Comparison of Model Order Reduction Methodologies for Thermal Problems

M. Salleras[1], T. Bechtold[2], L. Fonseca[3], J. Santander[3], E. B. Rudnyi[2], J. G. Korvink[2], S. Marco[1]

[1]Dept. d'Electrònica, Universitat de Barcelona, Martí i Franquès 1, 08028, Barcelona, SPAIN
[2]IMTEK, University of Freiburg, Georges-Köhler-Allee 103, 79110 Freiburg, GERMANY
[3]Centro Nacional de Microelectrónica, CNM-CSIC, Campus UAB, 08193, Bellaterra, SPAIN
email: msallera@el.ub.es, bechtold@imtek.de,

Abstract

In this paper two different methodologies for model order reduction for thermal problems are analyzed. These methodologies (Arnoldi algorithm and Proper Orthogonal Decomposition) face the problem from two different points of view. The first, deals with systemn matrices obtained from the space discretization of the system of partial differential equations (e.g. finite element method), which describe the relationship between the different nodes of the system, while Proper Orthogonal Decomposition deals with the input-output information of the system. Both methodologies have its own advantages and inconvenients, which will be discussed for a particular MEMS device.

1. Introduction

MEMS devices have a wide variety of actuation and sensing principles. In some of these MEMS devices temperature plays an important role in the dynamics of the system. Moreover the sensing or actuating part of some of these devices is thermally induced. Examples of such devices are thermo-pneumatic micropumps [1], microthrusters [2], microhotplate gas sensors [3], or thermopile based IR sensors [4].

In all these cases, modeling and simulation can address many important steps in the prototype design of the final device, such as optimizing sensor sensitivity, reducing device size, or improving power consumption. In some cases [2], optimal performance needs a real-time control of the input according to output magnitudes. In such a case, a feed-back loop with a model description of the system must be solved to predict the required new input values. It could also happen [1], that the device model must be included in the feed-back loop of a system-level simulation where other models describing other devices appear interacting with each other. These models could be obtained by a space discretization of the partial differential equations describing the physics involved. Different methods exist to discretize these equations obtaining a system of ordinary differential equations, but the most widely used is the finite element method. This method for practical applications tipically produces systems with 10^4-10^6 degrees of freedom. The computational cost for solving these huge systems is very expensive and it is quite common to spend several minutes (or hours) solving the whole system for a certain given input. This fact makes it prohibitive to include these full models in real-time control feed-back loops, or in system level simulations [1].

Model order reduction is a set of techniques which are focused on reducing the number of degrees of freedom. Indeed, there is a huge variety of techniques and different points of view to face this issue. Despite that, two main different groups can be distinguished.

The first group of techniques deals with the description of the system dynamics in the state-space formulation. They are based on mathematical transformations applied to the huge system matrices to obtain smaller matrices which with a small error should reproduce the dynamics of the whole system. The most basic technique of this group is Guyan reduction [5] which can only reproduce exactly the static behaviour of the original system, although complex techniques based on this reduction offer better performances. Recently, reduction techniques based on Padé or Padé-type [6] approximations of the transfer function via Lanczos [7] or Arnoldi [8] algorithms have gained great interest among engineers. The performance of these techniques is much better than Guyan techniques and a methodology insight is out of the scope of this paper. This methodology will only be introduced and its limitations explored in this paper. Arnoldi methodology has been done with mor4ansys [9,10], a software developed by IMTEK which is available in internet. The reader should address to the references for further details.

The second main group of techniques could be understood as a system identification techniques [11]. These techniques do not need the state-space description but an input-output data for a certain input must be supplied. This is the reason why they are mostly suited for extracting models of real systems from empirical data. The system itself is a blackbox for these methodologies, and the aim is to obtain a reduced order model which can describe adequately the input-output behaviour On the other hand, as the input-output data are obtained for a particulat input waveform, the accuracy of the model behaviour for different input waveforms must be explored. This group of techniques can even be splitted in two major sub-groups. One of them deals with RC-ladder identification techniques [12], and the other deals with proper orthogonal decomposition techniques. RC-ladder identification techniques are aimed to construct an RC circuit formed of Cauer networks stages which are interconnected forming a ladder [13]. The values of the different elements of the network are then optimized in order to fit the response of the different network nodes to (experimental or simulation) data in a least squares sense. These RC-ladder networks have the inherent limitation that they can only predict the thermal evolution in the

nodes where they were fit. Instead, Proper Orthogonal Decomposition (POD) and Arnoldi algorithm can predict the temperature evolution in all the nodes conforming the model. Proper Orthogonal Decomposition is mathematically related to well-known techniques as Principal Component Analysis (PCA), also known as Karhunen - Loève decomposition (KLD), and the singular value decomposition method (SVD) [14]. From a mathematical point of view, POD techniques provide an optimal (in least-square sense) low-order subspace to project the full system output. A critical point in POD is the selection of the excitation waveform, as well as the time stamps where the solution is computed. The excitation waveform must contain energy in the whole frequency range of interest. A typical choice, although not the only possibility, is to select a step input. For the time stamps (also known as snapshots), we selected logspaced points in the time axis in order to cover the system dynamics over a large range of time constants. It is important to note, that other selection can bring different efficiency in the reduction. This paper will be devoted to principal component analysis limitations and capabilities compared to Krylov methods. As a case study, these techniques will be applied to a thermopile based IR sensor [4].

2. IR Thermopile Sensor Structure

The basic structure of the thermopile based IR sensor, is a silicon bulk with a thermally isolated membrane. This membrane contains the hot junctions of the thermopiles, while the cold junctions are located in the opposite side of the membane, over the silicon bulk, in order to assure a maximum thermal isolation. The IR radiation is absorbed in a boron doped silicon rectangle located in the center of the membrane. In addition, a n$^+$polysilicon layer can be electrically heated for test or calibration purposes.
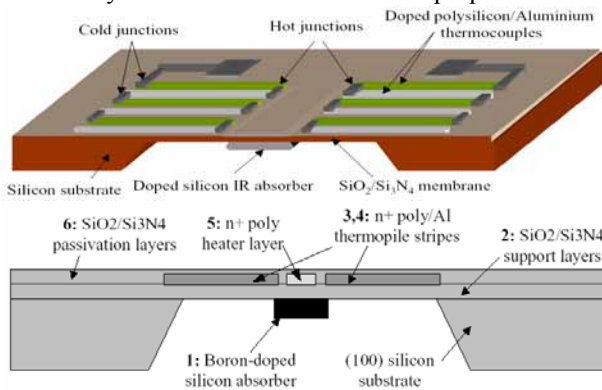

Figure 1: Thermopile IR sensor structure (top) and cross-sectional view (bottom) including, 1, isothermal radiation absorber; 2,6, SiO$_2$/Si$_3$N$_4$ support and passivation layers; 3, 4, poly-n+/Al thermopile stripes; 5, polysilicon heater layer.

In order to assure a high thermal isolation from the ambient, the thermopiles are located inside a silicon-oxide/silicon-nitride sandwich with a very low thermal conductivity (see figure 1). The sensor is fabricated with

CMOS compatible micromachining processes, in order to integrate on the same chip the needed electronics to amplify the output signal of the sensor. A total of 65+65 thermocouples are used to amplify the sensor sensitivity. A top view of the final device can be seen in figure 2.
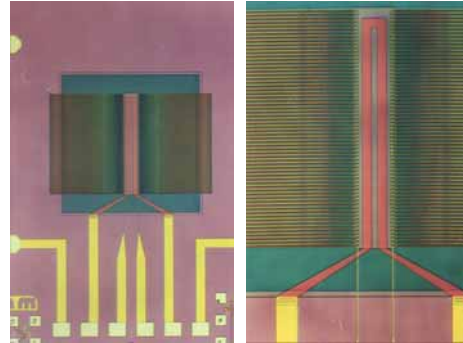

Figure 2: Top view photographs of the IR thermopile sensor.

3. FEM Modelling

The comercial software ANSYS® 8.1 has been used to build a finite element model of the device. A 3D model has been considered, with shell elements conforming the membrane, because the thermal gradients perpendicular to the membrane are assumed to be negligible [15]. Due to simmetry only one half of the device is modelled. The thermal properties of the different materials used in the finite element model are summarized in table 1. All the properties are considered temperature independent, and have been evaluated at T=300K.

Material	density ρ [Kg/m^3]	thermal conductivity K [W/(m·K)]	specific heat capacity C [J/(Kg·K)]
Silicon	2328	150	700
Si$_3$N$_4$	3180	24	880
SiO$_2$	2200	1.4	780
Aluminium	2692	235	900
Poly-n+	2328	30	754

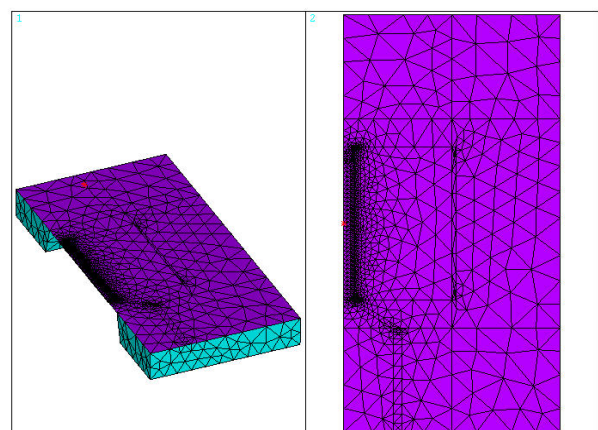Table 1: Material properties.

Figure 3: Meshing of the finite element model.

Air conduction and radiation have been included in the simulation. Despite that, because of the small

temperature change, the non-linearities introduced by the radiation are negligible. Initial conditions are applied for all the ndoes at T_0=300K. Dirichlet boundary conditions have been applied to the bottom of the silicon bulk at T=300K. Emissivity is set globally to ε=0.2, except for the absorber where ε=1 is assumed. A heat flux step of Φ=10W/m^2 is applied to consider the incident radiation. Finally a heat exchange of h=180W/(K·m^2) is considered. The obtained meshing consists of 1970 nodes and 8984 elements (see figure 3). The resulting stationary thermal distribution can be seen in figure 4.
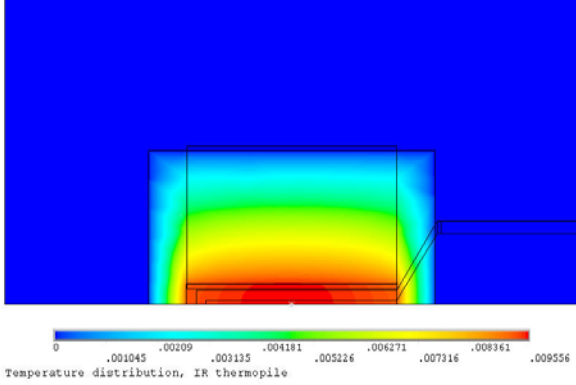

Figure 4: Stationary state temperature distribution for a heat flux step of Φ=10W/m^2.

The thermal transient simulation is carried out from $t=10^{-7}$s to t=1s, with 20 logspaced time points per decade, with a total number of 141 time stamps. From a practical point of view we are interested in the time evolution of the temperature difference between the hot junctions and the cold junctions. This is achieved by a volume weighted mean of the temperatures conforming the hot and cold junctions. For the finite element model this result is in figure 5. The maximum difference reached occurs for the stationary state and is 9.4mK.
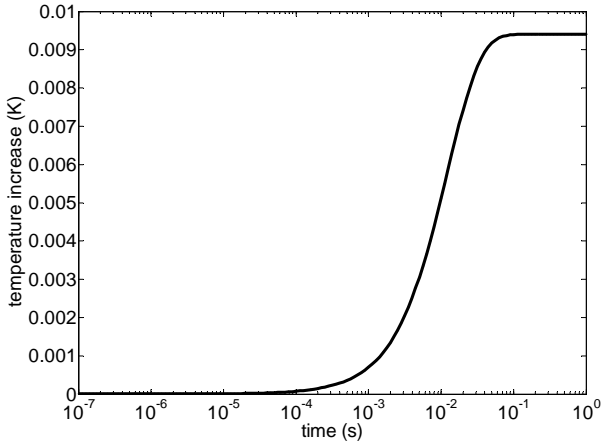

Figure 5: Time evolution of the difference T_{hot}-T_{cold} from the FEM model for a Φ=10W/m^2.

4. Model Order Reduction

A deep insight in the theoretical aspects of Principal Component Analysis and Arnoldi algorithm is out of the scope of this work. Instead a detailed analysis of the results for the case study will be discussed. Heat conduction equation is written as:

$$\frac{\partial T(\vec{x},t)}{\partial t} = \frac{1}{\rho C_P} \nabla \cdot \left(\kappa \cdot \nabla T(\vec{x},t) \right) + P \qquad (1)$$

where κ is the thermal conductivity, ρ the density, C_P the specific heat capacity, and P is the power applied. After a spatial discretization using the Finite Element Method, this equation leads to a set of ordinary differential equations (ODE) which can be written as:

$$\mathbf{CT'} + \mathbf{KT} = \mathbf{P} \qquad (2)$$

where C is the specific heat capacity matrix, K is the thermal conductivity matrix, P the applied power, and T the temperature vector. C and $K \in R^n \times R^n$ and P and $T \in R^n$, where n is the number of degrees of freedom resulting from the finite element model. Equation (2) is then a set of n first order ODE's which can be solved numerically but with a high computational cost. This equation can be stated in state space formalism as:

$$\mathbf{x'} = \mathbf{A} \cdot \mathbf{x} + \mathbf{B} \cdot \mathbf{u}$$
$$\mathbf{y} = \mathbf{C} \cdot \mathbf{x} + \mathbf{D} \cdot \mathbf{u} \qquad (3)$$

where $x \in R^n$ are the state variables, $u \in R^n$ the input vector and $y \in R^n$ the output vector. In order to write equation (3) as equation (2), we must set:

$$\mathbf{A} = -\mathbf{C}^{-1} \cdot \mathbf{K}$$
$$\mathbf{B} = \mathbf{C}^{-1}$$
$$\mathbf{C} = \mathbf{I}, \quad \mathbf{D} = \mathbf{0}$$

The basic idea of model order reduction is to describe the time variations of vector x with a low dimensional subspace as:

$$\mathbf{x} = \mathbf{M} \cdot \mathbf{r} + \boldsymbol{\varepsilon} \qquad (4)$$

where $M \in R^n \times R^k$ is the transformation matrix from the high-dimensional space to the low-dimensional space, with $k \ll n$, $r \in R^k$ is the reduced dimension state vector, and ε a small error which depends on the reduced order k. Replacing equation (4) in equation (3) leads to:

$$\mathbf{M}^T \cdot \mathbf{M} \cdot \mathbf{r'} = \mathbf{M}^T \cdot \mathbf{A} \cdot \mathbf{M} \cdot \mathbf{r} + \mathbf{M}^T \cdot \mathbf{B} \cdot \mathbf{u}$$
$$\tilde{\mathbf{y}} = \mathbf{C} \cdot \mathbf{M} \cdot \mathbf{r} + \mathbf{D} \cdot \mathbf{u} \qquad (5)$$

where in the top equation a left product by M^T has been applied. Considering that $M^T \cdot M = I$ equation (5) can be written as:

$$\mathbf{r'} = \hat{\mathbf{A}} \cdot \mathbf{r} + \hat{\mathbf{B}} \cdot \mathbf{u}$$
$$\tilde{\mathbf{y}} = \hat{\mathbf{C}} \cdot \mathbf{r} + \mathbf{D} \cdot \mathbf{u} \qquad (6)$$

where:

$$\hat{\mathbf{A}} = \mathbf{M}^T \cdot \mathbf{A} \cdot \mathbf{M} \quad \in \mathbb{R}^k \times \mathbb{R}^k$$

$$\hat{\mathbf{B}} = \mathbf{M}^T \cdot \mathbf{B} \quad \in \mathbb{R}^k \times \mathbb{R}^n \qquad (7)$$

$$\hat{\mathbf{C}} = \mathbf{C} \cdot \mathbf{M} \quad \in \mathbb{R}^n \times \mathbb{R}^k$$

are the reduced matrices in the state space description. It is important to remark that the input we apply and the solution we obtain are both high-dimensional vectors. After model reduction, the solution can be obtained in a reduced dimension space with minimum computational effort, but it is an approximation to the finite element model solution obtained in equation (3).

At this point both methodologies we are interested in will be applied to the case study. Arnoldi algorithm is applied to the constitutive matrices of the finite element model and as a result a transformation matrix between the high and the low-dimensional spaces is obtained [16]. Instead, principal component analysis is applied to the thermal evolution of all the nodes in the FEM model. This thermal evolution is computed as a transient solution of the finite element model at all the time stamps specified, in response to an incident irradiation flux step. As in the first case, the result is another transformation matrix with the same properties [14]. The resulting reduced systems of equations for each case can be written as equation (6) which we will solve to obtain approximations to the full system solution.

5. Results and Discussion

In order to compare the performance of the different methods, we have built three reduced models for each of the reduction techniques. The reduced systems orders are 5, 10 and 15. The full system and the reduced systems have been solved by means of the Backward Euler method. The chosen time points to evaluate the solution are the same we used in the finite element solution, from $t=10^{-7}$s to $t=1$s with 20 logspaced time points per decade. As a figure of merit to evaluate the goodness of the reduced model approximation, a mean least squares relative error has been used, defined as :

$$\varepsilon = \frac{1}{N} \sqrt{\sum_{i=1}^{N} \left(\frac{T_i^r - T_i^{FEM}}{T_i^{FEM}} \right)^2} \qquad (8)$$

where i corresponds for the different time stamps (or frequency points), T^r is the reduced order system solution and T^{FEM} is the full system solution.

To emphasize to good agreement of all the reduced models to the full system solution, in figure 6 the step response is plotted. In figure 7 we plot the difference of input response for reduced order systems respect full system solution. The y-scale has been adjusted to distinguish all the signals. As we can see the worst approximation corresponds to the PCA with order 5. While the best solutions correspond to PCA for orders 10 and 15. Arnoldi solutions have the same behaviour but all of them suffer from a very small offset for the stationary

solution. In relative magnitudes it corresponds to approximately 0.004%.
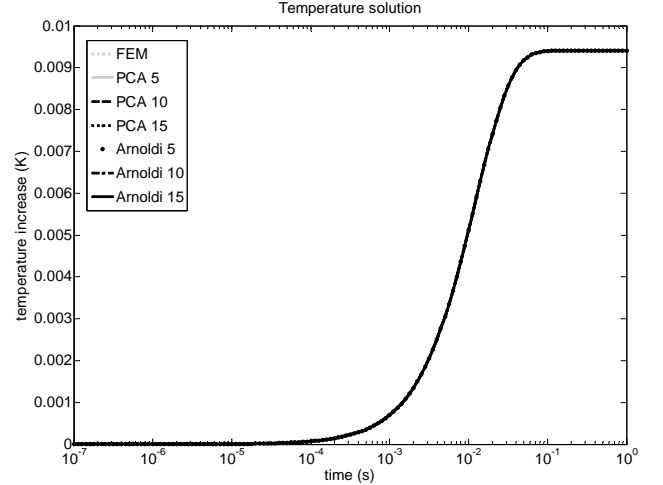

Figure 6: temperature increase versus time comparison.
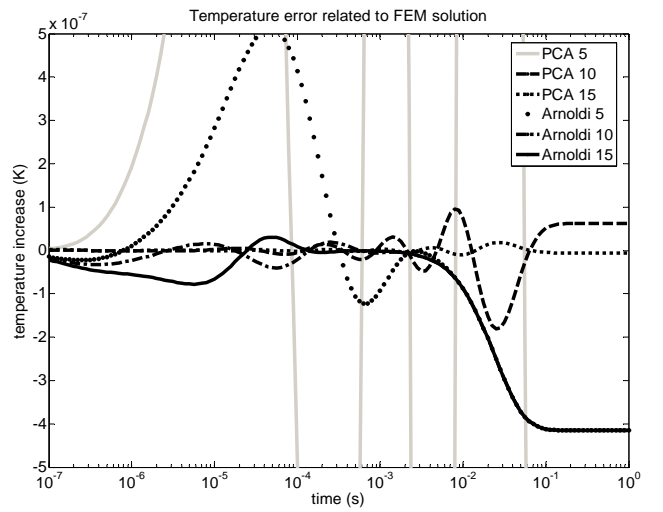

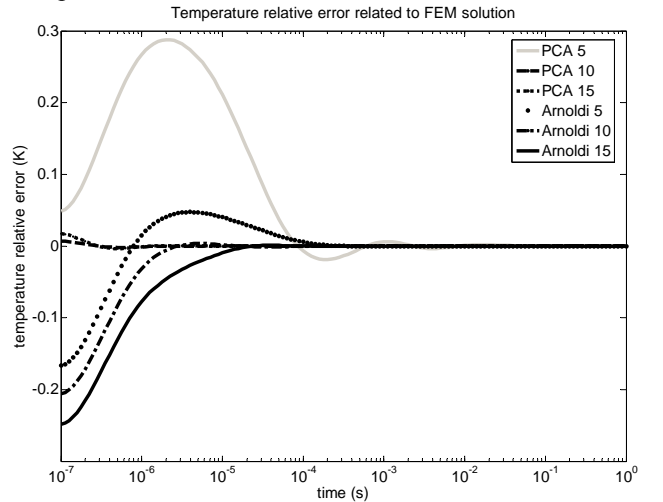Figure 7: temperature increase error versus time comparison.

Figure 8: temperature increase relative error versus time compared to full system solution for the different reduced order systems.

In figure 8 we have plotted the relative difference of the input responses for the reduced order systems. As it is expected the maximum relative errors occur where the signal is small, for short times. For Arnoldi reduced models the relative error for short times is around 20%, while PCA reduced models is below 2% except for the 5 order PCA model.

Due to the fact that we have the state space realization of the reduced systems, it is quite simpe to compute their frequency behaviour and compare it to the full system frequency behaviour. This will give us more information on how will be the system output for sinusoidal inputs.
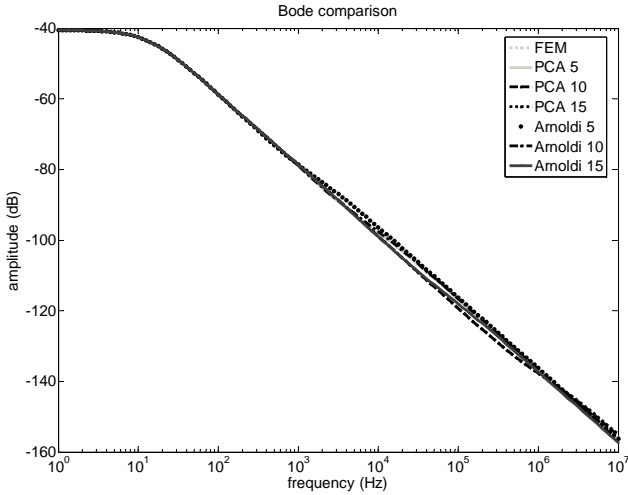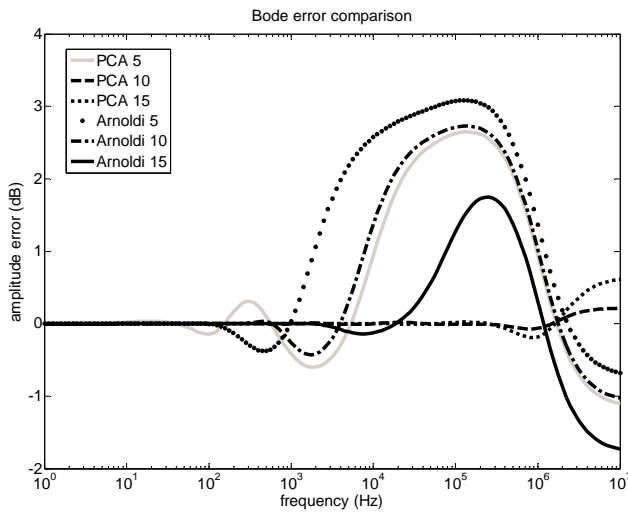

Figure 9: bode plot comparison.

Figure 10: Bode error diagram plot for the reduced systems related to the full system bode diagram.

The frequency response of the full and reduced systems have been evaluated from f=1Hz to f=10MHz, with 20 logspaced points per decade. Figure 9 shows this frequency behaviour for all the systems, and a good agreement is found for all the reduced systems. In order to evaluate more precisely these differences in figure 10 we plot the bode error diagrams for the different reduced systems. The worst agreement corresponds to the 5 order

Arnoldi system, whose discrepancies are noticeable for frequencies higher than 1kHz. The 5 order PCA and 10 order Arnoldi have a similar behaviour and their discrepancies are considerable around 10kHz. The 15 order Arnoldi system has an absolute error lower than 2dB which happens for frequencies higher than 100kHz. Finally, 10 and 15 order PCA systems have the best agreement with the full system frequency behaviour. Their maximum discrepancy occurs for the maximum frequency computed and is around 0.5dB.

Reduced system	ε time domain	ε frequency domain
PCA 5	$1.602 \cdot 10^{-2}$	$1.085 \cdot 10^{-4}$
PCA 10	$1.588 \cdot 10^{-6}$	$1.625 \cdot 10^{-7}$
PCA 15	$9.463 \cdot 10^{-6}$	$1.117 \cdot 10^{-6}$
Arnoldi 5	$1.905 \cdot 10^{-3}$	$2.162 \cdot 10^{-4}$
Arnoldi 10	$2.865 \cdot 10^{-3}$	$1.203 \cdot 10^{-4}$
Arnoldi 15	$5.033 \cdot 10^{-3}$	$3.216 \cdot 10^{-5}$

Table 2: Results summary: ε is the mean least squares relative error (see equation (8)) for the different reduced models in time domain and frequency domain.

As a summary of these results, in table 2 the mean least squares relative errors have been computed following equation (8) for time and frequency.

Principal component analysis as it has been presented here has two clear advantages with respect to Arnoldi algorithm. Its inputs are the step responses of the whole nodes of the finite element model, it can also be applied to experimental measurements, always that we have a large number of them. Moreover, when the data supplied comes from a finite element model, there is no restriction with respect to the linearity of the system which is being reduced. It is also possible to apply this technique to non-linear models. On the other hand, Arnoldi algorithm can only be applied to linear systems, although in many applications small errors are made when a nonlinear system is linearized around a suitable linearization point [16]. The restriction to linear systems in the Arnoldi algorithm must only be applied to the system matrices, no restriction appears when the nonlinearity is only present in the input. Another approach to avoid this problem when the system is weakly nonlinear is to apply Arnoldi algorithm to a quadratic approximation by a second order Taylor series expansion [17]. It is also true that Arnoldi algorithm is based on an expansion at a certain frequency, which in this case was f=0Hz. But there is no limitation to choose as an expansion point another frequency, for instance the working frequency of the device which is being reduced, in order to improve its frequency behaviour around this point. Even a multi-point expansion about several frequencies and restricting the reduced transfer function to match the first moments for all the frequencies has also been proposed, and known as rational-Krylov method [18].

6. Conclusions

A detailed comparison of two different order model reduction methodologies has been presented. In the case study presented, principal component analysis seems to produce a better agreement with full system, than the Krylov-based method. Despite that, PCA has an important additional computational cost, since the transient full system behaviour has to be solved in order to obtain the transformation matrix. Further studies must be made with more advanced Krylov-based techniques.

Acknowledgments

References

1. Carmona M., Marco S., Sieiro J., Ruiz O., Gómez-Cama J.M., Samitier J., "Modelling of Microsystems with Analog Hardware Description Languages," *Sensors and Actuators A*, Vol. 76 (1998), pp. 32-42.

2. Larangot B., Rossi C., Camps T., Berthold A., Pahm P.Q., Briand D., de Rooij N.F., Puig-Vidal M., Miribel-Català P., Montané E., López-Sánchez J., Samitier J., "Solid Propellant Micro-Rockets – Towards a new type of Power MEMS," *NanoTech*, Houston, TX, Sept. 2002.

3. Wöllenstein J., Böttner H., Plaza J.A., Cané C., Min Y., Tuller H.L. "A Novel Single Chip Thin Film Metal Oxide Array," *Sensors and Actuators B*, Vol. 93 (2003), pp. 350-355.

4. Calaza C., Meca E., Marco S., Moreno M., Samitier J., Fonseca L., Gràcia I., Cané C., "Assessment of the Final Metrological Characteristics of a MOEMS-Based NDIR Spectrometer Through System Modeling and Data Processing," *IEEE Sensors Journal*, Vol. 3, No. 5 (2003), pp. 587-594.

5. Friswell M.I., Penny J.E.T., Garvey S.D., "Using Linear Model Reduction to Investigate the Dynamics of Structures with Local Non-Linearities," *Mechanical Systems and Signal Processing*, Vol. 9, No. 3 (1995), pp. 317-328.

6. Feldmann P., Freund R.W., "Efficient Linear Circuit Analysis by Padé Approximation via the Lanczos Process," *IEEE Trans. Comput. Aid. Des. Integ. Circ. Syst.*, Vol. 14, No. 5 (1995), pp. 639-649.

7. Freund R.W., "Krylov-Subspace Methods for Reduced-Order Modeling in Circuit Simulation," *J. Comp. Appl. Math.*, Vol. 123 (2000), pp. 395-421.

8. Antoulas A.C., Sorensen D.C., "Approximation of Large-Scale Dynamical Systems: An Overview," *Int. J. Appl. Math. Comp. Sci.*, Vol. 11, No. 5 (2001), pp. 1093-1121.

9. www.imtek.uni-freiburg.de/simulation/mor4ansys/

10. Rundyi E.B., Korvink J.G., "Model Order Reduction for Large Scale Engineering Models Developed in ANSYS", *PARA'04, Workshop On State-of-the-art In Scientific Computing*, Technical University of Denmark, June 20-23, 2004.

11. Ljung. L., <u>System Identification - Theory for the user</u>, Prentice-Hall, (Englewood Cliffs NJ, 1987).

12. Rencz M., Székely V., "Dynamic Thermal Multiport Modeling of IC Packages," *IEEE Trans. Comp. Packag. Technol.*, Vol. 24, No. 4 (2001), pp. 596-604.

13. Palacín J., Salleras M., Puig M., Samitier J., Marco S., "Evolutionary Algorithms for Compact Thermal Modelling of Microsystems: Application to a Micro-Pyrotechnic Actuator," *Journal of Micromechanics and Microengineering*, Vol. 14 (2004), pp. 1-9.

14. Liang Y.C., Lee H.P., Lim S.P., Lin W.Z., Lee K.H., Wu C.G., "Proper Orthogonal Decomposition and its Applications – Part I: Theory," *Journal of Sound and Vibration*, Vol. 252, No. 3 (2002), pp. 527-544.

15. Elbel T., Lenggenhager R., Baltes H., "Model of Thermoelectric Radiation Sensors made by CMOS and Micromachining," *Sensors and Actuators A*, Vol. 35 (1992), pp. 101-106.

16. Rudnyi E.B., Korvink J.G., "Review: Automatic Model Reduction for Transient Simulation of MEMS-based Devices," *Sensors Update*, Vol. 11 (2002), pp. 3-33.

17. Chen Y., White J., " A Quadratic Method for Nonlinear Model Order Reduction," *International Conference on Modeling and Simulation of Microsystems*, San Diego, CA, March 2000.

18. Grimme E.J., Sorensen D.C., Van Dooren P., "Model Reduction of State Space Systems via an Implicitly Restarted Lanczos Method," *Numerical Algorithms*, Vol. 12 (1996), pp. 1-31.