DIMENSIONALITY REDUCTION OF FULL FIELDS BY THE PRINCIPAL COMPONENTS ANALYSIS

Christian Gogu[1,2], Raphael Haftka[2], Rodolphe Le Riche[1], Jérôme Molimard[1], Alain Vautrin[1]
[1] Ecole des Mines de Saint Etienne, 42023 Saint Etienne cedex 2, France
christian.gogu@gmail.com, leriche@emse.fr, molimard@emse.fr, vautrin@emse.fr
[2] University of Florida, 32611 Gainesville, FL, USA
haftka@ufl.edu

SUMMARY

The principal components analysis is applied to simulated full field displacement maps of 4569 measurement points thus allowing representation of any field within a certain domain as a linear combination of five basis vectors without losing significant information. A dimensionality reduction from 4569 to only 5 is thus achieved.

Keywords: optical full field methods, dimensionality reduction, principal component analysis

INTRODUCTION

Optical full field measurement methods for displacements or strains typically provide large quantities of information since each pixel of the image represents a measurement point. This has the great advantage of allowing to measure field heterogeneities. The sheer size of the data can however also pose problems in some situations.

Such a situation can arise if Bayesian identification needs to be applied to full field measurements. Bayesian identification of the elastic constants of a material differs from the usual identification schemes (such as finite element model updating) by accounting for uncertainties in the problems and also estimating the uncertainty in the identified material properties. Due to its statistical nature it handles probability distribution functions. In the case of full field methods we can have between thousands and millions of measurement points. Thousands-dimensional joint probability density functions for the measurements are however out of the realm of what can be numerically handled. A reduced dimensional representation of the measurements would thus be useful.

A possible method for achieving such a dimensionality reduction is the principal components analysis (PCA). This method has its origin in statistical analysis and depending on the field of application it is also known as proper orthogonal decomposition (POD) or the Karhunen-Loeve expansion.

The proper orthogonal decomposition method was initially used in computational fluid dynamics to represent complex, turbulent flows [1]. For a review of different uses of POD in computational fluid dynamics the reader can refer to [2].Other uses of POD include representing the structural response of non linear finite element under stochastic

loading [3] or real-time nonlinear mechanical modeling for interactive design and manufacturing [4]. POD was also applied to a multidisciplinary optimization problem of an aircraft wing [5].

The POD method seeks a reduced dimensional basis for the representation of the response (full fields in our case). This basis is constructed from a set of fields (often called snapshots) obtained with input parameters within a certain domain. The decomposition ensures that any field within the corresponding domain can be written as a linear combination of fixed POD fields (usually called POD modes). The accuracy increases with the number of POD modes used for the field representation. Typically we seek to represent the fields as a linear combination of less than a dozen POD modes while keeping reasonable accuracy in the reduced dimensional representation. It is important to note that for a full field with thousand measurement points this represents a dimensionality reduction in the field representation of one thousand to only a few dozen. The POD method can potentially allow to achieve this goal without loosing any significant amount of information.

The rest of the paper is organized as follows. In a first section we present the theoretical basis of proper orthogonal decomposition. We then present the simulated experiment for obtaining full fields. We then describe the dimensionality reduction problem applied to our specific case. Finally we present the proper orthogonal decomposition results.

PROPER ORTHOGONAL DECOMPOSITION

Let us consider a set of N vectors $\{U^i\}_{\mathrm{i=1..N}}$. A vector $U^i \in \mathbb{R}^n$, also called snapshot, can be the vector representation of displacement field for example. The aim of the proper orthogonal decomposition (POD) method is to construct an optimal reduced dimensional basis for the representation of the snapshots. For the POD method to work, it is necessary that the snapshot vectors have zero mean. If this is not the case the mean value needs to be subtracted from each vector.

We denote $\{\Phi_i\}_{i=1..K}$ the vectors of the orthogonal basis of the reduced dimensional representation of the snapshots. The POD method seeks to find the Φ_i that minimize the representation error:

$$\min \frac{1}{2}\sum_{i=1}^{N}\left\| U^i - \sum_{k=1}^{K}\alpha_{i,k}\Phi_k \right\|^2 \tag{1}$$

Because $\{\Phi_i\}_{i=1..K}$ is an orthogonal basis the coefficients $\alpha_{i,k}$ is given by the orthogonal projection of the snapshots onto the basis vectors:

$$\alpha_{i,k} = \left\langle U^i, \Phi_k \right\rangle \tag{2}$$

Eventually we have the following reduced dimensional representation $\widetilde{U}^i$ of the vectors of the snapshot set:

$$\widetilde{U}^i = \sum_{k=1}^{K}\alpha_{i,k}\Phi_k = \sum_{k=1}^{K}\left\langle U^i, \Phi_k \right\rangle \Phi_k \tag{3}$$

The achieved reduction in dimension is from N to K. The truncation order K needs to be selected such as to maintain a reasonably small error in the representations $\widetilde{U}^i$ of U^i. Selecting such a K is always problem specific.

The POD approach provides however a construction method for obtaining the optimal basis vectors that minimize the error defined in Eq. 1. This means that for a given truncation order we cannot find any other basis that better represents the snapshots subspace.

The basis $\{\Phi_i\}_{i=1..K}$ is constructed using the following matrix:

$$X = \begin{pmatrix} U_1^1 & \cdots & U_1^N \\ \vdots & \ddots & \vdots \\ U_n^1 & \cdots & U_n^N \end{pmatrix} \tag{4}$$

The vectors $\{\Phi_i\}_{i=1..K}$ are then obtained by the singular values decomposition of X, or equivalently by calculating the eigenvectors of the matrix XX^T. The singular values decomposition allows writing that:

$$X = \Phi \Sigma \Lambda^T \tag{5}$$

where Φ is the matrix of the column vectors Φ_i. Standard singular value decomposition routines (such as LAPACK [6]) typically provide the matrix Φ.

In the rest of the paper we will apply the POD decomposition to full field displacement measurements.

A truncation error criterion can be obtained by the following procedure. This criterion ε is defined for the sum of the error norms as shown in Equation 6.

$$\sum_{i=1}^{N} \left\| U^i - \sum_{k=1}^{K} \alpha_{i,k} \Phi_k \right\|^2 \leq \varepsilon \sum_{i=1}^{N} \left\| U^i \right\|^2 \tag{6}$$

With $\varepsilon = 1 - \dfrac{\sum_{i=1}^{K} \sigma_i^2}{\sum_{i=1}^{N} \sigma_i^2}$, where σ_i are the diagonal terms of matrix Σ.

SIMULATED EXPERIMENT

Experiment

The experiment we model is a tensile test on a composite plate with a hole. The laminate is made of graphite/epoxy with a stacking sequence of $[45,-45,0]_s$. The plate has the dimensions given in Figure 1. The applied tensile force is 600 N.

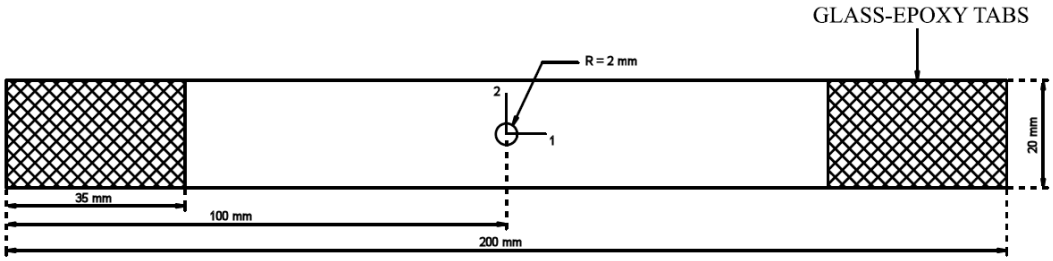


Figure 1: Specimen geometry. The specimen material is graphite/epoxy and the stacking sequence $[45,-45,0]_s$. The tensile force is 600 N.

Numerical model

The previous experiment is modeled using Abaqus® finite element software. A total of 8020 S4R elements were used. The virtual measurement is assumed to take place on a reference area of 2 mm x 2 mm at the center of the plate. This would be a typical area of an optical full field measurement (Moiré interferometry for example). The finite element mesh in the area of interest is represented in Figure 2 and the virtual measurements area highlighted in red.

To illustrate the applicability of the dimensionality reduction method we used here finite element fields, thus the term virtual measurement. Note however that this does not mean that only finite element fields could be represented in the reduced dimensional space. Once the POD modes are determined, measured fields can very well be projected onto these modes and expressed in the corresponding basis. This can even have some advantages, such as noise filtering as we will mention later.

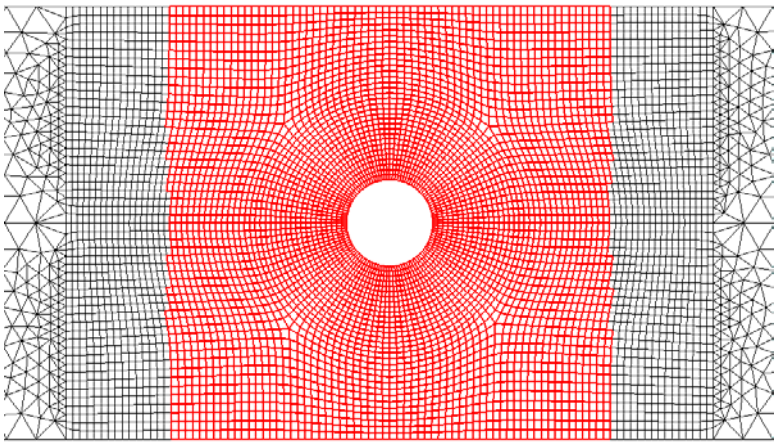


Figure 2: Finite element mesh. The area of the virtual measurements (2 mm x 2 mm) is highlighted.

DIMENSIONALITY REDUCTION PROBLEM

Problem statement

The general framework of the problem is in our case the following. We vary a certain number of model parameters such as elastic constants or plate dimensions and virtually measure each time the corresponding full field. The virtual measurements are considered here to be the displacement values at the 4569 nodes within the reference area (see Fig. 2).

If such a virtual measurement needs to be repeated a large number of times (for statistical sampling for example) it is not practical to describe the fields by the value of each pixel. A major reason is because if statistical methods (such as Bayesian identification) need to be used on the fields, thousands dimensional probability density functions are far outside the realm of what the statistical methods can handle.

The problem statement can then be formulated as follows. Can we find a reduced dimensional representation of the full fields for whatever combination of input parameters (elastic constants, plate dimensions in our case) within a certain domain?

To address this problem we propose to use the proper orthogonal decomposition method described in the first section.

POD implementation

For our plate with a hole virtual experiment we are interested in accounting for variations of the following parameters: ply elastic constants E_1, E_2, ν_{12}, G_{12} and ply thickness t. Accounting for variations in the elastic constants is typically of interest in material properties identification problems. We added here the ply thickness to illustrate a typical case for Bayesian identification. Bayesian identification allows to identify the elastic constants while accounting for other sources of uncertainty. The ply thickness was considered here to be one of such sources of uncertainty that potentially affects the identification. We will not carry out the Bayesian identification itself in this paper but for additional details on this method the reader can refer to [7],[8].

We are interested in variations of the parameters E_1, E_2, ν_{12}, G_{12} and t within the bounds given in Table 1. Again these could be typical values for an identification problem.

Table 1: Bounds on the input parameters of interest.

Parameter	E_1 (GPa)	E_2 (GPa)	ν_{12}	G_{12} (GPa)	t (mm)
Lower bound	126	7	0.189	3.5	0.12
Upper bound	234	13	0.351	6.5	0.18

We obtain the snapshots required for the POD approach by sampling 200 points within the bounds of Table 1. The points are obtained by Latin hypercube sampling, which consists in obtaining the 200 sample points by dividing the range of each parameter into 200 sections of equal marginal probability 1/200 and sampling once from each section.

Latin hypercube sampling typically ensures that the points are reasonably well distributed in the entire space.

At each of the 200 sampled points we run the finite element analysis and take a virtual full field measurement, which consists in the U and V displacement fields. Each of the 200 virtual measurements for U (same for V) is stored as a column vector and represents one snapshot used for the POD decomposition. The virtual measurement area covers 4569 finite element nodes so we obtain snapshots vectors of size 4569 x 1. The snapshots matrix X has then a size of 4569 x 200. Note that as mentioned in the POD theory section the snapshots need to have zero mean. In our case this was true for the U field but not for the V field, so we needed to subtract the mean value of each snapshot as shown in Eq. 7.

$$X = \begin{pmatrix} V_1^1 - \overline{V_1} & \cdots & V_1^N - \overline{V_N} \\ \vdots & \ddots & \vdots \\ V_n^1 - \overline{V_1} & \cdots & V_n^N - \overline{V_N} \end{pmatrix} \tag{7}$$

The POD modes of the 200 fields are then calculated using the singular value decomposition as shown in Equation 5.

Two snapshots (snapshot 1 and 199) that led to substantially different fields are illustrated next. The input parameters of the two snapshots are provided in Table 2.

Table 2: Input parameters for snapshots 1 and 199.

Parameter	E_1 (GPa)	E_2 (GPa)	v_{12}	G_{12} (GPa)	t (mm)
Snapshot 1	202.2	10.84	0.2142	4.989	0.1312
Snapshot 199	233.1	9.915	0.2153	5.952	0.1445

The fields of snapshot 1 are represented in Figure 3. Snapshot 199 has displacement and strain fields that are similar to those of snapshot 1. Among the notable differences are Max(U)= 0.0148 mm and Max(V)= 0.0105 mm for snapshot 199.

POD RESULTS

POD modes

In total we obtained 200 POD modes. The first six are represented graphically in Figures 4 and 5. Note that the POD decomposition is applied twice, once for the U fields and once for the V fields, thus the different modes for the two displacement components.

In Figure 6 we also provide the strain equivalents of the POD modes. Note that these are not calculated using the POD decomposition of the strain fields. Instead the displacement POD modes are derived to obtain the strain equivalents of the POD modes. In this study we are interested in the displacements POD decomposition so the strain equivalent POD modes are provided here mainly to facilitate understanding and qualitative comparison of the results. If one would work with strain fields rather than displacement fields than he should carry out the POD decomposition directly on the finite element strains which would lead to more accurate results.

In order to differentiate the displacement POD modes two additional steps were required. The displacements were interpolated on a 256 x 256 grid using cubic polynomials (Matlab gridata command). They were then numerically differentiated pixel by pixel. These two steps can introduce numerical derivation artifacts (as we will see later) so they are not advised if accurate strain POD modes are sought.

Note that we stop graphical representation at 5 strain equivalent POD modes (Figure 6) precisely because of numerical artifacts. For mode 5 some artifacts can already be noticed.

POD truncation

To recall, the POD decomposition allows to represent a displacement field obtained with any combination of input parameters within the bounds of Table 1 as a linear combination of the POD modes.

Even though we obtained 200 POD modes, typically substantially fewer POD modes are required to represent the fields with reasonable accuracy. In the present section we seek the truncation order that still allows keeping reasonable accuracy. Table 3 provides the truncation error criterion defined in Equation 6.

Table 3: Error norm truncation criterion.

K	2	3	4	5
ε for U fields	$2.439\ 10^{-7}$	$4.701\ 10^{-9}$	$7.280\ 10^{-11}$	$1.211\ 10^{-11}$
ε for V fields	$1.054\ 10^{-6}$	$2.900\ 10^{-9}$	$4.136\ 10^{-10}$	$3.517\ 10^{-11}$

Since ε is an overall error criterion, it is not easy to interpret in terms of error localization. Accordingly we will also seek to visualize what happens with the errors for individual snapshots. Figure 7 provides the truncation error maps in the reconstruction of snapshot 1 when using only the first 3, 4 and 5 POD modes. Similarly for snapshot 199 in Figure 8. We can note that already with 3 modes the maximum displacement error in the field is about 1000 times less than the maximum value of the field. The error further decreases by one to two orders of magnitudes when reaching 5 modes. Furthermore we can note that with 5 POD modes we start seeing artifacts in the error maps that are not likely to have a physical meaning. Since for the displacement maps we don't use any numerical derivation these artifacts are likely to be due to the finite

element discretization. We also note a slight asymmetry in the artifacts of the V displacement errors. This asymmetry is likely due to a slight asymmetry of the mesh. Indeed the finite element V displacements have their origin at the edge of the plate. In order to carry out the POD decomposition we had to subtract the mean values from the snapshots. In case of slight mesh dissymmetry the numerical estimate of the mean might be slightly off the true value which causes a dissymmetry in the POD modes and thus subsequent error maps. Note that the dissymmetry of the mesh is quite small (see Figure 2). The effects of the dissymmetry are noticeable only from mode 5 on, meaning that their corresponding error is of the same order of as the discretization error.

Finally in Figures 9 and 10 we provide the strain equivalent truncation error. This is to check that the POD modes have not missed any important features that have small characteristics on the displacement fields but larger characteristics on the strain maps. Note that the equivalent strain error maps were again obtained using numerical derivation thus exhibit more numerical artifacts. We note however that the same trend as for the displacements is seen as far as truncation error.

Truncating at K=5 means that the POD decomposition achieved a dimensionality reduction from 4569 to 5. The error maps show that this is achieved without loosing any significant displacement information. It is important to note that the obtained reduction does not depend on the number of measured points (4569 here). The reduction relies on expressing the fields as a linear combination of the determined POD modes. Whether describing each field and POD mode using 4569 or 1 Million points is irrelevant to the fact the field's variations can be expressed with good accuracy as a linear combination of only five modes.

Before concluding we also want to draw the reader's attention on the fact that the POD reduction acts also as a filter. This can have beneficial effects in some situations. The filter effect is obvious when analyzing the truncation error maps. When truncating at K=4 for example, the corresponding field reconstruction will obviously not include any of the higher modes. If the higher modes start representing mainly finite element discretization errors, then truncating at K=4 will filter out these errors. A different filtering can arise when constructing the POD modes using finite element results, then projecting noisy fields (such as actual measurements) on the POD basis. The projected representation will then filter out the noise. A more detailed investigation of the filtering effect will be carried out in future studies.

CONCLUSIONS

In the present article we were interested in a reduced dimensional representation of full field displacement maps from a plate with a hole. For our problem we were interested in the variations of five parameters (four elastic constants and one plate dimension) within +/-30% bounds. We want to represent any field stemming from within this domain in a reduced dimensional basis of only a few vectors (less than a dozen). This is achieved by using the proper orthogonal decomposition (POD) method. We used virtual experiments based on finite element results to calculate the POD modes required for the reduced dimensional representation. We then showed that using only the first five POD modes allows a representation of two of the field snapshots with an error of less than 0.1%. We thus achieved in our case a dimensionality reduction of 4569 to 5 without loosing any

significant information in the field's representation. If required much higher reductions are possible.

ACKNOWLEDGEMENTS

The authors would like to thank Gustavo H.C. Silva for his help with Python scripting.

References

1. Berkooz, G., Holmes, P., and Lumley, J.L., "The proper orthogonal decomposition in the analysis of turbulent flows", *Annu. Rev. Fluid Mech.*, 25, pp. 539–575, 1993.
2. Lucia, D.J., Beran, P.S. and Silva, W.A., "Reduced-order modeling: new approaches for computational physics", Progress in Aerospace Sciences, 40-1, pp. 51-117, 2004.
3. Schenk, C.A., Pradlwarter, H.J., Schuller, G.I., "Non-stationary response of large, non-linear finite element systems under stochastic loading", Comp & Struct., 83(14), pp.1086–1102, 2005.
4. Dulong J.L., Druesne F., Villon P., "A model reduction approach for real-time part deformation with nonlinear mechanical behavior", *International Journal on Interactive Design and Manufacturing*, 1(4), pp.1955–2513, 2007.
5. Coelho, R.F., Breitkopf, P., Knopf-Lenoir, C., Villon, P., "Bi-level Model Reduction for Coupled Problems", Stuctural and Multidisciplinary Optimization, (In Press), 2009
6. Anderson, E., Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen, "LAPACK User's Guide", (http://www.netlib.org/lapack/lug/lapack_lug.html), Third Edition, SIAM, Philadelphia, 1999.
7. Gogu, C., Haftka, R.T., Le Riche, R., Molimard, J., Vautrin, A. and Sankar, B.V., "Comparison between the basic least squares and the Bayesian approach for elastic constants identification", *Journal of Physics: Conference Series*, vol. 135, no. 1, pp. 012045, 2008.
8. Gogu, C., Haftka, R. T., Le Riche, R., Molimard, J., Vautrin, A., Sankar, B. V., "Bayesian Statistical Identification of Orthotropic Elastic Constants Accounting for Measurement and Modeling Errors", AIAA paper 2009-2258, *11th AIAA Non-Deterministic Approaches Conference*, Palm Springs, CA, May 2009.

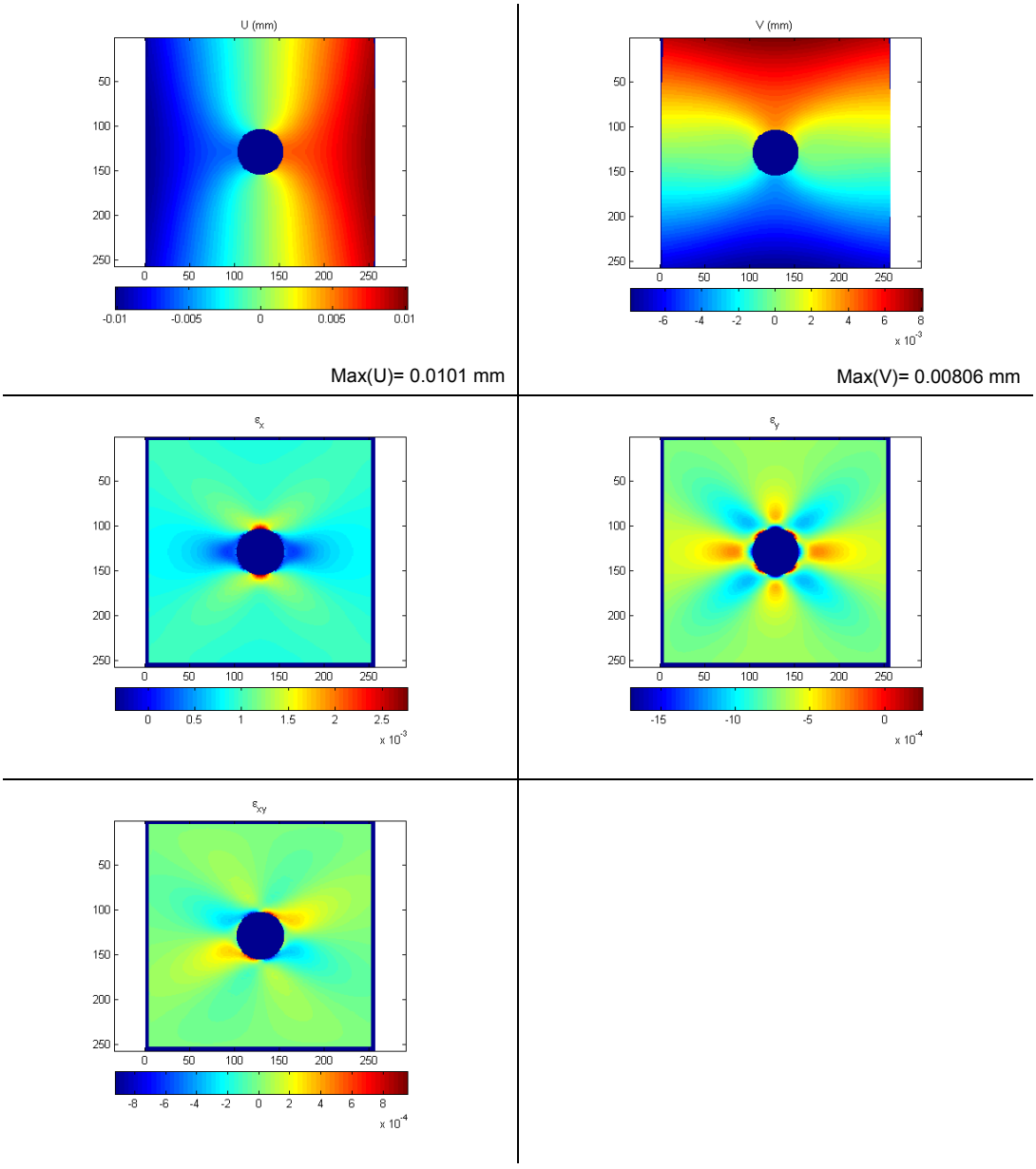


Figure 3: Displacement and strain maps for snapshot 1

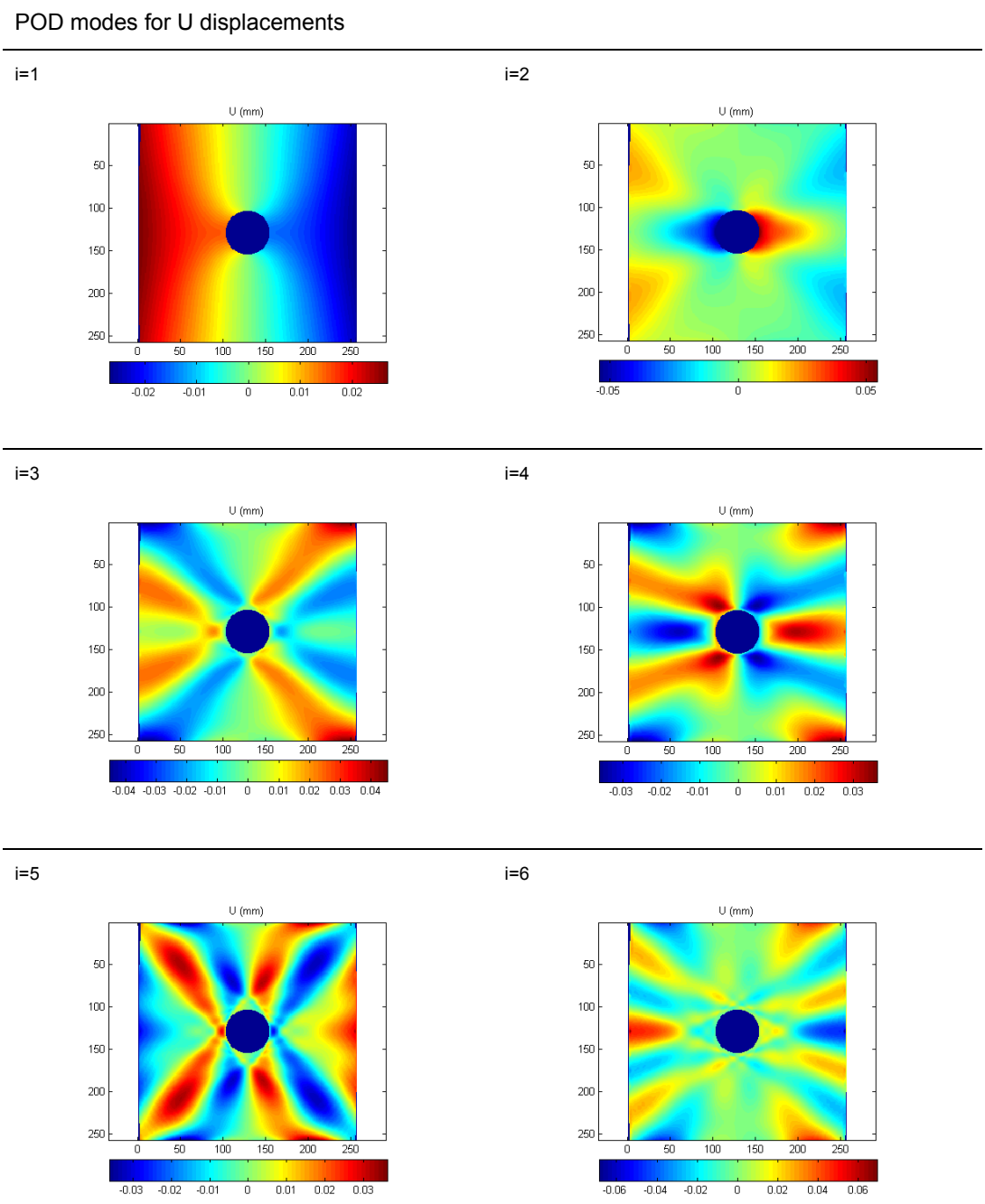


Figure 4: First six POD modes for U displacement fields

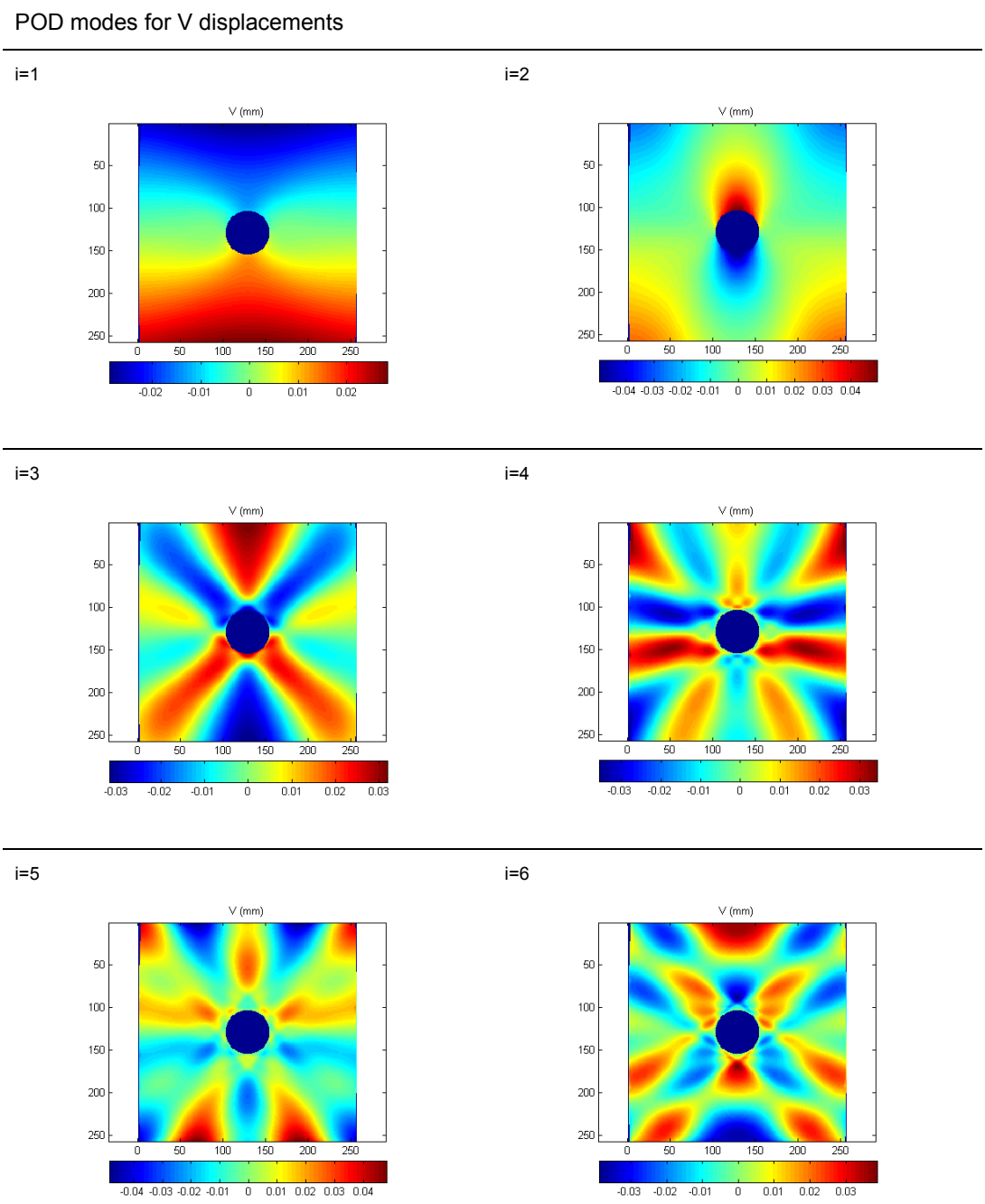
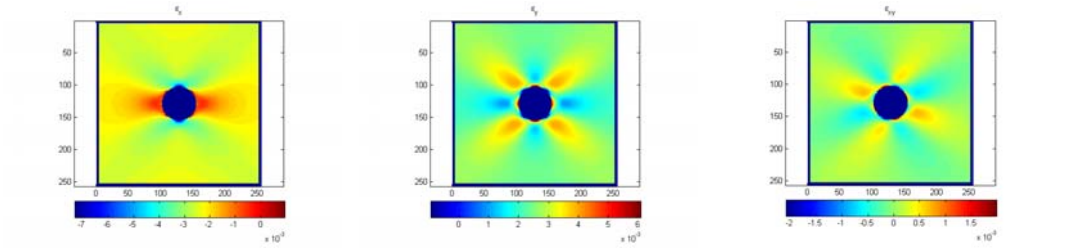


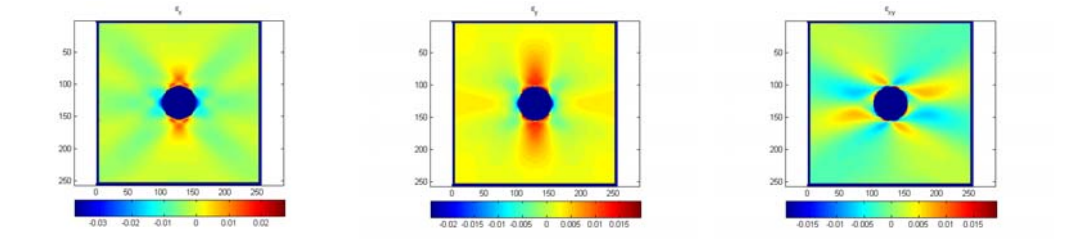
Figure 5: First six POD modes for V displacement fields

POD modes for ε

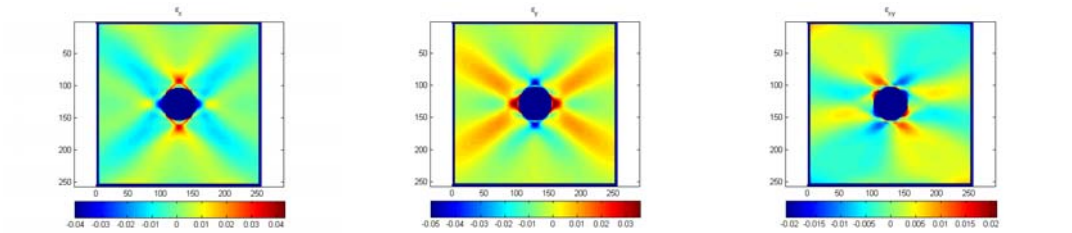
i=1



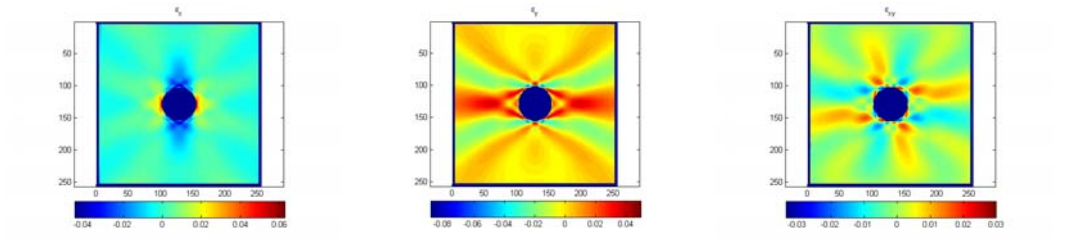
i=2



i=3



i=4



i=5

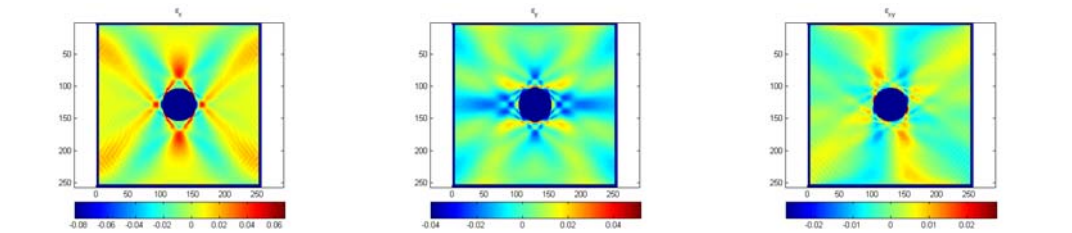
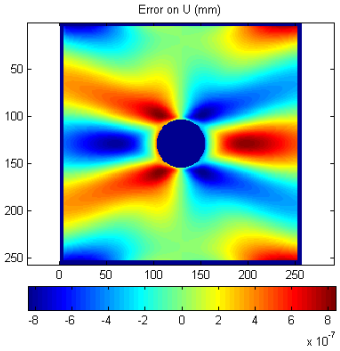
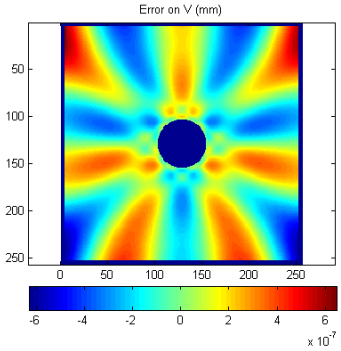


Figure 6: First five strain equivalent POD modes

K=3

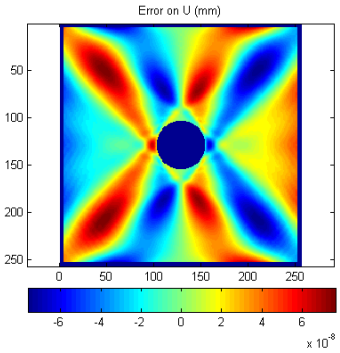


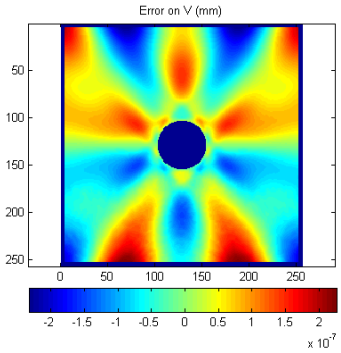


Max error: 8.991e-007 mm

Max error: 6.329e-007 mm

K=4

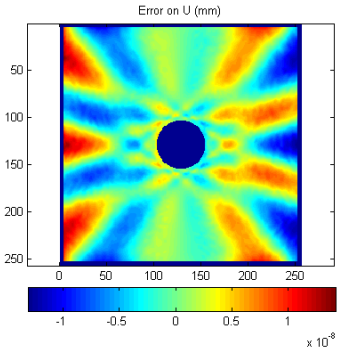


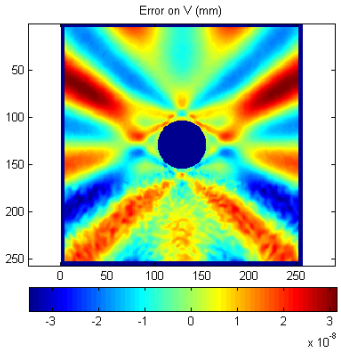


Max error: 9.479e-008 mm

Max error: 2.526e-007 mm

K=5



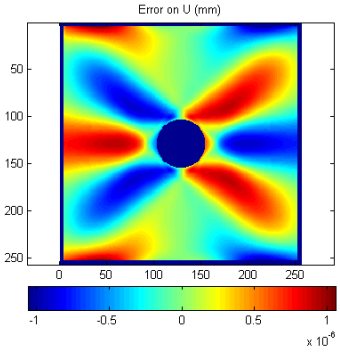


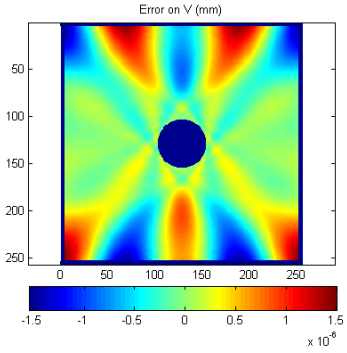
Max error: 1.698e-008 mm

Max error: 3.132e-008 mm

Figure 7: Displacements truncation error in snapshot 1 using 3, 4 and 5 modes. The maximum error can be compared to the maximum of the U and V fields, Max(U) = 0.0101 mm, Max(V)= 0.00806 mm.

K=3

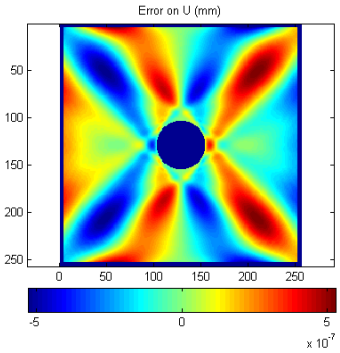


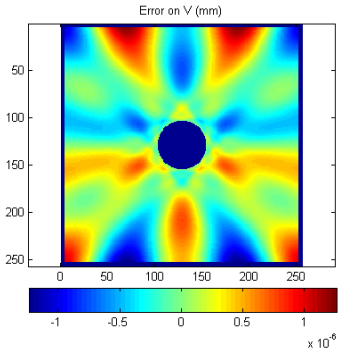


Max error: 1.353e-006 mm

Max error: 1.552e-006 mm

K=4

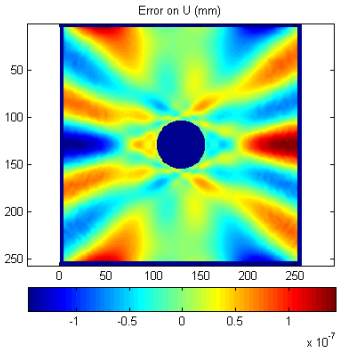


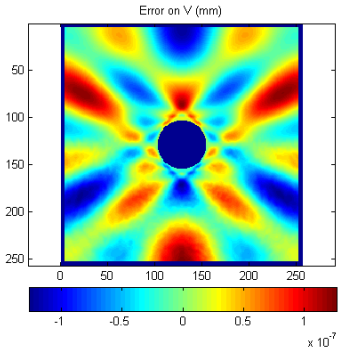


Max error: 6.929e-007 mm

Max error: 1.355e-006 mm

K=5





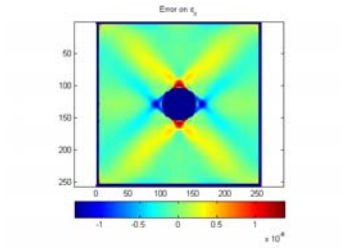
Max error: 1.623e-007 mm

Max error: 1.257e-007 mm

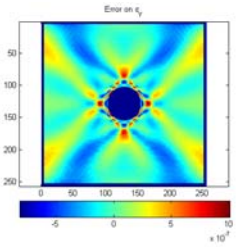
Figure 8: Displacements truncation error in snapshot 199 using 3, 4 and 5 modes. The maximum error can be compared to the maximum of the U and V fields, Max(U)= 0.0148 mm, Max(V)= 0.0105 mm.

Snapshot 1

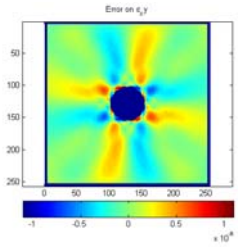
K=3



Max error: 1.397e-006

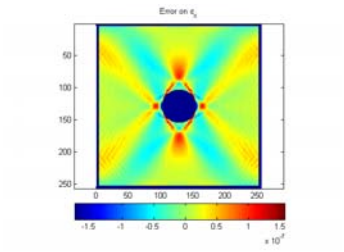


Max error: 6.728e-007

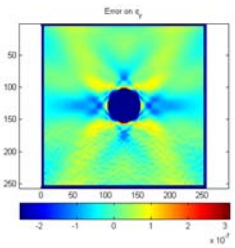


Max error: 8.579e-007

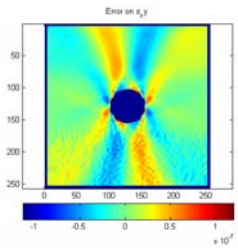
K=4



Max error: 1.567e-007

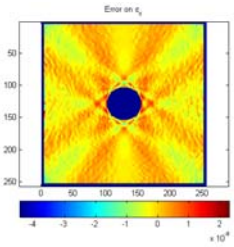


Max error: 3.119e-007

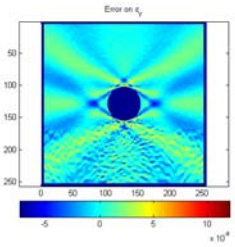


Max error: 1.201e-007

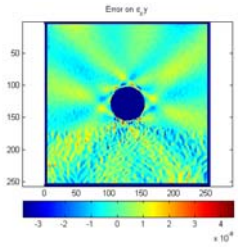
K=5



Max error: 2.328e-008



Max error: 1.206e-007

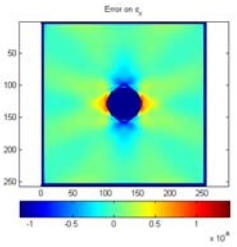


Max error: 4.482e-008

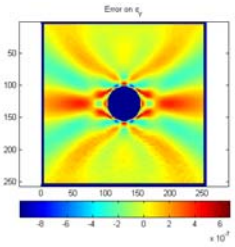
Figure 9: Strain equivalent truncation error in snapshot 1 using 3, 4 and 5 modes.

Snapshot 199

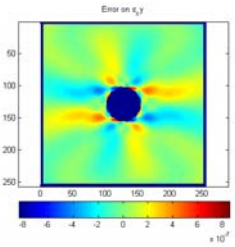
K=3



Max error: 1.377e-006

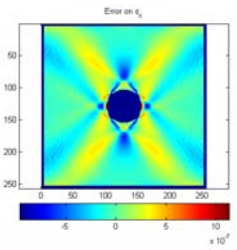


Max error: 1.007e-006

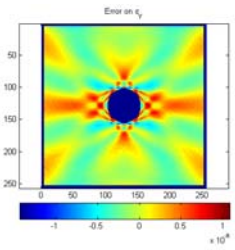


Max error: 1.097e-006

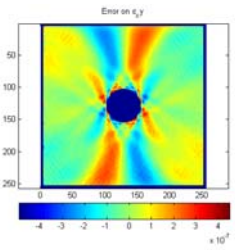
K=4



Max error: 1.138e-006

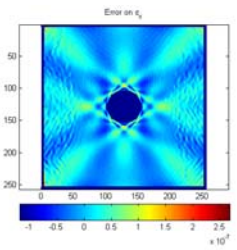


Max error: 1.082e-006

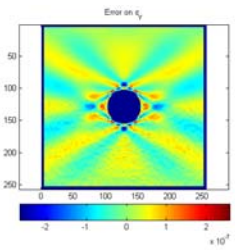


Max error: 4.508e-007

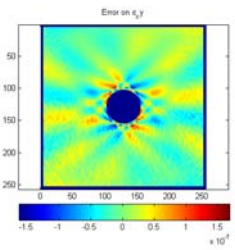
K=5



Max error: 2.683e-007



Max error: 2.589e-007



Max error: 1.689e-007

Figure 10: Strain equivalent truncation error in snapshot 199 using 3, 4 and 5 modes