# Linear Simple Cycle Reservoirs at the edge of Stability perform Fourier decomposition of the input driving signals

Robert Simon Fong,[1, a)] Boyu Li,[2, b)] and Peter Tiňo[1, c)]

[1)]*School of Computer Science, University of Birmingham, Birmingham, B15 2TT, UK*

[2)]*Department of Mathematical Sciences, New Mexico State University, Las Cruces, New Mexico, 88003, USA*

(Dated: 27 March 2025)

This paper explores the representational structure of linear Simple Cycle Reservoirs (SCR) operating at the edge of stability. We view SCR as providing in their state space feature representations of the input-driving time series. By endowing the state space with the canonical dot-product, we "reverse engineer" the corresponding kernel (inner product) operating in the original time series space. The action of this time-series kernel is fully characterized by the eigenspace of the corresponding metric tensor. We demonstrate that when linear SCRs are constructed at the edge of stability, the eigenvectors of the time-series kernel align with the Fourier basis. This theoretical insight is supported by numerical experiments.

---

[a)]Electronic mail: r.s.fong@bham.ac.uk

[b)]Electronic mail: boyuli@nmsu.edu

[c)]Electronic mail: p.tino@cs.bham.ac.uk

Recurrent Neural Networks (RNNs) are machine learning methods for modeling temporal dependencies in sequential data, but their training can be computationally demanding. Reservoir Computing (RC), a simplified subset of RNNs, circumvents this problem by fixing the internal dynamics of the network (the reservoir) and focusing training on the readout layer. Simple Cycle Reservoir (SCR) is a type RC model that stands out for its minimalistic design and proven capability to universally approximate a wide class of processes operating on time series data (namely time-invariant fading memory filters) even in the linear dynamics regime (and non-linear static readouts). We show that, interestingly, when linear SCR is constructed at the edge of stability, it implicitly represents the time series according to a well known and widely used technique of Fourier signal decomposition. This insight demonstrates that deep connections can exist between recurrent neural networks, classical signal processing techniques and statistics, paving the way for their enhanced understanding and innovative applications.

## I.  INTRODUCTION

Recurrent Neural Networks (RNNs) are input-driven parametric state-space machine learning models designed to capture temporal dependencies in sequential input data streams. They encode time series data into a latent state space, dynamically storing temporal information within state-space vectors.

Reservoir Computing (RC) models is a subset of RNNs that operate with a fixed, non-trainable input-driven dynamical system (known as the *reservoir*) and a static trainable readout layer producing model responses based on the reservoir activations. This design uniquely simplifies the training process by concentrating adjustments solely on the readout layer (thus avoiding back-propagating the error information backwards through time), leading to enhanced computational efficiency. The simplest implementations of RC models include Echo State Networks (ESNs)[Jae01,MNM02,TD01,LJ09].

Simple Cycle Reservoirs (SCR) represent a specialized class of RC models characterized by a single degree of freedom in the reservoir construction (modulo the state space dimensionality), structured through uniform ring connectivity and binary input weights with an

aperiodic sign pattern. Recently, SCRs were shown to be universal approximators of time-invariant dynamic filters with fading memory over $\mathbb{C}$ and $\mathbb{R}$ respectively in[LFT24,FLT24], making them highly suitable for integration in photonic circuits for high-performance, low-latency processing[BVdSBV22,LBFM+17,HVK+20].

Understanding the intricacies of SCRs in depth is essential. In this work, we employ the kernel view of linear ESNs introduced in[Tin20], in which the state-space 'reservoir' representation of (potentially left-infinite) input sequences is treated as a feature map corresponding to the given reservoir model through the associated reservoir kernel. For linear reservoirs, the canonical dot product of two input sequences' feature representations is analytically expressible as the (semi-)inner product of the sequences themselves. The corresponding metric tensor reveals the representational structure imposed by the reservoir on the input sequences, in particular in the metric tensor's eigenspace containing dominant projection axes (time series 'motifs') and scaling ('importance' factors).

To assess the "richness" of linear SCR state-space representations,[Tin20] proposed analyzing the relative area of motifs under the Discrete Fourier Transform (DFT). It was observed that the richness of these representations collapses at the edge of stability when the spectral radius $\rho$ of the dynamic coupling matrix equals to 1.

In this paper we theoretically analyze the collapse of motif richness at the edge of stability and show that when $\rho = 1$ the SCR kernel motifs correspond to Fourier basis. We begin by reviewing the notion of SCRs[RT10], kernel view of ESNs[Tin20], and Reservoir Motif Machines[TFL24] in Section II. The contribution of this paper, in the subsequent sections, are outlined as follows:

1. In Section IV, we show in $\mathbb{C}$ that motifs of linear SCR at the edge of stability are harmonic functions.

2. In Section V, we show in $\mathbb{R}$ that $n$ dimensional linear SCR has $\lceil \frac{n}{2} \rceil$ symmetric motifs and $\lfloor \frac{n}{2} \rfloor$ skew-symmetric motifs.

3. In Section VI, we combine the results of the previous two sections and demonstrate numerically that in $\mathbb{R}$, the motifs of linear SCR at edge of stability are exactly the columns of real Fourier basis matrix.

4. Finally in Section VII, we conclude the paper with numerical experiments supporting

our findings.

## II.  SIMPLE CYCLE RESERVOIR AND ITS TEMPORAL KERNEL

Let $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ be a field. We first formally define the principal object of our study - parametrized linear driven dynamical system with a (possibly) non-linear readout.

**Definition II.1.** A **linear reservoir system** over $\mathbb{K}$ is is the triplet $R := (\mathbf{W}, \mathbf{V}, h)$ where the **dynamic coupling** $\mathbf{W} \in \mathbb{M}_{n \times n}(\mathbb{K})$ is an $n \times n$ matrix over $\mathbb{K}$, the **input-to-state coupling** $\mathbf{V} \in \mathbb{M}_{n \times m}(\mathbb{K})$ is an $n \times m$ matrix, and the state-to-output mapping (**readout**) $h : \mathbb{K}^n \to \mathbb{K}^d$ is a (trainable) continuous function.

The corresponding linear dynamical system is given by:

$$\begin{cases} \mathbf{x}_t & = \mathbf{W}\mathbf{x}_{t-1} + \mathbf{V}\mathbf{c}_t \\ \mathbf{y}_t & = h(\mathbf{x}_t) \end{cases} \tag{II.1}$$

where $\{\mathbf{c}_t\}_{t \in \mathbb{Z}_-} \subset \mathbb{K}^m$, $\{\mathbf{x}_t\}_{t \in \mathbb{Z}_-} \subset \mathbb{K}^n$, and $\{\mathbf{y}_t\}_{t \in \mathbb{Z}_-} \subset \mathbb{K}^d$ are the external inputs, states and outputs, respectively. We abbreviate the dimensions of $R$ by $(n, m, d)$.

We make the following assumptions for the system:

1. $W$ is assumed to be strictly **contractive**. In other words, its operator norm $\|W\| < 1$. The system (II.1) thus satisfies the fading memory property (FMP)[LFT24].

2. We assume the input stream is $\{\mathbf{c}_t\}_{t \in \mathbb{Z}_-}$ is **uniformly bounded**. In other words, there exists a constant $M$ such that $\|\mathbf{c}_t\| \leq M$ for all $t \in \mathbb{Z}_-$.

The contractiveness of $W$ and the uniform boundedness of input stream imply that the images $\mathbf{x} \in \mathbb{K}^n$ of the inputs $\mathbf{c} \in (\mathbb{K}^m)^{\mathbb{Z}_-}$ under the linear reservoir system live in a compact space $X \subset \mathbb{K}^n$. With slight abuse of mathematical terminology we call $X$ a **state space**.

**Definition II.2.** Let $\mathbf{C} = [c_{ij}]$ be an $n \times n$ matrix. We say $\mathbf{C}$ is a **permutation matrix** if there exists a permutation $\sigma$ in the symmetric group $S_n$ such that

$$c_{ij} = \begin{cases} 1, & \text{if } \sigma(i) = j, \\ 0, & \text{if otherwise.} \end{cases}$$

We say a permutation matrix $\mathbf{C}$ is a **full-cycle permutation**[Nota] if its corresponding permutation $\sigma \in S_n$ is a cycle permutation of length $n$. Finally, a matrix $\mathbf{W} = \rho \cdot \mathbf{C}$ is called a **contractive full-cycle permutation** if $\rho \in (0, 1)$ and $\mathbf{C}$ is a full-cycle permutation.

The idea of simple cycle reservoir was presented in[RT10] as a reservoir system with a very small number of design degrees of freedom, yet retaining performance capabilities of more complex or (unnecessarily) randomized constructions. In fact, it can be shown that even with such a drastically reduced design complexity, SCR models are universal approximators of fading memory filters[LFT24,FLT24].

**Definition II.3.** A linear reservoir system $R = (\mathbf{W}, \mathbf{w}, h)$ with dimensions $(n, m, d)$ is called a **Simple Cycle Reservoir (SCR)** [Notb] if

1. $\mathbf{W}$ is a contractive full-cycle permutation, and

2. $\mathbf{w} \in \mathbb{M}_{n \times m}(\{-1, 1\})$.

One possibility to understand inner representations of the input-driving time series forming inside the reservoir systems is to view the reservoir state space as a temporal feature space of the associated reservoir kernel[Tin20,GGO22]. Consider a linear reservoir system $R = (\mathbf{W}, \mathbf{w}, h)$ over $\mathbb{K}$ with dimensions $(n, 1, d)$ operating on univariate input.

Let $\tau > n$ denote the length of the look back window and consider two sufficiently long time series of length $\tau > n$,

$$\mathbf{u} = (u(-\tau + 1), u(-\tau + 2), \ldots, u(-1), u(0))$$
$$=: (u_1, u_2, \ldots, u_\tau) \in \mathbb{K}^\tau$$

and

$$\mathbf{v} = (v(-\tau + 1), v(-\tau + 2), \ldots, v(-1), v(0))$$
$$=: (v_1, v_2, \ldots, v_\tau) \in \mathbb{K}^\tau$$

we consider the reservoir states reached upon reading them (with zero initial state) their feature space representations[Tin20]:

$$\phi(\mathbf{u}) = \sum_{j=1}^{\tau} u_j \mathbf{W}^{\tau - j} \mathbf{w}, \quad \phi(\mathbf{v}) = \sum_{j=1}^{\tau} v_j \mathbf{W}^{\tau - j} \mathbf{w}.$$

The canonical dot product (reservoir kernel)

$$K(\mathbf{u}, \mathbf{v}) = \langle \phi(\mathbf{u}), \phi(\mathbf{v}) \rangle$$

can be written in the original time series space as a semi-inner product $\langle \mathbf{u}, \mathbf{v} \rangle_Q = \mathbf{u}^\top \mathbf{Q} \mathbf{v}$, where

$$Q_{i,j} = \mathbf{w}^\top \left( \mathbf{W}^\top \right)^{i-1} \mathbf{W}^{j-1} \mathbf{w}. \tag{II.2}$$

Since the (semi-)metric tensor $\mathbf{Q} \in \mathbb{M}_{\tau \times \tau}(\mathbb{K})$ is symmetric and positive semi-definite, it admits the following eigen-decomposition:

$$\mathbf{Q} = \mathbf{M} \Lambda_Q \mathbf{M}^\top, \tag{II.3}$$

where $\Lambda_Q := \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_{N_m})$ is a diagonal matrix consisting of non-negative eigenvalues of $\mathbf{Q}$ with the corresponding eigenvectors $\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_{N_m} \in \mathbb{K}^\tau$ (columns of $\mathbf{M}$). The $N_m := \operatorname{rank}(Q) \le N \le \tau$ eigenvectors of $\mathbf{M}$ with positive eigenvalues are called the *motifs* of $R$. We have:

$$K(\mathbf{u}, \mathbf{v}) = \left( \Lambda_Q^{\frac{1}{2}} \mathbf{M}^\top \mathbf{u} \right)^\top \left( \Lambda_Q^{\frac{1}{2}} \mathbf{M}^\top \mathbf{v} \right).$$

In particular, the reservoir kernel is a canonical dot product of time series projected onto the motif space spanned by $\{\mathbf{m}_i\}_{i=1}^{N_m}$:

$$K(\mathbf{u}, \mathbf{v}) = \langle \tilde{\mathbf{u}}, \tilde{\mathbf{v}} \rangle,$$

where

$$\begin{aligned}
\tilde{\mathbf{u}} &= \Lambda_Q^{\frac{1}{2}} \mathbf{M}^\top \mathbf{u} \tag{II.4} \\
&= \begin{bmatrix} \lambda_1^{\frac{1}{2}} \cdot \langle \mathbf{m}_1, \mathbf{u} \rangle \\ \vdots \\ \lambda_{N_m}^{\frac{1}{2}} \cdot \langle \mathbf{m}_{N_m}, \mathbf{u} \rangle \end{bmatrix} \\
&= \left( \lambda_i^{\frac{1}{2}} \cdot \langle \mathbf{m}_i, \mathbf{u} \rangle \right)_{i=1}^{N_m} \in \mathbb{M}_{N_m}(\mathbb{K}).
\end{aligned}$$

**Reservoir Motif Machine** (RMM)[TFL24] is a predictive model motivated by the kernel view of linear Echo State Networks described above. By projecting the $\tau$-blocks of input time series onto the reservoir motif space given by $\operatorname{span}(\{\mathbf{m}_i\}_{i=1}^{N_m})$, RMM captures temporal and structural dynamics in a computationally efficient feature map.

In particular, rather than relying on the motif weights $\lambda_i^{\frac{1}{2}}$ determined by the reservoir $R$ in Equation (II.4), RMM introduces a set of adaptable motif coefficients, denoted as $C := \{c_i \in \mathbb{R}\}_{i=1}^{N_m}$, to define its feature map as follows:

$$\varphi(\mathbf{u}; C) = \begin{bmatrix} c_1 \cdot \langle \mathbf{m}_1, \mathbf{u} \rangle \\ \vdots \\ c_{N_m} \cdot \langle \mathbf{m}_{N_m}, \mathbf{u} \rangle \end{bmatrix}$$
$$= (c_i \cdot \langle \mathbf{m}_i, \mathbf{u} \rangle)_{i=1}^{N_m} \in \mathbb{R}^{N_m}. \tag{II.5}$$

This feature map is used to train a predictive model, such as linear regression or kernel-based methods, directly in the motif space.

**Remark II.4.** To streamline the theoretical analysis of SCR kernels, in line with[Tin20], we assume that the length of the look-back window (past horizon) $\tau$ is an integer multiple of the dimension of the state space $n$, i.e. there exists $k \in \mathbb{N}_+$ such that $\tau = k \cdot n$. In particular, denoting by $\mathbf{m}$ a SCR motif calculated with $\tau = n$, it is shown in[Tin20] that when $\tau = k \cdot n$, the corresponding motif is a concatenation of $k$ copies of $\mathbf{m}$, scaled by $\rho^{l \cdot n}$, $l = 0, 1, 2, ..., k - 1$. Recall that $\rho$ is the spectral radius of the dynamic coupling $W$. Hence, to study SCR motifs with past horizon $\tau = k \cdot n$, for any $k \in \mathbb{N}_+$, it is sufficient to study only the base case $\tau = n$.

## III.   ON THE EDGE OF STABILITY OF SIMPLE CYCLE RESERVOIRS

Having defined the reservoir temporal kernel, one can ask how "representationally rich" is the associated feature space (span of the reservoir motifs). To quantify the 'richness' of the reservoir feature space of a linear reservoir system over $\mathbb{R}$,[Tin20] proposed the following procedure:

Consider a linear reservoir system $R = (\mathbf{W}, \mathbf{w}, h)$ with dimensions $(n, 1, d)$ over $\mathbb{R}$. Suppose $\mathbf{W}$ has spectral radius $\rho$. Recall from Remark II.4 that: to study the SCR motif structure of, it is sufficient to consider the past horizon $\tau = n$. The motif matrix $\mathbf{M} \in \mathbb{M}_{n \times n}(\mathbb{R})$ is constructed according to Equation (II.3) from the matrix $\mathbf{Q}$ (metric tensor of the inner product of the reservoir kernel).

First, Discrete Fourier Transform (DFT) is applied to the kernel motifs (columns of $\mathbf{M}$), considering only those with motif weights upto a threshold of $10^{-2}$ of the highest motif

weight. This yields an $n \times n'$ matrix of Fourier coefficients over $\mathbb{C}$ with $n' \leq n$. These Fourier coefficients are then collected in a (multi)set $Z := \{z_k\}_{k=1}^{n \cdot n'}$.

To evaluate the diversity and spread of the Fourier coefficients in the complex plane,[Tin20] proposed calculating the coarse-grained area occupied by $Z \subseteq \mathbb{C}$. In particular, the box $[-7, 7]^2$ in the complex plane is partitioned into a grid of cells (following[Tin20] we use side length 0.05). The **relative area** covered by $Z \subseteq \mathbb{C}$ is defined as the ratio of the number of cells visited by the coefficients $z_k \in Z$ to the total number of cells in the grid. An example of the distribution of Fourier coefficients $Z$ of linear SCR with $n = 97$ at $\rho = 0.9813, 0.999, 1$ and $\mathbf{w}$ being the first $n$ digits of binary expansion of $\pi$ is presented in Figure 1[Notc].
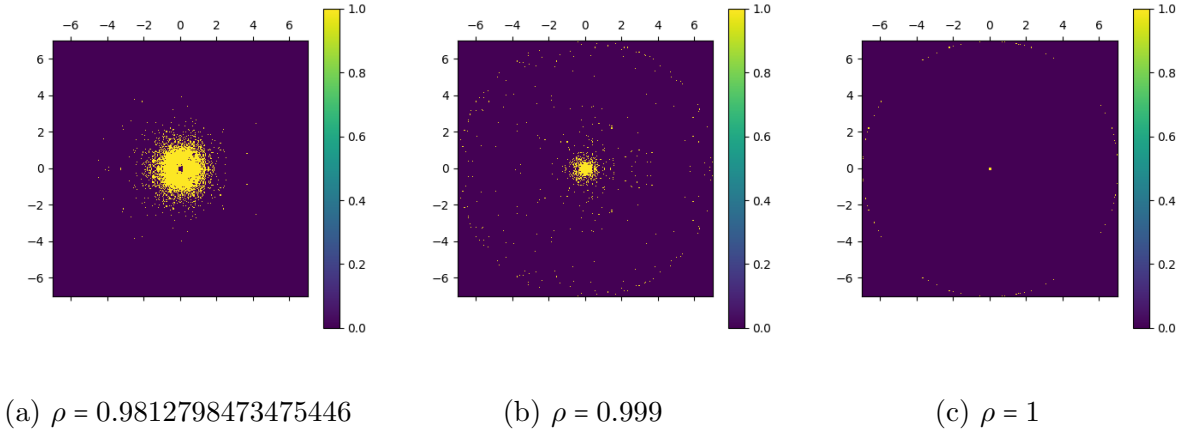


(a) $\rho = 0.9812798473475446$      (b) $\rho = 0.999$      (c) $\rho = 1$

FIG. 1: Example of Fourier coefficient of linear SCR at $\rho = 0.9812798473475446, 0.999, 1$ respectively. $0.9812798473475446$ in particular is where the relative area peaks in Figure 2.

Replicating the experiment in[Tin20] in Figure 2, we observe that the "richness" of the motif space of SCR, as measured by relative area under the current setup, increases as the spectral radius approaches approximately $\rho \approx 0.9813$. Beyond this point, the measure sharply declines, aligning with the results for a randomly generated reservoir. This decline is also observed in is also shown in Figure 2, as $\rho$ increases from approximately 0.98 (the peak of Figure 2) to the edge of stability at $\rho = 1$.

In this paper, we aim to investigate this 'collapse of representational richness'[Notd] of SCR models at the edge of stability $\rho = 1$. In particular, we will show that at the edge of stability, the motifs of linear SCR are (sampled) harmonic functions.

We now derive properties of the motifs of linear SCR in both the complex real domains:
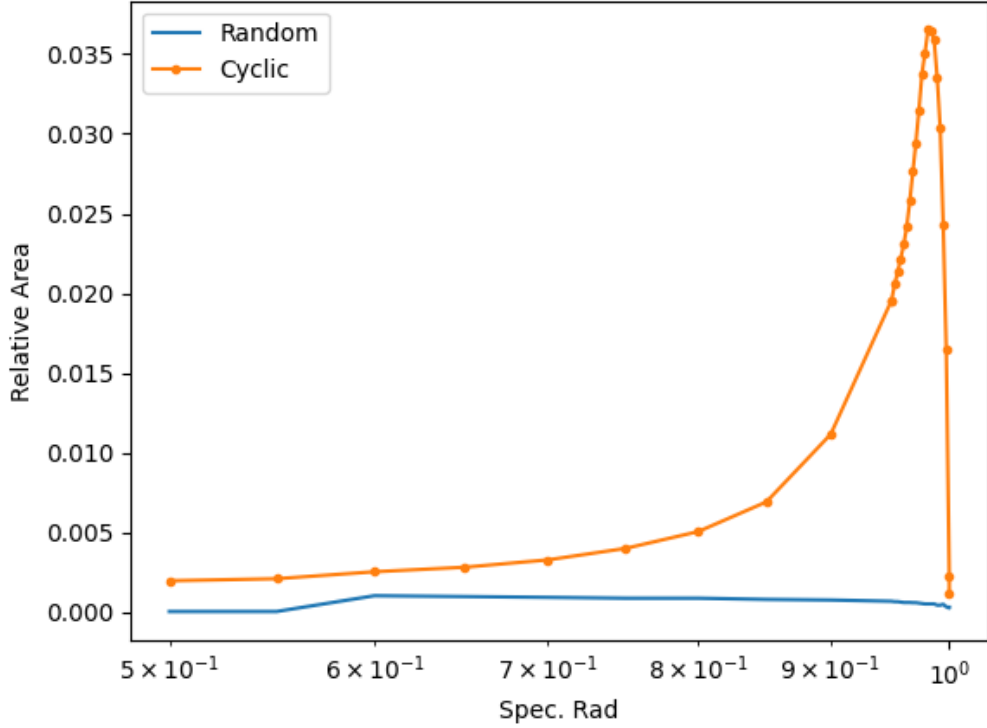
FIG. 2: Relative area of Linear SCR and Randomly generated Reservoir with respect to
the spectral radius

1. In Section IV, we show in $\mathbb{C}$ that the motifs are harmonic functions. Following the approach of[LFT24], we begin with unitary dynamical coupling and progress to cyclic permutations.

2. In Section V, we demonstrate in $\mathbb{R}$ that the number of symmetric motifs is $\lceil \frac{n}{2} \rceil$, while the number of skew-symmetric motifs is $\lfloor \frac{n}{2} \rfloor$. In line with[FLT24], we start with orthogonal coupling and then move onto cyclic permutation.

Combining these two results, we conclude that at $\rho = 1$, the motifs alternate between real and imaginary components of the first $n$ elements of the Fourier bases, which correspond to cosines and sines, respectively. We supplement our theoretical findings with numerical simulations of the motif space of SCR in the next section, which then lead to the numerical experiments in the final section.

9

## IV.  UNIT SPECTRAL RADIUS SCR IMPLIES HARMONIC MOTIFS IN COMPLEX DOMAIN

We first show that, in the complex domain $\mathbb{C}$, the motifs of SCR can be derived explicitly. In particular, in this section we set $\mathbb{K} = \mathbb{C}$ and show that when the spectral radius $\rho = 1$, the motifs of linear SCR are *harmonic*, i.e. they are precisely the Fourier basis (columns of the Fourier matrix).

Consider a linear reservoir system $R = (\mathbf{W}, \mathbf{w}, h)$ over $\mathbb{C}$ with dimensions $(n, 1, d)$ and $\mathbf{W}$ of spectral radius $\rho$. Let $\mathbf{Q}_\rho$ denote the metric tensor of the reservoir kernel (Eq. (II.2)).

In the spirit of[LFT24], we begin by considering a more general setting of $\mathbf{W} = \rho\mathbf{U}$, where $\mathbf{U}$ is a unitary matrix in $\mathbb{M}_{n \times n}(\mathbb{C})$ (i.e. $\mathbf{U}\mathbf{U}^* = \mathbf{U}^*\mathbf{U} = \mathbf{I}$) and $\rho \in (0, 1]$. We then move to the special case where $\mathbf{U} = \mathbf{C}$ is a cyclic permutation. Since $\mathbf{U}$ is unitary, its eigenvalues all have magnitude 1. We let its eigenvalues be $\{s_j \in \mathbb{C} : 1 \leq j \leq n\}$ with corresponding eigenvectors $\{\xi_j : 1 \leq j \leq n\}$, and we know each $|s_j| = 1$.

By construction:

$$\mathbf{Q}_\rho = [Q_{\rho\ (i,j)}]_{1 \leq i,j \leq \tau} = [\mathbf{w}^*\mathbf{W}^{*(i-1)}\mathbf{W}^{j-1}\mathbf{w}]_{1 \leq i,j \leq \tau}.$$

The matrix $\mathbf{Q}_\rho$ is a $\tau \times \tau$ matrix. Denote

$$\mathbf{X}_\rho = [\mathbf{W}^{*(i-1)}\mathbf{W}^{j-1}]_{1 \leq i,j \leq \tau},$$

and,

$$\hat{\mathbf{w}} = \begin{bmatrix} \mathbf{w} & & & \\ & \mathbf{w} & & \\ & & \ddots & \\ & & & \mathbf{w} \end{bmatrix}.$$

Notice that $\mathbf{X}_\rho$ is an $\tau n \times \tau n$ matrix while $\hat{\mathbf{w}}$ is an $\tau n \times \tau$ matrix. Then by construction, $\mathbf{Q}_\rho = \hat{\mathbf{w}}^*\mathbf{X}_\rho\hat{\mathbf{w}}$. Notice that since $\mathbf{W} = \rho\mathbf{U}$ and $\mathbf{U}$ is unitary, we can rewrite $\mathbf{X}_\rho$ as follows:

$$\mathbf{X}_\rho = [\rho^{i+j-2}\mathbf{U}^{j-i}]_{1 \leq i,j \leq \tau} = \begin{bmatrix} \mathbf{I} & \rho\mathbf{U} & \rho^2\mathbf{U}^2 & \cdots & \rho^{\tau-1}\mathbf{U}^{\tau-1} \\ \rho\mathbf{U}^* & \rho^2\mathbf{I} & \rho^3\mathbf{U} & & \vdots \\ \rho^2\mathbf{U}^{*2} & \rho^3\mathbf{U}^* & \rho^4\mathbf{I} & & \\ \vdots & & & \ddots & \\ \rho^{\tau-1}\mathbf{U}^{*(\tau-1)} & & & & \rho^{2(\tau-1)}\mathbf{I} \end{bmatrix}$$

Finally, we let $\lambda = 1 + \rho^2 + \rho^4 + \cdots + \rho^{2(\tau-1)} = \begin{cases} \tau, & \text{if } \rho = 1 \\ \frac{1-\rho^{2\tau}}{1-\rho^2}, & \text{if } \rho < 1 \end{cases}$.

With the ultimate goal of computing the eigen-decomposition of $\mathbf{Q}_\rho$, we first characterize the eigenvalues and eigenvectors of $\mathbf{X}_\rho$. For the eigenvalues we first observe the following:

**Lemma IV.1.** *The matrix $\mathbf{X}_\rho$ satisfies $\mathbf{X}_\rho^2 = \lambda \mathbf{X}_\rho$.*

*Proof.* Multiplying the $i$-th row of $\mathbf{X}_\rho$ with its $j$-th column, we obtain:

$$\sum_{k=0}^{\tau-1} \rho^{i+k-1} \mathbf{U}^{-(i-k-1)} \cdot \rho^{j+k-1} \mathbf{U}^{j-k-1} = \left( \sum_{k=0}^{\tau-1} \rho^{2k} \right) \rho^{i+j-2} \mathbf{U}^{j-i} = \lambda \rho^{i+j-2} \mathbf{U}^{j-i}.$$

This proves the desired equality. $\qquad\square$

As a result, we can now fully characterize the *eigenvalues* of $\mathbf{X}_\rho$.

**Corollary IV.2.** *The eigenvalues of $\mathbf{X}_\rho$ are either $0$ or $\lambda$. Moreover, the multiplicity of the eigenvalue $\lambda$ is $n$.*

*Proof.* Let $\xi$ be an eigenvector of $\mathbf{X}_\rho$ with eigenvalue $\kappa$. Then $\mathbf{X}_\rho \xi = \kappa \xi$, and thus

$$\mathbf{X}_\rho^2 \xi = \kappa^2 \xi = (\lambda \mathbf{X}_\rho) \xi = \kappa \lambda \xi.$$

This implies that either $\kappa = 0$ or $\lambda$.

Moreover, $Tr(\mathbf{X}_\rho) = n(1 + \rho^2 + \cdots + \rho^{2(\tau-1)}) = n\lambda$. But the trace also equals the sum of all the eigenvalues. Since the eigenvalues of $\mathbf{X}_\rho$ can only be $0$ or $\lambda$, we conclude that the multiplicity of the eigenvalue $\lambda$ is precisely $n$. $\qquad\square$

We now turn to characterize the non-zero *eigenvectors* of $\mathbf{X}_\rho$. We will then use these vectors to compute the motif matrix of $\mathbf{Q}_\rho$. Recall $\xi_j$ are the eigenvectors of $\mathbf{U}$ with the corresponding eigenvalues $s_j$. For each $j$, define

$$\hat{\xi}_j = \begin{bmatrix} \xi_j \\ \rho s_j^{-1} \xi_j \\ \vdots \\ \rho^{\tau-1} s_j^{-(\tau-1)} \xi_j \end{bmatrix} \in \mathbb{C}^{n\tau}$$

**Lemma IV.3.** *Each $\hat{\xi}_j$ is an eigenvector of $\mathbf{X}_\rho$ with eigenvalue $\lambda$. Moreover, $\left\{ \frac{1}{\sqrt{\lambda}} \hat{\xi}_j \right\}_{j=1}^{n}$ is an orthonormal set.*

*Proof.* First, since $\mathbf{U}$ is unitary, we have $\mathbf{U}\xi_j = s_j\xi_j$ and $\mathbf{U}^*\xi_j = s_j^{-1}\xi_j$. Therefore, $\mathbf{U}^k\xi_j = s_j^k\xi_j$ and $\mathbf{U}^{*k}\xi_j = s_j^{-k}\xi_j$ for all non-negative integers $k$, or, stated more compactly, $\mathbf{U}^k\xi_j = s_j^k\xi_j$ for all integers $k$.

Expanding $\mathbf{X}_\rho\hat{\xi}_j$, we observe that its $i$-th block entry equals:

$$\sum_{k=0}^{n-1} \rho^{i+k-1}\mathbf{U}^{-(i-k-1)}\rho^k s_j^{-k}\xi_j = \sum_{k=0}^{n-1} \rho^{i+2k-1}s_j^{-i+1}\xi_j = \left(\sum_{k=0}^{n-1}\rho^{2k}\right)\rho^{i-1}s_j^{-(i-1)}\xi_j = \lambda\rho^{i-1}s_j^{-(i-1)}\xi_j.$$

This is precisely $\lambda$ times the $i$-th block entry of $\hat{\xi}_j$. This proves that $\mathbf{X}_\rho\hat{\xi}_j = \lambda\hat{\xi}_j$.

Since each $\xi_j$ is a unit vector and each $|s_j| = 1$, we have that $\|\hat{\xi}_j\|^2 = \sum_{k=0}^{n-1}\rho^{2k} = \lambda$. Hence, $\frac{1}{\sqrt{\lambda}}\hat{\xi}_j$ is a vector of norm 1. Finally, eigenvectors $\{\xi_j\}$ of any unitary matrix form an orthonormal basis, so $\langle\xi_i, \xi_j\rangle = \delta_{ij}$. This implies:

$$\langle\hat{\xi}_i, \hat{\xi}_j\rangle = \sum_{k=0}^{n-1}\rho^{2k}s_i^{-k}(\overline{s_j})^{-k}\langle\xi_i, \xi_j\rangle,$$

which is 0 when $i \neq j$. This proves that $\{\hat{\xi}_j\}$, $j = 1, \ldots, n$, are pairwise orthogonal. $\qquad\square$

Since all the other eigenvalues of $\mathbf{X}_\rho$ are 0, we immediately have the desired eigen-decomposition of the matrix $\mathbf{X}_\rho$ as:

$$\mathbf{X}_\rho = \frac{1}{\lambda}\sum_{j=1}^n \hat{\xi}_j\hat{\xi}_j^*$$

Recall an $n \times n$ full-cycle permutation matrix is given by:

$$\mathbf{C} := \begin{bmatrix} 0 & \cdots & \cdots & \cdots & 1 \\ 1 & 0 & & \cdots & 0 \\ 0 & 1 & 0 & \vdots & \vdots \\ \vdots & 0 & \ddots & 0 & \vdots \\ 0 & \vdots & 0 & 1 & 0 \end{bmatrix}$$

For the rest of the section, we set $\rho = 1$ **and $\mathbf{W} = \mathbf{C}$ is a full-cycle permutation** and we will now compute eigen-decomposition of the metric tensor $\mathbf{Q} = \mathbf{Q}_1$ of the corresponding reservoir kernel.

From elementary matrix analysis, we know the eigenvalues of a full-cycle permutation $\mathbf{C}$ are precisely the $n$-th root of unities $\{\omega_j = e^{\frac{2\pi ij}{n}}, 1 \leq j \leq n\}$. Its normalized eigenvectors for

each eigenvalue $\omega_j$ is given by the Fourier basis:

$$\xi_j = \frac{1}{\sqrt{n}} \begin{bmatrix} 1 \\ \omega_j \\ \omega_j^2 \\ \vdots \\ \omega_j^{n-1} \end{bmatrix} \in \mathbb{C}^n, \quad j = 1, \ldots, n. \tag{IV.1}$$

We will now compute the eigenvalues and eigenvectors of $\mathbf{Q}$ explicitly. First, we see that

$$\mathbf{Q} = \hat{\mathbf{w}}^* \mathbf{X}_1 \hat{\mathbf{w}},$$

where $\mathbf{X}_1 = \frac{1}{n} \sum_{j=1}^n \hat{\xi}_j \hat{\xi}_j^*$. Define $\xi_j^* \mathbf{w} = \langle \mathbf{w}, \xi_j \rangle = d_j$.

While, following Remark II.4, we focused on the case where $\tau = n$, the results of this section up to this point (Lemma IV.1, Corollary IV.2, and Lemma IV.3) hold for general $\tau$.

**Theorem IV.4.** *Consider a linear SCR system $R = (\mathbf{C}, \mathbf{w}, h)$ over $\mathbb{C}$ with dimensions $(n, 1, d)$ and with the full cycle permutation matrix $\mathbf{C}$ as its dynamical coupling. Let $\mathbf{Q}$ denote metric tensor of the reservoir kernel under the past horizon $\tau = n$. Then, the eigenvectors $\xi_j$, $j = 1, 2, ..., n$ of $\mathbf{C}$ (eq. (IV.1)) are also eigenvectors of $\mathbf{Q}$ with the corresponding eigenvalues equal to the squared projections of the input coupling vector $\mathbf{w}$ onto $\xi_j$, $|d_j|^2 = |\xi_j^* \mathbf{w}|^2$. In other words, $R$ has motifs $\xi_j$, with motif weights $|d_j| = |\xi_j^* \mathbf{w}| = |\langle \mathbf{w}, \xi_j \rangle|$, $j = 1, 2, ..., n$.*

*Proof.* Define an $n \times \tau$ matrix,

$$\mathbf{A} = \frac{1}{\sqrt{n}} \begin{bmatrix} \hat{\xi}_1^* \\ \hat{\xi}_2^* \\ \vdots \\ \hat{\xi}_n^* \end{bmatrix} \hat{\mathbf{w}}$$

Then $\mathbf{Q} = \mathbf{A}^* \mathbf{A}$, and we have:

$$\mathbf{A} = \begin{bmatrix} d_1 & \omega_1^{-1} d_1 & \cdots & \omega_1^{-(\tau-1)} d_1 \\ d_2 & \omega_2^{-1} d_2 & \cdots & \omega_2^{-(\tau-1)} d_2 \\ \vdots & \vdots & \vdots & \vdots \\ d_n & \omega_n^{-1} d_n & \cdots & \omega_n^{-(\tau-1)} d_n \end{bmatrix} = \mathbf{D} \cdot \mathbf{F}$$

Here, by the assumption that $\tau = n$, we can define two $n \times n$ matrices $\mathbf{D} = \mathrm{diag}\{d_1, d_2, \cdots, d_n\}$ and $\mathbf{F} = [\omega_i^{-(j-1)}]$. Note that $\mathbf{F}^*$ is the discrete Fourier transform matrix, so we know

$\mathbf{F}^*\mathbf{F} = \mathbf{F}\mathbf{F}^* = n\mathbf{I}$. Therefore,

$$\mathbf{Q} = \left(\frac{1}{\sqrt{n}}\mathbf{F}^*\right)\operatorname{diag}\{|d_1|^2, |d_2|^2, \cdots, |d_n|^2\}\left(\frac{1}{\sqrt{n}}\mathbf{F}\right).$$

Here, $\left(\frac{1}{\sqrt{n}}\mathbf{F}\right)$ is a unitary matrix. The $j$-th column of $\left(\frac{1}{\sqrt{n}}\mathbf{F}^*\right)$, which is precisely $\xi_j$, is the eigenvector for $|d_j|^2$ $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

**Remark IV.5.** When $\tau \neq n$, we can still reach a similar decomposition as above, but the matrix $\mathbf{F} = \left[\omega_i^{-(j-1)}\right]$ is an $n \times \tau$ matrix. When $\tau$ is not an integer multiple of $n$ in particular, its rows may not be orthogonal to each other in general, so it is hard to characterize the eigenvectors of $\mathbf{Q}$. However, when $\tau$ is an integer multiple of $n$, the rows of $\mathbf{F}$ are orthogonal to each other. In this case, the eigenvectors of $\mathbf{Q}$ is given by the $n$ of these $\tau$ dimensional vectors $(\omega_i^{-j})_{j=0}^{\tau-1}$.

As a result, in the case when $\rho = 1$ and the matrix $\mathbf{W} = \mathbf{C}$ is a full-cycle permutation, the motif matrix for $\mathbf{Q}$ consists of $\xi_j$, which is precisely the discrete Fourier transform matrix. In other words, the columns of the motif matrix are precisely the Fourier basis.

**Remark IV.6.** Notice that $|d_j|^2 = |d_{n-j}|^2$, so that the eigenvalues and eigenvectors of the motif matrix come in pairs. Here, $\xi_j$ and $\xi_{n-j} = \overline{\xi_j}$ share the same eigenvalue $|d_j|^2 = |d_{n-j}|^2$.

# V.  MOTIFS OF ORTHOGONAL DYNAMICS WITH UNIT SPECTRAL RADIUS IN REAL DOMAIN

In this section we return to the real domain, i.e. $\mathbb{K} = \mathbb{R}$ and show that at unit spectral radius, the motifs of SCR consists of a fixed number of symmetric and skew symmetric vectors. As in[FLT24], we begin by deriving properties of the motif space of linear reservoir systems with orthogonal dynamical coupling and then move on to the special case of cyclic permutation.

Let $\mathbf{W} \in \mathbb{M}_{n \times n}(\mathbb{R})$ be the dynamical coupling matrix of a reservoir system $R = (\mathbf{W}, \mathbf{w}, h)$ over $\mathbb{R}$. Suppose $\mathbf{W}$ is orthogonal with spectral radius $\rho = 1$. We now show that the matrix corresponding to the reservoir kernel $\mathbf{Q} := \mathbf{Q}_{\rho=1}$ is Toeplitz.

$\mathbf{Q}$ is Toeplitz if and only if $\mathbf{Q}_{ij} = \mathbf{Q}_{i+1,j+1}$ for all $i, j = 1, \ldots, n-1$. By construction of $\mathbf{Q}$, this is satisfied if and only if:

$$\mathbf{w}^\top \left(\mathbf{W}^\top\right)^{i-1} \mathbf{W}^{j-1}\mathbf{w} = \mathbf{w}^\top \left(\mathbf{W}^\top\right)^i \mathbf{W}^j\mathbf{w}.$$

Now, orthogonality of $\mathbf{W}$ implies $\mathbf{W}^\top\mathbf{W} = \mathbf{I}$. Without loss of generality assume that $i \geq j$. Then

$$(\mathbf{W}^\top)^{i-1}\mathbf{W}^{j-1} = (\mathbf{W}^\top)^{i-j} = (\mathbf{W}^\top)^i\,\mathbf{W}^j,$$

showing that $\mathbf{Q}$ is indeed Toeplitz.

Let $\mathbf{J}$ denote the exchange matrix with 1 on the antidiagonal and 0 everywhere else:

$$\mathbf{J} := \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 1 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

**Lemma V.1.** *Symmetric Toeplitz square matrices are centrosymmetric.*

*Proof.* Let $\mathbf{T}$ denote a symmetric Toeplitz $\tau \times \tau$ matrix. Let $\{t_k\}_{k=0}^{\tau-1}$ denote the generating sequence of $\mathbf{T}$ such that $\mathbf{T}$ is expressed as:

$$\mathbf{T} = \begin{bmatrix} t_0 & t_1 & \cdots & t_{\tau-2} & t_{\tau-1} \\ t_1 & t_0 & \cdots & t_{\tau-3} & t_{\tau-2} \\ t_2 & t_1 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \cdots & t_0 & t_1 \\ t_{\tau-1} & t_{\tau-2} & \cdots & t_1 & t_0 \end{bmatrix}.$$

Consider:

$$\mathbf{T}_{\tau+1-i,\tau+1-j} = t_{\tau+1-i-\tau-1+j}$$

$$= t_{j-i} = t_{i=j} = \mathbf{T}_{ij}.$$

The shows that $\mathbf{T}$ satisfies the definition of centro-symmetry. The second last equality is given by symmetry of $\mathbf{T}$ and the last equality is due to $\mathbf{T}$ being Toeplitz.

$\square$

**Corollary V.2.** *Let $\mathbf{W}$ be the dynamical coupling matrix of a reservoir system $R = (\mathbf{W}, \mathbf{w}, h)$. If $\mathbf{W}$ is orthogonal with spectral radius $\rho = 1$, then $\mathbf{Q}$ is symmetric centrosymmetric.*

*Proof.* $\mathbf{Q}$ is symmetric by construction. Moreover, when $\mathbf{W}$ is orthogonal with spectral radius 1, $\mathbf{Q}$ is Toeplitz and hence centrosymmetric by lemma V.1. $\square$

**Definition V.3.** Let $\mathbb{K}$ be an arbitrary field. A $\tau$-dimensional vector $v$ over $\mathbb{K}$ is **symmetric** if:

$$\mathbf{J}v = v.$$

Similarly, $v$ is **skew-symmetric** if:

$$\mathbf{J}v = -v.$$

By[CB76]: symmetric centrosymmetric matrices of order $\mathbf{n}$ admits an orthonormal basis of eigenvectors with:

1. $\lceil \frac{n}{2} \rceil$ symmetric eigenvectors,

2. $\lfloor \frac{n}{2} \rfloor$ skew-symmetric eigenvectors

## A. Motif structure of cyclic permutation dynamics with unit spectral radius

Following Remark II.4, we focus on the case where $\tau = n$ below. In the special case when the dynamical coupling is a cyclic permutation, the corresponding matrix $\mathbf{Q}$ is *circulant*. A circulant matrix is a specific type of Toeplitz matrix in which each row is a cyclic shift of the previous row, with elements shifted one position to the right.

The cyclic permutation can equivalently be expressed as the linear map:

$$\mathbf{C} : \mathbb{R}^n \to \mathbb{R}^n$$
$$(a_k)_{k=1}^n \mapsto (b_k)_{k=1}^n := \left( a_{k-1 \pmod n} \right)_{k=1}^n \tag{V.1}$$

Then by construction we have:

**C.1** $\mathbf{C}^\alpha \cdot \mathbf{C}^\beta = \mathbf{C}^{\alpha+\beta \pmod n}$ .

**C.2** $\left( \mathbf{C}^\top \right)^i = \mathbf{C}^{n-i \pmod n}$

**Theorem V.4.** *Let $\mathbf{W} \in \mathbb{M}_{n \times n}(\mathbb{R})$ be the dynamical coupling matrix of a reservoir system $R = (\mathbf{W}, \mathbf{w}, h)$. If $\mathbf{W}$ is a cyclic permutation with spectral radius $\rho = 1$. Then the metric tensor $\mathbf{Q}_1 := \mathbf{Q}_{\rho=1}$ of the reservoir kernel is circulant. Moreover, there exists an orthogonal basis such that $R$ has $\lceil \frac{n}{2} \rceil$ symmetric motifs and $\lfloor \frac{n}{2} \rfloor$ skew-symmetric motifs.*

*Proof.* Let $r_i$ denote the $i^{th}$ row of the matrix $\mathbf{Q} := \mathbf{Q}_1$. It suffices to show that $r_i = \mathbf{C}^{i-1} \cdot r_1$. By construction, for $j = 1, \ldots, n$:

$$(r_1)_j = \mathbf{w}^\top \mathbf{C}^{j-1} \mathbf{w}$$

$$(r_i)_j = \mathbf{w}^\top \left(\mathbf{C}^\top\right)^{i-1} \mathbf{C}^{j-1} \mathbf{w}$$

By **C.1** and **C.2**, the $j^{th}$ component of $r_i$ can be rewritten as:

$$(r_i)_j = \mathbf{w}^\top \left(\mathbf{C}^\top\right)^{i-1} \mathbf{C}^{j-1} \mathbf{w}$$
$$= \mathbf{w}^\top \mathbf{C}^{n-i+1} \mathbf{C}^{j-1} \mathbf{w} = \mathbf{w}^\top \mathbf{C}^{j-i \pmod n} \mathbf{w} \tag{V.2}$$

By Equation V.1 and **C.1**, the $j^{th}$ component of $\mathbf{C}^{i-1} \cdot r_1$ can be written as:

$$\mathbf{C}^{i-1} \left(r_1\right)_j = \mathbf{w}^\top \mathbf{C}^{j-(i-1)-1 \pmod n} \mathbf{w} = \mathbf{w}^\top \mathbf{C}^{j-i \pmod n} \mathbf{w}$$

which coincides with Equation V.2 and thus $\mathbf{Q}_1$ is circulant.

Circulant matrices are Toeplitz. Since $\mathbf{Q}_1$ is symmetric circulant, it is symmetric centrosymmetric of order $n$. Thus by[CB76], $\mathbf{Q}_1$ admits an orthonormal basis spanning the space of space of motif consisting of $\lceil \frac{n}{2} \rceil$ symmetric eigenvectors of $\mathbf{Q}_1$ and $\lfloor \frac{n}{2} \rfloor$ skew-symmetric eigenvectors of $\mathbf{Q}_1$. $\qquad\square$

# VI. LINEAR SCR AT UNIT SPECTRAL RADIUS IS FOURIER TRANSFORM

This section bridges the theoretical results and numerical experiments by explicitly constructing the Fourier basis matrix corresponding to a linear SCR over $\mathbb{R}$. We present numerical simulations to validate our theoretical findings and outline the components of the numerical experiments in the final section.

Combining the results of the previous sections, we conclude that: At unit spectral radius, the motifs of linear SCR over $\mathbb{R}$ with $n$ neurons of look back window $\tau = n$ are:

**R.1** Harmonic, i.e. they correspond to the Fourier basis (Theorem IV.4).

**R.2** There are $\lceil \frac{n}{2} \rceil$ symmetric vectors (cosines) and $\lfloor \frac{n}{2} \rfloor$ skew-symmetric vectors (sines) with positive eigenvalues. (Theorem V.4).

We now demonstrate the explicit construction of the Fourier basis matrix corresponding to the motifs of a linear SCR at unit spectral radius.

For a linear reservoir system over $\mathbb{R}$, the motifs must also be real. By conditions **R.1** and **R.2**, each motif must then correspond to either the real *or* imaginary parts of the Fourier basis. More precisely, for $k = 0, \ldots, \lceil \frac{n}{2} \rceil$ and $i = 0, \ldots, \tau - 1$ (The ceiling function accounts for Theorem V.4.), the real Fourier basis matrix $\mathbf{F} = \mathbf{F}[i, j] \in \mathbb{M}_{\tau \times n}(\mathbb{R})$ is defined as

$$
\begin{aligned}
\mathbf{F}[i, 2k] &= \sqrt{\frac{2}{\tau}} \cos \frac{2\pi k i}{\tau} \quad \ldots \quad \text{Even columns} \\
\mathbf{F}[i, 2k + 1] &= \sqrt{\frac{2}{\tau}} \sin \frac{2\pi k i}{\tau} \quad \ldots \quad \text{Odd columns}
\end{aligned}
\tag{VI.1}
$$

We claim that the motifs of a linear SCR are exactly the first $n$ columns of $\mathbf{F}$.

In Figure 3 and Figure 4, we present the Fourier analysis of motifs of the SCR at unit spectral radius when $\tau = n = 97$.

In Figure 4, we present examples of 4 randomly chosen motifs and their corresponding Fourier basis in $\mathbf{F}$. While some motifs and their corresponding Fourier basis are shifted by a fixed phase of $\pi$ or reflected over the $x$-axis, their Fourier spectra align as illustrated in Figure 3. Furthermore, Figure 3, we observe that the eigenvalues come in pairs, as discussed in IV.6. The motifs may not be strictly harmonic in the classical sense due to the coarseness of the discretization grid (non-integer division of period) (See Remark VI.1 below).
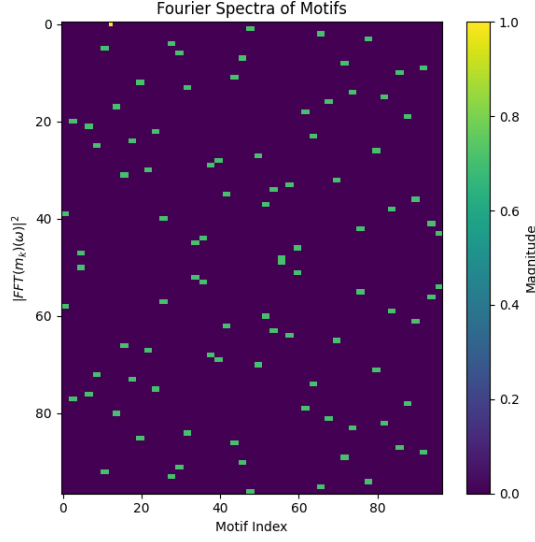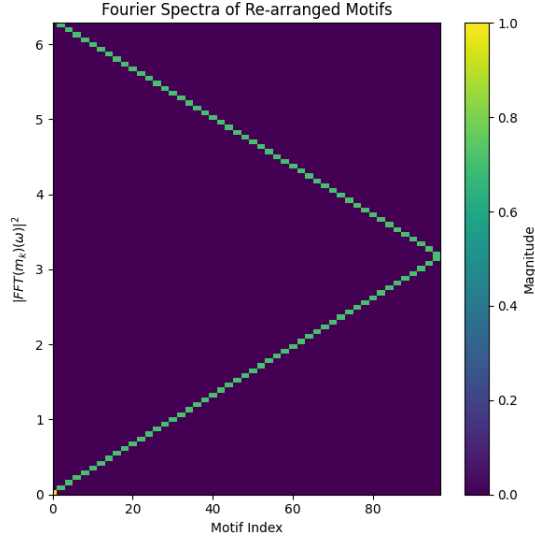
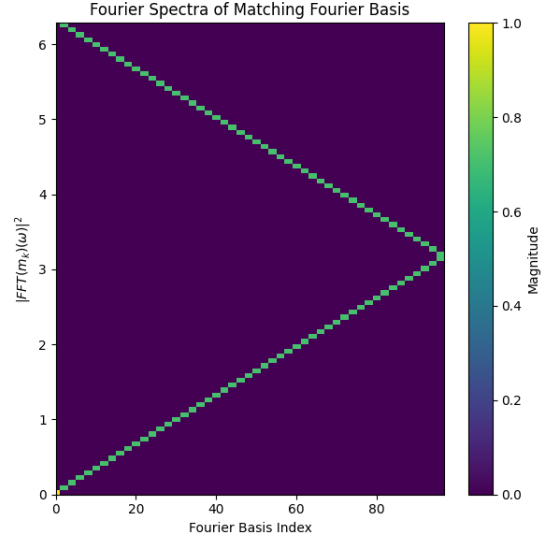(a) Column-wise DFT of motifs of linear SCR with unit spectral radius.
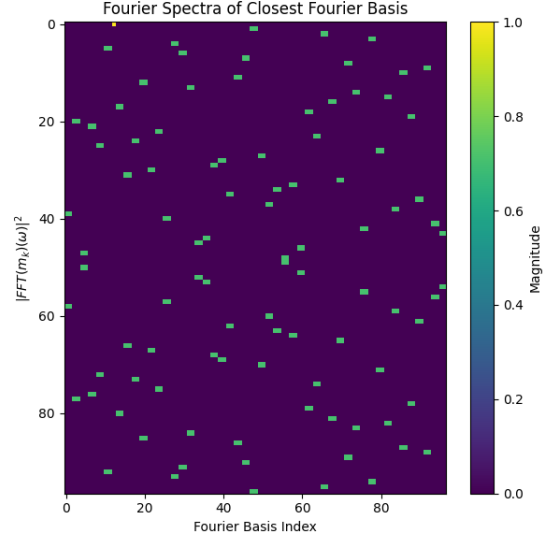
(b) Column-wise DFT of first $n$ columns of **F**.

(c) Columns of Fig 3a rearranged according to largest Fourier spectra.

(d) Columns of Fig 3b rearranged with same indices as Fig 3c.

FIG. 3: Column-wise DFT of motifs of linear SCR with $\rho = 1$ and the column-wise DFT of **F**. The first row shows the Fourier spectra in the original form and the second row has their columns rearranged with the same shuffling indices.
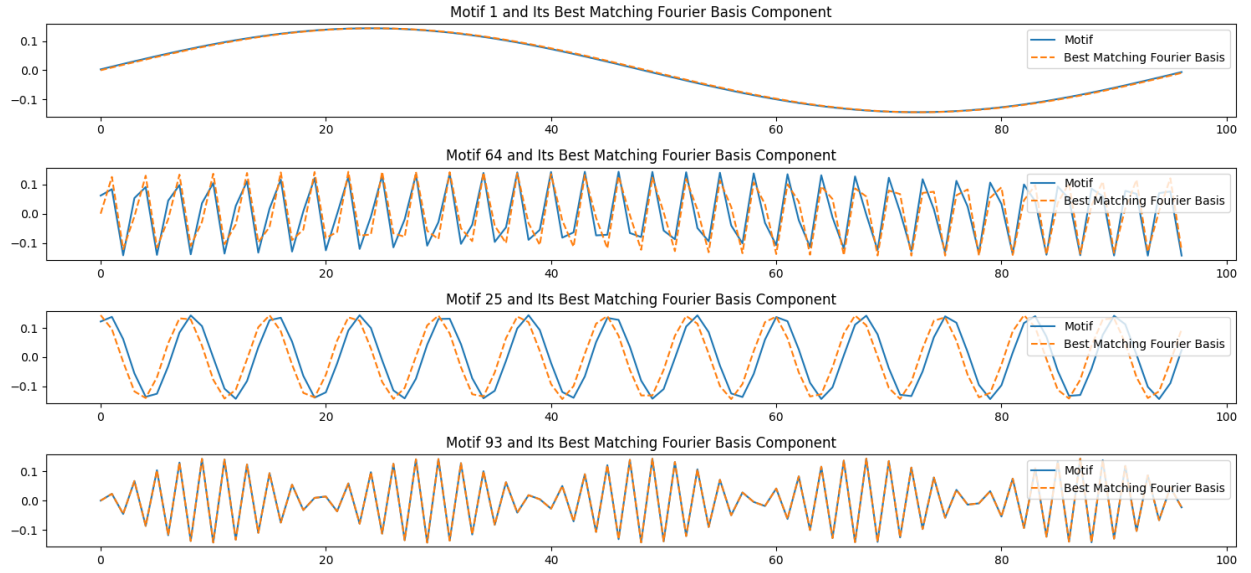
FIG. 4: Example of 4 randomly chosen motifs and their corresponding Fourier basis.
Notice some are off by a phase of $\pi$.

**Remark VI.1.** In practice, motifs represent harmonic functions sampled at specific frequencies. This explains why certain motifs may not visually appear as harmonic functions (e.g., motif 93 in Figure 4), yet their Fourier spectra reveal characteristics consistent with harmonic functions (see Figure 3). The effect of sampling is further demonstrated in Figure 5, which compares the 93rd Fourier basis generated according to Equations VI.1 at two different sampling sizes.

On the first plot the sine curve is being sampled a high frequency with 500 sampling points; whereas the sampling points are reduced to $97 = n = \tau$ on the second plot. On the third plot we see that they are generated by the exact same function but with different sampling frequency. Notice that the first plot appears to be harmonic but the second one does not.

Therefore, While this function may not visually resemble a harmonic function at the default sampling size of $n = 97$, increasing the sampling size to 500 makes the function progressively align with the visual properties of a harmonic function.

The reason the Fourier spectra align with those of harmonic functions in the computational process is due to the matching of sampling frequencies within the DFT algorithm.
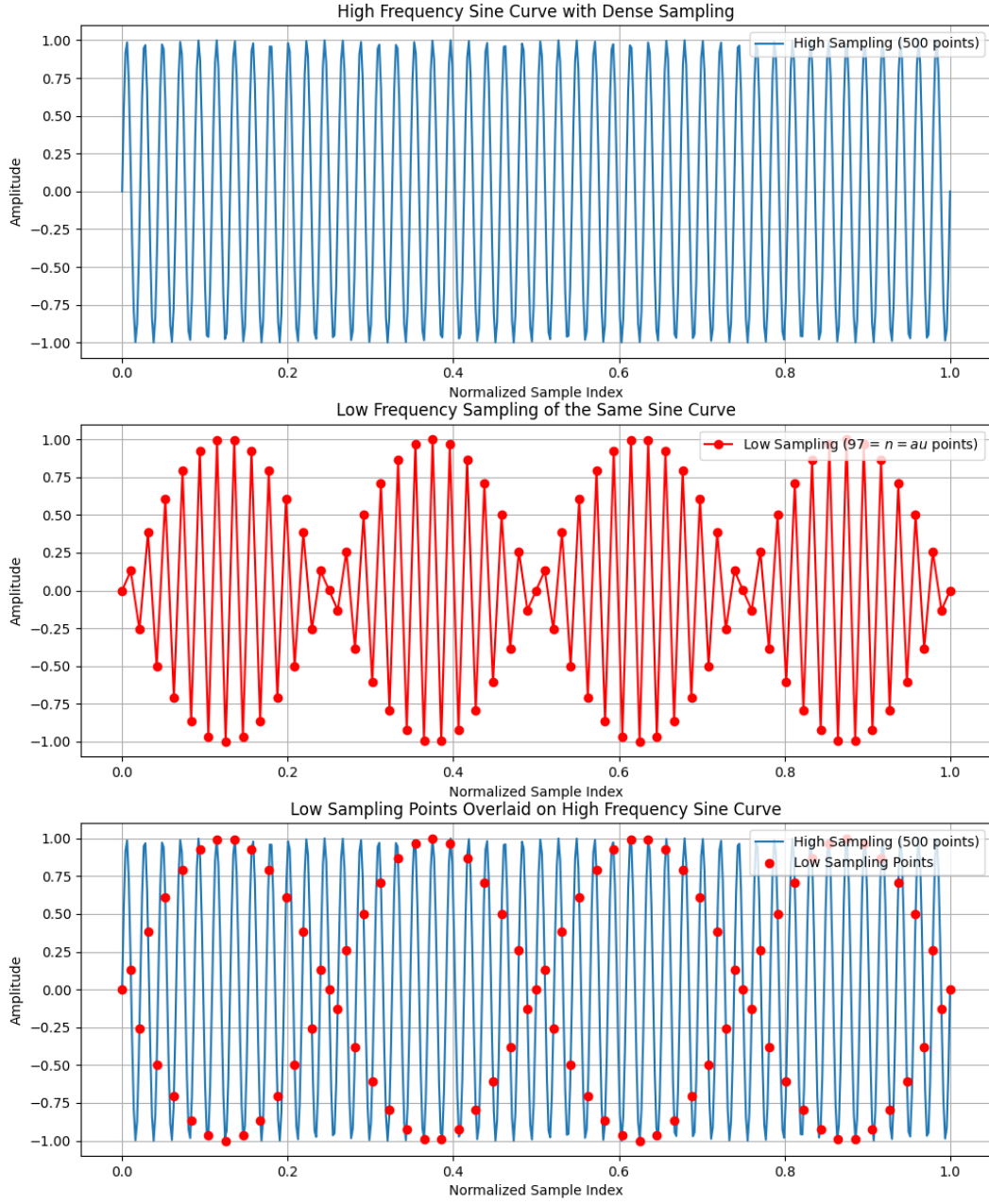
FIG. 5: Plots of Fourier basis number 93 constructed under Equations VI.1 sampled under two different frequencies.

## VII.   ILLUSTRATIVE EXAMPLES ON TIME SERIES FORECASTING

We conclude the paper with numerical experiments illustrating our findings. We compare time series forecasting results using the so-called Reservoir Motif Machines (RMM)[TFL24] and the linear SCR. RMM is a simple time series forecasting method based on the feature space representation of time series derived from linear reservoirs and has demonstrated remarkable predictive performance, even surpassing that of more complex transformer models on univariate time series forecasting tasks. A brief exposition of RMM is presented in Section II.

We use RMM because it allows us to explicitly define the feature space representation by imposing a set of motifs, rather than having the feature space representation implicitly defined as in classical ESNs, such as SCR. This provides an ideal platform to showcase our theoretical findings: that the feature space representation of the motif space of a linear SCR is the same as that defined by the Fourier basis matrix $\mathbf{F}$ in Equation VI.1.

For reproducibility of the experiments, all experiments are CPU-based and are performed on Apple M3 Max with 128GB of RAM. The source code and data of the numerical analysis is openly available at `https://github.com/Lampertos/motif_Fourier`.

We compare the prediction results on univariate time series forecasting across the following models:

1. Lin-RMM with SCR motifs and unit spectral radius,

2. Lin-RMM with Fourier basis motifs as discussed in Section VI, and

3. Linear SCR with unit spectral radius.

It is essential to emphasize that this experiment is not intended to showcase the predictive capability of the models, but rather to highlight the similarities in feature space representation between the linear SCR model over $\mathbb{R}$ and the Fourier basis motifs introduced in Section VI. Consequently, hyperparameters are kept constant throughout the experiments to underscore this feature space comparison.

The fixed set of hyperparameters for all experiments are as follows: $r_{\mathrm{in}} = 0.05$ for input weights, $n = 97$ for the number of reservoir neurons, and $\tau = 2n = 194$ for the look-back window length. All models are trained using ridge regression with a ridge coefficient of $10^{-3}$. The prediction horizon is set to 168.

## A.   Datasets

To facilitate the comparison of our results with the state-of-the art, we have used the same datasets and the same experimental protocols used in the recent time series forecasting papers[ZZP+20]. Those are briefly described below for the sake of completeness.

*a.   ETT*   The Electricity Transformer Temperature dataset[Note]consists of measurements of oil temperature and six external power-load features from transformers in two regions of China.  The data was recorded for two years, and measurements are provided either hourly (indicated by 'h') or every 15 minutes (indicated by 'm').  In this paper we used `oil temperature` of the `ETTh1`, `ETTm1` dataset for univariate prediction with train/validation/test split being 12/4/4 months.

*b.   ECL*   The Electricity Consuming Load[Notf] consists of hourly measurements of electricity consumption in kWh for 321 Portuguese clients during two years.  In this paper we used client `MT 320` for univariate prediction.  The train/validation/test split is 15/3/4 months.

*c.   Weather*   The Local Climatological Data (LCD) dataset[Notg] consists of hourly measurements of climatological observations for 1600 weather stations across the US during four years. The dataset was used for univariate prediction of the `Wet Bulb Celcius` variable.

## B.   Discussion

From Figure 6 and Figure 7 we see that there's virtually no difference between Lin-RMM with unit spectral radius SCR motifs and Lin-RMM with Fourier basis. This confirms our observations in the previous sections. Furthermore Figure 6 also confirms that Lin-RMM with unit spectral radius SCR motifs (red bar) has superior performance against classical SCR with unit spectral radius, this affirms the studies in[TFL24].

Amongst RMM's we compare the MSE loss in Figure 7.  Notice that the difference between the MSE loss between RMM under Fourier motifs and SCR motifs are around $1e-13$ which is negligible compared to the MSE loss of the models against standardized input signals.
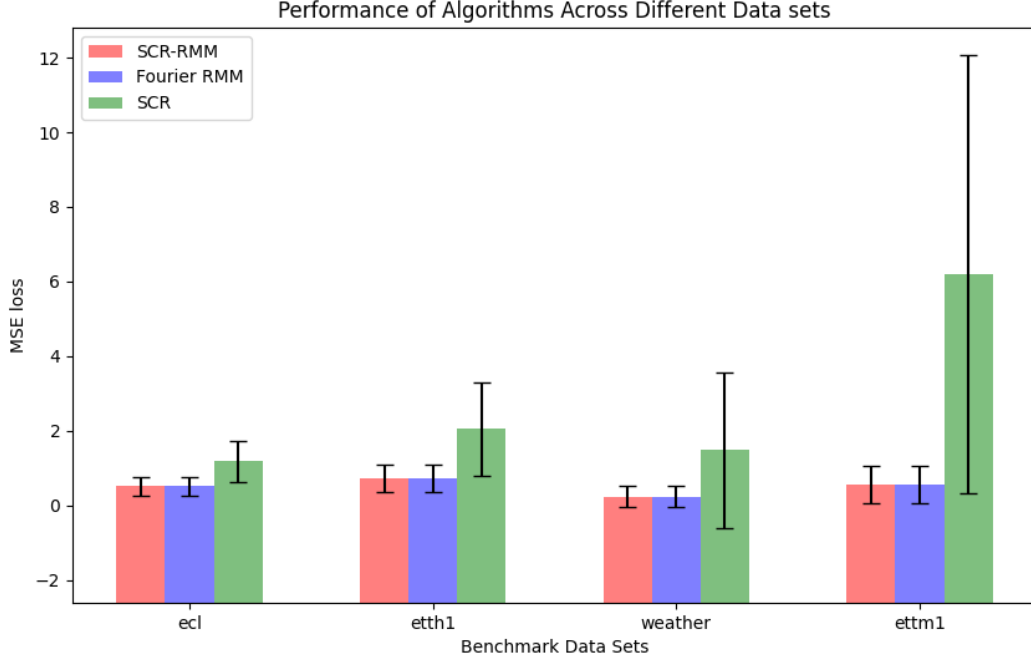
FIG. 6: Statistics of MSE loss of Fourier RMM, unit SCR RMM and unit SCR across different data sets with fixed prediction horizon 168.
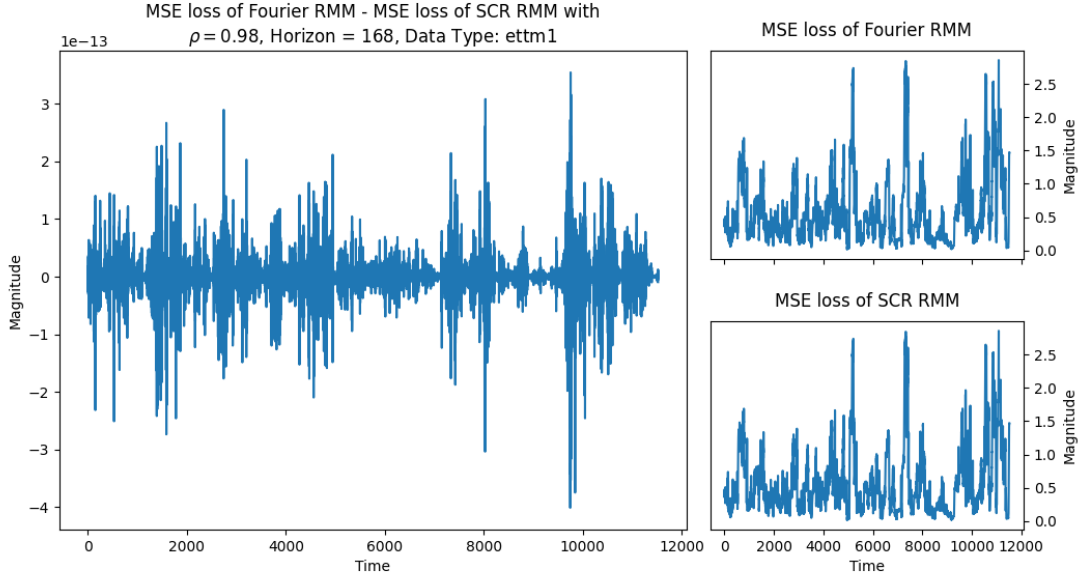


FIG. 7: Comparison of MSE loss between Fourier RMM and unit SCR RMM in ETTm2.

## VIII.   CONCLUSION

Linear recurrent neural networks (RNN), such as Echo State Networks (ESN) can be thought of as providing feature representations of the input-driving time series in their state space[Tin20,GGO22]. By endowing the feature (state) space with the canonical dot-product, one can "reverse engineer" the corresponding inner product in the space of time series (time series kernel) that is defined through the the dot product in the RNN state space[Tin20]. This in turn helps to shed light on the inner representational schemes employed by the RNN to process the input-driving time series. In particular, the induced (semi-)inner product in the time series space can be theoretically analyzed through eigen-decomposition of the corresponding metric tensor. The eigenvectors (time series motifs) define the projection basis of the induced feature space and the (decay of) eigenvalues its dominant subspace and effective dimensionality. The induced time series kernels by the Simple Cycle Reservoir (SCR) models were shown to be superior (in terms of dimensionality, motif variability, and memory) to several alternative ESN constructions[Tin20].

In this paper we have shown a rather surprising result: When SCR is constructed at the edge of stability, the basis of its induced time series feature space correspond to the well known and widely used basis for signal decomposition - namely the Fourier basis.

This insight also explains the reduction in relative area covered by Fourier representations of SCR motifs observed by[Tin20] at the edge of stability. Our results imply that the feature space representation of a linear SCR at unit spectral radius effectively performs a weighted projection onto the Fourier basis.

This observation is supported by numerical experiments, in which we compared the time series forecasting accuracy of Lin-RMM with the motif space defined by a linear SCR at unit spectral radius and the Fourier basis, respectively.

# REFERENCES

[BVdSBV22] Ian Bauwens, Guy Van der Sande, Peter Bienstman, and Guy Verschaffelt. Using photonic reservoirs as preprocessors for deep neural networks. *Frontiers in Physics*, 10:1051941, 2022.

[CB76] A. Cantoni and P. Butler. Eigenvalues and eigenvectors of symmetric centrosymmetric matrices. *Linear Algebra and its Applications*, 13(3):275–288, 1976.

[FLT24] Robert Simon Fong, Boyu Li, and Peter Tiňo. Universality of real minimal complexity reservoir. *arXiv preprint arXiv:2408.08071*, 2024.

[GGO22] Lukas Gonon, Lyudmila Grigoryeva, and Juan-Pablo Ortega. Reservoir kernels and volterra series. *arXiv preprint arXiv:2212.14641*, 2022.

[HVK+20] Krishan Harkhoe, Guy Verschaffelt, Andrew Katumba, Peter Bienstman, and Guy Van der Sande. Demonstrating delay-based reservoir computing using a compact photonic integrated chip. *Optics express*, 28(3):3086–3096, 2020.

[Jae01] H. Jaeger. The "echo state" approach to analysing and training recurrent neural networks. Technical report gmd report 148, German National Research Center for Information Technology, 2001.

[LBFM+17] Laurent Larger, Antonio Baylón-Fuentes, Romain Martinenghi, Vladimir S Udaltsov, Yanne K Chembo, and Maxime Jacquot. High-speed photonic reservoir computing using a time-delay-based architecture: Million words per second classification. *Physical Review X*, 7(1):011015, 2017.

[LFT24] Boyu Li, Robert Simon Fong, and Peter Tiňo. Simple Cycle Reservoirs are Universal. *Journal of Machine Learning Research*, 25(158):1–28, 2024.

[LJ09] M. Lukosevicius and H. Jaeger. Reservoir computing approaches to recurrent neural network training. *Computer Science Review*, 3(3):127–149, 2009.

[MNM02] W. Maass, T. Natschlager, and H. Markram. Real-time computing without stable states: a new framework for neural computation based on perturbations. *Neural Computation*, 14(11):2531–2560, 2002.

[RT10] Ali Rodan and Peter Tiňo. Minimum complexity echo state network. *IEEE transactions on neural networks*, 22(1):131–144, 2010.

[TD01] P. Tiňo and G. Dorffner. Predicting the future of discrete sequences from fractal representations of the past. *Machine Learning*, 45(2):187–218, 2001.

[TFL24] Peter Tiňo, Robert Simon Fong, and Roberto Fabio Leonarduzzi. Predictive modeling in the reservoir kernel motif space. In *2024 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8, 2024.

[Tin20] Peter Tino. Dynamical systems as temporal feature spaces. *J. Mach. Learn. Res.*, 21:44–1, 2020.

[ZZP+20] Haoyi Zhou, Shanghang Zhang, Jieqi Peng, Shuai Zhang, Jianxin Li, Hui Xiong, and Wan Zhang. Informer: Beyond efficient transformer for long sequence time-series forecasting. In *AAAI Conference on Artificial Intelligence*, 2020.

[Nota] Also called left circular shift or cyclic permutation in the literature.

[Notb] We note that the assumption on the aperiodicity of the sign pattern in $V$ is not required for this study.

[Notc] In particular, $\rho = 0.9812798473475446$ is where the relative area peaks at Figure 2.

[Notd] This notion of "Representational richness" essentially refers to the extent to which motifs can capture a wide range of frequency components. In particular, the distribution of Fourier coefficients measures the frequency and spectrum that the motifs cover. At $\rho = 1$, motifs are harmonic (as we will demonstrate in Section IV), and therefore represented by just 2 Fourier coefficients. However, for $\rho < 1$ and close to the "phase transition" in the area covered by the motifs' Fourier coefficients, each motif has a multitude of Fourier coefficients. In this (very vague) sense, the set of such motifs is "representationally richer" than the set of harmonic motifs. At $\rho = 1$, while we get the **first** $n$ elements of Fourier basis, we may not get a sufficient span of frequencies to capture the input signal, and this is what we meant by the "loss of representational richness".

[Note] https://github.com/zhouhaoyi/ETDataset.

[Notf] https://archive.ics.uci.edu/dataset/321/
electricityloaddiagrams20112014.

[Notg] https://www.ncei.noaa.gov/data/local-climatological-data/.