COMISIÓN NACIONAL DE BANCOS Y SEGUROS

WORKING PAPER:

Dynamic financial processes identification using sparse regressive reservoir computers

AUTHORS

Fredy Vides
Idelfonso B. R. Nogueira
Evelyn Flores
Lendy Banegas

January 18, 2024

Dynamic financial processes identification using sparse regressive reservoir computers

Fredy Vides,[1, 2, a)] Idelfonso B. R. Nogueira,[3, b)] Evelyn Flores,[1, c)] and Lendy Banegas[1, d)]
[1)]*Department of Statistics and Research, National Commission of Banks and Insurance Companies of Honduras, Honduras*
[2)]*Scientific Computing Innovation Center, Universidad Nacional Autónoma de Honduras, Honduras*
[3)]*Department of Chemical Engineering, Norwegian University of Science and Technology, Norway*

(Dated: January 18, 2024)

In this document, we present key findings in structured matrix approximation theory, with applications to the regressive representation of dynamic financial processes. Initially, we explore a comprehensive approach involving generic nonlinear time delay embedding for time series data extracted from a financial or economic system under examination. Subsequently, we employ sparse least-squares and structured matrix approximation methods to discern approximate representations of the output coupling matrices. These representations play a pivotal role in establishing the regressive models corresponding to the recursive structures inherent in a given financial system. The document further introduces prototypical algorithms that leverage the aforementioned techniques. These algorithms are demonstrated through applications in approximate identification and predictive simulation of dynamic financial and economic processes, encompassing scenarios that may or may not exhibit chaotic behavior.

The intricate dynamics inherent in financial processes often pose challenges for accurate modeling and prediction. Nonetheless, the synergy of sparse representation techniques with Nonlinear Regressive Reservoir Computers (NRRCs) proves advantageous in modeling financial processes dynamics. Firstly, this approach excels in capturing the intricate nonlinear dynamics of financial data. NRRCs, adept at modeling complex relationships between input and output data, coupled with sparse representation, effectively identify the key dynamic components, ensuring more accurate and precise modeling of underlying dynamics. Secondly, the methodology promotes efficient data utilization. NRRCs, capable of learning from a relatively small dataset, align well with the limited scope and complexity of financial processes data. By pinpointing crucial variables, the approach enhances modeling efficiency, conserving time and resources. Thirdly, the approach exhibits flexibility and adaptability. NRRCs swiftly respond to changing conditions, making them ideal for the dynamic nature of financial processes. The amalgamation of NRRCs with sparse representation facilitates the identification of changes in the underlying structure, enabling prompt adjustments to the model. In conclusion, integrating sparse representation techniques with time series models employing nonlinear regressive reservoir computers yields several advantages for financial processes dynamics modeling. It ensures accurate modeling of complex dynamics, optimizes data utilization, and provides adaptability to evolving conditions.

I. INTRODUCTION

Regressive models and reservoir computers are robust computational tools for the identification and simulation of financial and economic systems[2]. In recent years, a new class of architectures, termed next-generation reservoir computers, has emerged[7]. In this study, we delve into the intrinsic network architecture associated with these reservoir computers, which significantly contribute to data dimensionality reduction. This architecture also facilitates the parametric identification processes by leveraging the matrix structural constraints induced by the network architecture. The document outlines key aspects of the theory and algorithms pertaining to the computation of specific types of regressive reservoir computers. The focus of this study is on reservoir computers, the architecture of which can be approximated by either linear or nonlinear regressive 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 regressive models, to the simulation of dynamic financial processes. 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 financial and economic dynamic models identification based on the ideas presented in §III and §IV has been developed and is available in[15].

Even though, the applications of the structure preserving function approximation technology developed as part of the work reported in this document can range from numerical modeling of cyber-physical systems[16], to climate simulation[10]. We will focus on applications to financial processes identification in this paper.

Financial processes have become complex systems where several dynamic entities constantly communicate and affect each other. Hence, the financial processes identification has become a critical aspect of modern finance. The identified

[a)]Electronic mail: fredy.vides@unah.edu.hn
[b)]Electronic mail: idelfonso.b.d.r.nogueira@ntnu.no
[c)]Electronic mail: evelyn.flores@cnbs.gob.hn
[d)]Electronic mail: lendy.banegas@cnbs.gob.hn

models can become a helpful tool for institutions to analyze and predict financial trends, manage risk, and make informed investment decisions (Bodie et al., 2014). However, the complexity and uncertainty of financial markets make these tasks challenging. Financial processes often exhibit nonlinear and complex behavior, which makes it difficult to model and identify the underlying dynamics (Cont, 2001). Traditional linear models may fail to capture the intricate relationships between variables, leading to inaccurate predictions and suboptimal decision-making.

Despite the challenges posed by the factors described above, data quality, and market efficiency, machine learning techniques offer promising solutions for improving the accuracy and utility of financial models. Machine learning has been used to identify the relationship between the key financial ratios that characterize a firm's financial position. For instance, Dixon, Klabjan, and Bang's[5] work applies deep learning to predict financial market movements. The authors use a classification approach to predict financial market movements. Their findings suggest that deep learning algorithms can provide valuable insights and predictions about financial market movements, outperforming traditional methods. Sirignano and Cont[13] propose a deep learning model to identify the dynamics of price formation of a high-frequency limit order book. Their model was able to capture universal features of price formation across different markets, highlighting the potential of machine learning to model complex financial systems.

Overall, the recent literature suggests that machine learning has significant potential in modeling financial data. These techniques are increasingly utilized to capture complex patterns, make accurate predictions, and optimize decision-making in the financial domain. However, it is still an open issue to be investigated. In this scenario, this work also contributes to the field of financial data identification by applying the proposed tools in this context leading to a better understanding of the underlying financial processes addressed here.

A prototypical algorithm for the computation of sparse structured recursive models based on the ideas presented in §III, is presented in §IV. Some numerical simulations of financial processes based on 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 $d_p(n)=n(n^p-1)/(n-1)+1$. 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 values[8] (§2.5.3) denoted by the expressions $s_j(A)$ for $j=1,\ldots,\min\{m,n\}$. We will write $\mathrm{rk}_\delta(A)$ to denote the number

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

For a nonzero matrix $A\in\mathbb{R}^{m\times n}$, the symbol A^+ will be used to denote the pseudoinverse[8] (§5.5.4) of A.

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

$$\mathbf{x}_L(t)=\begin{bmatrix}\mathbf{x}_L(t)[1]^\top & \mathbf{x}_L(t)[2]^\top & \cdots & \mathbf{x}_L(t)[n]^\top\end{bmatrix}^\top\in\mathbb{R}^{nL},$$

with

$$\mathbf{x}_L(t)[j]=\begin{bmatrix}x(t-L+1)[j]\\ x(t-L+2)[j]\\ \vdots\\ x(t-1)[j]\\ x(t)[j]\end{bmatrix}\in\mathbb{R}^L$$

for $1\leq j\leq n$, where $x(s)[j]$ 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 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 $\Pi_p:\mathbb{R}^n\to\mathbb{R}^{n^p}$ that is determined by the expression $\Pi_p(x):=x^{\otimes p}$, for each $x\in\mathbb{R}^n$. Given two matrices $X=[x_{i,j}]$, $Y=[y_{i,j}]$ in $\mathbb{R}^m$, we will write $X\odot Y$ to denote the operation corresponding to their Hadamard product $X\odot Y:=\left[x_{i,j}y_{i,j}\right]\in\mathbb{R}^m$.

For any matrix $A\in\mathbb{R}^{m\times n}$, we will denote by $\mathrm{colsp}(A)$ the columns space of the matrix A. Given a list $A_1,A_2,\ldots,A_m$ 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.

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 $\hat{X} \in \mathbb{R}^{n\times p}$ is an approximate sparse representation of X with respect to A, or a sparse representation of X for short, if $\|\hat{X}A - XA\|_F \leq C\delta$ for some $C > 0$ that does not depend on δ, and $\hat{X}$ has fewer nonzero entries than X.

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 $P^2 = P = P^\top$. 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. STRUCTURED DYNAMIC TRANSFORMATION MODEL IDENTIFICATION

Given two discrete-time dynamic systems determined by two time series $\{x(t)\}_{t\geq 1}$ and $\{y(t)\}_{t\geq 1}$, respectively. We will study the identification process of maps determined by the expression

$$y(t) = \mathscr{F}(x(t)) + r(t), \tag{III.1}$$

where $\{r(t)\}_{t\geq 1}$ denotes the sequence of residual errors determined for each $t \geq 1$ by $r(t) := \|x(t) - \mathscr{F}(x(t))\|$ for some suitable norm $\|\cdot\|$.

A. Low-rank approximation and sparse linear least squares solvers

In this section, some low-rank approximation methods with applications to the solution of sparse linear least squares problems are presented.

Definition III.1. *Given $\delta > 0$ and a matrix $A \in \mathbb{C}^{m\times n}$, we will write $\mathrm{rk}_\delta(A)$ to denote the nonnegative integer determined by the expression*

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

where the numbers $s_j(A)$ represent the singular values corresponding to an economy-sized singular value decomposition of the matrix A.

Lemma III.2. *We will have that $\mathrm{rk}_\delta\left(A^\top\right) = \mathrm{rk}_\delta(A)$ for each $\delta > 0$ and each $A \in \mathbb{C}^{m\times n}$.*

Proof. Given an economy-sized singular value decomposition

$$U \begin{bmatrix} s_1(A) & & & \\ & s_2(A) & & \\ & & \ddots & \\ & & & s_{\min\{m,n\}}(A) \end{bmatrix} V = A$$

we will have that

$$V^\top \begin{bmatrix} s_1(A) & & & \\ & s_2(A) & & \\ & & \ddots & \\ & & & s_{\min\{m,n\}}(A) \end{bmatrix} U^\top = A^\top$$

is an economy-sized singular value decomposition of $A^\top$. This implies that

$$\mathrm{rk}_\delta\left(A^\top\right) = \sum_{j=1}^{\min\{m,n\}} H_\delta(s_j(A)) = \mathrm{rk}_\delta(A)$$

and this completes the proof. □

Lemma III.3. *Given $\delta > 0$ and $A \in \mathbb{C}^{m\times n}$ we will have that $\mathrm{rk}_\delta(A) \leq \mathrm{rk}(A)$.*

Proof. We will have that $\mathrm{rk}(A) = \sum_{j=1}^{\min\{m,n\}} H_0(s_j(A)) \geq \sum_{j=1}^{\min\{m,n\}} H_\delta(s_j(A)) = \mathrm{rk}_\delta(A)$. This completes the proof. □

Theorem III.4. *Given $\delta > 0$ and $y, x_1, \ldots, x_m \in \mathbb{C}^n$, let*

$$X = \begin{bmatrix} | & | & & | \\ x_1 & x_2 & \cdots & x_m \\ | & | & & | \end{bmatrix}.$$

If $\mathrm{rk}_\delta(X) > 0$ and if we set $r = \mathrm{rk}_\delta(X)$ and $s_{n,m}(r) = \sqrt{r(\min\{m,n\} - r)}$ then, there are a rank r orthogonal projector Q, r vectors $x_{j_1}, \ldots, x_{j_r} \in \{x_1, \ldots, x_m\}$ and r scalars $c_1 \ldots, c_r \in \mathbb{C}$ such that $\|X - QX\|_F \leq (s_{n,m}(r)/\sqrt{r})\delta$, and $\|y - \sum_{k=1}^r c_k x_{j_k}\| \leq \left(\sum_{k=1}^r |c_k|^2\right)^{\frac{1}{2}} s_{n,m}(r)\delta + \|(I_n - Q)y\|$.

Proof. Let us consider an economy-sized singular value decomposition $USV = A$. If u_j denotes the j-column of U, let Q be the rank $r = \mathrm{rk}_\delta(A)$ orthogonal projector determined by the expression $Q = \sum_{j=1}^r u_j u_j^*$. It can be seen that

$$\begin{aligned} \|X - QX\|_F^2 &= \sum_{j=r+1}^{\min\{m,n\}} s_j(X)^2 \\ &\leq (\min\{m,n\} - r)\delta^2 = \frac{s_{n,m}(r)^2}{r}\delta^2. \end{aligned}$$

Consequently, $\|X - QX\|_F \leq \frac{s_{n,m}(r)}{\sqrt{r}}\delta$.

Let us set.

$$\hat{X} = \begin{bmatrix} | & | & & | \\ \hat{x}_1 & \hat{x}_2 & \cdots & \hat{x}_m \\ | & | & & | \end{bmatrix} = QX$$

$$\hat{X}_y = \begin{bmatrix} | & | & & | & | \\ \hat{x}_1 & \hat{x}_2 & \cdots & \hat{x}_m & \hat{y} \\ | & | & & | & | \end{bmatrix} = Q\begin{bmatrix} X & y \end{bmatrix}$$

Since by lemma III.3 $\mathrm{rk}(X) \geq \mathrm{rk}_\delta(X)$, we will have that $\mathrm{rk}(\hat{X}) = r = \mathrm{rk}_\delta(X) > 0$, and since we also have that $\hat{x}_1, \ldots, \hat{x}_m, \hat{y} \in \mathrm{span}(\{u_1, \ldots, u_r\})$, there are r linearly independent $\hat{x}_{j_1}, \ldots, \hat{x}_{j_r} \in \{\hat{x}_1, \ldots, \hat{x}_m\}$ such that

$\text{span}(\{u_1,\ldots,u_r\}) = \text{span}(\{\hat{x}_{j_1},\ldots,\hat{x}_{j_r}\})$, this in turn implies that $\hat{y} \in \text{span}(\{\hat{x}_{j_1},\ldots,\hat{x}_{j_r}\})$ and there are $c_1,\ldots,c_r \in \mathbb{C}$ such that $\hat{y} = \sum_{k=1}^{r} c_k \hat{x}_{j_k}$. It can be seen that for each $z \in \{x_1,\ldots,x_m\}$

$$\|z - Qz\| \leq \|X - QX\|_F \leq \frac{s_{n,m}(r)}{\sqrt{r}}\delta,$$

and this in turn implies that

$$\begin{aligned}\left\|y - \sum_{k=1}^{r} c_k x_{j_k}\right\| &= \left\|y - \sum_{k=1}^{r} c_k x_{j_k} - \left(\hat{y} - \sum_{k=1}^{r} c_k \hat{x}_{j_k}\right)\right\| \\ &= \left\|y - \sum_{k=1}^{r} c_k x_{j_k} - Q\left(y - \sum_{k=1}^{r} c_k x_{j_k}\right)\right\| \\ &\leq \left(\sum_{k=1}^{r} |c_k|^2\right)^{\frac{1}{2}} s_{n,m}(r)\delta + \|(I_n - Q)y\|.\end{aligned}$$

This completes the proof. □

As a direct implication of theorem III.4 one can obtain the following corollary.

Corollary III.5. *Given $\delta > 0$, $A \in \mathbb{C}^{m\times n}$ and $y \in \mathbb{C}^m$. If $\text{rk}_\delta(A) > 0$ and if we set $r = \text{rk}_\delta(A)$ and $s_{n,m}(r) = \sqrt{r(\min\{m,n\} - r)}$ then, there are $x \in \mathbb{C}^n$ and a rank r orthogonal projector Q that does not depend on y, such that $\|Ax - y\| \leq \|x\| s_{n,m}(r)\delta + \|(I_m - Q)y\|$ and x has at most r nonzero entries.*

Proof. Let us set $x = \mathbf{0}_{n,1}$ and $a_j = A\hat{e}_{j,n}$ for $j = 1,\ldots,n$. Since $r = \text{rk}_\delta(A) > 0$ and $s_{n,m}(r) = \sqrt{r(\min\{m,n\} - r)}$, by theorem III.4 we will have that there is a rank r orthogonal projector Q such that $\|A - QA\|_F \leq (s_{n,m}(r)/\sqrt{r})\delta$, and without loss of generality r vectors $a_{j_1},\ldots,a_{j_r} \in \{a_1,\ldots,a_n\}$ and r scalars $c_1\ldots,c_r \in \mathbb{C}$ with $j_1 \leq j_2 \leq \cdots \leq j_r$ (reordering the indices j_k if necessary), such that $\|y - \sum_{k=1}^{r} c_k a_{j_k}\| \leq \left(\sum_{k=1}^{r} |c_k|^2\right)^{\frac{1}{2}} s_{n,m}(r)\delta + \|(I_m - Q)y\|$. If we set $x_{j_k} = c_k$ for $k = 1,\ldots,r$, we will have that $\|x\| = \left(\sum_{k=1}^{r} |c_k|^2\right)^{\frac{1}{2}}$ and $Ax = \sum_{k=1}^{r} x_{j_k} a_{j_k} = \sum_{k=1}^{r} c_k a_{j_k}$. Consequently, $\|Ax - y\| \leq \|x\| s_{n,m}(r)\delta + \|(I_m - Q)y\|$. This completes the proof. □

The results and ideas presented in this section can be translated into a sparse linear least squares solver algorithm described by algorithm A.1 in §IV.

B. Sparse structured nonlinear regressive model identification

Given time series data sets $\Sigma_x = \{x(t)\}_{t\geq 1}$ and $\Sigma_y = \{y(t)\}_{t\geq 1}$ in R^n corresponding to the orbits of two discrete-time dynamic financial systems of interest, let us consider the problem of identifying a map $\mathscr{T}$ relating the time series data according to the expression

$$y(t) = \mathscr{T}(x(t)) + r(t), \tag{III.2}$$

where $r(t)$ is some suitable small residual term defined as in (III.1). One may need to preprocess the time series data before proceeding with the approximate representation of a suitable evolution operator. For this purpose, given some prescribed suitable integer $L > 0$, one can consider the time series $\mathscr{D}_L(\Sigma_x)$ and $\mathscr{D}_L(\Sigma_y)$ determined by the expressions.

$$\begin{aligned}\mathscr{D}_L(\Sigma_x) &= \{\mathbf{x}_L(t)\}_{t\geq L} \\ \mathscr{D}_L(\Sigma_y) &= \{\mathbf{y}_L(t)\}_{t\geq L}\end{aligned}$$

For the dilated time series $\mathscr{D}_L(\Sigma_x)$ and $\mathscr{D}_L(\Sigma_y)$, the identification process corresponding to the relation (III.20), can be translated into the approximate solution of equations of the form

$$\mathbf{y}_L(t) = \tilde{\mathscr{T}}(\mathbf{x}_L(t)), \tag{III.3}$$

for $t \geq L$. Where $\tilde{\mathscr{T}}$ is the mapping to be approximately identified.

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

$$\eth_p(x) := \begin{bmatrix}\Pi_1(x)\\ \Pi_2(x)\\ \vdots \\ \Pi_p(x)\\ 1\end{bmatrix} = \begin{bmatrix}x^{\otimes 1}\\ x^{\otimes 2}\\ \vdots \\ x^{\otimes p}\\ 1\end{bmatrix} \tag{III.4}$$

Given integers $p, L > 0$, and two orbits $\Sigma_x = \{x_t\}_{t\geq 1}$ and $\Sigma_y = \{y_t\}_{t\geq 1}$ in $\mathbb{R}^n$, corresponding to two related dynamic financial processes of interest. For finite samples $\Sigma_N^x = \{x_t\}_{t=1}^T \subset \Sigma_x$ and $\Sigma_N^y = \{y_t\}_{t=1}^T \subset \Sigma_y$, let us consider the matrices:

$$\begin{aligned}\mathbf{H}_L^{(0,p)}(\Sigma_T^x) &= \begin{bmatrix}\eth_p(\mathbf{x}_L(L)) & \cdots & \eth_p(\mathbf{x}_L(T))\end{bmatrix} \\ \mathbf{H}_L^{(1)}(\Sigma_T^y) &= \begin{bmatrix}\mathbf{y}_L(L) & \cdots & \mathbf{y}_L(T)\end{bmatrix}\end{aligned} \tag{III.5}$$

The mapping identification mechanism used in this study for dilated systems of the form (III.3), will be approximately described by the expression:

$$\mathbf{y}_L(t) = \hat{\mathscr{T}}(\mathbf{x}_L(t)) = W\eth_p(\mathbf{x}_L(t)),\ t \geq L, \tag{III.6}$$

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 the techniques and ideas previously presented in this section, the matrix W in (III.6) can be estimated by approximately solving the matrix equation

$$W\mathbf{H}_L^{(0,p)}(\Sigma_T^x) = \mathbf{H}_L^{(1)}(\Sigma_T^y). \tag{III.7}$$

The devices described by (III.6) are called regressive reservoir computers (RRC) in this paper.

For any given integers $L, n, p > 0$. Taking advantage of the maps $\eth_p$, one can find an 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 $R_{p,L}(n)^+ R_{p,L}(n)\eth_p(x) \approx \eth_p(x)$ 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.6. *Given $\varepsilon \in \mathbb{R}^+$ and $L, n, p \in \mathbb{Z}^+$. There are an integer $0 < \rho_p(n) < d_p(n)$ and a sparse matrix $R_{p,L}(n) \in \mathbb{R}^{\rho_p(n)\times d_p(n)}$ with $d_p(n)$ nonzero entries, such that $\|R_{p,L}(n)^+ R_{p,L}(n)\eth_p(x) - \eth_p(x)\| \leq \sqrt{d_p(nL)}\varepsilon$ 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 $\{\hat{x}_1,...,\hat{x}_{nL},\hat{x}_{nL+1}\} \subset \mathbb{R}$ such that for each $1 \leq m \leq p$ and every pair on index sets $\{i_1,\ldots,i_m\},\{j_1,\ldots,j_m\} \subset \{1,\ldots,nL+1\}$:

$$\prod_{k=1}^{m} \hat{x}_{j_k} = \prod_{k=1}^{m} \hat{x}_{i_k}, \Leftrightarrow \exists \sigma \in \mathfrak{S}_{nL-1} : i_k = \sigma(j_k), \forall 1 \leq k \leq m. \tag{III.8}$$

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

One can choose a suitable scale factor $\nu > 0$, and we can now set $\mathbf{y} = \nu \begin{bmatrix} \hat{x}_1 & \hat{x}_2 & \cdots & \hat{x}_{nL} \end{bmatrix}^\top$, $d = d_p(n)$ and

$$\tilde{\mathbf{x}} = \begin{bmatrix} \tilde{x}_1 & \cdots & \tilde{x}_d \end{bmatrix}^\top := \eth_p(\mathbf{y}). \tag{III.9}$$

If in addition we consider the following assignments:

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

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

Then, by (III.4), (III.8) and (III.9) for each $j = 2,\ldots,d$ one can find $1 \leq k_1(j),\ldots,k_{n_j}(j) \leq d$ such that

$$|\tilde{x}_j - \tilde{x}_{k_m(j)}| \leq \varepsilon \tag{III.10}$$

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

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

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

From the definition of $R_{p,L}(n)$, it can be seen that $R_{p,L}(n)R_{p,L}(n)^T$ is a diagonal matrix in $\mathbb{R}^{\rho_p(n)\times\rho_p(n)}$ determined by the following expression.

$$R_{p,L}(n)R_{p,L}(n)^\top = \begin{bmatrix} \frac{1}{\hat{n}_1} & 0 & \cdots & \cdots & 0 \\ 0 & \frac{1}{\hat{n}_2} & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \frac{1}{\hat{n}_{\rho_p(n)-1}} & 0 \\ 0 & \cdots & \cdots & 0 & \frac{1}{\hat{n}_{\rho_p(n)}} \end{bmatrix} \tag{III.11}$$

Here, for $1 \leq j \leq \rho_p(n)$, each $\hat{n}_j$ is equal to the number of nonzero entries in the i-th row of $R_{p,L}(n)$. Furthermore, we will have that $\hat{n}_j \geq 1$ for each $1 \leq j \leq \rho_p(n)$, since by (III.10) each row of $R_{p,L}(n)$ has at least one nonzero entry.

By (III.11) and from the definition of $R_{p,L}(n)$, it can be seen that $R_{p,L}(n)R_{p,L}(n)^\top$ is invertible, and that the Moore-Penrose pseudoinverse $R_{p,L}(n)^+$ of $R_{p,L}(n)$ is determined by the expression:

$$\begin{aligned} R_{p,L}(n)^+ &= R_{p,L}(n)^\top \left(R_{p,L}(n)R_{p,L}(n)^\top\right)^{-1} \\ &= \begin{bmatrix} \hat{e}_{1,d} & \sum_{l=1}^{n_2} \hat{e}_{k_l(2),d} & \cdots & \sum_{l=1}^{n_{r-1}} \hat{e}_{k_l(r-1),d} & \hat{e}_{d,d} \end{bmatrix} \end{aligned} \tag{III.12}$$

with $r = \rho_p(n)$.

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 for each $1 \leq j \leq d_p(n)$:

$$|(R_{p,L}(n)^+ R_{p,L}(n))\eth_p(x))[j] - \eth_p(x)[j]| \leq 2\varepsilon$$

Consequently,

$$\|R_{p,L}(n)^+ R_{p,L}(n)\eth_p(x) - \eth_p(x)\| \leq 2\sqrt{d_p(n)}\varepsilon.$$

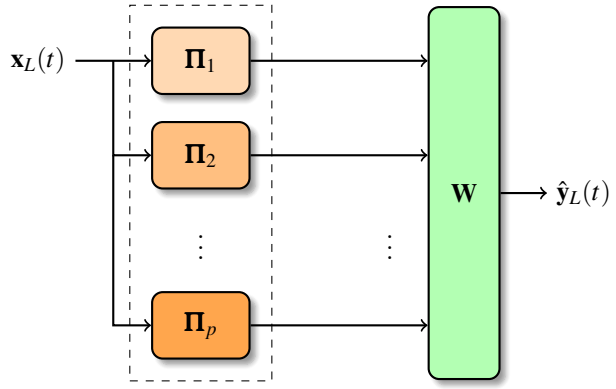
This completes the proof. □

In order to reduce to computational effort corresponding to the solution of (III.7), using the matrix $R_{p,L}(n)$ described by Theorem III.6, one can obtain an approximate reduced representation of (III.7) determined by the expression.

$$\bar{W}R_{p,L}(n)\mathbf{H}_L^{(0,p)}(\Sigma_T^x) = \mathbf{H}_L^{(1)}(\Sigma_T^y) \tag{III.13}$$

The architecture of the regressive reservoir computers considered in this study was inspired by next generation reservoir computers[7].

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



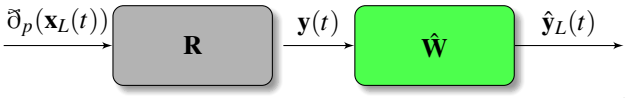
(III.14)

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

$$\begin{aligned} \mathbf{W}(\Pi_1(\mathbf{x}_L(t)),\ldots,\Pi_p(\mathbf{x}_L(t))) &:= \tilde{W}\begin{bmatrix} \Pi_1(\mathbf{x}_L(t)) \\ \vdots \\ \Pi_p(\mathbf{x}_L(t)) \end{bmatrix} + c_W \\ &= \begin{bmatrix} \tilde{W} & c_W \end{bmatrix} \eth_p(\mathbf{x}_L(t)) \end{aligned}$$

and where the matrix $W = \begin{bmatrix} \tilde{W} & c_W \end{bmatrix}$ is determined by (III.7).

The structure of the generic block $\mathbf{W}$ ub (III.14) can be factored in the form



(III.15)

The layers $\mathbf{R}$ and $\hat{\mathbf{W}}$ of the device (III.15) are determined by the expressions

$$\mathbf{R}(\mathbf{x}) = \hat{R}\mathbf{x},$$
$$\hat{\mathbf{W}}(\mathbf{y}) = W\mathbf{y}$$

for any pair of suitable vectors $\mathbf{x},\mathbf{y}$. Where W is a sparse representation of an approximate solution to (III.13) and $\hat{R}$ is determined by Theorem III.6.

Using the reservoir computer models described by (III.6), (III.14) and (III.15), we can compute approximate representations of the mappings that satisfy (III.3) using the expression

$$\begin{aligned}\hat{\mathscr{T}}(\mathbf{x}_L(t)) &:= \hat{K}\left(\hat{\mathbf{W}}\circ\mathbf{R}\circ\eth_p(\mathbf{x}_L(t)\right)\\ &= \hat{K}W\hat{R}\eth_p\left(\mathbf{x}_L(t)\right)),\end{aligned} \tag{III.16}$$

for each $t \geq L$, with

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

Furthermore, we can use the identified RRC model $\hat{\mathscr{T}}$ to simulate the behavior $y_t = \mathscr{T}(x_t)$ of the system described by (III.20) for $L \leq t \leq \tau$, by performing the operation:

$$\mathbf{T}\left(\mathbf{x}_L(t)\right) := \hat{K}\hat{\mathscr{T}}(\mathbf{x}_L(t)) = \hat{K}W\eth_p\left(\mathbf{x}_L(t)\right)), \tag{III.17}$$

for some suitable $\tau > 0$.

Theorem III.7. *Given $\delta > 0$, two integers $p, L > 0$, a sample $\Sigma_T = \{x_t\}_{t=1}^{T}$ from a dynamic financial 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.13) *with $R_{p,L}(n)$ and $r_p(n)$ determined by Theorem III.6. If $r = \mathrm{rk}_\delta(R_{p,L}(n)\mathbf{H}_L^{(0,p)}(\Sigma_T)) > 0$, then there is a sparse representation $\hat{W} \in \mathbb{R}^{nL\times\rho_p(nL)}$ of $\bar{W}$ with at most $r\rho_p(nL)$ nonzero entries such that*

$$\|\hat{W}R_{p,L}(n)\mathbf{H}_L^{(0,p)}(\Sigma_T) - \bar{W}R_{p,L}(n)\mathbf{H}_L^{(0,p)}(\Sigma_T)\|_F \leq K\delta, \tag{III.18}$$

for $K = \sqrt{nL(\min\{\rho_p(nL), T-L\}-r)}(\sqrt{r}\|\hat{W}\|_F + \|\bar{W}\|_F)$, where $\rho_p(nL)$ is the integer described by Theorem III.6.

Proof. Let us set $H = R_{p,L}(n)\mathbf{H}_L^{(0,p)}(\Sigma_T)^{\top}$ and $Y = H\bar{W}^{\top}$. It suffices to prove that there is a sparse representation $\hat{W} \in \mathbb{R}^{nL\times\rho_p(nL)}$ with at most $r\rho_p(nL)$ nonzero entries such that

$$\|H\hat{W}^{\top} - Y\|_F \leq K\delta.$$

Since we have that

$$\begin{aligned}\mathrm{rk}_\delta(H) &= \mathrm{rk}_\delta\left(\left(R_{p,L}(n)\mathbf{H}_{L,G}^{(0,p)}(\Sigma_T)\right)^{\top}\right)\\ &= \mathrm{rk}_\delta(R_{p,L}(n)\mathbf{H}_L^{(0,p)}(\Sigma_T)) > 0\end{aligned}$$

by Lemma III.2. By Corollary III.5, if we set $r = \mathrm{rk}_\delta(H)$ and $\alpha = \sqrt{r(\min\{\rho_p(nL), T-L\}-r)}$. We will have that there is a rank r orthogonal projector Q such that for each $j = 1,\ldots,nL$, there is $\hat{v}_j \in \mathbb{R}^{nL}$ with at most r nonzero entries, for which $\|H\hat{v}_j - Y\hat{e}_{j,M}\| \leq \alpha\|\hat{v}_j\|\delta + \|(I_{T-L} - Q)Y\hat{e}_{j,nL}\|$. Consequently, if we set

$$\hat{W} = \begin{bmatrix} | & & | \\ \hat{v}_1 & \cdots & \hat{v}_{nL} \\ | & & | \end{bmatrix}^{\top}$$

we will have that $\hat{W}$ has at most nrL nonzero entries and

$$\begin{aligned}\|H\hat{W}^{\top} - Y\|_F^2 &= \sum_{j=1}^{nL}\|H\hat{v}_j - Y\hat{e}_{j,nL}\|^2\\ &\leq M(\alpha\|\hat{W}\|_F\delta + \|(I_{T-L} - Q)Y\|_F)^2,\end{aligned}$$

and this in turn implies that,

$$\|H\hat{W}^{\top} - Y\|_F \leq \sqrt{nL}(\alpha\|\hat{W}\|_F\delta + \|(I_{T-L} - Q)H\|_F\|\bar{W}\|_F). \tag{III.19}$$

By (III.19) and by Theorem III.4 we will have that

$$\begin{aligned}\|H\hat{W}^{\top} - Y\|_F &\leq \sqrt{nL}(\alpha\|\hat{W}\|_F\delta + (\alpha/\sqrt{r})\|\bar{W}\|_F\delta)\\ &= \alpha\sqrt{(nL/r)}(\sqrt{r}\|\hat{A}\|_F + \|A\|_F)\delta = K\delta.\end{aligned}$$

This completes the proof. □

1. *Sparse structured nonlinear autoregressive model identification*

Given some time series data $\Sigma \subset \mathbb{R}^n$ corresponding to an orbit determined by the difference equation

$$x(t+1) = \mathscr{A}(x(t)), \tag{III.20}$$

for some discrete-time dynamic financial model $(\hat{\mathbf{\Sigma}}, \mathscr{T})$ to be identified. One can use the methods presented in §III B to identify the mapping $\mathscr{S}$, by considering the RRC model identification determined by the problem

$$y(t) = \mathscr{A}(x(t)),$$

for the time series $\Sigma_x := \{x(t)\}_{t\geq 1}$ and $\Sigma_y := \{y(t)\}_{t\geq 1}$ in $\mathbb{R}^n$, with $y(t) := x(t+1)$ for each $t \geq 1$.

IV. ALGORITHMS

The sparse model identification methods presented in §III A can be translated into prototypical algorithms that will be presented in this section, some programs for data reading and writing, synthetic signals generation, and predictive simulation are also included as part of the **DyNet** tool-set available in[15].

A. Sparse linear least squares solver and structured assembling matrix identification algorithms

As an application of the results and ideas presented in §III A one can obtain a prototypical sparse linear least squares solver algorithm like algorithm A.1.

The least squares problems $c = \arg\min_{\hat{c}\in\mathbb{C}^K}\|\hat{A}\hat{c} - y\|$ to be solved as part of the process corresponding to algorithm A.1 can be solved with any efficient least squares solver available in the language or program where the sparse linear least squares solver algorithm is implemented. For the Python version of algorithm A.1 the function `lstsq` is implemented.

Algorithm A.1 SLRSolver: Sparse linear least squares solver algorithm

Data: $A \in \mathbb{C}^{m\times n}$, $Y \in \mathbb{C}^{m\times p}$, $\delta > 0, N \in \mathbb{Z}^+$, $\varepsilon > 0$
Result: $X = \mathbf{SLRSolver}(A, Y, \delta, N, \varepsilon)$

1. Compute economy-sized SVD $USV = A$
2. Set $s = \min\{m,n\}$
3. Set $r = \mathrm{rk}_\delta(A)$
4. Set $U_\delta = \sum_{j=1}^{r} U\hat{e}_{j,s}\hat{e}_{j,s}^*$
5. Set $T_\delta = \sum_{j=1}^{r} (\hat{e}_{j,s}^* S \hat{e}_{j,s})^{-1} \hat{e}_{j,s}\hat{e}_{j,s}^*$
6. Set $V_\delta = \sum_{j=1}^{r} \hat{e}_{j,s}\hat{e}_{j,s}^* V$
7. Set $\hat{A} = U_\delta^* A$
8. Set $\hat{Y} = U_\delta^* Y$
9. Set $X_0 = V_\delta^* T_\delta \hat{Y}$
10. **for** $j = 1, \ldots, p$ **do**
    Set $K = 1$
    Set error $= 1 + \delta$
    Set $c = X_0 \hat{e}_{j,p}$
    Set $x_0 = c$
    Set $\hat{c} = \begin{bmatrix} \hat{c}_1 & \cdots & \hat{c}_n \end{bmatrix}^\top = \begin{bmatrix} |\hat{e}_{1,n}^* c| & \cdots & |\hat{e}_{n,n}^* c| \end{bmatrix}^\top$
    Compute permutation $\sigma : \{1,\ldots,n\} \to \{1,\ldots,n\}$ such that: $\hat{c}_{\sigma(1)} \geq \hat{c}_{\sigma(2)} \geq \cdots \geq \hat{c}_{\sigma(n)}$
    Set $N_0 = \max\left\{\sum_{j=1}^{n} H_\varepsilon\left(\hat{c}_{\sigma(j)}\right), 1\right\}$
    while $K \leq N$ **and** error $> \delta$ **do**
        Set $x = \mathbf{0}_{n,1}$
        Set $A_0 = \sum_{j=1}^{N_0} \hat{A}\hat{e}_{\sigma(j),n}\hat{e}_{j,N_0}^*$
        Solve $c = \arg\min_{\tilde{c}\in\mathbb{C}^{N_0}} \|A_0\tilde{c} - \hat{Y}\hat{e}_{j,p}\|$
        for $k = 1, \ldots, N_0$ **do**
            Set $x_{\sigma(k)} = \hat{e}_{k,N_0}^* c$
        end for
        Set error $= \|x - x_0\|_\infty$
        Set $x_0 = x$
        Set $\hat{c} = \begin{bmatrix} \hat{c}_1 & \cdots & \hat{c}_n \end{bmatrix}^\top = \begin{bmatrix} |\hat{e}_{1,n}^* x| & \cdots & |\hat{e}_{n,n}^* x| \end{bmatrix}^\top$
        Compute permutation $\sigma : \{1,\ldots,n\} \to \{1,\ldots,n\}$ such that: $\hat{c}_{\sigma(1)} \geq \hat{c}_{\sigma(2)} \geq \cdots \geq \hat{c}_{\sigma(n)}$
        Set $N_0 = \max\left\{\sum_{j=1}^{n} H_\varepsilon\left(\hat{c}_{\sigma(j)}\right), 1\right\}$
        Set $K = K + 1$
    end while
    Set $x_j = x$
11. **end for**
12. Set $X = \begin{bmatrix} | & | & & | \\ x_1 & x_2 & \cdots & x_p \\ | & | & & | \end{bmatrix}$

return X

In this section, we focus on the applications of the structured matrix approximation methods presented in SIII, to dynamical financial systems identification via regressive reservoir computers.

Algorithm A.2 Compression matrix computation algorithm

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

1. Choose nL pseudorandom numbers $\hat{x}_1, \ldots, \hat{x}_{nL} \in \mathbb{R}$ from $N(0,1)$
2. Set $\mathbf{y} = \nu \begin{bmatrix} \hat{x}_1 & \hat{x}_2 & \cdots & \hat{x}_{nL} \end{bmatrix}^\top$
3. Set $d = d_p(n)$
4. Set $\tilde{\mathbf{x}} = \begin{bmatrix} \tilde{x}_1 & \cdots & \tilde{x}_d \end{bmatrix}^\top := \eth_p(\mathbf{y})$
5. Choose a pseudorandom number $\alpha \in N(0,1)$;
6. Set $\tilde{x}_d := \alpha$
7. Set $R = e_{1,d}^\top$
8. **for** $j = 2, \ldots, d$ **do**
    Find $1 \leq k_1, \ldots, k_{n_j} \leq d$ such that $|\tilde{x}_j - \tilde{x}_{k_m}| \leq \varepsilon$, for each $1 \leq m \leq n_j$
    if $k_1 = j$ **then**
        Set $R_0 := (1/n_j) \sum_{l=1}^{n_j} \hat{e}_{k_l,d}^T$
        Set $R := \begin{bmatrix} R \\ R_0 \end{bmatrix}$
    end if
9. **end for**
10. Set $R_{p,L}(n) := R$

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

B. Structured coupling matrix identification algorithm

Given a discrete-time dynamic financial model $(\Sigma, \mathscr{T})$ and a structured data sample $\Sigma_T \subset \Sigma$, we can apply Algorithm A.2 and Algorithm A.1, in order to compute the output coupling matrix that can be used to obtain an approximate representation of the evolution operator $\mathscr{T}$, corresponding to the orbit Σ. For this purpose, one can use the following Algorithm.

V. NUMERICAL SIMULATIONS AND APPLICATIONS

In this section, we will present some numerical simulations computed using the **DyNet** toolset available in[15], which was developed as part of this project. The toolset consists of a collection of Python 3.10.4 programs for structured sparse identification and numerical simulation of discrete-time dynamical financial systems.

The numerical experiments documented in this section were performed with Python 3.10.4. All the programs written for real-world data reading, synthetic data generation, and sparse model identification as part of this project are available at[15].

The numerical simulations described in this section were conducted on an Ubuntu 18.04.6 LTS server system. This system operates on a virtual machine within Hyper-V, equipped with 16 vCores of an Intel(R) Xeon(R) Gold 6238 CPU, running at 2.10 GHz (2095 MHz), and with 64 GB of RAM.

Algorithm B.1 RRC Model: RRC model identification

Data: $\Sigma_N^x = \{x_t\}_{t=1}^T, \Sigma_N^y = \{y_t\}_{t=1}^T \subset \mathbb{R}^n$
Result: OUTPUT COUPLING AND COMPRESSION MATRICES: $\hat{W}, \bar{W}, R_{p,L}(n)$

1. Choose or estimate the lag value L using auto-correlation function based methods
2. Set a tensor order value p
3. Compute compression matrix $R_{p,L}(n)$ applying Algorithm A.2
4. Compute matrices:

$$\mathbf{H}_0 := \mathbf{H}_L^{(0,p)}(\Sigma_T^x)$$
$$\mathbf{H}_1 := \mathbf{H}_L^{(1)}(\Sigma_T^y)$$

5. Approximately solve:

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

for $\hat{W}$ applying Algorithm A.1

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

To asses prediction errors for the experiments considered in the following sub-sections, we will use a root-mean-square error estimate determined for any given vector time series sample $\Sigma_T(x) := \{x(t)\}_{1\leq t\leq T} \subset \mathbb{R}^n$ and the set of predictions $\Sigma_T(\hat{x}) := \{\hat{x}(t)\}_{1\leq t\leq T}$ determined by some model under consideration by the following expression.

$$RMSE(x,\hat{x}) := \frac{1}{\sqrt{nT}}\left\|\mathbf{H}_1^{(1)}(\Sigma_T(x)) - \mathbf{H}_1^{(1)}(\Sigma_T(\hat{x}))\right\|_F \quad \text{(V.1)}$$

A. Sparse autoregressive reservoir computers for dynamical nonlinear financial system behavior identification

In this section, we focus on conducting numerical simulations to examine the behavior of a financial system modeled by a nonlinear dynamical system. These simulations aim to explore the intricate relationships among the interest rate (IR), investment demand (ID), and price index (PI) under two distinct scenarios. The governing equations of the model are as follows:

$$\begin{aligned}
\dot{x}_1 &= x_3 + (x_2 - s)x_1,\\
\dot{x}_2 &= 1 - cx_2 - x_1^2,\\
\dot{x}_3 &= -x_1 - ex_3,\\
x_1(0) &= x_0, x_2(0) = y_0, x_3(0) = z_0.
\end{aligned} \quad \text{(V.2)}$$

Here, x_1, x_2, and x_3 denote the interest rate, investment demand, and price index, respectively.

As observed in[2], systems of the form (V.2) can exhibit, among others behavior types, chaotic and eventually approximately periodic dynamic behavior depending on the configuration of parameters and initial conditions considered for (V.2).

1. *Chaotic behavior identification*

For $s = 3, c = 0.1, e = 1$, let us consider the initial conditions $x_0 = 2, y_0 = 3, z_0 = 2$. For this configuration, one can obtain synthetic time series data $\Sigma_{12000} \subset \mathbb{R}^3$ obtained by applying a fourth-order adaptive numerical integration method to (V.2) for the configuration determined by the previous choice of parameters, obtaining an orbit's samples set Σ_{12000} whose elements are uniformly distributed with respect to the time interval [0,120].

The training orbit's data set corresponding to the first 50% of the data in Σ_{12000}, together with the remaining data used for model validation, are illustrated in Figure 1. The factorization for the output coupling matrix $W = \hat{W}R$ determined by Theorems III.6 and III.7 are illustrated in Figure 2.

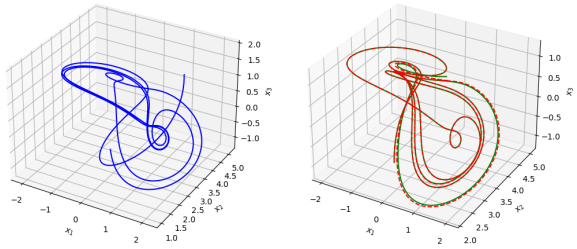


Figure 1: Training orbits data (left), validation orbits data (right). The green line corresponds to validation data, and the red dotted line corresponds to the model's predictions.

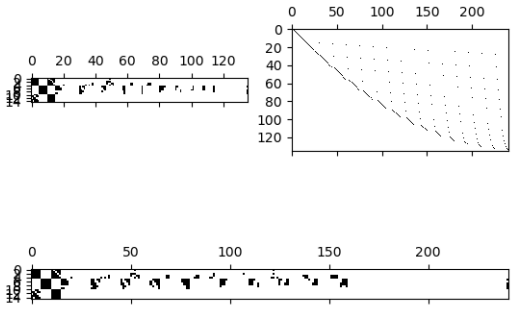


Figure 2: Matrix factors $\hat{W}$ (top-left) and R (top-right), output coupling matrix $W = \hat{W}R$ (bottom).

2. *Eventually approximately periodic behavior identification*

For $s = 0.5, c = 0.1, e = 0.1$, let us consider the initial conditions $x_0 = 1, y_0 = 1, z_0 = 1$. For this configuration, one can obtain synthetic time series data $\Sigma_{12000} \subset \mathbb{R}^3$ obtained by applying a fourth-order adaptive numerical integration method to (V.2) for the configuration determined by the previous choice of parameters, obtaining an orbit's samples set Σ_{12000} whose elements are uniformly distributed with respect to the time interval [0,120].

The training orbit's data set corresponding to the first 6.67% of the data in Σ_{12000}, together with the remaining data used for model validation, are illustrated in Figure 3. The factorization for the output coupling matrix $W = \hat{W}R$ determined by Theorems III.6 and III.7 are illustrated in Figure 4.

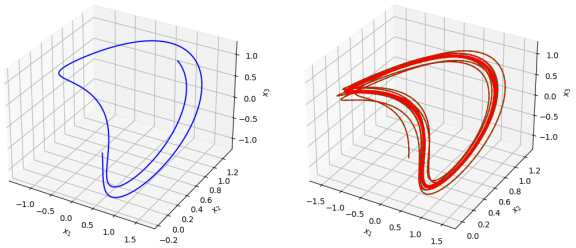


Figure 3: Training orbits data (left), validation orbits data (right). The green line corresponds to validation data, and the red dotted line corresponds to the model's predictions.

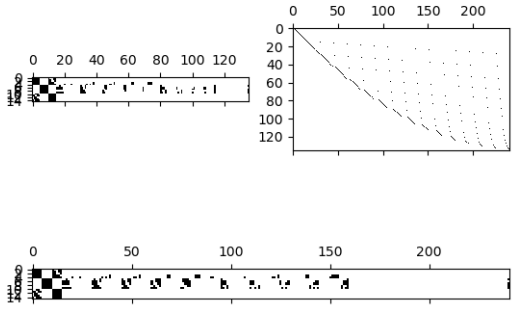


Figure 4: Matrix factors $\hat{W}$ (top-left) and R (top-right), output coupling matrix $W = \hat{W}R$ (bottom).

3. *Learning interest rates with sparse regressive reservoir computers*

In this section, for financial systems described by (V.2) we will consider the problem corresponding to the identification and simulation of the interest rate signals x_1, when the signals x_2, x_3 are known.

The models considered in this section are determined by (III.6), (III.14) and (III.15), and can be described by expressions of the form:

$$\begin{bmatrix} \hat{x}_1(t-1) \\ \hat{x}_1(t) \end{bmatrix} := \hat{W}R_{2,2}(2)\eth_2 \left(\begin{bmatrix} x_2(t-1) \\ x_2(t) \\ x_3(t-1) \\ x_3(t) \end{bmatrix} \right) \tag{V.3}$$

The first case under consideration corresponds to the identification of the interest rate when the system described by (V.2) exhibits chaotic behavior, the signals and model parameters corresponding to this system identification process are illustrated in Figure 5.

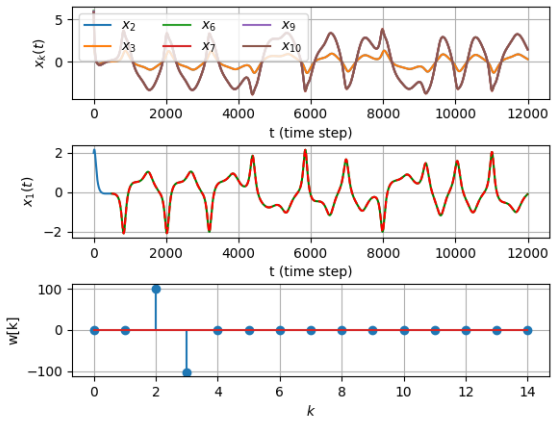


Figure 5: Chaotic interest rate identification.

When a financial system described by (V.2) exhibits an eventually periodic behavior, one can use models of the form (V.3) to learn the behavior of the interest rates, the related signals and model parameters are illustrated in Figure.

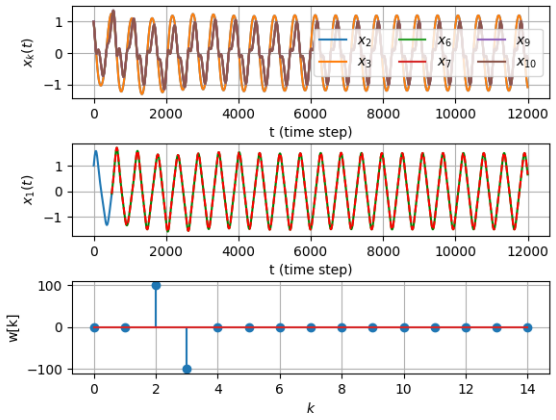


Figure 6: Periodic interest rate identification.

The computational setting used for all the experiments performed in this section is documented in the Python 3.10.4 program `FDSExperiment.py` in[15] that can be used to replicate these results.

B. Forecasting and Assessing Financial Margin Dynamics for Honduran Commercial Banks with structured SRRC Models

In Honduras, the banking system plays a crucial role in the country's economy by serving as a financial intermediary. It facilitates the flow of financial resources from savers to borrowers among different economic agents. The operational efficiency of banks in this intermediation process is evident through lower costs in resource allocation. This efficiency positively influences savings, investments, and ultimately the country's economic performance. Conversely, inefficiency in allocating resources, characterized by high intermediation costs, can hinder financial deepening and thus impede economic growth.

The financial system is continually evolving due to technological advances, growing competition, and the modernization of financial transactions. Despite these changes, the primary function of banks remains the intermediation of financial resources. This enduring role underscores the importance of banks in managing the flow of funds between different economic agents, crucial for sustaining the dynamics of the economy.

Considering the impact of the financial margin on both banks' pricing and risk-taking, it is evident that this margin influences not just the distribution of resources but also market competitiveness[1]. As a result, the evaluation and analysis of the financial intermediation margin, employing advanced methods and techniques, have become pivotal in banking supervision. This focus allows for the strengthening of the regulatory framework, enhancement of supervisory processes in risk management, and the formulation and implementation of policies aimed at boosting the competitiveness and efficiency of the banking system.

The financial intermediation margin is often used as an efficiency indicator in banking activities. In accordance with this study, the margin is defined as the difference between the interest rates banks charge for lending to various agents and the rates they offer on the deposits they receive. This spread essentially reflects the cost efficiency of a bank in managing its lending and deposit activities.

In this study, to obtain more precise estimates of the rates charged and paid by banking institutions[4], synthesized signals were used, derived from anonymized data. These were obtained by calculating the financial intermediation margin using the monthly financial statements of 15 commercial banks, emphasizing ex-post rates. The margin was calculated as the difference between the ratio of financial income to financial assets and the ratio of financial expenses to financial liabilities. This calculation is represented in the following equation:

$$\mathbf{\Phi}(t)[i] = \frac{If(t)[i]}{Af(t)[i]} - \frac{Gf(t)[i]}{Pf(t)[i]} \tag{V.4}$$

Where:

- $If(t)[i]/Af(t)[i]$ represents the financial income to financial assets for bank i at step t, and
- $Gf(t)[i]/Pf(t)[i]$ represents the financial expenses to financial liabilities for bank i at time t.

It is important to recognize that high levels of the financial intermediation margin can indicate market inefficiency, resulting in elevated intermediation costs and discouraging savings and investment activities. However, it should also be noted that reductions in the margin do not automatically signify efficiency improvements[11]. Moreover, increased levels of the intermediation margin may be indicative of systemic issues, such as limited competition, perceived credit risk, and inefficiencies in operating costs[12].

Let us consider the synthesized time series data $\Sigma_{80}(\mathbf{\Phi}) := \{\mathbf{\Phi}(t) : 0 \leq t \leq 79\}$, based on privacy-protecting and behavior-preserving transformations of the anonymized financial margins dynamics data of 15 Honduran commercial banks. Where for each t, the i-th entry $\mathbf{\Phi}(t)[i]$ of each vector $\mathbf{\Phi}(t)$ corresponds to the financial margin at time-step t of the i-th bank.

Let us consider the sample $\Sigma_{36}(\mathbf{\Phi}) := \{\mathbf{\Phi}(t) : 0 \leq t \leq 35\}$, and let us write $\mathbf{C}_{36}(\mathbf{\Phi})$ to denote the Spearman's correlation matrix corresponding to $\Sigma_{36}(\mathbf{\Phi})$ represented graphically with the heat map that is illustrated in Figure 7.

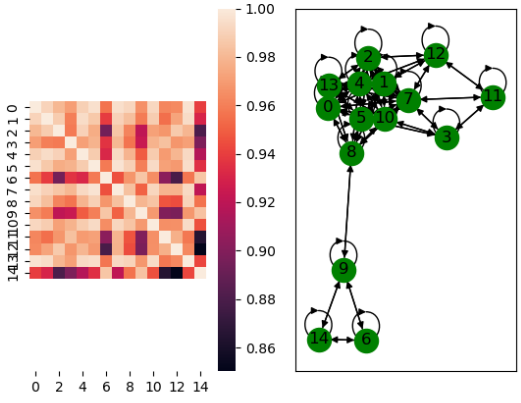


Figure 7: Heat map corresponding to $\mathbf{C}_{36}(\mathbf{\Phi})$ (left), relational graph $G_{36}(\mathbf{\Phi})) := (V_G, E_G)$ (right).

Let $G_{36}(\mathbf{\Phi})) := (V_G, E_G)$ denote the relational graph illustrated in Figure 7, that corresponds to the adjacency matrix $\mathbf{A}_{36}(\mathbf{\Phi})$ whose i,j-entries $\mathbf{A}_{36}(\mathbf{\Phi})[i,j]$ are determined by the i,j-enries $\mathbf{C}_{36}(\mathbf{\Phi})[i,j]$ of the correlation matrix $\mathbf{C}_{36}(\mathbf{\Phi})$, according to the rule:

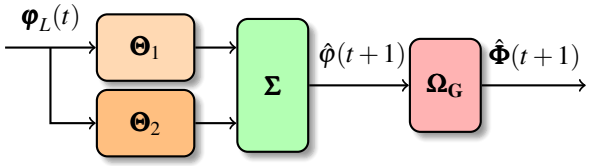
$$\mathbf{A}_{36}(\mathbf{\Phi})[i,j] := \begin{cases} 1, & |\mathbf{C}_{36}(\mathbf{\Phi})[i,j]| \geq |\widehat{\mathbf{C}_{36}(\mathbf{\Phi})}| \\ 0, & |\mathbf{C}_{36}(\mathbf{\Phi})[i,j]| < |\widehat{\mathbf{C}_{36}(\mathbf{\Phi})}| \end{cases} \tag{V.5}$$

where $|\widehat{\mathbf{C}_{36}(\mathbf{\Phi})}|$ denotes the median of the absolute values of the entries of $\mathbf{C}_{36}(\mathbf{\Phi})$.

One can apply the techniques in[6,9] to identify the dynamic factors based on the spectral decomposition of $\mathbf{C}_{36}(\mathbf{\Phi})$ that can be used to explain more than 90% of the variation of the financial margins time series data under consideration.

1. *Dynamic factors and sparse autoregressive reservoir computers*

For the particular case considered in this section, it is possible to identify one dominant dynamic factor $\{\varphi(t) : 0 \leq t \leq 79\}$ that explains more than 90% of the variation of the financial margins time series data $\Sigma_{80}(\mathbf{\Phi})$, and its dynamics and contribution to the dynamic behavior of $\Sigma_{80}(\mathbf{\Phi})$ can be approximately identified using a combination of SRRC models that extends the ideas presented in[14], and is determined by the device depicted in the diagram:



(V.6)

Where $\mathbf{\Theta}_1$ and $\mathbf{\Theta}_2$ are SRRC models with embedding degree 1 and 2, respectively, trained using the sample $\Sigma_{36}(\mathbf{\Phi})$ to approximately identify the mapping $\mathbf{\Theta}(\boldsymbol{\varphi}_L(t)) = \varphi(t+1)$.

The block $\mathbf{\Omega}_{\mathbf{G}}$ can be identified using a sample $\Sigma_{43}(\boldsymbol{\phi}) := \{\boldsymbol{\phi}(t) : 0 \leq t \leq 42\}$, where the i-th component $\boldsymbol{\phi}(t)[i]$ of each $\boldsymbol{\phi}(t)$ corresponds to an estimate of $\varphi(t+1)$ obtained by applying the i-th model $\mathbf{\Theta}_i$ to $\boldsymbol{\varphi}_L(t)$, for $i = 1, 2$.

To identity the model block $\mathbf{\Sigma}$ one needs to solve a structured matrix equation of the form:

$$\hat{\mathbf{S}} \begin{bmatrix} | & & | \\ \boldsymbol{\phi}(0) & \cdots & \boldsymbol{\phi}(42) \\ | & & | \end{bmatrix} = \begin{bmatrix} \varphi(0) & \cdots & \varphi(42) \end{bmatrix} \tag{V.7}$$

for $\hat{\mathbf{S}}$. And to identify the model block one needs to solve the structured matrix equation:

$$\mathbf{G}\mathbf{W}\hat{\mathbf{S}} \begin{bmatrix} | & & | \\ \boldsymbol{\phi}(0) & \cdots & \boldsymbol{\phi}(42) \\ | & & | \end{bmatrix} = \begin{bmatrix} | & & | \\ \mathbf{\Phi}(0) & \cdots & \mathbf{\Phi}(42) \\ | & & | \end{bmatrix} \tag{V.8}$$

for $\mathbf{G}$. Here, the matrix $\mathbf{W}$ has been computed as part of the dynamic factor identification process, and the structure of the matrix $\mathbf{G}$ is determined by the relational graph $G_{36}(\mathbf{\Phi})$), in the sense that $\mathbf{G}$ has to satisfy (V.8) and the constraint:

$$\mathbf{A}_{36}(\mathbf{\Phi}) \odot \mathbf{G} = \mathbf{G}$$

The graphical representation of the identified coupling matrices corresponding to the model described in (V.6) are presented in Figure 8.

The identified dynamic factor is illustrated in Figure 9.

From the 15 financial margins dynamics of commercial banks that have been identified using model (V.6), three have been chosen for the visualization of the corresponding residual error distributions in Figure 10.

C. Regressive identification of financial margins signals

The models of the form (III.14) can also be used to learn the relation betwwen the behavior of the financial margin signal $\mathbf{\Phi}(t)[i]$ of a given bank i, and the financial margins signals of all other banks except the i-th bank.

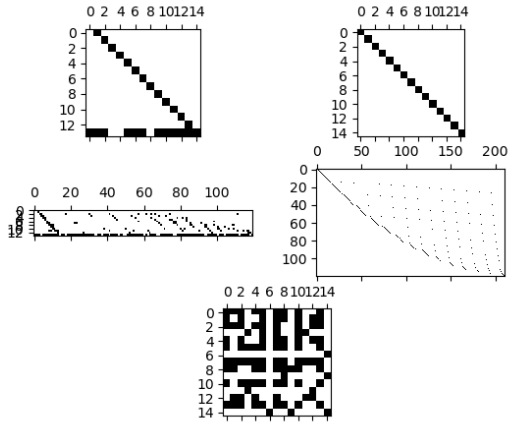


Figure 8: Output coupling matrix decompositions corresponding to: model $\mathbf{\Theta}_1$ (top), model $\mathbf{\Theta}_2$ (middle), and the output coupling matrix $\mathbf{G}$ for the model block $\mathbf{\Omega}_{\mathbf{G}}$. (bottom)

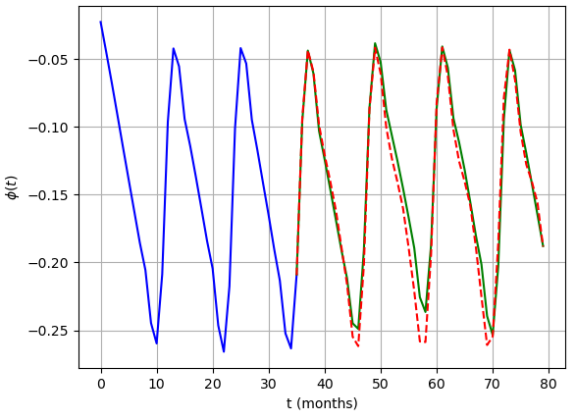


Figure 9: Identified dynamic factor. The blue line corresponds to training data, the green line to reference data, and red dashed line to the identified dynamic behavior data.

For the simulations documented in this section we have considered the financial margins signals $\mathbf{\Phi}(t)[1]$ and $\mathbf{\Phi}(t)[15]$, the SRRC models used for their identification are described as follows.

$$\begin{bmatrix} \mathbf{\Phi}(t-1)[1] \\ \mathbf{\Phi}(t)[1] \end{bmatrix} := \hat{W} R_{2,2}(14) \eth_2 \left(\begin{bmatrix} \mathbf{\Phi}(t-1)[2] \\ \mathbf{\Phi}(t)[2] \\ \vdots \\ \mathbf{\Phi}(t-1)[15] \\ \mathbf{\Phi}(t)[15] \end{bmatrix} \right)$$

$$\begin{bmatrix} \mathbf{\Phi}(t-1)[15] \\ \mathbf{\Phi}(t)[15] \end{bmatrix} := \hat{W} R_{2,2}(14) \eth_2 \left(\begin{bmatrix} \mathbf{\Phi}(t-1)[1] \\ \mathbf{\Phi}(t)[1] \\ \vdots \\ \mathbf{\Phi}(t-1)[14] \\ \mathbf{\Phi}(t)[14] \end{bmatrix} \right) \tag{V.9}$$

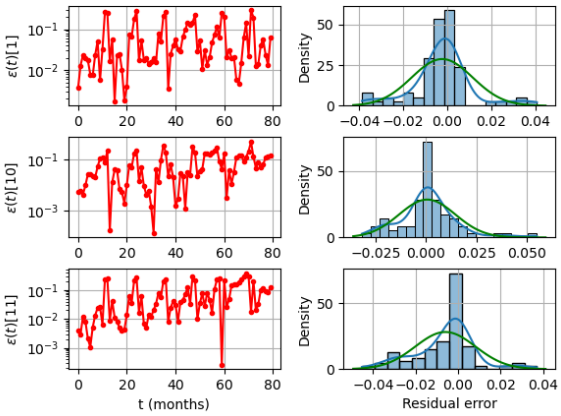


Figure 10: Identified financial margin dynamics for three commercial banks (left), the corresponding residual error densities; where the bars represent the empirical densities, and the blue and green lines correspond to the kernel density estimate and normal approximation of the error densities, respectively (right).

For each $t \geq 2$, let us denote by $\mathbf{x}(t)$ the vector output signal in $\mathbb{R}^{435}$ determined by $R_{2,2}(14)$ and $\eth_2$ for any suitable input signal $\mathbf{u}(t) \in \mathbb{R}^{28}$, at time step t, according to the following expression.

$$\mathbf{x}(t) := R_{2,2}(14)\eth_2(\mathbf{u}(t)) \tag{V.10}$$

The approximation errors and model parameters corresponding to this system identification processes are illustrated in Figure 11.

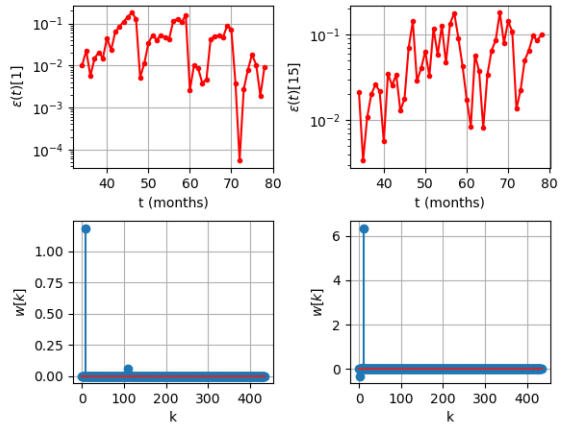


Figure 11: Financial margins identification.

The data and modeling techniques discussed in these sections reveal that within the Honduran commercial banking system, there is a banking institution with one of the highest intermediation margins. However, this institution also exhibits the second highest default rate. This situation arises from the specific market segment in which the institution engages in credit placement, characterized by a higher credit risk.

The institution in question, due to its perception of risk, sets notably higher active interest rates compared to other banks in the system. Conversely, the bank with the lowest rate of non-performing loans has one of the smallest financial margins in the commercial banking system.

Moreover, in line with one of the postulates of post-Keynesian theory[3], which posits that credit placement by commercial banks is driven not by the availability of savings but by the anticipated profitability of companies that ensures repayment of the loans, it becomes clear that a primary goal of Honduran banks, akin to any business, is the maximization of profits.

Aligned with the aforementioned insights, high financial intermediation margins in the Honduran banking sector could indicate inefficiencies and an accumulation of risks. Considering the profit expectations of the institutions, the financial margin needs to be sufficiently large to cover operating expenses and achieve the desired profit level.

Banks, aiming for higher returns, often extend credit to sectors of the economy associated with higher risk levels. This increased risk perception leads to higher interest rates, subsequently elevating the cost of loanable funds. This situation particularly affects those banks whose deposit base is insufficient to satisfy credit demand, compelling them to rely on interbank loans, which in turn escalates their financial costs.

This combination of factors influences the approximate synchronization that one can observe for the financial margins of commercial banks in the Honduran financial system. This is indicative of a dominant dynamic factor explaining more than 90% of data variation. In terms of competition, banks with the largest market shares in both loans and deposits, accounting for over 80% of the market, set interest rates based on their profit expectations and operating expenses, considering the monetary policy stance. This behavior is noted by other banks, compelling them to emulate it. This replication is primarily driven by an increase in the cost of funding (interbank rates), which is then passed on to the rates to cover the additional costs, thereby continuing to generate profits and maintaining their market share in both credit and deposit sectors.

The computational setting used for the experiments performed in this section is documented in the Python program Python program `FinancialMarginsDynamics.py` in[15] that can be used to replicate these experiments.

VI. CONCLUSIONS

The results in §III A and §III B in the form of algorithms like the ones described in §IV, can be effectively used for the sparse structured identification of financial dynamical models that can be used to compute data-driven predictive and prescriptive numerical simulations.

The integration of Sparse Regressive Reservoir Computing (SRRC) models into the forecasting and assessment of financial margin dynamics has shown remarkable promise, especially in contexts with limited training data. The sparse representations crucial for identifying the models' parameter

matrices make SRRC models exceptionally adept at working with time series that have a relatively low volume of available training data. This attribute is particularly valuable in the financial sector, where data scarcity can often be a challenge. In the Honduran commercial banking sector, for instance, SRRC models have efficiently managed to capture and analyze the dynamics of financial margins despite the constraints of data availability.

From the perspective of regulatory bodies, such as the National Commission of Banks and Insurance Companies of Honduras, the inherent nature of reservoir computing in these models is a significant advantage. SRRC models are tailored to capture complex and nonlinear interactions between financial variables, offering deep insights into the interdependencies and influences within the banking system. This capability is crucial for regulatory oversight, as it aids in understanding the subtleties of market behavior and risk factors. The models provide a robust analytical tool for monitoring, regulation, and policy-making, ensuring that regulatory bodies are equipped with accurate and comprehensive analyses to oversee and guide the banking sector effectively.

VII. FUTURE DIRECTIONS

The extension of sparse RRC modeling techniques to equivariant system identification will be studied in future communications. Further implementations of the structured sparse model identification algorithms presented in this document to compute data-driven dynamic general equilibrium models will be the subject of future communications.

DATA AVAILABILITY

The programs and data that support the findings of this study will be openly available in the DyNet-CNBS repository, reference number[15], in due time.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

ACKNOWLEDGMENT

The structure preserving matrix computations needed to implement the algorithms in §IV, were performed with Python 3.10.4, with the support and computational resources of the National Commission of Banks and Insurance Companies (**CNBS**) of Honduras. The views expressed in the article do not necessarily represent the views of the National Commission of Banks and Insurance Companies of Honduras.

REFERENCES

[1] Niluthpaul Sarker Anupam Das Gupta1 and Mohammad Rifat Rahman. Relationship among cost of financial intermediation, risk, and efficiency: Empirical evidence from bangladeshi commercial banks. *DCogent Economics Finance*, 2021.

[2] Rajat Budhiraja, Manish Kumar, Mrinal K. Das, Anil Singh Bafila, and Sanjeev Singh. A reservoir computing approach for forecasting and regenerating both dynamical and time-delay controlled financial system behavior. *PLOS ONE*, 16(2):1–24, 02 2021.

[3] Jesús C. and José Luís . El margen financiero de la banca comercial en méxico 1995-2005. *Quivera. Revista de Estudios Territoriales*, 9:155–169, 2007.

[4] Jean François Clevy and Rossana Díaz. Determinantes del spread bancario en nicaragua. Technical report, Banco Central de Nicaragua, 2005.

[5] Matthew Dixon, Diego Klabjan, and Jin Hoon Bang. Classification-based financial markets prediction using deep neural networks, 2017.

[6] Mario Forni, Marc Hallin, Marco Lippi, and Lucrezia Reichlin. The Generalized Dynamic-Factor Model: Identification and Estimation. *The Review of Economics and Statistics*, 82(4):540–554, 11 2000.

[7] Daniel J. Gauthier, Erik Bollt, Aaron Griffith, and Wendson A. S. Barbosa. Next generation reservoir computing. *Nature Communications*, 12(1):5564, Sep 2021.

[8] Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, Baltimore, 3rd edition, 1996.

[9] M. Kryshko. Data-rich dsge and dynamic factor models. IMF Working Papers WPIEA2011216, International Monetary Fund, 2011.

[10] Balasubramanya T. Nadiga. Reservoir computing as a tool for climate predictability studies. *Journal of Advances in Modeling Earth Systems*, 13(4):e2020MS002290, 2021. e2020MS002290 2020MS002290.

[11] DAVID RICARDO PINEDA. Determinantes del spread bancario en honduras. Technical report, Banco Central de Honduras, 2010.

[12] Tigran Poghosyan. Financial intermediation costs in low income countries: The role of regulatory, institutional, and macroeconomic factors. *Economic Systems*, 37(1):92–110, 2013.

[13] Justin Sirignano and Rama Cont. Universal features of price formation in financial markets: perspectives from deep learning, 2018.

[14] Fredy 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.

[15] Fredy Vides. Dynet-cnbs: Network dynamics identification and reservoir computing tools with financial applications. 2023. https://github.com/FredyVides/DyNet-CNBS.

[16] Ye 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.