrete-time dynamical system whose behavior can be approximately described by a model of the form (III.12), that can be computed by approximately solving an equation of the form (III.9) for $p = 1$ and some suitable integers $T, L > 0$ with $T > L$. If we consider any sample $\Sigma_T = \{\mathbf{x}_t\}_{t=1}^T \subset \Sigma$. We will have that an approximate matrix solvent \mathcal{L}_W of equation:

$$\mathcal{L}_W \mathbf{H}_{L,G}^{(0,1)}(\Sigma_T) = \mathbf{H}_{L,G}^{(1)}(\Sigma_T) \quad (\text{III.19})$$

will *approximately* satisfy the condition

$$\mathcal{L}_W \mathfrak{D}_1((g \otimes I_L)\mathbf{x}) = \Pi_1((g \otimes I_L)\mathbf{x}), \quad (\text{III.20})$$

for each $\mathbf{x} \in \text{colsp}(\mathbf{H}_{L,G}^{(0,1)}(\Sigma_T))$ and each $g \in G$.

Given $L > 0$, one can use matrices of the form (III.19) to express local linear approximants $\hat{\mathbf{T}}$ of the dilated evolution operator $\tilde{\mathcal{T}}$ corresponding to a given system of the form (III.4), that is based on an orbit $\Sigma \subset \mathbb{R}^n$ determined by (III.3), as follows.

$$\hat{\mathbf{T}}(\mathbf{x}_t) = \hat{K} \mathcal{L}_W \mathbf{x}_L(t), t \geq L. \quad (\text{III.21})$$

From here on, the matrix \mathcal{L}_W is called a linear approximate representation of the equivariant evolution operator for Σ .

Given $\varepsilon > 0$. When a linear approximate representation \mathcal{L}_W of the equivariant evolution operator for Σ provides a good approximation of the evolution operator under study. Given some suitable integer lag value $L > 0$ such that there is an approximate linear model \mathcal{L}_W for Σ of the form (III.21) with small residuals $r_t = \|\mathbf{x}_L(t+1) - \mathcal{L}_W(\mathbf{x}_L(t))\|_\infty$ for each $t \geq L$. Let $s = S + L - 1$. By applying a Krylov subspace approach¹³ (§6.1), one can find a matrix $W_k \in \mathbb{R}^{d_p(nL) \times k}$ whose columns form an orthonormal basis of the subspace $\mathcal{K}_{r_p(n)} \subset \mathbb{R}_p^d(nL)$ determined by the expression

$$\mathcal{K}_{r_p(n)} = \text{span}(\{\mathbf{x}_L(s), \mathcal{L}_W \mathbf{x}_L(s), \dots, \mathcal{L}_W^{r_p(n)-1} \mathbf{x}_L(s)\}),$$

for some suitable $r_p(n) \in \mathbb{Z}^+$. If in addition, $\Lambda(W_k^* \mathcal{L}_W W_k) \subset \mathbf{S}^1$ and if there is a unitary U_W such that $\|\mathcal{U}_W \mathbf{x}_L(t) - \mathcal{L}_W \mathbf{x}_L(t)\| \leq \varepsilon$ for each $t \geq s$. Using⁹ (Algorithm 12.4.1) one can compute a *local representation* of \mathcal{U}_W with respect to $\text{colsp}(W_k)$ using the following expression.

$$\mathcal{U}_W := \mathcal{U}_K + I_{nL} - W_k W_k^\top \quad (\text{III.22})$$

with

$$\mathcal{U}_K = U_L V_L$$

where

$$U_L S_L V_S = W_k^\top \mathbf{H}_{L,G}^{(1)}(\Sigma_T) \mathbf{H}_{L,G}^{(0,1)}(\Sigma_T)^\top W_k$$

corresponds to the singular value decomposition⁹ (§2.5.3) of $W_k^\top \mathbf{H}_{L,G}^{(1)}(\Sigma_T) \mathbf{H}_{L,G}^{(0,1)}(\Sigma_T)^\top W_k$.

IV. ALGORITHMS

In this section we focus on the applications of the structured matrix approximation methods presented in SIII, to reservoir computer models for equivariant dynamical systems.

A. Structured assembling matrix identification algorithm

B. Structured coupling matrix identification algorithm

Given a matrix representation $G_N \subset \mathbb{U}(n)$ of a finite group G , a G -equivariant system (Σ, \mathcal{T}) and a structured data sample $\Sigma_T \subset \Sigma$, we can apply Algorithm 1 and¹⁸ (Algorithm 1), in order to to compute the otuput coupling matrix that can be used to obtain an approximate representation of the evoulution operator \mathcal{T} , corresponding to the orbit Σ . For this purpose one can use the following Algorithm.

V. NUMERICAL SIMULATIONS

In this section we will present some numerical simulations computed using the **SPORT** toolset available in²⁰, that was

Algorithm 1: Compression matrix computation algorithm

Data: $n, p, L \in \mathbb{Z}^+, v, \varepsilon \in \mathbb{R}^+$.
Result: COMPRESSION MATRIX FACTOR: $R_{p,L}(n)$

- 0: Choose nL pseudorandom numbers $\hat{x}_1, \dots, \hat{x}_{nL} \in \mathbb{R}$ from $N(0, 1)$;
- 1: Set $\mathbf{y} = v [\hat{x}_1 \ \hat{x}_2 \ \dots \ \hat{x}_{nL}]^\top$;
- 2: Set $d = d_p(n)$;
- 3: Set $\tilde{\mathbf{x}} = [\tilde{x}_1 \ \dots \ \tilde{x}_d]^\top := \bar{\delta}_p(\mathbf{y})$;
- 4: Choose a pseudorandom number $\alpha \in N(0, 1)$;
- 5: Set $\tilde{x}_d := \alpha$;
- 6: Set $R = e_{1,d}^\top$;
- 7: **for** $j = 2, \dots, d$ **do**
 - 7.0: Find $1 \leq k_1, \dots, k_{n_j} \leq d$ such that $|\tilde{x}_j - \tilde{x}_{k_m}| \leq \varepsilon$, for each $1 \leq m \leq n_j$;
 - 7.1: **if** $k_1 = j$ **then**
 - 7.1.0: Set $R_0 := (1/n_j) \sum_{l=1}^{n_j} \hat{e}_{l,d}^T$;
 - 7.1.1: Set $R := \begin{bmatrix} R \\ R_0 \end{bmatrix}$;
 - 7.2: **end**
- 8: Set $R_{p,L}(n) := R$;

return $R_{p,L}(n)$

Algorithm 2: ERRC Model: ERRC model identification

Data: $\Sigma_N = \{x_t\}_{t=1}^T \subset \mathbb{R}^n$, $\pi : G \rightarrow \mathbb{O}(n)$
Result: Output coupling and compression matrices: $\hat{W}, \tilde{W}, R_{p,L}(n)$

- 0.0: Choose or estimate the lag value L using auto-correlation function based methods;
- 0.1: Set a tensor order value p ;
- 0.2: Compute compression matrix $R_{p,L}(n)$ applying Algorithm 1;
- 1: Compute matrices:

$$\mathbf{H}_0 := \mathbf{H}_{L,G}^{(0,p)}(\Sigma_T)$$

$$\mathbf{H}_1 := \mathbf{H}_{L,G}^{(1)}(\Sigma_T)$$

2: Approximately solve:

$$\hat{W} (R_{p,L}(n) \mathbf{H}_0) = \mathbf{H}_1$$

for \hat{W} applying¹⁸ (Algorithm 1);

3: Set:

$$\tilde{W} = \frac{1}{|G|} \sum_{g \in G} \pi_{l,L}(g)^\top \hat{W} R_{p,L}(n) \pi_{r,L,p}(g)$$

return $\hat{W}, \tilde{W}, R_{p,L}(n)$

developed as part of this project, the toolset consists of a collection of programs written in Matlab, Julia and Python that can be used for sparse identification and numerical simulation of dynamical systems.

The numerical experiments documented in this section were performed with Python 3.8.10. Some Python programs were used to generate synthetic data used for the system identification processes considered in this paper. All the programs written for synthetic data generation and sparse model identification as part of this project are available at²⁰.

The numerical simulations reported in this section were computed on a Linux Ubuntu Server 20.04 PC equipped with an Intel Xeon E3-1225 v5 (8M Cache, 3.30 GHz) processor and with 40GB RAM.

The computational setting used for the experiments considered in this section is documented as part of the Python program Experiment.py in²⁰, that can be used to replicate this experiments presented in this document.

A. Identification of Lorenz systems with ERRC

In this section a reservoir computer model corresponding to a pair of attractors of a Lorenz system of the form.

$$\begin{aligned} \frac{dx}{dt} &= \sigma(y - x) \\ \frac{dy}{dt} &= x(\rho - z) - y \\ \frac{dz}{dt} &= xy - \beta z \end{aligned} \tag{V.1}$$

The system (V.1) is \mathbb{Z}_2 -equivariant¹⁵, and for the configuration used for this experiment, the matrix representation of the corresponding group of symmetries $\mathbb{Z}_2 = \langle r | r^2 = e \rangle$ is determined by the following assignment.

$$r \mapsto r_\rho = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

1. Chaotic attractor

Let us consider the chaotic attractor corresponding to the system (V.1) for $\alpha = 10$, $\beta = 8/3$ and $\rho = 28$. The synthetic signals corresponding to the data sample $\Sigma_{2400} \subset \mathbb{R}^3$ that will be used for system identification have been computed with an explicit time integration scheme, using lsoda from the FORTRAN library odepak via the Python function odeint, with the Python program LorenzSystem.py in²⁰. The signals corresponding to the identified/predicted dynamical behavior data $\Sigma_{1600}^{(P)} \subset \mathbb{R}^3$ are visualized in figure 1.

We will apply algorithm 2 along the lines of theorem III.3 to identify model (V.1), using a RRC with embedding map $\bar{\delta}_2 : \mathbb{R}^{3L} \rightarrow \mathbb{R}^{d(3L)}$ for $L = 5$.

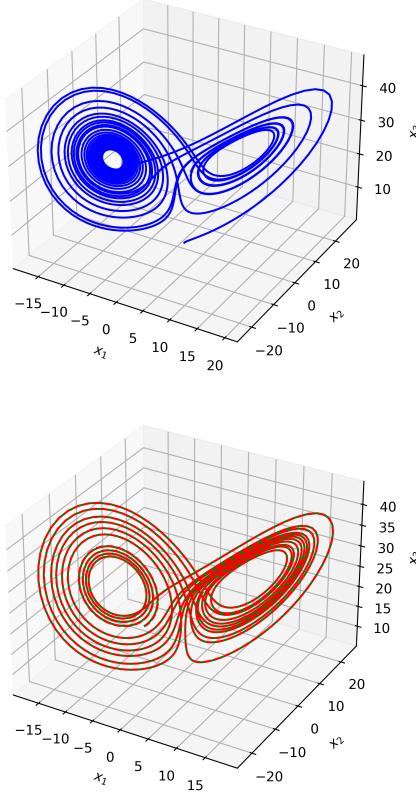


Figure 1: 3D graphical representation of: training data (left), predicted behavior (right).

The prediction errors in the ℓ_∞ -norm are shown in figure 2.

The structured coupling matrix corresponding to the equivariant evolution operator identified by algorithm is represented graphically in Figure 3.

2. Periodic attractor

Let us consider the periodic attractor corresponding to the system (V.1) for $\alpha = 10$, $\beta = 8/3$ and $\rho = 99.65$. The synthetic signals corresponding to the data sample $\Sigma_{200} \subset \mathbb{R}^3$ that will be used for system identification have been computed with an explicit time integration scheme, using lsoda from the FORTRAN library odepak via the Python function odeint, with the Python program LorenzSystem.py in²⁰. The signals corresponding to the identified/predicted dynamical behavior data $\Sigma_{1600}^{(P)} \subset \mathbb{R}^3$ are visualized in figure 4.

We will apply algorithm 2 along the lines of theorem III.3 to identify model (V.1), using a RRC with embedding map $\tilde{\mathcal{O}}_2 : \mathbb{R}^{3L} \rightarrow \mathbb{R}^{d(3L)}$ for $L = 5$.

The prediction errors in the ℓ_∞ -norm are shown in figure 5.

The structured coupling matrix corresponding to the equivariant evolution operator identified by algorithm is repre-

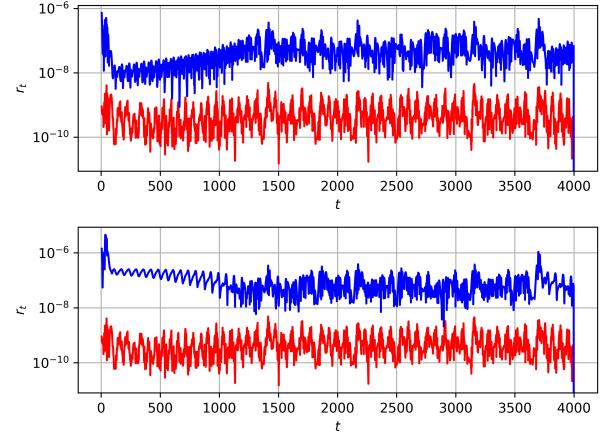


Figure 2: Prediction errors in the ℓ_∞ -norm.

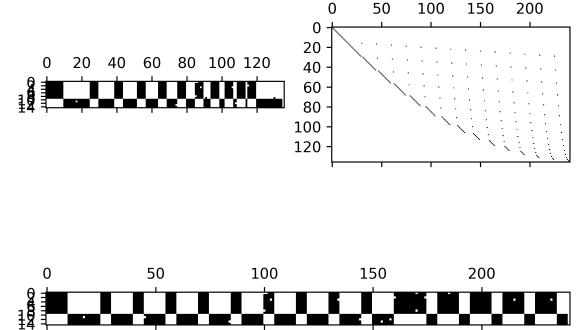


Figure 3: Graphical representation of coupling matrix and its corresponding structured factors.

sented graphically in Figure 6.

B. Identification of a Hamiltonian system with symmetries using ERRC

Let us consider a Hamiltonian system determined by the following initial value problem.

$$\begin{aligned} \frac{dq}{dt} &= p^3 - p, \\ \frac{dp}{dt} &= q^3 - q, \\ q(0) &= 1, p(0) = 0 \end{aligned} \tag{V.2}$$

As observed in¹⁵ the system (V.2) is K_4 -equivariant. For the configuration used for this experiment, the matrix representation of the corresponding group of symmetries

$$K_4 = \langle \gamma_1, \gamma_2 \mid \gamma_1^2 = \gamma_2^2 = (\gamma_1 \gamma_2)^2 = e \rangle$$

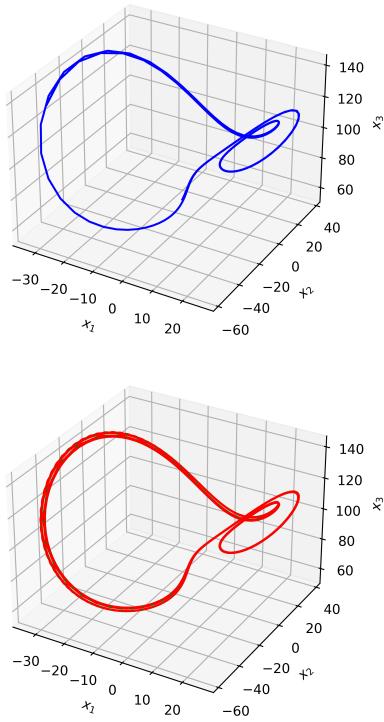


Figure 4: 3D graphical representation of: training data (left), predicted behavior (right).

is determined by the following assignments.

$$\begin{aligned}\gamma_1 &\mapsto \gamma_{1,p} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\ \gamma_2 &\mapsto \gamma_{2,p} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}\end{aligned}$$

The synthetic signals corresponding to the data sample $\Sigma_{600} \subset \mathbb{R}^3$ that will be used for system identification have been computed with an explicit time integration scheme, using lsoda from the FORTRAN library odepack via the Python function odeint, with the Python program HamiltonianSystem.py in²⁰. The model was trained using an embedding map ∂_p of order $p = 3$, with 15% of the synthetic reference data.

The reference synthetic signal data and the corresponding identified signals are illustrated in Figure 7.

The prediction errors in the ℓ_∞ -norm are shown in figure 8.

The structured coupling matrix corresponding to the equivariant evolution operator identified by algorithm is represented graphically in Figure 9.

The computational setting used for the experiments performed in this section is documented in the Python programs Experiment.py and HamiltonianSystem.py in²⁰ that can be used to replicate these experiments.

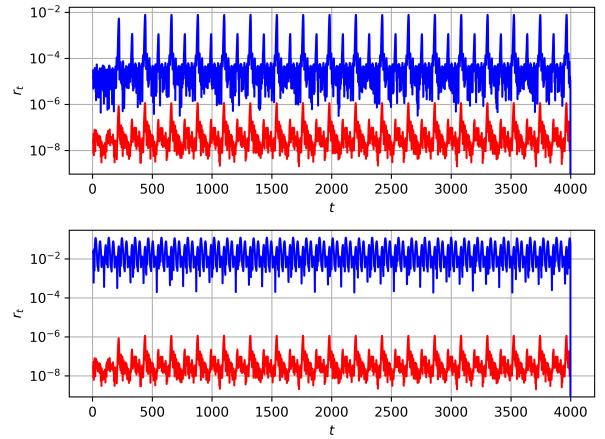


Figure 5: Prediction errors in the ℓ_∞ -norm.

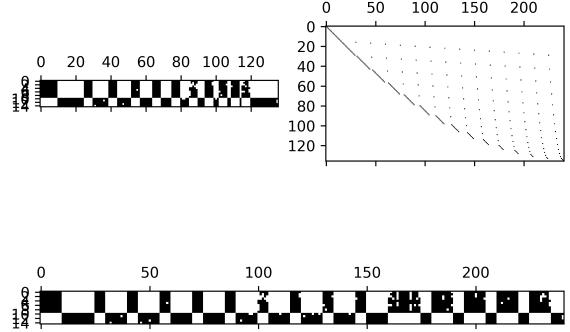


Figure 6: Graphical representation of coupling matrix and its corresponding structured factors.

C. Identification of a network of Duffing oscillators with symmetries via ERRC

In this section a network of three coupled Duffing oscillators with the following configuration

$$\frac{dx_i}{dt} = y_i, \quad i \in \{1, 2, 3\} \quad (\text{V.3})$$

$$\frac{dy_i}{dt} = \sigma y_j - x_i(\beta + \alpha^2 x_i) + \sum_{ij} \eta_{ij}(x_i - x_j), \quad i \in \{1, 2, 3\}$$

$$x_1(0) = 8, x_2(0) = 7, x_3(0) = 4,$$

$$y_1(0) = 15, y_2(0) = 14, y_3(0) = 9$$

is identified, for $\alpha = 1, \beta = -36, \sigma = 0$ and with all the coupling strengths η_{ij} equal to $1/5$. The configuration of these experiment is based on the example II.2 considered in¹⁴, an important part of the motivation for the study of this types of networks of oscillators comes from interesting applications in engineering and biological cybernetics like the ones presented in^{6, 12} and⁵. Since all the coupling strengths η_{ij} are equal to $1/5$, as established in¹⁴ the system (V.3) will be D_3 -equivariant, and for the configuration used for this experiment, the matrix representation of the corresponding group of symmetries $D_3 = \langle r, \kappa | r^3 = \kappa^2 = e, \kappa r \kappa = r^{-1} \rangle$ is deter-

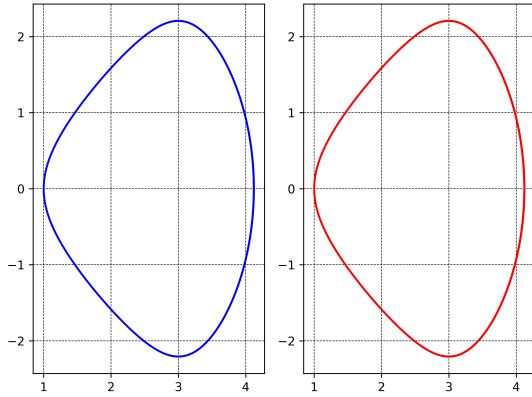


Figure 7: 2D graphical representation of: training data (left), predicted behavior (right).

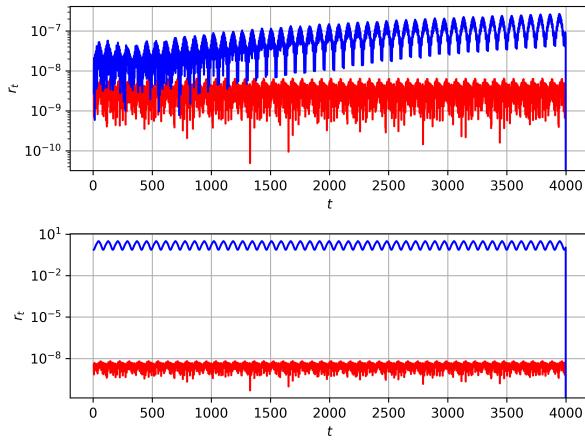


Figure 8: Prediction errors in the ℓ_∞ -norm.

mined by the following assignments.

$$r \mapsto r_{\rho} = I_2 \otimes \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

$$\kappa \mapsto \kappa_{\rho} = I_2 \otimes \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

The synthetic signals corresponding to the data sample $\Sigma_{2400} \subset \mathbb{R}^3$ that will be used for system identification have been computed with an explicit time integration scheme, using lsoda from the FORTRAN library odepak via the Python function odeint, with the Python program DuffingOscillators.py in²⁰. The model was trained using an embedding map $\bar{\partial}_p$ of order $p = 2$, with 30% of the synthetic reference data.

The reference synthetic signal data and the corresponding identified signals are illustrated in Figure 10.

The prediction errors in the ℓ_∞ -norm are shown in figure 11.

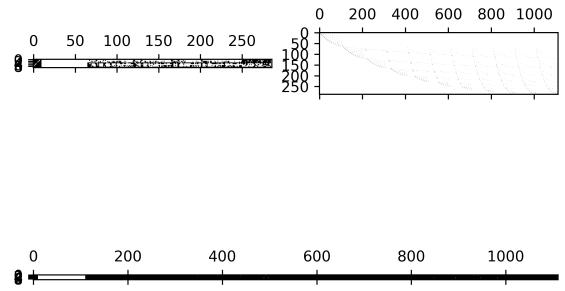


Figure 9: Graphical representation of coupling matrix and its corresponding structured factors.

The strucuted coupling matrix corresponding to the equivariant evolution operator identified by algorithm is represented graphically in Figure 12.

The computational setting used for the experiments performed in this section is documented in the Python programs Experiment.py and DuffingOscillators.py in²⁰ that can be used to replicate these experiments.

D. Identification of Rössler systems with TERRC

In this section a reservoir computer model corresponding to a pair of attractors of a Rössler system of the form.

$$\begin{aligned} \frac{dx}{dt} &= -y - z \\ \frac{dy}{dt} &= x + ay \\ \frac{dz}{dt} &= b + z(x - c) \end{aligned} \tag{V.4}$$

For the experiments performed in this section we will consider a trivial symmetry group for (V.4), and for the configuration used for this experiment, the matrix representation of the corresponding trivial group of symmetries $E = \langle e \rangle$ is determined by the following assignment.

$$e \mapsto e_{\rho} = I_3$$

1. Chaotic attractor

Let us consider the chaotic attractor corresponding to the system (V.1) for $a = 0.2$, $b = 0.2$ and $c = 5.7$. The synthetic signals corresponding to the data sample $\Sigma_{2400} \subset \mathbb{R}^3$ that will be used for system identification have been computed with an explicit time integration scheme, using lsoda from the FORTRAN library odepak via the Python function odeint, with the Python program LorenzSystem.py in²⁰. The signals corresponding to the identified/predicted dynamical behavior data $\Sigma_{1600}^{(P)} \subset \mathbb{R}^3$ are visualized in figure 13.

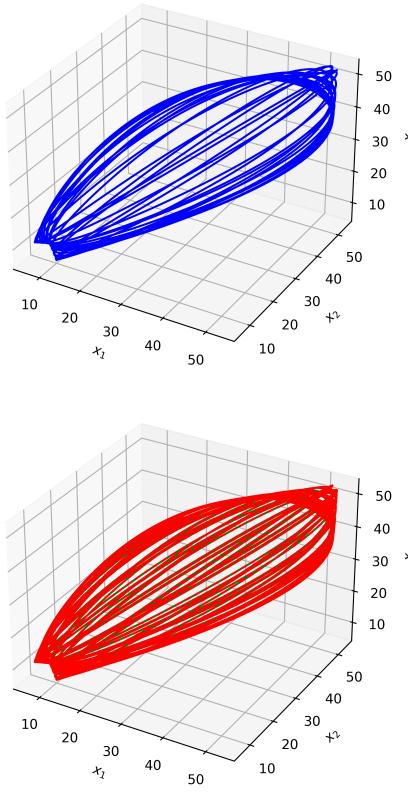


Figure 10: 3D graphical representation of: training data (left), predicted behavior (right).

We will apply algorithm 2 along the lines of theorem III.3 to identify model (V.4), using a RRC with embedding map $\tilde{\Omega}_2 : \mathbb{R}^{3L} \rightarrow \mathbb{R}^{d(3L)}$ for $L = 5$.

The prediction errors in the ℓ_∞ -norm are shown in figure 14.

The strucuted coupling matrix corresponding to the equivariant evolution operator identified by algorithm is represented graphically in Figure 15.

2. Periodic attractor

Let us consider the periodic attractor corresponding to the system (V.1) for $a = 0.1$, $b = 0.1$ and $c = 4$. The synthetic signals corresponding to the data sample $\Sigma_{300} \subset \mathbb{R}^3$ that will be used for system identification have been computed with an explicit time integration scheme, using lsoda from the FORTRAN library odepack via the Python function odeint, with the Python program LorenzSystem.py in²⁰. The signals corresponding to the training and predicted dynamical behavior data $\Sigma_{1600}^{(P)} \subset \mathbb{R}^3$ are visualized in figures 16 and 17, respectively.

We will apply algorithm 2 along the lines of theorem III.3

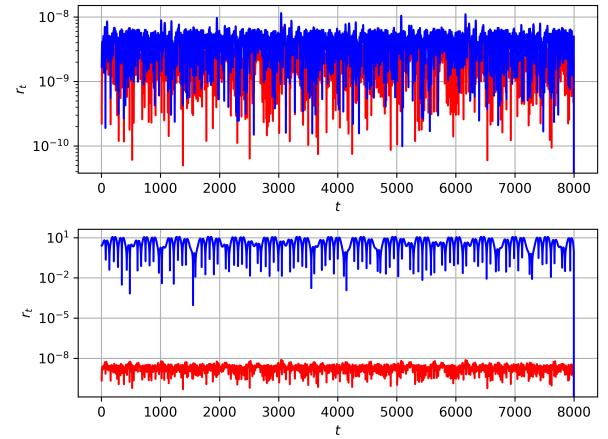


Figure 11: Prediction errors in the ℓ_∞ -norm.

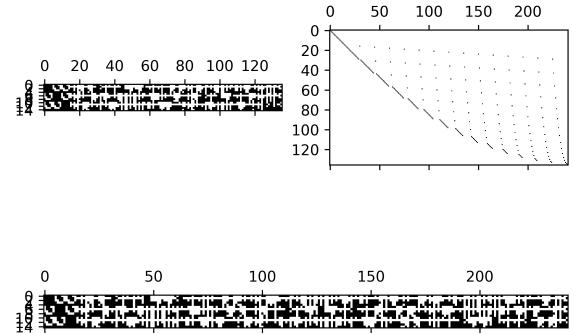


Figure 12: Graphical representation of coupling matrix and its corresponding structured factors.

to identify model (V.1), using a RRC with embedding map $\tilde{\Omega}_2 : \mathbb{R}^{3L} \rightarrow \mathbb{R}^{d(3L)}$ for $L = 5$.

The prediction errors in the ℓ_∞ -norm are shown in figure 18.

The strucuted coupling matrix corresponding to the equivariant evolution operator identified by algorithm are represented graphically in Figures 19 and 20.

The spectra of the linear and unitary approximants of the evolution operator for this attractor are illustrated in Figures 21

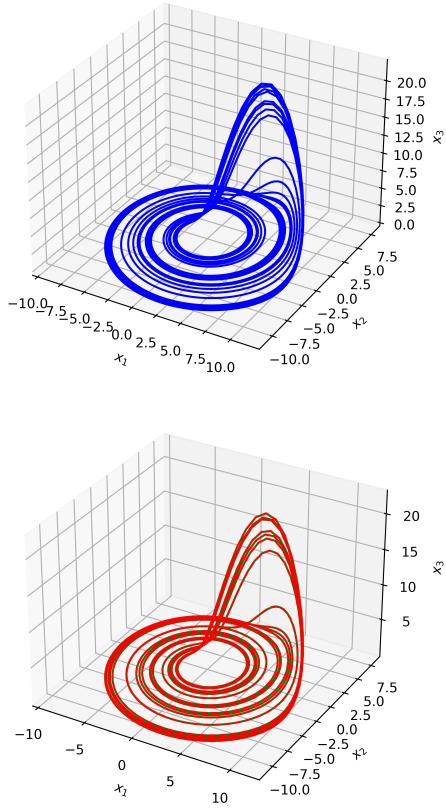


Figure 13: 3D graphical representation of: training data (left), predicted behavior (right).

E. Identification of pairs of coupled Van Der Pol oscillators using TERRC

In this section we will compute a reservoir computer model for a pair of coupled Van Der Pol oscillators of the form.

$$\begin{aligned} \frac{dx_1}{dt} &= -y_1, \\ \frac{dx_2}{dt} &= -y_2, \\ \frac{dy_1}{dt} &= -x_1 - \alpha y_1 \left(\frac{y_1^2}{3} - 1 \right) + \beta + \varepsilon y_1^2 + \gamma (2y_1 - y_2), \\ \frac{dy_2}{dt} &= -x_2 - \alpha y_2 \left(\frac{y_2^2}{3} - 1 \right) + \beta + \varepsilon y_2^2 + \gamma (2y_2 - y_1) \end{aligned} \quad (\text{V.5})$$

For the experiments performed in this section we will consider a trivial symmetry group for (V.4), and for the configuration used for this experiment, the matrix representation of the corresponding trivial group of symmetries $E = \langle e \rangle$ is determined by the following assignment.

$$e \mapsto e_{\rho} = I_3$$

Let us consider the attractor corresponding to the system (V.5) for $\alpha = 1.0, \beta = 2.0, \gamma = -0.5, \delta = -0.5, \varepsilon = 0.5$, with $x_1(0) = 2, x_2(0) = 1, y_1(0) = 0, y_2(0) = 0$. The synthetic signals corresponding to the data sample $\Sigma_{2500} \subset \mathbb{R}^3$ that will

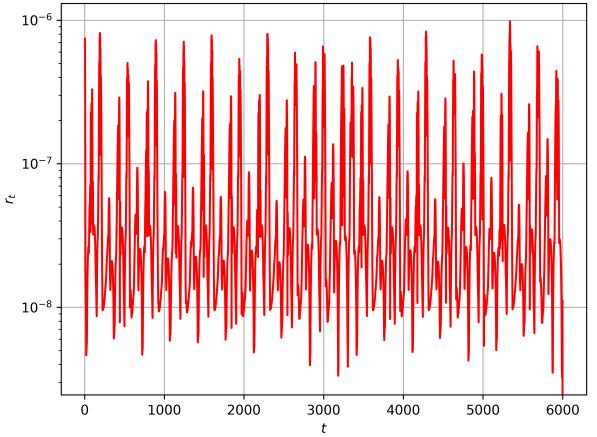


Figure 14: Prediction errors in the ℓ_∞ -norm.

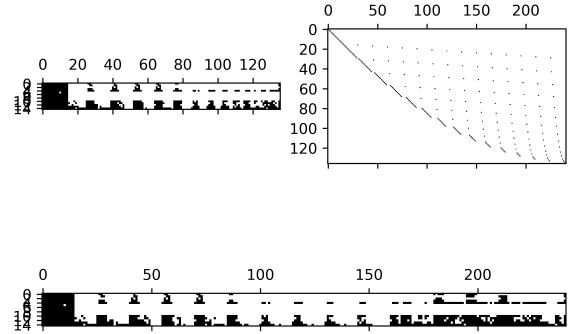


Figure 15: Graphical representation of coupling matrix and its corresponding structured factors.

be used for system identification have been computed with an explicit time integration scheme, using lsoda from the FORTRAN library odepak via the Python function odeint, with the Python program LorenzSystem.py in²⁰. The signals corresponding to the identified/predicted dynamical behavior data $\Sigma_{1600}^{(P)} \subset \mathbb{R}^3$ are visualized in figure 22.

We will apply algorithm 2 along the lines of theorem III.3 to identify model (V.4), using a RRC with embedding map $\partial_2 : \mathbb{R}^{3L} \rightarrow \mathbb{R}^{d(3L)}$ for $L = 5$.

The prediction errors in the ℓ_∞ -norm are shown in figure 23.

The strucuted coupling matrix corresponding to the equivariant evolution operator identified by algorithm is represented graphically in Figure 24.

F. Approximation Errors

The approximation root mean square errors (**RMSE**) are summarized in table I.

It is appropriate to mention that the root mean square errors can present little fluctuations as one performs several numer-

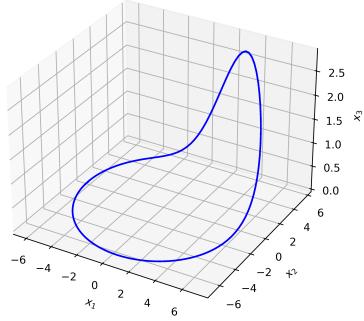


Figure 16: 3D graphical representation of training data.

ical simulations, due primarily to the nature of the neural-network models, as the linear components tend to present very low or no variability from simulation to simulation.

VI. CONCLUSION

The ERRC model computation techniques presented in this document can be efectively implemented for identification and simulation of equivariant discrete-time dynamical systems.

FUTURE WORK

The extension of this techniques to model reduction and anomaly detection will be studied in future communications. The connections of the results in §III A to the solution of problems related to controllability and realizability of finite-state systems in classical and quantum information and automata theory in the sense of^{2-4,17}, will be further explored.

Some applications of the algorithms in §IV to data-driven discovery of cyber-physical systems, will be presented in future communications. Further applications of RRC models to to industrial automation and building information modeling (BIM) technologies, will be further explored as well.

Table I: RMSE

Experiment	RMSE $_{\infty}$
Experiment 1	0.07861432
Experiment 2	0.02571805
Experiment 3	2.63349747e-05
Experiment 4	0.00040892
Experiment 5	0.0041668
Experiment 6.1	3.69513991e-05
Experiment 6.2	0.00933797
Experiment 6.3	0.00586315
Experiment 7	0.02229667

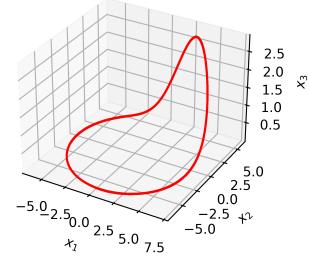
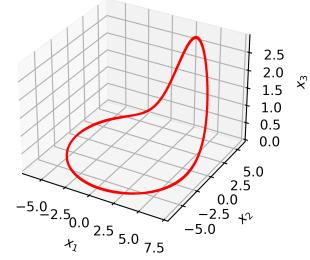
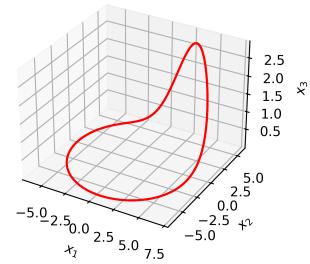


Figure 17: 3D graphical representation of predicted behavior.

ACKNOWLEDGMENT

The structure preserving matrix computations needed to implement the algorithms in §IV, were performed using Python 3.8.10 with the support and computational resources of the Scientific Computing Innovation Center (**CICC-UNAH**) of the National Autonomous University of Honduras under the research project PI-063-DICIHT.

The author wishes to thank Terry Loring, Marc Rieffel, Alex Cerjan, José García, Esteban Segura and Carlos Vargas for interesting and useful discussions.

DATA AVAILABILITY

The programs and data that support the findings of this study are openly available in the SDSI repository, reference number²⁰. The Python programs that support the findings of this study are openly available in the SPORT repository, reference number²⁰.

REFERENCES

¹Bahar Azari and Deniz Erdogmus. Equivariant deep dynamical model for motion prediction. In Gustau Camps-Valls, Francisco J. R. Ruiz, and Isabel

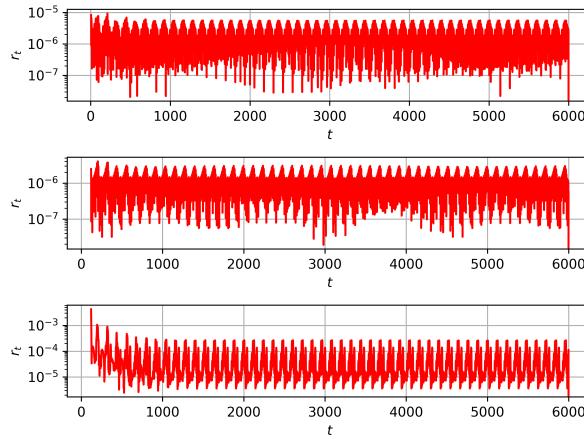


Figure 18: Prediction errors in the ℓ_∞ -norm.

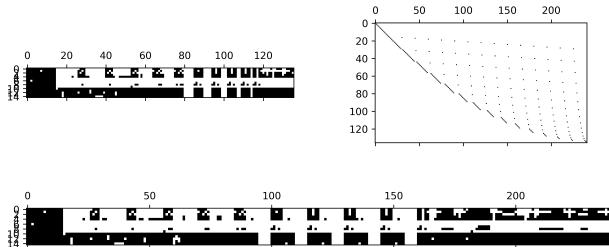


Figure 19: Graphical representation of coupling matrix and its corresponding structured factors.

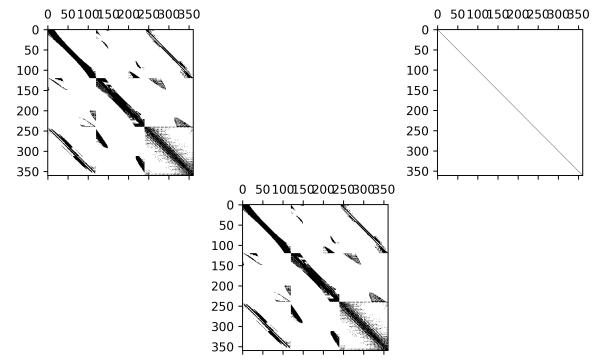


Figure 20: Graphical representation of coupling matrix and its corresponding structured factors.

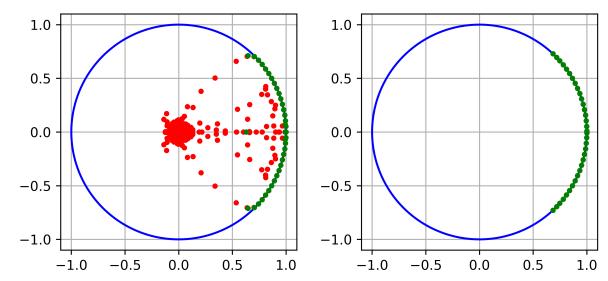


Figure 21: Graphical representation of the spectra of the linear and unitary approximants of the evolution operator.

- Valera, editors, *Proceedings of The 25th International Conference on Artificial Intelligence and Statistics*, volume 151 of *Proceedings of Machine Learning Research*, pages 11655–11668. PMLR, 28–30 Mar 2022.
- ²A. M. Bloch, R. W. Brockett, and C. Rangan. Finite controllability of infinite-dimensional quantum systems. *IEEE Transactions on Automatic Control*, 55(8):1797–1805, Aug 2010.
- ³R. Brockett and A. Willsky. Finite group homomorphic sequential system. *IEEE Transactions on Automatic Control*, 17(4):483–490, August 1972.
- ⁴R. W. Brockett. Reduced complexity control systems. *IFAC Proceedings Volumes*, 41(2):1 – 6, 2008. 17th IFAC World Congress.
- ⁵J. J. Collins and I. N. Stewart. Coupled nonlinear oscillators and the symmetries of animal gaits. *Journal of Nonlinear Science*, 3(1), 1993.
- ⁶J. J. Collins and I. N. Stewart. A group-theoretic approach to rings of coupled biological oscillators. *Biological Cybernetics*, 71(2), 1994.
- ⁷Marc Finzi, S. Stanton, Pavel Izmailov, and A. Wilson. Generalizing convolutional neural networks for equivariance to lie groups on arbitrary continuous data. In *ICML*, 2020.
- ⁸Daniel J. Gauthier, Erik Boltt, Aaron Griffith, and Wendson A. S. Barbosa. Next generation reservoir computing. *Nature Communications*, 12(1):5564, Sep 2021.
- ⁹Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, Baltimore, 3rd edition, 1996.
- ¹⁰Balasubramanya T. Nadiga. Reservoir computing as a tool for climate predictability studies. *Journal of Advances in Modeling Earth Systems*, 13(4):e2020MS002290, 2021. e2020MS002290 2020MS002290.
- ¹¹Marc A. Rieffel. Actions of finite groups on C^* -algebras. *Mathematica Scandinavica*, 47(1):157–176, 1980.
- ¹²Fabio Della Rossa, Louis Pecora, Karen Blaha, Afroza Shirin, Isaac Klickstein, and Francesco Sorrentino. Symmetries and cluster synchronization in multilayer networks. *Nature Communications*, 11(1), 2020.
- ¹³Yousef Saad. *Numerical Methods for Large Eigenvalue Problems*. Society for Industrial and Applied Mathematics, 2011.

¹⁴Anastasiya Salova, Jeffrey Emenheiser, Adam Rupe, James P. Crutchfield, and Raissa M. D’Souza. Koopman operator and its approximations for systems with symmetries. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 29(9):093128, 2019.

¹⁵Subhrajit Sinha, Sai Pushpak Nandanoori, and Enoch Yeung. Koopman operator methods for global phase space exploration of equivariant dynamical systems. *IFAC-PapersOnLine*, 53(2):1150–1155, 2020. 21st IFAC World Congress.

¹⁶Benjamin Steinberg. *Representation Theory of Finite Groups: An Introductory Approach*. Springer, 2012.

¹⁷D. C. Tarraf. An input-output construction of finite state ρ/μ approximations for control design. *IEEE Transactions on Automatic Control*, 59(12):3164–3177, Dec 2014.

¹⁸Fredy Vides. Sparse system identification by low-rank approximation. *CoRR*, abs/2105.07522, 2021.

¹⁹Fredy Vides. Computing semilinear sparse models for approximately eventually periodic signals. *IFAC-PapersOnLine*, 55(20):205–210, 2022. 10th Vienna International Conference on Mathematical Modelling MATHMOD 2022.

²⁰Fredy Vides. SPORT: Structure preserving operator representation toolset for Python. 2022. <https://github.com/FredyVides/SPORT>.

²¹Ye Yuan, Xiuchuan Tang, Wei Zhou, Wei Pan, Xiuting Li, Hai-Tao Zhang, Han Ding, and Jorge Goncalves. Data driven discovery of cyber physical systems. *Nature Communications*, 10(1):4894, Oct 2019.

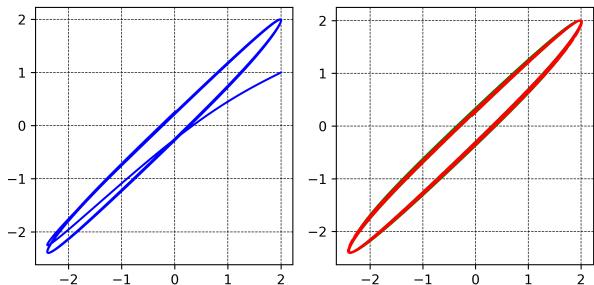


Figure 22: 2D graphical representation of: training data (left), predicted behavior (right).

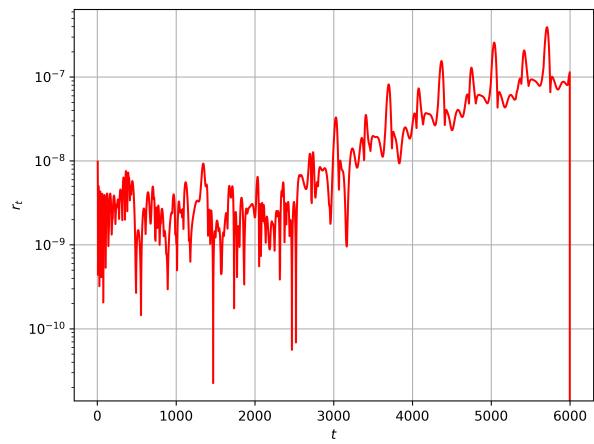


Figure 23: Prediction errors in the ℓ_∞ -norm.

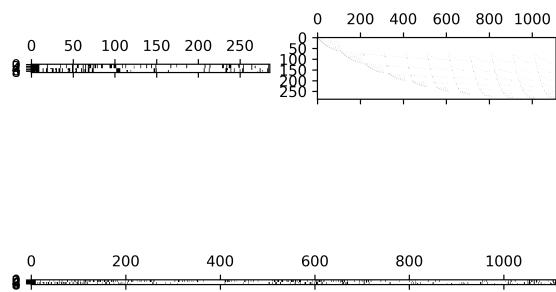


Figure 24: Graphical representation of coupling matrix and its corresponding structured factors.