Scientific Bulletin, Series A: Applied Mathematics and Physics, ISSN 1223-7027, vol.64, issue no.2, pp. 35-44, University "Politehnica" of Bucharest, 2002.

# Implicit moments matching through Krylov subspaces projections methods. Arnoldi methods

N. POPESCU-BODORIN*

**MSC(2000)**: 93A15, 93B40, 65L99
**Key words**: reduced-order modeling techniques, moments matching, Padé approximation, Krylov subspaces projections methods, Arnoldi methods.

***Rezumat.*** *Această lucrare se doreşte a fi prima dintr-o serie de studii ale autorului asupra metodelor de reducere a ordinului cu suprapunere implicită de momente. Scopul ei este :*
*i) să descrie sistemul redus şi relațiile matriceale între sistemul initial şi cel redus obținut prin algoritmii iterativi Arnoldi şi Block-Arnoldi.*
*ii) să propună 'Padé via Arnoldi' şi 'Padé via Block-Arnoldi' ca metode de reducere a ordinului cu suprapunere implicită de momente Markov.*
*iii) să fie premergătoare expunerii metodelor Lanczos, Block-Lanczos, Two-Step Lanczos şi a Metodelor Raționale.*

*Abstract. This paper is intended to be the first step of the author in a series of studies on reduced-order modeling and implicit moments matching methods. Its purpose is to:*
*i) describe the reduced system and the matriceal relation between the initial system and the reduced system that results from Arnoldi and Block-Arnoldi iterrative procedures.*
*ii) suggest Padé via Arnoldi and Padé via Block-Arnoldi as order-reducing and implicit moments matching techniques.*
*iii) lead to the study of Lanczos, Block-Lanczos, Two-step Lanczos and of Rational methods*

## Introduction

The practical problem that constitues the basis of this paper is to describe the functionning of passive interconnecting networks in high frequency conditions.
Assuming the existence of a unique solution for electromagnetic field, initial and limit conditions, if we apply one of the **FIT**( Finite Integration Tehnique)**, FDTD**( Finite Diferences Time Domain)**, FEM**( Finite Element Method)**, BEM**( Boundary Element

*The author is student of Master Courses 'Geometry and Dynamic Systems', Mathematics Department n°.1, 'Politehnica' University, Bucharest (2002).

Scientific Bulletin, Series A: Applied Mathematics and Physics, ISSN 1223-7027, vol.64, issue no.2, pp. 35-44, University "Politehnica" of Bucharest, 2002.

Method) discretization methods we obtain the following equations for a time-invariant linear dynamic large scale N-dimensional MIMO system as in [1]:

$$\begin{cases} \mathbf{C}\dot{\mathbf{x}}(t) = -\mathbf{G}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) = \mathbf{E}^{\mathbf{T}}\mathbf{x}(t) \end{cases} \tag{1}$$

where :

i) the dimension of the matrices are :$\mathbf{C},\mathbf{G}\in\mathbf{R}^{NxN}$ ; $\mathbf{B}\in\mathbf{R}^{Nxm}$ ; $\mathbf{E}\in\mathbf{R}^{Nxp}$ ;

ii) the time-dependent vectorial functions are : inputs $\mathbf{u}(t)\in\mathbf{R}^{m}$ , outputs $\mathbf{y}(t)\in\mathbf{R}^{p}$ , state vector $\mathbf{x}(t)\in\mathbf{R}^{N}$ ;

iii) the matrix $\mathbf{C}$ is singular but $\mathbf{G}+s_0\mathbf{C}$ is a regular pencil (i.e. $\exists$ $s_0\in\mathbf{C}$ so that matrix $\mathbf{G}+s_0\mathbf{C}$ is numerically nonsingular)

Like in [2], applying the Laplace transformation $g(s) = \int_0^\infty g(t)e^{-st}dt$

to system (1) and assuming $\mathbf{x}(0)=0$ we obtain the following equivalent equations in the frequency domain :

$$\begin{cases} s\mathbf{C}\ \mathbf{x}(s) = -\mathbf{G}\ \mathbf{x}(s) + \mathbf{B}\ \mathbf{u}(s) \\ \mathbf{y}(s) = \mathbf{E}^{\mathbf{T}}\mathbf{x}(s) \end{cases} \tag{2}$$

If we eliminate the state vector $\mathbf{x}$ in (2) we obtain the input-output equation $\mathbf{y}(s)=\mathbf{H}(s)\mathbf{u}(s)$ where the transfer function $\mathbf{H}$ is given by $\mathbf{H}(s)=\mathbf{E}^{T}(\mathbf{G}+ s_0\ \mathbf{C})^{-1}\mathbf{B}$

The main idea of related works [1] and [4] is the following: given an initial large scale system there should be found another linear system of reduced order whose input-output behavior is a good approximation of the first one's. Consequently the reduced model of the initial system (1) is

$$\begin{cases} \mathbf{C}_n\dot{\mathbf{x}}(t) = -\mathbf{G}_n\mathbf{x}(t) + \mathbf{B}_n\mathbf{u}(t) \\ \mathbf{y}(t) = \mathbf{E}_n^{T}\mathbf{x}(t) \end{cases} \text{ respectively } \begin{cases} s\mathbf{C}_n\mathbf{x}(s) = -\mathbf{G}_n\mathbf{x}(s) + \mathbf{B}_n\mathbf{u}(s) \\ \mathbf{y}(s) = \mathbf{E}_n^{T}\mathbf{x}(s) \end{cases} \tag{3),(4}$$

having the matriceal dimensions reduced to $\mathbf{C}_n,\mathbf{G}_n\in\mathbf{R}^{nxn}$ , $\mathbf{B}_n\in\mathbf{R}^{nxm}$ , $\mathbf{E}_n\in\mathbf{R}^{nxp}$ and satisfying the following conditions :

i) the order of the reduced model is smaller (preferably much smaller) than the order of the initial system : n << N

ii) it preserves the properties of stability, controlability, observability şi pasivity of the initial system;

iii) the reduced model is Padé-approximant for the initial system, which means that $\mathbf{H}_n(s) = \mathbf{E}_n(\mathbf{G}_n + s\mathbf{C}_n)^{-1}\mathbf{B}_n$ must be a Padé approximation of $\mathbf{H}$ .All this will be

dealt with in detail in the following chapter.

## 1. Preliminaries

### 1.1. Padé approximation.

Let **H**(s) be the transfer function of the initial system and consider its Taylor expansion around the complex value $\mathbf{s_0}$ : $\mathbf{H}(s) = \sum_{0}^{\infty} \mathbf{M}_i (s - s_0)^i$ (5)

The matriceal coeficients $\mathbf{M}_i \in \mathbf{R}^{pxm}$ are the Markov moments of **H(s)** corresponding to the complex value $s_0$. The matriceal function $\mathbf{H}_n(s) \in \mathbf{R}^{pxm}$ is a **q**-order Padé approximation of **H**(s) if their Taylor expansions are identical for the first **q** coeficients. I.e :

$$\mathbf{H}_n(s) = \mathbf{H}(s) + \theta((s - s_0)^q) = \sum_{i=0}^{q-1} \mathbf{M}_i (s - s_0) \quad (6)$$

### 1.2.Markov moments expresions for MIMO systems.

Consider :

i) $s_0 \in \mathbf{C}$ the Taylor expansion point for the transfer function **H**(s) (5).

ii) $\mathbf{F}_1, \mathbf{F}_2 \in \mathbf{C}^{NxN}$ two nonsingular matrices that are a factorization of $\mathbf{G}+\mathbf{s}_0\mathbf{C}$ matrix

$$\mathbf{G} + s_0\mathbf{C} = \mathbf{F}_1\mathbf{F}_2 \quad (7)$$

Let us denote : $\mathbf{A} = -\mathbf{F}_1^{-1}\mathbf{C}\mathbf{F}_2^{-1}$; $\mathbf{R} = \mathbf{F}_1^{-1}\mathbf{B}$ ; $\mathbf{L} = \mathbf{F}_2^{-T}\mathbf{E}$ ; $s = s_0 + \sigma$ ; (8)

hence **H**(s) can be rewritten as : $\mathbf{H}(s) = \mathbf{L}^T(\mathbf{I}-(s - s_0)\mathbf{A})^{-1}\mathbf{R} = \mathbf{L}^T(\mathbf{I}-\sigma\mathbf{A})^{-1}\mathbf{R}$ (9)

and consequently : $\mathbf{H}(s) = \mathbf{L}^T(\sum_{i=0}^{\infty} A^i (s - s_0)^i)\mathbf{R} = \sum_{i=0}^{\infty} \mathbf{L}^T \mathbf{A}^i \mathbf{R} \sigma^i$ (10)

where the matriceal coeficients $\mathbf{M}_i = \mathbf{L}^T\mathbf{A}^i\mathbf{R}$ are Markov moments.

The relation (10) proves that describing the input-output behaviours of system (2) around $s=s_0$ is equivalent with describing the input-output behavior around s = 0 of the system:

$$\begin{cases} (\mathbf{I} - \sigma\mathbf{A})\mathbf{x}(\sigma) = \mathbf{R}\mathbf{u}(\sigma) \\ \mathbf{y}(\sigma) = \mathbf{L}^T\mathbf{x}(\sigma) \end{cases}$$

### 1.3.Markov moments expresions for SISO systems.

Consider the linear dynamic **SISO** system : $\begin{cases} s\mathbf{C}\mathbf{x}(s) = -\mathbf{G}\mathbf{x}(s) + \mathbf{b}\mathbf{u}(s) \\ \mathbf{y}(s) = \mathbf{e}^T\mathbf{x}(s) \end{cases}$ (11)

Scientific Bulletin, Series A: Applied Mathematics and Physics, ISSN 1223-7027, vol.64, issue no.2, pp. 35-44, University "Politehnica" of Bucharest, 2002.

whose transfer function is $\mathbf{H}(s) = \mathbf{e}^T(\mathbf{G}+s\mathbf{C})^{-1}\mathbf{b}$. (12)

Having $s_0 \in \mathbf{C}$, şi $\mathbf{F}_1$, $\mathbf{F}_2 \in \mathbf{C}^{NxN}$ with the same meaning as in **MIMO** case

and denoting: $\mathbf{A} = -\mathbf{F}_1^{-1}\mathbf{C}\mathbf{F}_2^{-1}$ ; $\mathbf{r} = \mathbf{F}_1^{-1}\mathbf{b}$ ; $\mathbf{l} = \mathbf{F}_2^{-T}\mathbf{e}$ ; $s = s_0 + \sigma$ (13)

the transfer function $\mathbf{H}(s)$ can be rewritten as

$$\mathbf{H}(s_0+\sigma) = \mathbf{l}^T(\sigma\,\mathbf{C} + \mathbf{G} + s_0\mathbf{C})^{-1}\mathbf{b} = \mathbf{l}^T(\mathbf{I} - \sigma\mathbf{A})^{-1}\mathrm{r} = \sum_{k=0}^{\infty}\mathbf{l}^T\mathbf{A}^k\mathbf{r}\sigma^k = \sum_{k=0}^{\infty} m_k\sigma^k . \quad (14)$$

Consequently, in this particular case the Markov moments are real numbers $\mathbf{m}_k = \mathbf{l}^T\mathbf{A}^k\mathbf{r}$ , where $k \geq 0$. Relation (14) proves that describing the input-output behaviours of system (11) around $s=s_0$ is equivalent with describing the input-output behavior around $s = 0$ of

the system: $$\begin{cases} (\mathbf{I} - \sigma\mathbf{A})\mathbf{x}(\sigma) = \mathbf{r}\mathbf{u}(\sigma) \\ \mathbf{y}(\sigma) = \mathbf{l}^T\mathbf{x}(\sigma) \end{cases} \quad (15)$$

## 2. Reducing order methods based on Krylov subspaces projections

### 2.1.Arnoldi-SISO method.

(with Classic-Gram-Schmidt orthogonalization method, having a single starting vector)

Briefly :

Having as a starting point the SISO system

$$\begin{cases} \mathbf{C}\dot{\mathbf{x}}(t) = -\mathbf{G}\mathbf{x}(t) + \mathbf{b}\mathbf{u}(t) \\ \mathbf{y}(t) = \mathbf{l}^T\mathbf{x}(t) \end{cases} \quad \text{respectively} \quad \begin{cases} s\mathbf{C}\mathbf{x}(s) = -\mathbf{G}\mathbf{x}(s) + \mathbf{b}\mathbf{u}(s) \\ \mathbf{y}(s) = \mathbf{l}^T\mathbf{x}(s) \end{cases} \quad (16)$$

Arnoldi method sets an orthogonal matrix $\mathbf{V_n} \in \mathbf{R}^{\mathbf{Nxn}}$ which determines with each iteration n a reduced n-dimensional model of the initial system :

$$\begin{cases} \mathbf{C}_n\dot{\mathbf{x}}(t) = -\mathbf{G}_n\mathbf{x}(t) + \mathbf{b}_n\mathbf{u}(t) \\ \mathbf{y}(t) = \mathbf{l}_n^T\mathbf{x}(t) \end{cases} \quad \text{respectively} \quad \begin{cases} s\mathbf{C}_n\mathbf{x}(s) = -\mathbf{G}_n\mathbf{x}(s) + \mathbf{b}_n\mathbf{u}(s) \\ \mathbf{y}(s) = \mathbf{l}_n^T\mathbf{x}(s) \end{cases} \quad (17)$$

where $\mathbf{G}_n = \mathbf{V}_n^T\mathbf{G}\mathbf{V}_n$ ; $\mathbf{C}_n = \mathbf{V}_n^T\mathbf{C}\mathbf{V}_n$ ; $\mathbf{b}_n = \mathbf{V}_n^T\mathbf{b}$ ; $\mathbf{l}_n = \mathbf{V}_n^T\mathbf{l}$ ; (18)

#### 2.1.1. The Arnoldi-SISO algorithm.

Consider **A**, **r** as defined in (13) with $\mathbf{F_1}$, $\mathbf{F_2}$ being the trivial factorization of $\mathbf{(G+s_0C)}$ **:** $\mathbf{F}_1\mathbf{F}_2 = (\mathbf{G}+\mathbf{s}_0\mathbf{C})\mathbf{I}_N$ and $\mathbf{K}_m(\mathbf{A},\mathbf{r})$=span$\{\mathbf{r}, \mathbf{Ar}, \mathbf{A}^2\mathbf{r}, \ldots, \mathbf{A}^{m-1}\mathbf{r}\}$ as the m-order Krylov subspace generated by **A** and **r**. The following iterrative algorithm [3] sets up an orthonormal basis for $\mathbf{K}_m(\mathbf{A},\mathbf{r})$ :

Scientific Bulletin, Series A: Applied Mathematics and Physics, ISSN 1223-7027, vol.64, issue no.2, pp. 35-44, University "Politehnica" of Bucharest, 2002.

1. $\mathbf{v}_1 = \dfrac{\mathbf{r}}{\|\mathbf{r}\|}$
2. For j=1, 2, … , m Do
    3. For $i = \overline{1,j} : h_{ij} = \langle \mathbf{A}\mathbf{v}_j, \mathbf{v}_i \rangle$
    4. $\bar{\mathbf{v}}_{j+1} = \mathbf{A}\mathbf{v}_j - \sum_{i=1}^{j} h_{ij}\mathbf{v}_i$
    5. $h_{j+1,j} = \|\bar{\mathbf{v}}_{j+1}\|$
    6. If $h_{j+1,j} = 0$ then Stop
    7. $\mathbf{v}_{j+1} = \dfrac{\bar{\mathbf{v}}_{j+1}}{\|\bar{\mathbf{v}}_{j+1}\|}$

8. End Do.

$\mathbf{V}_n$ being the matrix with columns $\mathbf{v}_1$ , $\mathbf{v}_2$ , … , $\mathbf{v}_n$ and considering the relations in steps 3, 4, 7, we obtain the equalities :

$$\begin{cases} \forall i \in \overline{1,n} : \|\mathbf{v}_i\| = 1 \\ \forall i,j \in \overline{1,n} : \langle \mathbf{v}_i, \mathbf{v}_j \rangle = \mathbf{v}_i^T \mathbf{v}_j = \delta_{ij} = \begin{cases} 1; i = j \\ 0; i \neq j \end{cases} \\ \forall j \in \overline{1,n} : \mathbf{A}\mathbf{v}_j = \sum_{i=1}^{j} h_{ij}\mathbf{v}_i + h_{j+1,j}\mathbf{v}_{j+1} \end{cases} ; \begin{cases} \mathbf{V}_n^T \mathbf{V}_n = \mathbf{I}_n \\ \mathbf{A}\mathbf{V}_n = \mathbf{V}_n \mathbf{H}_n + \mathbf{v}_{n+1} h_{n+1,n} \mathbf{e}_n^T \\ \mathbf{V}_n^T \mathbf{A} \mathbf{V}_n = \mathbf{H}_n \end{cases} . \quad (19),(20)$$

where the matrix $\mathbf{H}_n = [h_{ij}]_{1 \le i,j \le n}$ is a Hessenberg one.

### 2.1.2 The properties of the linear operator $\mathbf{P} = \mathbf{V}_n \mathbf{V}_n^T$ ; $\mathbf{P}:\mathbf{R}^N \rightarrow \mathbf{K}_n(\mathbf{A},\mathbf{r})$.

i) $\mathbf{P}^2 = \mathbf{P}$.

ii) Let $\mathbf{v}_i$ , $i \in \overline{1,n}$, be an arbitrary column of matrix $\mathbf{V}_n$ (an arbitrary vector of the basis $\mathbf{V}_n$ of the Krylov subspace $\mathbf{K}_n(\mathbf{A},\mathbf{r})$=spancol($\mathbf{V}_n$) that was generated by the first **n** iterrations of the algorithm).

Scientific Bulletin, Series A: Applied Mathematics and Physics, ISSN 1223-7027, vol.64, issue no.2, pp. 35-44, University "Politehnica" of Bucharest, 2002.

Then : $$\mathbf{P}\mathbf{v}_i = \mathbf{V}_n\mathbf{V}_n^T\mathbf{v}_i = \mathbf{V}_n\begin{bmatrix}\mathbf{v}_1^T\mathbf{v}_i\\ \mathbf{v}_2^T\mathbf{v}_i\\ \dots\dots\\ \mathbf{v}_n^T\mathbf{v}_i\end{bmatrix} = \mathbf{V}_n\mathbf{e}_i = \mathbf{v}_i \ ;$$

hence : $\mathbf{P}|_{\mathbf{K}_n(\mathbf{A},\mathbf{r})} = \mathbf{I}_N$ ; (21)

iii) $\forall \mathbf{w} \in \mathbf{K}_n^{\perp}(\mathbf{A},\mathbf{r}) : \mathbf{P}\mathbf{w} = 0$ ;

Therefore $\mathbf{P} = \mathbf{V}_n\mathbf{V}_n^T$ is the orthogonal projector onto $\mathbf{K}_n(\mathbf{A},\mathbf{r})$.

### 2.1.3 The reduced-order model realization through the Arnoldi algorithm and Padé connexion.

The following will demonstrate that the relations (13), (14), (20), (21) in themselves determine an n-dimensional reduced model of the initial system.

$$\mathbf{V}_n^T\mathbf{A}\mathbf{V}_n = \mathbf{H}_n \Rightarrow -\mathbf{V}_n^T(\mathbf{G}+s_0\mathbf{C})^{-1}\mathbf{C}\mathbf{V}_n = \mathbf{H}_n \Rightarrow -\mathbf{V}_n^T(\mathbf{G}+s_0\mathbf{C})^{-1}\mathbf{V}_n\mathbf{V}_n^T\mathbf{C}\mathbf{V}_n = \mathbf{H}_n \Rightarrow$$

$$\Rightarrow -(\mathbf{V}_n^T(\mathbf{G}+s_0\mathbf{C})\mathbf{V})^{-1}\mathbf{V}_n^T\mathbf{C}\mathbf{V}_n = \mathbf{H}_n \Rightarrow -(\mathbf{V}_n^T\mathbf{G}\mathbf{V}_n + s_0\mathbf{V}_n^T\mathbf{C}\mathbf{V}_n)^{-1}\mathbf{V}_n^T\mathbf{C}\mathbf{V}_n = \mathbf{H}_n \Rightarrow$$

$$\Rightarrow \mathbf{A}_n = -(\mathbf{G}_n + s_0\mathbf{C}_n)^{-1}\mathbf{C}_n = \mathbf{H}_n$$

where $\mathbf{G}_n = \mathbf{V}_n^T\mathbf{G}\mathbf{V}_n$ şi $\mathbf{C}_n = \mathbf{V}_n^T\mathbf{C}\mathbf{V}_n$.

On the other hand $\forall k \in \overline{1,n}$ :

$$m_k = \mathbf{l}^T\mathbf{A}^k\mathbf{r} = \mathbf{l}^T\mathbf{V}_n\mathbf{H}_n^k\mathbf{V}_n^T\mathbf{r} = (\mathbf{V}_n^T\mathbf{l})^T\mathbf{H}_n^k\mathbf{V}_n^T(\mathbf{G}+s_0\mathbf{C})^{-1}\mathbf{b} =$$

$$= (\mathbf{V}_n^T\mathbf{l})^T\mathbf{H}_n^k(\mathbf{V}_n^T\mathbf{G}\mathbf{V}_n + s_0\mathbf{V}_n^T\mathbf{C}\mathbf{V}_n)^{-1}\mathbf{V}_n^T\mathbf{b} = \mathbf{l}_n^T\mathbf{H}_n^k(\mathbf{G}_n + s_0\mathbf{C}_n)^{-1}\mathbf{b}_n$$

where $\mathbf{l}_n = \mathbf{V}_n^T\mathbf{l}$, $\mathbf{b}_n = \mathbf{V}_n^T\mathbf{b}$, $\mathbf{r}_n = (\mathbf{G}_n + s_0\mathbf{C}_n)^{-1}\mathbf{b}_n$, $\overline{\mathbf{m}}_k = \mathbf{l}_n^T\mathbf{A}_n^k\mathbf{r}_n$.

So, $m_k = \mathbf{l}^T\mathbf{A}^k\mathbf{r} = \mathbf{l}_n^T\mathbf{A}_n^k\mathbf{r}_n = \overline{m}_k$.

In conclusion, we proved the following theorem :

**Theorem 1**(Padé via Arnoldi) **:** The setting up of the matrix $\mathbf{V}_n$ , by which the Arnoldi iterrative process applied to the initial system (16) results in a reduced n-dimensional model (17) after n iterrations , is equivalent with finding a local n-order Padé approximation of the transfer function of the initial system , and also with an orthogonal projection of the initial system onto Krylov subspace $\mathbf{K}_n(\mathbf{A},\mathbf{r})$ .

**Remark 1 :**The proof of the above theorem is not influenced by the choice of particular factorization $\mathbf{F}_1$ ,$\mathbf{F}_2$ of regular pencil $\mathbf{G}+s_0\mathbf{C}$ ,but it is obviously influenced by the choice of $s_0$ . Consequently :

i) while in exact arithmetics the choice of factorization influences the form of intermediary systems (**A, r, l**) and ($\mathbf{A}_n$, $\mathbf{r}_n$, $\mathbf{l}_n$) but does not affect Markov moments matching, in real arithmetics (during implementation) , computing effort and time as well as Markov moments matching errors vary with the choice of factorization;

ii) the theorem has a local character and it depends on (a frequency band around) $s_0$. It will be very useful in practice to have the answer for «How wide could the frequency band around $s_0$ be so that the reduced model can be a good approximation of the initial system?» or for «How many points are relevant in a certain frequency band and which are those?».

**Remark 2 :**In order to result in an n-dimensional reduced model the algorithm must not have a breakdown at step 6. If the algorithm does stop at iterration j then the degree of **r** relative to **A** is j and $\mathbf{K}_n(\mathbf{A},\mathbf{r})= \mathbf{K}_j(\mathbf{A},\mathbf{r})$, $\forall n>j$ .[3]

### 2.2. Block-Arnoldi-MIMO method.

(with Classic-Block-Gram-Schmidt orthogonalization, having multiple starting vectors)

Briefly :

Having as a starting point the MIMO system

$$\begin{cases}\mathbf{C}\dot{\mathbf{x}}(t) = -\mathbf{G}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t)\\ \mathbf{y}(t) = \mathbf{L}^{\mathrm{T}}\mathbf{x}(t); \mathbf{B} \in \mathbf{R}^{\mathrm{N}\times \mathrm{m}}\end{cases} \quad \text{respectively} \quad \begin{cases}s\mathbf{C}\mathbf{x}(s) = -\mathbf{G}\mathbf{x}(s) + \mathbf{B}\mathbf{u}(s)\\ \mathbf{y}(s) = \mathbf{L}^{\mathrm{T}}\mathbf{x}(s)\end{cases} \tag{22}$$

Block-Arnoldi method sets an orthogonal matrix $\mathbf{U}_k\in\mathbf{R}^{\mathrm{Nxkm}}$ which determins with each iteration k a reduced n-dimensional model (n=km) of the initial system:

$$\begin{cases}\mathbf{C}_{\mathrm{n}}\dot{\mathbf{x}}(t) = -\mathbf{G}_{\mathrm{n}}\mathbf{x}(t) + \mathbf{B}_{\mathrm{n}}\mathbf{u}(t)\\ \mathbf{y}(t) = \mathbf{L}_{\mathrm{n}}^{\mathrm{T}}\mathbf{x}(t)\end{cases} \quad \text{respectively} \quad \begin{cases}s\mathbf{C}_{\mathrm{n}}\mathbf{x}(s) = -\mathbf{G}_{\mathrm{n}}\mathbf{x}(s) + \mathbf{B}_{\mathrm{n}}\mathbf{u}(s)\\ \mathbf{y}(s) = \mathbf{L}_{\mathrm{n}}^{\mathrm{T}}\mathbf{x}(s)\end{cases} \tag{23}$$

where $\mathbf{G}_{\mathrm{n}} = \mathbf{U}_{\mathrm{k}}^{\mathrm{T}}\mathbf{G}\mathbf{U}_{\mathrm{k}}$, $\mathbf{C}_{\mathrm{n}} = \mathbf{U}_{\mathrm{k}}^{\mathrm{T}}\mathbf{C}\mathbf{U}_{\mathrm{k}}$, $\mathbf{B}_{\mathrm{n}} = \mathbf{U}_{\mathrm{k}}^{\mathrm{T}}\mathbf{B}$, $\mathbf{L}_{\mathrm{n}} = \mathbf{U}_{\mathrm{k}}^{\mathrm{T}}\mathbf{L}$, n=km. (24)

#### 2.2.1. Block-Arnoldi-MIMO algorithm.

Block-Arnoldi method is the obvious generalization of the Arnoldi-SISO method , but it is able to handle m starting vectors ($\mathbf{B}\in\mathbf{R}^{\mathrm{Nxm}}$).

Consider **A**, **R** defined in (7) (8) with $\mathbf{F}_1$, $\mathbf{F}_2$being the trivial factorization of $\mathbf{G}+s_0\mathbf{C}$

$$\mathbf{F}_1\,\mathbf{F}_2 = (\mathbf{G}+s_0\mathbf{C})\mathbf{I}_{\mathrm{N}}$$

and $\mathbf{K}_k(\mathbf{A},\mathbf{R})$=span{$\mathbf{R}$, $\mathbf{AR}$, $\mathbf{A}^2\mathbf{R}$, … , $\mathbf{A}^{k-1}\mathbf{R}$} being the k-order block-Krylov subspace generated by **A** and **R**.

The following iterrative algorithm [3] sets up an orthonormal basis of $\mathbf{K}_k(\mathbf{A},\mathbf{R})$ :

1. Compute Q-R factorization of $\mathbf{R}$ : $\mathbf{R} = \mathbf{V}_1\mathbf{H}_{1,0}$
2. For j=1, 2, … , k Do
   3. For $i = \overline{1,j} : \mathbf{H}_{ij} = \mathbf{V}_i^T\mathbf{A}\mathbf{V}_j$
   4. $\overline{\mathbf{V}}_{j+1} = \mathbf{A}\mathbf{V}_j - \sum_{i=1}^{j}\mathbf{V}_i\mathbf{H}_{ij}$
   5. Compute Q-R factorization of $\overline{\mathbf{V}}_{j+1} : \overline{\mathbf{V}}_{j+1} = \mathbf{V}_{j+1}\mathbf{H}_{j+1,j}$
6. EndDo.

In the first step the algorithm sets up the matrix $\mathbf{V}_1$ having the property that

$$\text{spancol}(\mathbf{V}_1) = \text{spancol}(\mathbf{R}) .$$

The initial set of vectors $\mathbf{V}_1 = [\mathbf{v}_1, \mathbf{v}_2, .....\mathbf{v}_m]$ is an orthogonal basis of $\mathbf{K}_0(\mathbf{A},\mathbf{R})$

Vector-wise, with each iterration j, the algorithm generates the set of vectors $\mathbf{V}_{j+1} = [\mathbf{v}_{mj+1}, \mathbf{v}_{mj+2}, .....\mathbf{v}_{m(j+1)}]$ with the following properties :

i) $\forall k \in \overline{1,m} : \|\mathbf{v}_{mj+k}\| = 1$

ii) $\forall k,p \in \overline{1,m} : \langle \mathbf{v}_{mj+k}, \mathbf{v}_{mj+p}\rangle = \mathbf{v}_{mj+k}^T\mathbf{v}_{mj+p} = \delta_{kp}$

iii) $\forall k \in \overline{1,m}; \forall p \in \overline{1,mj} : \langle \mathbf{v}_{mj+k}, \mathbf{v}_p\rangle = \mathbf{v}_{mj+k}^T\mathbf{v}_p = 0$

iv) $\{\mathbf{v}_1, \mathbf{v}_2, ....., \mathbf{v}_m, .....\mathbf{v}_{mj}, \mathbf{v}_{mj+1}, \mathbf{v}_{mj+2}, .....\mathbf{v}_{m(j+1)}\}$ is an orthogonal basis of $\mathbf{K}_j(\mathbf{A},\mathbf{R})$

Matrix-wise, with each iterration j, the algorithm generates an orthogonal matrix $\mathbf{V}_{j+1}$ that is orthogonal to all matrices $\mathbf{V}_1$, $\mathbf{V}_2$, …. ,$\mathbf{V}_j$ , previously generated. In short, matrices $\mathbf{V}_1$, $\mathbf{V}_2$, …. ,$\mathbf{V}_{j+1}$ , have the following properties :

$$\begin{aligned} &\text{i) } \mathbf{V}_i^T\mathbf{V}_i = \mathbf{I}_m \\ &\text{ii) } \forall i \neq p; i,p \in \overline{1,j+1} : \mathbf{V}_i^T\mathbf{V}_p = \mathbf{O}_m \\ &\text{iii) } [\mathbf{V}_1, \mathbf{V}_2, ....., \mathbf{V}_{j+1}] \text{ is a basis of } \mathbf{K}_{j+1}(\mathbf{A},\mathbf{R}) . \end{aligned} \tag{25}$$

Let us consider the matrices $\mathbf{U}_k = [\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_k]$ , $\mathbf{H}_k = [\mathbf{H}_{ij}]_{1\le i,j\le k}$ , $\mathbf{H}_{ij} = \mathbf{O}_m$ , $\forall i>j+1$ and $\mathbf{E}_k$ the matrix formed by the last m columns in $\mathbf{I}_{km}$ . We notice that the following relations occur :

$$\begin{cases} \text{spancol}(\mathbf{U}_k) = \mathbf{K}_k(\mathbf{A},\mathbf{R})\,;\ \mathbf{A}\mathbf{U}_k = \mathbf{U}_k\mathbf{H}_k + \mathbf{V}_{k+1}\mathbf{H}_{k+1}\mathbf{E}_k^T\,; \\ \mathbf{U}_k^T\mathbf{U}_k = \mathbf{I}_{km}\,;\ \mathbf{U}_k^T\mathbf{A}\mathbf{U}_k = \mathbf{H}_k \in \mathbf{R}^{mk\times mk} \end{cases} \tag{26}$$

Matrix $\mathbf{H}_k$ is not Hessenberg as in the Arnoldi-SISO case but it is banded-Hessenberg, having m subdiagonals instead of one.

### 2.2.2. The properties of the linear operator $\mathbf{P} = \mathbf{U}_k\mathbf{U}_k^T$ ; $\mathbf{P}:\mathbf{R}^N \rightarrow \mathbf{K}_k(\mathbf{A},\mathbf{R})$.

i) $\mathbf{P}^2 = \mathbf{P}$ .

ii) Let $\mathbf{v}_i$ , $i\in\overline{1,n}$, be an arbitrary column of matrix $\mathbf{U}_k$ (an arbitrary vector of the basis block-Krylov subspace $\mathbf{K}_k(\mathbf{A}, \mathbf{R})$ = spancol($\mathbf{U}_k$) that was generated by the first k iterrations of the algorithm ).

Then : $\mathbf{P}\mathbf{v}_i = \mathbf{U}_k\mathbf{U}_k^T\mathbf{v}_i = \sum_{j=1}^{k}\mathbf{V}_j\,\mathbf{V}_j^T\mathbf{v}_i = \mathbf{V}_i\,\mathbf{V}_i^T\mathbf{v}_i = \mathbf{V}_i\mathbf{e}_i = \mathbf{v}_i$ ;

hence : $\mathbf{P}\,|_{\mathbf{K}_k(\mathbf{A},\mathbf{R})} = \mathbf{I}_N$ . (27)

iii) $\forall \mathbf{w} \in \mathbf{K}_k^{\perp}(\mathbf{A},\mathbf{R}) : \mathbf{P}\mathbf{w} = 0$ ;

In conclusion, $P = \mathbf{U}_k\mathbf{U}_k^T$ is the orthogonal projector onto $\mathbf{K}_k(\mathbf{A}, \mathbf{R})$.

### 2.2.3. The reduced-order model realization through Block-Arnoldi algorithm and Padé connexion.

The following will demonstrate that relations (26), (27), in themselves determine an n-dimensional reduced model of the initial system.

$$\mathbf{U}_k^T\mathbf{A}\mathbf{U}_k = \mathbf{H}_k \Rightarrow -\mathbf{U}_k^T(\mathbf{G}+s_0\mathbf{C})^{-1}\mathbf{C}\mathbf{U}_n = \mathbf{H}_k \Rightarrow -\mathbf{U}_k^T(\mathbf{G}+s_0\mathbf{C})^{-1}\mathbf{U}_k\mathbf{U}_k^T\mathbf{C}\mathbf{U}_k = \mathbf{H}_k$$

$$\Rightarrow -(\mathbf{U}_k^T(\mathbf{G}+s_0\mathbf{C})\mathbf{U}_k)^{-1}\mathbf{U}_k^T\mathbf{C}\mathbf{U}_k = \mathbf{H}_k \Rightarrow -(\mathbf{U}_k^T\mathbf{G}\mathbf{U}_k + s_0\mathbf{U}_k^T\mathbf{C}\mathbf{U}_k)^{-1}\mathbf{U}_k^T\mathbf{C}\mathbf{U}_k = \mathbf{H}_k$$

$$\Rightarrow \mathbf{A}_n = -(\mathbf{G}_n + s_0\mathbf{C}_n)^{-1}\mathbf{C}_n = \mathbf{H}_k$$

where $\mathbf{G}_n = \mathbf{U}_k^T\mathbf{G}\mathbf{U}_k$ and $\mathbf{C}_n = \mathbf{U}_k^T\mathbf{C}\mathbf{U}_k$ . On the other hand ,

$$\forall j\in\overline{1,n} : \mathbf{M}_j = \mathbf{L}^T\mathbf{A}^j\mathbf{R} = \mathbf{L}^T\mathbf{U}_k\mathbf{H}_k^j\mathbf{U}_k^T\mathbf{R} = (\mathbf{U}_k^T\mathbf{L})^T\mathbf{H}_k^j\mathbf{U}_k^T(\mathbf{G}+s_0\mathbf{C})^{-1}\mathbf{B} =$$

$$= (\mathbf{U}_k^T\mathbf{L})^T\mathbf{H}_k^j(\mathbf{U}_k^T\mathbf{G}\mathbf{U}_k + s_0\mathbf{U}_k^T\mathbf{C}\mathbf{U}_k)^{-1}\mathbf{U}_k^T\mathbf{B} = \mathbf{L}_n^T\mathbf{H}_k^j(\mathbf{G}_n + s_0\mathbf{C}_n)^{-1}\mathbf{B}_n$$

where : $\mathbf{L}_n = \mathbf{U}_k^T\mathbf{L}$ , $\mathbf{B}_n = \mathbf{U}_k^T\mathbf{B}$ , $\mathbf{R}_n = (\mathbf{G}_n + s_0\mathbf{C}_n)^{-1}\mathbf{B}_n$ , $\overline{\mathbf{M}}_j = \mathbf{L}_n^T\mathbf{A}_n^j\mathbf{R}_n$ ;

hence $\mathbf{M}_j = \mathbf{L}^T \mathbf{A}^j \mathbf{R} = \mathbf{L}_n^T \mathbf{A}_n^j \mathbf{R}_n = \overline{\mathbf{M}}_j$.

In conclusion, we proved the following theorem :

**Theorem 2** (Padé via Block-Arnoldi)**:** The setting up of the matrix $\mathbf{U}_k$, by which the Block-Arnoldi iterrative process applied to the initial system (22) results in a reduced n-dimensional (n=km) model (23) after k iterrations, is equivalent with finding a local n-order Padé approximation of the transfer function of the initial system , and also with an orthogonal projection of the initial system onto Krylov subspace $\mathbf{K}_k(\mathbf{A},\mathbf{R})$

In analogy with the **Theorem 1**, Remark1 is also applicable here.

**Remark 3 :** The above described algorithm is general and does not consider the situation in which the initial columns (or the current cluster of columns) that generates the Block-Krylov subspace are (or could become) linear dependent. These cases require a built-in (exact or inexact) deflation procedure in the algorithm as in [4].

## Conclusions

This paper shows how implicit moments matching can be done by applying the Arnoldi and Block-Arnoldi algorithms to SISO and MIMO systems respectively.

The theoretical results presented here depend neither on factorization (7) nor on the choice of the (classic or modified) orthogonalization procedure implemented in the algorithms.

As the space alloted for this paper here is limited, we shall present the Lanczos, Block- Lanczos and Two-Step Lanczos methods in another issue.

For the same reason we shall discuss the similarities and the differences between those types of algorithms in a following paper.

**Acknowledgements**

The author wish to thank Prof. PhD. Valeriu Prepeliță for carefully reading this text.

## REFERENCES

[1] P. Feldmann , Roland Freund, *Efficient Linear Circuit Analysis by Padé approximation via the Lanczos process*, In proceedings of EURO-DAC '94 with EURO-VHDL '94, Grenoble, France, pages 170-175, IEEE Computer Society Press, 1994.

[2] Valeriu Prepeliță, Tiberiu Vasilache, Mona-Mihaela Doroftei, *Control Theory*, University "Politehnica" of Bucharest , 1997.

[3] Yousef Saad, *Iterative Methods for Sparse Linear Systems*, PWS Publishing Company, 1996.

[4] Roland Freund, Martin Gutknecht, Noël M. Nachtingal, *An implementation of Look-Ahead Lanczos Algorithm for non-hermitian matrices*, IEEE Computer Society Press, 1993.