UNIVERSITÀ DEGLI STUDI DI PALERMO

Meccanica dei Materiali e delle Strutture
Vol. 3 (2012), no.1, pp. 25-32
ISSN: 2035-679X
Dipartimento di Ingegneria Civile, Ambientale, Aerospaziale, Dei Materiali
DICAM

NON-ORTHOGONAL MODAL REPRESENTATION OF THE WIND-INDUCED PRESSURE FIELD ON BLUFF BODIES

Luigi Carassale[*], Michela Marrè Brunenghi[†]

[*] Dipartimento di Ingegneria Civile, Chimica e dell'Ambiente (DICCA)
Università di Genova
Via Montallegro 1, 16145 Genova, Italy
e-mail: luigi.carassale@unige.it

[†] Dipartimento di Ingegneria Civile, Chimica e dell'Ambiente (DICCA)
Università di Genova
Via Montallegro 1, 16145 Genova, Italy
e-mail: michela.marre.brunenghi@unige.it

(Ricevuto 10 Giugno 2012, Accettato 10 Ottobre 2012)

Key words: Aerodynamics, pressure fields, POD, PCA, ICA, coherent structures.

Parole chiave: Aerodinamica, campi di pressione, POD, PCA, ICA, strutture coerenti.

Abstract. *The concept of coherent structure is often invoked for the qualitative analysis of the wind-induced pressure field acting on bluff bodies. Coherent structures are traditionally extracted from the measured data through the Proper Orthogonal Decomposition (or Principal Component Analysis, PCA), but serious concerns on their meaningfulness have been risen. In an attempt to mitigate these problems, the Independent Component Analysis (ICA) has been introduced and demonstrated some potentials to extract coherent structures that are consistent with the observed phenomenon in term of some features that are defined in this paper. Considering the case of the pressure measured on a square-base prism immersed in a turbulent boundary layer with a non-symmetrical configuration, PCA and ICA are applied to extract coherent structures. The role of the model order is carefully investigated.*

Sommario. *Il concetto di struttura coerente è spesso impiegato nell'analisi qualitativa del campo di pressione indotto dal vento in prossimità della superficie di corpi tozzi. Le strutture coerenti vengono tradizionalmente estratte dai dati sperimentali attraverso la Proper Orthogonal Decomposition (o Principal Component Analysis, PCA), ma i dubbi sulla loro interpretazione fisica sono piuttosto diffusi. Allo di chiarire questo problema è stata introdotta la Independent Component Analysis (ICA), che ha dimostrato ampie potenzialità nell'estrarre pattern coerenti con i fenomeni osservati. Nel presente lavoro sono state applicate le tecniche statistiche di espansione modale PCA ed ICA ad un campo di pressione originato da uno strato limite turbolento intorno ad un prisma a sezione quadrata. L'enfasi dell'analisi è concentrata sull'effetto della dimensione del modello ridotto impiegato per analizzare i dati.*

1 INTRODUCTION

The use of high-frequency pressure scanners in bluff-body aerodynamics can usually provide accurate estimations of both local and global wind-induced actions that can be used for design purpose. Besides, the analysis of the recorded time histories may reveal precious information on the qualitative behaviour of the flow, detecting the signature of the vortexes generated in the neighbourhood of the body and possibly propagate due to advection effects. This signature is identified through the observation of highly-correlated regions that appear recurrently in the pressure field and are referred to as coherent structures (CS) of the pressure field, in analogy with the definition of this term traditionally employed as qualitative descriptor of the flow velocity field[1].

The concept of CS has been intimately related to the statistical technique traditionally used for their estimation, the Proper Orthogonal Decomposition (POD[2]), which, in the discrete-space context, is often referred to as Principal Component Analysis (PCA). This technique represents the random fluctuation of the pressure field as a sum of deterministic pressure distributions called modes, modulated by random coefficients (depending on time) called Principal Components (PC). PCA has a hierarchical nature in terms of energy: the first PCA mode represents the best (in the mean square sense) mono-variate approximation of the pressure field and is usually interpreted as the dominant CS or the signature of the dominant flow phenomenon acting on the body surface[3]. Due to its mean-square optimality, PCA has been widely applied to identify Reduced-Order Models (ROM) of the measured data and to formulate qualitative interpretations of the observed phenomena. In numerous applications PCA demonstrated its ability in extracting CSs that resemble the observed physical phenomena[4], but in many other circumstances serious concerns on their meaningfulness have been risen[5]. In general terms, a meaningful coherent structure should represent a typical or recurrent configuration of the pressure field and, if two different physical phenomena (characterized by different statistical properties take place) at the same time, they should be represented by two different CSs.

The main source of PCA limitations to extract meaningful CSs has been identified in the orthogonality of its modes that, from a physical point of view, does not have any justification. To overcome this constraint, the use of non-orthogonal modal expansions based on the Independent Component Analysis (ICA[6]) have been adopted[7,8] demonstrating that, at least in the applications discussed in these references, the above requirements for the meaningfulness of the CSs are well fulfilled.

ICA is usually formulated assuming that the measured data derive from a generative model in which unknown sources, said Independent Components (IC), are instantaneously mixed by an unknown matrix referred to as mixing matrix. ICs and mixing matrix are estimated from the data exploiting the condition that the ICs are mutually statistically independent. This kind of hypothesis is obviously correct when ICA is used to solve the Blind Source Separation (BSS) problem, in which the existence of independent (though unknown) sources is postulated, but may be questionable when ICA is applied as a feature-recognition tool for the extraction of the CSs. In this case, the concept of virtual source[9] may be invoked, but the issue of estimating the appropriate number of these sources may become critical. This is a major problem in synthesizing an ICA model from measured data and is related to the different behavior that PCA and ICA representations have as the model order is increased. In particular, when working with PCA, the passage from a model with order n to a model with order $n + 1$ only involves the introduction of a new mode, without changing the modes from 1 to n. On the contrary, increasing the order of an ICA model produces, in general, the modification of all the modes involved in the representation. Indeed, even if the ICA model spans the same data subspace as the PCA model of the same order, ICA modes are determined by an optimization problem, whose result may change as fresh information are added increasing the model order. For this reason, it is clear that the model order should be selected as large enough to represent the relevant part of the observation, but not larger.

In the present paper PCA and ICA are applied to derive a qualitative description of the flow past a square-base prism immersed in a turbulent boundary layer at 15° angle of incidence. This condition is selected because it produces a very rich dynamics of the flow field. On the lateral face exposed to the wind the flow reattaches creating an oscillating recirculation bubble, while on the opposite lateral face the boundary layer is fully separated. This particular case study is used to discuss the ability of PCA and ICA to extract CSs that are meaningful in the sense defined above. The role of the model order in ICA representation is carefully investigated and some indicators able to guide its selection are proposed.

2 COHERENT STRUCTURES AND MODAL REPRESENTATIONS

The concept of CS is intimately related to the modal expansion adopted to represent the pressure field. In this sense, since an infinite class of modal representations can be formulated, an infinite set of different CSs can be associated to the same random field. This section describes the formulation of PCA and ICA with particular emphasis on their ability of synthesizing Reduced-Order Models (ROMs); furthermore it shows the application of these techniques to the qualitative study of the pressure field generated on a square-base prism immersed in a turbulent boundary layer. The experiment has been carried out in the wind tunnel of the University of Genova in which the model (height h = 40 cm, aspect ratio 5:1) has been oriented according to the angle of incidence α = 15°. The mean wind velocity profile corresponds to a power law with exponent 0.22; at the top level of the model the mean wind velocity is U = 12.2 m/s, the intensity of the longitudinal turbulence is I_u = 13% and its integral length scale is L_u = 40 cm. The pressure field is observed along a mono-dimensional domain constituted by an instrumented ring of N = 28 pressure taps located at z = 0.6h and connected through short tubes to a 32-channel PSI pressure scanner mounted inside the model. For this purpose, let $\mathbf{q}(t)=[q_1(t),\ldots,q_N(t)]^{\mathrm{T}}$ be an N-variate random process containing the zero-mean fluctuation of the measurements deriving from the pressure taps; t is the discretized time.

The configuration chosen for the present discussion is particularly interesting since it corresponds to the angle in which the lift coefficient of a square cylinder in smooth flow changes its slope due to the reattachment of the boundary layer on the Windward Lateral Face (WLF), while on the Leeward Lateral Face (LLF) the boundary layer remains completely separated[10]. In the present three-dimensional condition, with turbulent incoming flow, the situation is surely more complicated but some traces of this deep difference in the flow past the two lateral faces is expected to appear in the statistical representation of the pressure data.

The observed phenomenon is assumed to be statistically stationary and ergodic, thus time averaging is used as a surrogate of ensemble averaging to estimate any statistical quantity.

2.1 Principal Component Analysis (PCA)

Let $\mathbf{C_{qq}}$ be the zero-time-lag covariance matrix of $\mathbf{q}$ estimated from the data as:

$$\mathbf{C_{qq}} = \mathrm{E}\left[\mathbf{q}(t)\mathbf{q}(t)^{\mathrm{T}}\right] \tag{1}$$

in which the statistical expectation E[·]. According to PCA, $\mathbf{q}(t)$ is represented by the modal expansion:

$$\mathbf{q}(t) = \sum_{k=1}^{N} \boldsymbol{\phi}_k x_k(t) = \boldsymbol{\Phi}\mathbf{x} \tag{2}$$

where the vectors $\boldsymbol{\phi}_k$ ($k = 1,\dots,N$), usually called modes, are the eigenvectors of $\mathbf{C}_{\mathbf{qq}}$, i.e. the solutions of the problem:

$$\mathbf{C}_{\mathbf{qq}}\boldsymbol{\phi}_k = \lambda_k \boldsymbol{\phi}_k \qquad (k = 1,\dots,N) \tag{3}$$

The eigenvectors are mutually orthogonal, are conventionally normalized to have unit norm and are assembled column-wise in the matrix $\boldsymbol{\Phi}$. The coefficients x_k are referred to as the Principal Components (PC) of the process and are assembled in the vector $\mathbf{x} = [x_1,\dots,x_N]^{\mathrm{T}}$; they are mutually uncorrelated and their variance is provided by the eigenvalues λ_k. Eigenvectors and PCs are enumerated in such a way that their corresponding eigenvalues are sorted in decreasing order.

For practical purpose, a truncated version of Eq. (2) is defined through the partial sum:

$$\mathbf{q}^{(n)} = \sum_{k=1}^{n} \boldsymbol{\phi}_k x_k = \boldsymbol{\Phi}^{(n)}\mathbf{x}^{(n)} \tag{4}$$

where $\boldsymbol{\Phi}^{(n)}$ and $\mathbf{x}^{(n)}$ contains the first $n \le N$ PCA modes and PCs, respectively. The vector $\mathbf{q}^{(n)}$ represents the best (in the mean square sense) n-variate approximation of $\mathbf{q}$. For this reason Eq. (4) is often used to synthesize ROMs, calibrating the model order n in such a way to retain the significant part of the observation and cancel the noise.

The PCs can be calculated from the data through a simple orthogonal projection, i.e.:

$$\mathbf{x}^{(n)}(t) = \boldsymbol{\Phi}^{(n)^{\mathrm{T}}}\mathbf{q}^{(n)}(t) = \boldsymbol{\Phi}^{(n)^{\mathrm{T}}}\mathbf{q}(t) \tag{5}$$

In the context of bluff-body aerodynamics, the use of PCA as a feature-recognition tool to identify typical or recurrent pressure patterns, possibly associated to some specific physical phenomenon, has been experimented by several authors[11].Sometimes the results of the analyses appeared encouraging, but there is a general opinion that PCA modes is unable to identify physically-meaningful CSs[5]. A major reason of this failure is believed to derive from the orthogonality constraint of the modal shapes that is intrinsically present in PCA[7-8] and has no physical justification.

Figure 1 shows the first 6 PCA modes extracted from the data. The relative importance of each mode in the representation of the pressure field is quantified through the quantity

$$\Sigma_n = \left(\frac{\mathrm{E}\left[\left\| \mathbf{q}^{(n)} \right\|^2 \right]}{\mathrm{E}\left[\left\| \mathbf{q} \right\|^2 \right]} \right)^{0.5} \tag{6}$$

reported in Figure 1 for reference; the 6$^{\text{th}}$-order ROM defined through these modes retains about 93% of the Standard Deviation (SD) of the full-order dataset. Mode 1 represents a pressure distribution acting on the windward face and provides about 67% of the SD of $\|\mathbf{q}\|$.

The pressure on the LLF is represented, with a similar spatial distribution, by modes 2, 3 and 4. All the reported modes provide a contribution to the pressure field acting on the WLF, with higher-order modes representing details characterized by a smaller spatial scale. This latter behaviour is encountered quite commonly in PCA applications and is essentially due to the orthogonality constraint that is implicitly involved in the definition of the modes. Besides, it is worth noting that, even if the pressure fields acting on the two lateral faces have very different statistical properties (i.e. different probability distribution, spectral properties, propagation velocity), several modes provide contributions on both the lateral faces (modes 2, 3, 4).

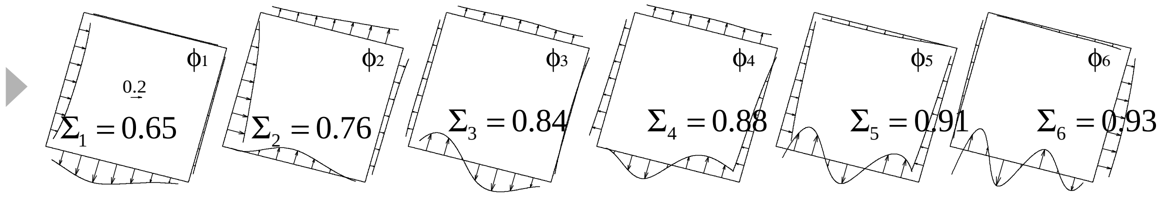

Figure 1: PCA modes $\boldsymbol{\phi}_1$ to $\boldsymbol{\phi}_6$.

2.2 Independent Component Analysis (ICA)

ICA is usually formalized assuming that the observed data $\mathbf{q}(t)$ are produced by a generative model of the kind[6]:

$$\mathbf{q}(t) = \mathbf{A}\mathbf{s}(t) \tag{7}$$

where s is a vector of n ≤ N statistically independent sources sk (k = 1,…,n) said independent components (IC) and A is an N×n full-rank matrix referred to as mixing matrix. The ICA model postulates the existence of an unknown number n of unknown statistically independent sources that are statically (i.e. with no memory) mixed by an unknown mixing matrix. This formulation is clearly tailored on the Blind Source Separation (BSS) problem, whose solution can be obtained through three conceptual steps: 1 – the number of sources n is identified as the number of non-zero eigenvalues of $\mathbf{C}_{\mathbf{qq}}$; 2 – the mixing matrix $\mathbf{A}$ is obtained imposing that the sources s_k are statistically independent; 3 – the ICs are calculated by solving the linear system (7), which is possible since $\mathbf{q}$ is in the range of $\mathbf{A}$ for any t.

In case the IC mixture is contaminated by additive noise Eq. (7) holds in an approximate sense and can be rewritten as:

$$\mathbf{q}^{(n)}(t) = \mathbf{A}\mathbf{s}(t) \tag{8}$$

where n is estimated by counting the eigenvalues of $\mathbf{C}_{\mathbf{qq}}$ that are above a suitable threshold and $\mathbf{q}^{(n)}$ is a reduced version of $\mathbf{q}$ obtained through Eq. (4).

Under the condition that all (but possibly one) ICs have non-Gaussian distribution, the ICA problem as formulated above has a solution, which is unique up to scaling and permutation of the Ics[6]. To avoid permutation and scaling indeterminacy, the ICs are conventionally assumed as unit variance and are enumerated in such a way the columns $\mathbf{a}_k$ of the matrix $\mathbf{A}$ have decreasing norm.

The ICA model (Eq. (8)) is analogous to the representation formula offered by PCA (Eq. (4)), with the difference that the columns $\mathbf{a}_k$ of the matrix $\mathbf{A}$ are, in general, non-orthogonal and that the ICs s_k are now statistically independent (instead of simply uncorrelated like the PCs x_k). Due to this analogy, ICA can be used as tool to estimate CSs from experimental datasets[7-8]. When ICA is applied for this purpose, the situation becomes more complicated than for the BSS problem since the ICs are usually not explicitly present in the observed physical system and the concept of virtual source[9] must be invoked. Virtual sources, however, are not necessarily statistically independent thus the ICA problem should be reformulated as: given the observed data $\mathbf{q}$, find a suitable number n of virtual sources that are as much statistically independent as possible and provide an n^{th}-order approximation of $\mathbf{q}$ through Eq. (8). Also in this case the choice of the model order n may be based on the eigenvalues of $\mathbf{C}_{\mathbf{qq}}$, however the brute application of this criterion can be troublesome.

The use of ICA as a feature-recognition tool for the analysis of random pressure fields is interpreted as the search for a set of CSs, which evolve in time with amplitudes that are as much statistically independent as possible.

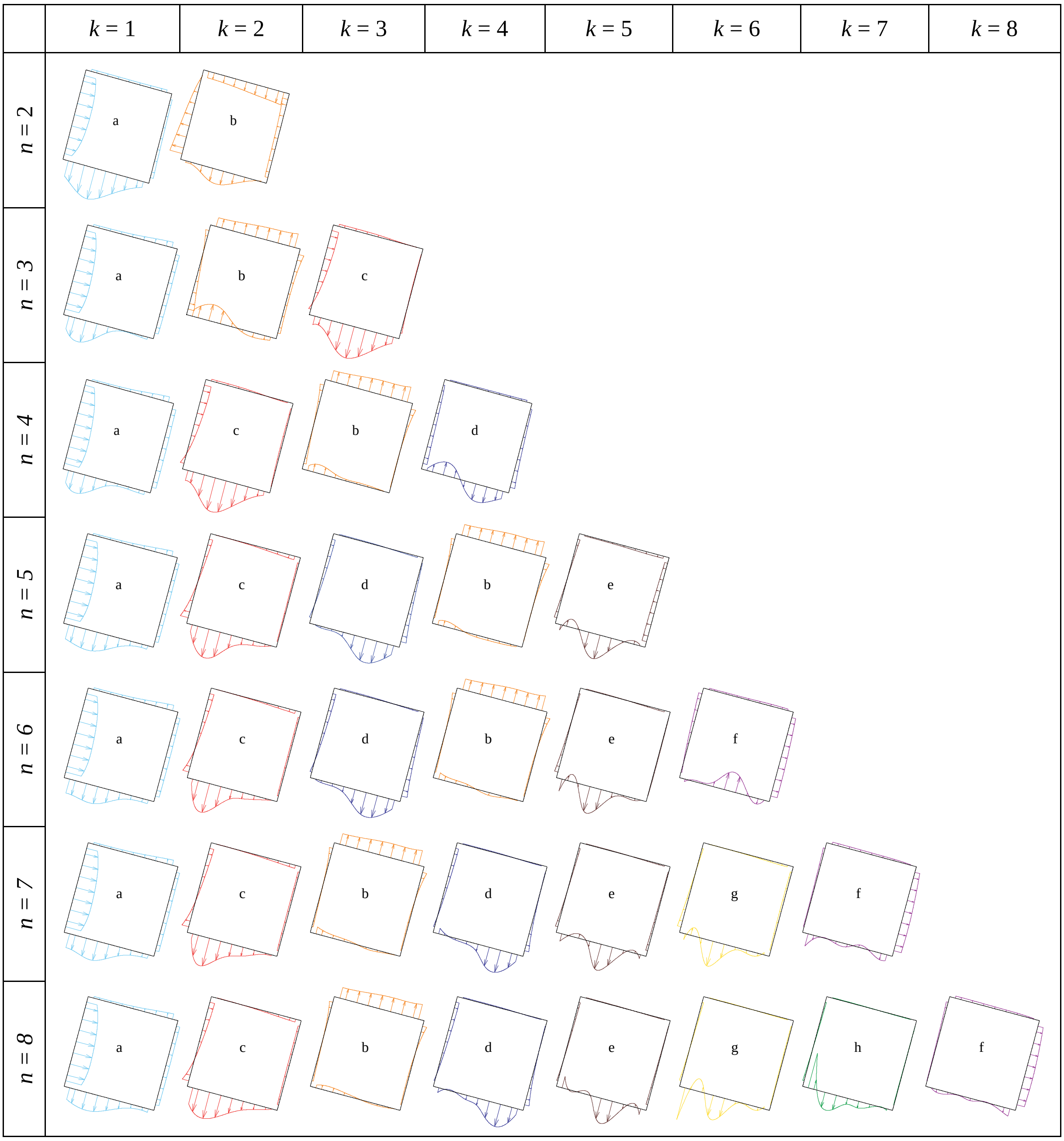

Figure 2: ICA modes for model orders n = 2 to 8

It has been observed that, working with ICA representations, the choice of the model order n can be more critical than for PCA. To investigate this issue, Figure 2 shows the ICA modes estimated for model orders n from 2 to 8. For each model, modes are plotted row-wise, sorted with respect to their norm and labeled with a letter (from a to h) on the basis of their shape. As the model order increases some modes (modes a, b and f) stabilize on a particular shape, while other modes split into different modes (modes c, d, e, g, h). The former behavior belongs to the modes representing the pressure field on the windward face (mode a), on the LLF (mode b) and on the leeward face (mode f); the latter behavior is shown by the modes involved in the representation of the pressure on the WLF (modes c, d, e, g, h).

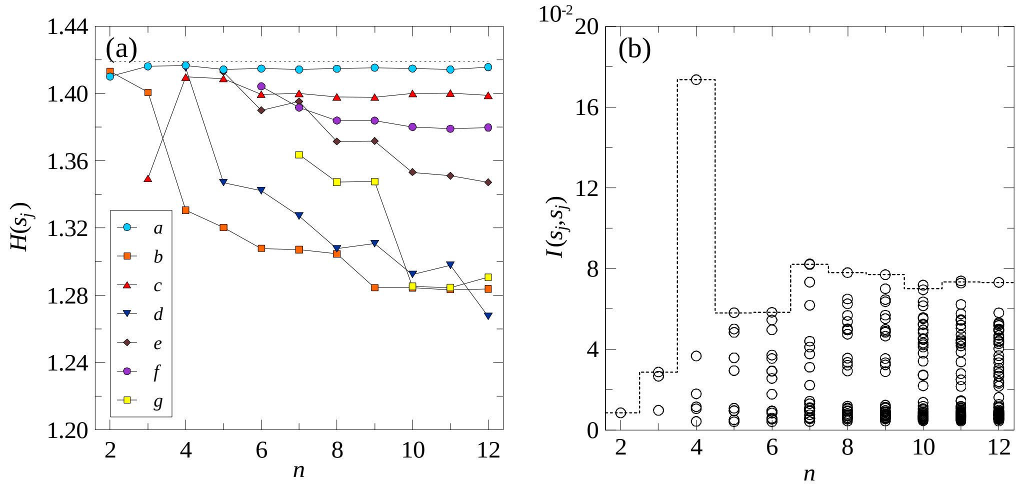

Figure 3: Entropy of the ICs (a), mutual information between pairs of ICs (b) as functions of the model order.

Figure 3 describes a complementary aspect of this evolution. Figure 3a shows the differential entropy of the ICs corresponding to ICA modes a to g, defined as:

$$H(s_k) = -\int_{-\infty}^{\infty} p_{s_k}(\xi)\log\left[p_{s_k}(\xi)\right]\mathrm{d}\xi \qquad (k = a,\dots,g) \tag{9}$$

where p_{s_k} is the pdf of the k^{th} IC. As the model order increases, the entropy of the ICs tends to decrease, indicating an increment of their non-Gaussianity (the maximum entropy corresponding to the Gaussian distribution is reported with dotted line). On the other hand, the reduction of entropy reflects a reduction of the information content carried by each IC, which, in some sense, tends to distill a specific feature of the observed phenomenon. Figure 3b shows, for different model orders, the mutual information between pairs of ICs defined as:

$$I(s_j, s_k) = \int_{\mathbb{R}^2} p_{s_j s_k}(\xi,\eta)\log\left[\frac{p_{s_j s_k}(\xi,\eta)}{p_{sj}(\xi)\,p_{s_k}(\eta)}\right]\mathrm{d}\xi\,\mathrm{d}\eta \qquad (j,k = 1,\dots,n) \tag{10}$$

where $p_{s_j s_k}$ is the joint pdf of the j^{th} and k^{th} ICs. The mutual information can be interpreted as a measure of the statistical dependency of two random variables, or from an information theory view point, it measures the overlapping of the information carried by s_j and s_k. The mutual information can be used as a guide to select the model order, for example excluding those orders that produce IC pairs with large mutual information (like n = 4 in the present application), and privileging the orders for which the mutual information of IC pairs are clustered about small values (like n = 7, 8 in the present application).

3 CONCLUSIONS

PCA and ICA have been applied to estimate the CSs of the wind-induced pressure field acting on a square-base prism immersed in a turbulent boundary layer with a non-symmetrical configuration. The focus of the discussion has been on understanding whether or not these CSs are meaningful representing separately (i.e. in different modes) parts of the pressure field with distinct statistical characteristics and on the problem of the sub-space dimension. In particular, the following conclusions can be proposed:

1. PCA is unable to separate the pressure fluctuation acting on different parts of the prism cross section, even if they have very different statistical characteristics; on the contrary, ICA separates in different modes the pressure field acting on the four faces of the prism.
2. Up to the model order 8, ICA use a large number of modes to represent the pressure acting on the WLF, while on the other faces a single mode is employed. This behavior is determined by the weaker spatial correlation present on the WLF.
3. The choice of the most appropriate model order for ICA may be critical; the tracking of the mode stability for different model orders, as well as the study of entropy and mutual information of the ICs are a valuable guides for the order selection.

REFERENCES

[1] P. Holmes, J.L. Lumley and G. Berkooz, *Turbulence: Coherent Structures, Dynamical Systems and Symmetry*, Cambridge University Press, (1996)

[2] J.L. Lumley, *Stochastic Tools in Turbulence*, Academic Press, (1970).

[3] H. Kikuchi, Y. Tamura, H. Ueda and K. Hibi, Dynamic wind pressure acting on a tall building model - proper orthogonal decomposition, *J. Wind Engng. Ind. Aerodyn.*, **69-71**, 631–646 (1997).

[4] A. Kareem and J.E. Cermak, Pressure fluctuations on a square building model in boundary-layer flows, *J. Wind Engng. Ind. Aerodyn.*, **16**, 17–41 (1984).

[5] Y. Tamura, S. Suganuma, H. Kikuchi and K. Hibi, Proper orthogonal decomposition of random wind pressure field, *J. Fluids & Struct.*, **13**, 1069–1095 (1999).

[6] A. Hyvärinen, J. Karhunen and E. Oja, *Independent Component Analysis*, John Wiley and Sons, (2001).

[7] L. Carassale, Analysis of aerodynamics pressure measurements by dynamic coherent structures, *Probabilistic Engineering Mechanics*, **28**, 66-74, (2012).

[8] L. Carassale and M. Marrè-Brunenghi, Statistical analysis of wind-induced pressure fields: a methodological perspective, *J. Wind Engng. Ind. Aerodyn.*, **99**, 700-710, (2011).

[9] W. Zhou and D. Chelidze, Blind source separation based vibration mode identification, *Mechanical Systems and Signal Processing*, **21**, 3072-3087, (2007).

[10] R.F. Huang, B.H. Lin and S.C. Yen, Time-averaged topological flow patterns and their influence on vortex shedding of a square cylinder in crossflow at incidence, *J. Fluids & Struct.*, **26**, 406–429, (2010).

[11] A. Kareem and C.M. Cheng, Pressure and force fluctuations on isolated roughened circular cylinders of finite height in boundary layer flows, *J. Fluids & Struct.*, **13**, 907–933, (1999).