Paper C

Energy-scale Aware Feature Extraction for Flow Visualization

Armin Pobitzer[1], Murat Tutkun[2,3], Øyvind Anreassen[2], Raphael Fuchs[4], Ronald Peikert[4], and Helwig Hauser[1]

[1]Department of Informatics, University of Bergen, Norway
[2]Norwegian Defence Research Establishment, Norway
[3]Université Lille Nord de France, France
[4]ETH Zürich, Switzerland

C

Abstract

In the visualization of flow simulation data, feature detectors often tend to result in overly rich response, making some sort of filtering or simplification necessary to convey meaningful images. In this paper we present an approach that builds upon a decomposition of the flow field according to dynamical importance of different scales of motion energy. Focusing on the high-energy scales leads to a reduction of the flow field while retaining the underlying physical process. The presented method acknowledges the intrinsic structures of the flow according to its energy and therefore allows to focus on the energetically most interesting aspects of the flow. Our analysis shows that this approach can be used for methods based on both local feature extraction and particle integration and we provide a discussion of the error caused by the approximation. Finally, we illustrate the use of the proposed approach for both a local and a global feature detector and in the context of numerical flow simulations.

This article was published in Computer Graphics Forum, 30(3):771–780, 2011.

1 Introduction

Fluid flow is one of the most common phenomena occurring in nature and industrial processes. One may think of air flow around vehicles, fluid flow through turbines, blood flow in vessels, or weather phenomena driven by atmospheric flows. Detailed numerical flow simulations and carefully executed experimental measurements provide accesses to high-quality, time-dependent vector field data.

Even though today's technology and tools, together with available computer power, enable us to visualize large vector fields directly, it often is not interesting, perhaps not necessary either, to visualize every detail one can find in the data. Often, the aim is to find and extract certain features of flow fields instead, using appropriate feature extraction techniques, as surveyed by Post et al. [162].

As the availability of computational time and power increase, numerical simulations are steadily carried out for more complex configurations. For such complex simulations, in general, feature extraction methodologies also produce somewhat dense outputs, which makes interpretation of the data using visualization tools more difficult. Removal of the unimportant details enhancing the complexity in our visualizations is, therefore, needed.

A number of different approaches concerning removal of the details and simplification of the output of feature detectors have been investigated extensively in the past. Most of them have geometric measures (e.g., length, area, volume of structures, or the reciprocal distance) as decision criteria [170, 21, 77]. These methods are in the tradition of classical image processing [52]. The main drawback of these methods is associated with the fact that a proper assessment on what is removed and what is actually retained in the data is very difficult to perform. The relation between geometric filtering and intrinsic properties of the flow is not clear at all.

Flow fields, in particular turbulent flows, are formed by different scales of motion. Exploring the properties of these scales are crucial in order to understand underlying dynamics. The most common way of investigating these different scales of motion is to characterize them by their turbulence kinetic energy. This is due to the fact that turbulence essentially is a cascade of kinetic energy, extracted from mean flow at the largest scales and dissipated into internal heat at the smallest scales. Therefore, identifying and describing the scales carrying significant amount of energy is important in order to enhance our ability to understand, predict, and control turbulence.

Once the features connected to the large energetic scales are found, temporal and spatial evolution of these features can be investigated. Even though there

are several ways to find and visualize the large scales of motions, quantification of the observations is found to be difficult. On the other hand, *Proper Orthogonal Decomposition* (POD), which is detailed in Sec. 3, is proven to be an unbiased and mathematical way of describing turbulent scales of motion [123]. In this way, it is possible to describe the scales of motion, or the so-called modes, and sort them according to their energy content. Decomposing the flow field according to its inherent energy levels also guarantees a protection against oversimplification when the removal of unimportant details is needed.

In fluid dynamics, POD is a well-established method, in particular within the turbulence community, which originally was introduced to the field by Lumley in 1967 [123]. One of the strengths of the method is its independence of user-defined parameters and thresholds. The decomposition is solely data-driven and does not require any prior knowledge about the different scales of motion and their energy-scale relation.

Once the decomposition has been carried out, the original field can be represented as a combination of different orthogonal basis functions, which are sorted according to the *turbulence kinetic energy* carried by each of these functions. The reconstruction of the velocity field can be carried out for a chosen subset of the basis functions, as well. In this case, the reconstruction can be performed, for example, using the modes containing the largest amount of energy, if the purpose is to remove high frequency - low energy scales. This may sound similar to low-pass filtering in one way, however, the cut-off chosen in POD is purely physical and has no unwanted and unexpected effects on the actual data. As recently discussed by Velte et al. [216], low-pass filters, due to the convolution of system functions with the actual signal, affect both time and length scales of the flow. This certainly is an undesirable situation.

The approach proposed in this paper leads to an energy-scale aware extraction of features, by first breaking the field into its energy-components and then applying conventional feature extraction methods. This differs from the usual approach, which first extracts features directly from the flow, and then condenses the output. Since the reduction in complexity in the proposed methodology in this paper is realized according to the dynamical importance of different scales and energy, it ensures that the feature detection essentially captures the characteristic structures embedded into the whole field. Fig. 1 compares possible pipelines following the standard procedure (left branch) to our approach (right branch).

Since the analysis of time-dependent flow has recently received a significant amount of attention, we investigate also some examples to demonstrate that our approach is indeed applicable for not only local feature detectors, but also to the methodologies based on particle integration.

Dette er en postprint-versjon / This is a postprint version.
DOI til publisert versjon / DOI to published version: 10.1111/j.1467-8659.2011.01926.x

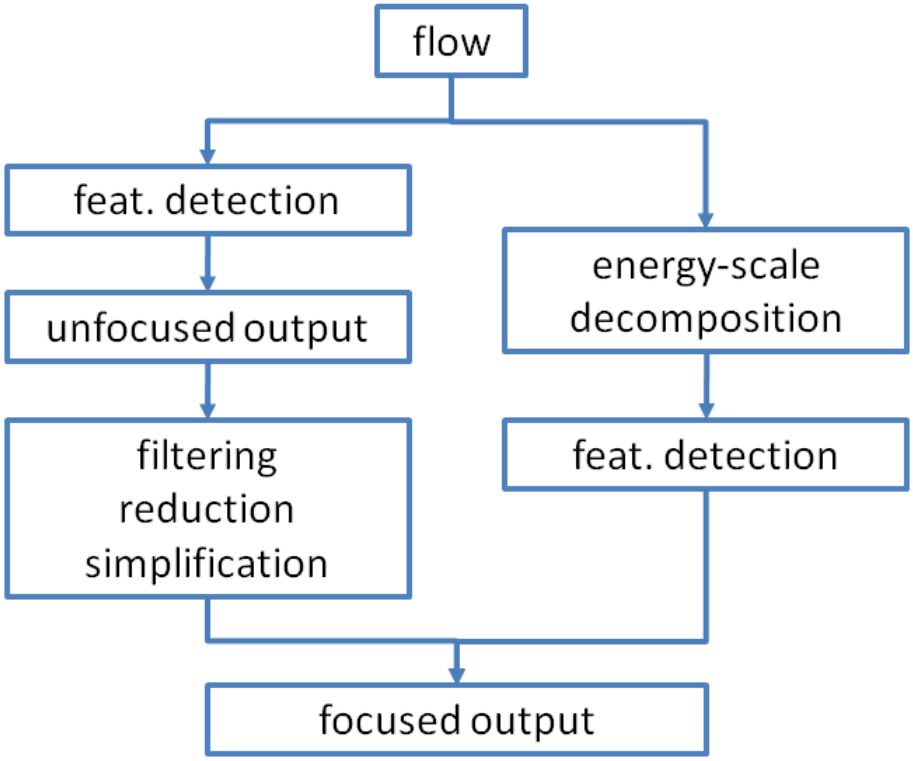


Figure 1: Comparison of the feature extraction pipeline following classical approaches to the left and the pipeline according to our approach to the right

The main contributions of this paper are:

- Utilization of the of POD for flow visualization in combination with feature extraction,
- Detailed analysis of both local and integrative error of the condensed fields,
- A way to link energy-scales and flow features.

The paper is structured as follows: first we give an overview of related work, then we discuss the theoretical foundations of the POD. In the subsequent sections we explain how POD can be used for energy-scale aware feature extraction and present results along with an error analysis. Finally, we discuss the results obtained and future work.

This paper is a cooperation between visualization researchers at the University of Bergen, Norway and ETH Zurich, Switzerland and physicists at the

Applied Fluid Mechanics group at the Norwegian Defence Research Establishment (FFI), Kjeller, Norway.

2 Related work

Proper orthogonal decomposition has been used extensively by the turbulence community in order to break the flow into different scales of motion. The decomposition results in a description of the flow field using a set of orthogonal functions, which are called eigenfunctions (also called eigenmodes or basis functions). POD is very efficient in extracting the characteristic scales of motions according to their energy content. Furthermore, the result is optimal in the sense that the first mode has the largest amount of energy, and second mode contains the second largest amount of energy, and so on.

The successful application of POD dates back to the 1980s, when POD was used to analyze a high Reynolds number, axis-symmetric mixing layer by Glauser [48] and low Reynolds number turbulent pipe flow by Herzog [76]. It has been shown in these studies that the orthogonal decomposition was very efficient at organizing the data such that the first POD mode contained about 40% of the turbulence kinetic energy, whereas the second and third POD modes contain another 40% of the energy. Since then, POD has been extensively utilized for different turbulent flows and configurations, as summarized by Tutkun et al. [215]. In recent years, POD and its capability of capturing the essential dynamics by the least possible number of scales have opened new research directions, such as low-order dynamical system modeling, flow control, and data compression/storage.

Most of the existing work on flow decomposition for visualization purposes focuses on the Helmholtz-Hodge decomposition. By this decomposition, a vector field is split uniquely into an irrotational component which is orthogonal to the boundary and a divergence-free component which is tangential to the boundary. It is also possible to split off a third component which is both irrotational and divergence-free. This methodology has been used in visualization for computing flow topology in both 2D [160] and 3D [209] and also for the 3D meshless (SPH) case [153]. Wiebel et al. used the Helmholtz-Hodge decomposition in order to analyze the deviations from a potential flow [228]. An overview of partition-based flow visualization methods is given by Salzbrunn et al. [175].

The motivation for a *simplification of vector fields* can be either data compression or complexity reduction. Complexity reduction is usually done in order to get a simpler and clearer visualization. In both cases, of course, an important goal is to preserve the important structures within the field, while removing the unnecessary details. There are two basic approaches to vector field sim-

Dette er en postprint-versjon / This is a postprint version.
DOI til publisert versjon / DOI to published version: 10.1111/j.1467-8659.2011.01926.x

plification: The first one is based on mesh reduction techniques such as edge collapses or vertex removals and the reduction of the point set in a meshless representation. The second one, on the other hand, is based on modifying data values, e.g., with local or global filtering.

Dey et al. [23] proposed a method, based on vertex removal and Delaunay triangulation, to simplify the underlying mesh. They allow simplification steps controlled by an pointwise error threshold. Instead of such a numerical pointwise error treshold method, a topological criterion is often chosen for guiding simplification. De Leeuw and van Liere [19] proposed to compute the set of critical points and to collapse groups of them to single critical points. Simplification in their method is possible if the group lies within a small neighborhood and the removed points consists of pairs of saddles and non-saddles. After collapsing the critical points, a new vector field is synthesized in the respective neighborhood. Tricoche et al. [210] presented a similar approach, where a cancellation of pairs of critical points is conducted instead of collapsing a neighborhood. Their simplification criterion is the existence of a common separatrix, connecting the two critical points. Theisel et al. [198] defined an importance criterion for critical points by their persistence under iterated Laplacian smoothing. This is then used for applying edge collapse operations under the constraint that important critical points and separatrices are not removed. Laney et al. [104] show how discrete Mores theory and combinatorial vector fields can be used to achieve topological simplification in the context of scalar fields. This simplification is based on a measure called *persistence* (which differs from the work of Theisel et al. discussed above). Morse theory is also applicable to vector field topology. Recently Reininghaus et al. [164] presented an approximative algorithm that removes the previously limited applicability to real live data. In the context of vector field topology, it is important to mention that this is a streamline-based view on the flow and that if the extraction of additional features is desired, it still would have to be based on the original field.

One important class of features in the context of fluid dynamics are vortical structures. Cucitore et al. [16] review Eulerian detectors (Hunt's Q, λ_2, swirl, and others) and suggest a non-local measure of swirl, based on trajectories to extract vortices. Jiang et al. [82] search for trajectories rotating about a common axis to verify the existence of a vortex, while Sadarjoen and Post [167] compute curvature centers of trajectories. The method introduced by Lugt [122] requires a vortex to be a portion of the fluid moving around a common axis. As an indicator for such a structure, the author proposes closed or spiralling pathlines. Haller [61] describes vortices through the stability of manifold structures which are related to fluid trajectories. The M_z criterion [61] can be considered as an accumulation of a local measure, which is based on the strain tensor along a trajectory. Haller [61] (see also Sahner et al. [173])

therefore adds up all time steps along the trajectory at which the particle is classified to belong to a vortex.

Another class of interesting features in flows is separation and coherent motion structures. The finite-time Lyapunov exponent (FTLE), as described by Haller [60], can be used to measure separation of trajectories in time dependent flows and to extract the *Lagrangian Coherent Structures* (LCS). Sadlo and Peikert [168] extract ridges from 3D FTLE fields efficiently. Garth et al. [42] present an efficient approximation and show as that 3D FTLE might be approximated by 2D FTLE in selected cross-sections. Green et al. [53] discuss the application of Lyapunov exponents for the extraction and visualization of vortices.

In recent work by Olcay et al. [138], the influence of noise and the spatio-temporal resolution of the velocity field on the extracted LCS is investigated. The authors show that a coarse resolution can significantly influence the location of a LCS. Smoothing the field is shown to have the same effect. Spatial noise can have a significant effect on single realizations of the LCS, but the mean location remains near the LCS extracted from the unperturbed field. Lagrangian smoothing has been shown to be better than a purely steady analysis by Shi et al. [186] and by Fuchs et al. [41].

3 The proper orthogonal decomposition

The *Proper Orthogonal Decomposition* is based on the two-point correlation tensor. Lumley's original formulation [123] is usually referred to as the *classical POD*, which is both time- and space-continuous. Data sets that are obtained through numerical simulations are usually studied using *snapshot POD* methodology, which can be considered as time-discrete formulation of the classical POD, introduced by Sirovich [189]. We restrict ourselves to the presentation of this method. Let

$$\mathbf{u}(\mathbf{x}, t) \tag{1}$$

denote the velocity field and

$$\mathbf{u}_n(\mathbf{x}) = \mathbf{u}(\mathbf{x}, t_n) \tag{2}$$

the n-th snapshot of the field. Then the POD of the function space spanned by N snapshots is an orthonormal basis (ONB) formed of N functions $\sigma^i \in L^2$ such that each σ^i fulfills

$$\sigma^i = \underset{\sigma \in \mathrm{L}^2 \setminus \langle \sigma^1, \ldots, \sigma^{i-1} \rangle}{\arg\max} \frac{1}{N} \sum_{n=1}^{N} \langle \sigma, \mathbf{u}_n \rangle^2 \tag{3}$$

Dette er en postprint-versjon / This is a postprint version.
DOI til publisert versjon / DOI to published version: 10.1111/j.1467-8659.2011.01926.x

Figure 2: Flow in a T-junction and its POD modes, constructed from 100 snapshots. From left to right: the actual flow (inflow from the left and top), its first and 6$^{\text{th}}$ POD mode. The background color is the norm of the point wise contribution of the respective mode to the reconstruction in the respective time step. We see that the mode with the highest energy content gives a simpler instantaneous flow pattern that the flow itself and reveals structures buried under other scales. The arrows are scaled according to norm of the respective vectors in the single modes

Dette er en postprint-versjon / This is a postprint version.
DOI til publisert versjon / DOI to published version: 10.1111/j.1467-8659.2011.01926.x

$\langle \mathbf{f}, \mathbf{g} \rangle = \int_\Omega \langle \mathbf{f}(\mathbf{x}), \mathbf{g}(\mathbf{x}) \rangle_{\mathbb{R}^3} \, \mathrm{d}\lambda^3(\mathbf{x})$ denotes the usual scalar product on L^2 function space (quadratically Lebesgue-integrable functions). Notice that the maximum values in the previous expression are descending as i increases. The obtained basis functions are usually referred to as *modes*. Given such an *orthonormal basis* (ONB) $\left(\sigma^i\right)_{i=1}^N$, the single snapshots have the following representation in this new basis

$$\mathbf{u}_n(\mathbf{x}) = \sum_{i=1}^{N} \left\langle \mathbf{u}_n, \sigma^i \right\rangle \sigma^i(\mathbf{x}) \tag{4}$$

The square of this coefficient represents the amount of energy of the flow field in "direction" of σ^i at time t_n. The total amount of energy in a time step can be obtained by summing over the index i. Further details can be found in Pedersen's thesis [144].

The strength of the POD is that it is a parameter-free method. The basis functions (and hence any linear combination of them) always fulfill the given boundary conditions and inherit the properties of the original field, such as divergence-freeness (cf. Berkooz [4] and Berkooz et al. [5]).

The flow is a T-junction, together with some of its POD modes is shown in Fig 2. One inlet is in horizontal direction, another one in vertical direction. An obstacle is placed under the vertical inlet. The fluid flows through the horizontal inlet first, while the inflow from the top begins after some time. The data set consists of 35781 vertices organized in a Cartesian grid. We decompose the field using POD based on all 100 snapshots. To the left the instantaneous flow in a cross section is shown, the color of the background reveals the velocity magnitude. The picture in the middle shows the first mode of the POD. We see a clearly simplified flow pattern, revealing structures that have not been so clearly visibly in the original flow field. To the right, the 6^{th} mode of the same POD is shown. As it is explicitly visible, the complexity increases with an increasing mode number, such as the enhancement of small vortical structures. Of course, the importance of the mode decreases as its order increases. This can be seen by the color code on the background of each of these figures. The color is assigned by the pointwise norm of the terms used in a reconstruction. The first POD mode, shown in the middle, recovers almost everything shown in the left figure, which is the full field. The 6^{th} mode does not have a significant contribution. All arrows are scaled according to the respective vector norm.

It is worthwhile keeping in mind that the single POD modes are static fields, i.e., only limited information about the temporal dynamics of the system can be retrieved directly from them.

Dette er en postprint-versjon / This is a postprint version.
DOI til publisert versjon / DOI to published version: 10.1111/j.1467-8659.2011.01926.x

3.1 Practical computation

The defining equation (3) for the POD is reformulated as an eigenvalue problem of the dimension of the number of grid cells (cf. [144]). Since this number is typically several orders of magnitude bigger than the number of available snapshots, we use an alternative formulation that leads to an eigenvalue problem of the dimension of the number of snapshots [144]. This reduced eigenvalue problem is then given by

$$\sum_{m=1}^{N} C_{n,m} a_m^{(i)} = \lambda_{(i)} a_n^{(i)} \qquad \text{with } C_{n,m} = \frac{1}{N} \langle \mathbf{u}_n, \mathbf{u}_m \rangle \tag{5}$$

The modes can be computed as follows

$$\sigma^i = \frac{\sum_{n=1}^{N} a_n^i \mathbf{u}_n}{|| \sum_{n=1}^{N} a_n^i \mathbf{u}_n ||} \tag{6}$$

Notice that this formulation requires the reordering of the σ_i according to the associated λ_i-values in order to guarantee the descending energy content. The fulfillment of the ONB condition follows directly from the algorithm. The relative energy content of the i^{th} mode with respect to the total energy is given by

$$\frac{\lambda_i}{\sum_{i=1}^{N} \lambda_i} \tag{7}$$

3.2 Global flow field approximation using POD

The original velocity field can be reconstructed using the basis functions obtained by solving the POD equation. Since the recovery of the full field can be performed by summing the relevant basis functions and their associated coefficients, an approximated velocity field can be constructed using, for instance, only the modes with large amount of energy. This is achieved by choosing the maximal index i_p such that $(\sum_{i=1}^{i_p} \lambda_i)/(\sum_{i=1}^{N} \lambda_i) \leq p$ for a desired p. Hence,

$$\mathbf{u}_p(\mathbf{x}, t) = \sum_{i=1}^{i_p} \left\langle \mathbf{u}_t, \sigma^i \right\rangle \sigma^i(\mathbf{x}) \tag{8}$$

is an approximation to the original field, based on the i_p most dominant motion energy-scales that capture $p \cdot 100$ percent of the total turbulence kinetic energy for the full field.

4 Energy-scale aware feature extraction

In order to focus on energetically important features, instead of applying the feature detector to the flow field in the first place and then coping with the unfocused output, we propose the following approach:

1. Decomposition of the original flow field using POD,
2. Reconstruction of the field by selecting the most dominant modes in terms of energy using eq. (8),
3. Utilization of a feature detector.

The main advantage of using this method is related to the fact that the extracted features are the features which are associated with the dynamically most important scales of motion in the flow.

Since the proposed approach is based on an approximation of the whole field, it is necessary to quantify the error as induced by the truncation that is applied during the reconstruction. In the context of feature detection, both local errors, and integrative (or global) errors need to be addressed. It is important to point out that POD operates in L^2, i.e., there is no guaranteed bound for the local error. This means that, in theory, even an approximation that contains almost 100% of the energy, can have comparably large local errors. The situation for the integrative error is even less clear, since we know from theory that small local error along the integration path can accumulate to a large error in the final particle position [57].

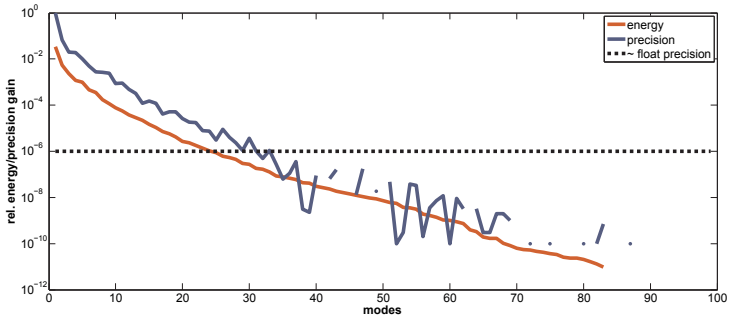
5 Results

In this section we illustrate the utilization of the POD methodology for energy-scale aware feature extraction on one local (vorticity) and one integration based method (Finite-time Lyapunov exponent - FTLE).

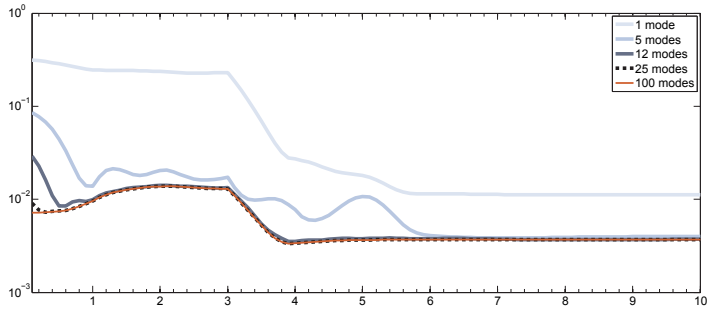
Vorticity thresholding: This is a classic method to detect vortical flow behavior. Vorticity ω is defined as the curl of the flow field $\mathbf{v}$, i.e. $\omega = \nabla \times \mathbf{v}$. Vorticity has the appeal that it is a Galilean invariant measure and simplicity used in a number of other vortex detectors [83]. Hence, investigating the impact of POD on vorticity gives indications regarding its applicability and implications for a whole class of vortex detectors.

Finite-Time Lyapunov Exponent: Let $\mathbf{v}$ be any flow field and $\varphi_{t_0}^T$ its flow map defined by $\varphi_{t_0}^T(\mathbf{x}_0) := \mathbf{x}(T)$, provided that $\mathbf{x}$ is the solution of the initial

DOI til publisert versjon / DOI to published version: 10.1111/j.1467-8659.2011.01926.x



(a) Energy and error statistics for n-th order approximation



(b) Temporal error evolution for selected approximations

Figure 3: (a) The relative energy and precision gain of the i-th order approximation compared to the $(i-1)$st approximation. The precision is calculated using the time averaged relative L^2-error of the respective approximations. (b) The function graph of the L^2-error plotted against time for several approximations

value problem $\dot{\mathbf{x}}(t) = \mathbf{v}(\mathbf{x}, t), \quad \mathbf{x}(t_0) = \mathbf{x}_0$. Then the (maximum) finite-time Lyapunov exponent (FTLE) λ is defined as $\lambda := \ln ||\nabla \varphi_{t_0}^T(\mathbf{x})||/|T - t_0|$.

We apply these methods on two different data sets from numerical flow simulations: The simulation of flow in a T-junction and the *direct numerical simulation* (DNS) of turbulent channel flow.

5.1 Flow in a T-junction

We apply the proposed approach to the simulation of a flow in a complex T-junction with two inlets and one outlet. See Sec. 3 and Fig. 2 for descriptions of the data set and flow geometry, respectively. The red curve in Fig. C.3(a) shows

Dette er en postprint-versjon / This is a postprint version.
DOI til publisert versjon / DOI to published version: 10.1111/j.1467-8659.2011.01926.x

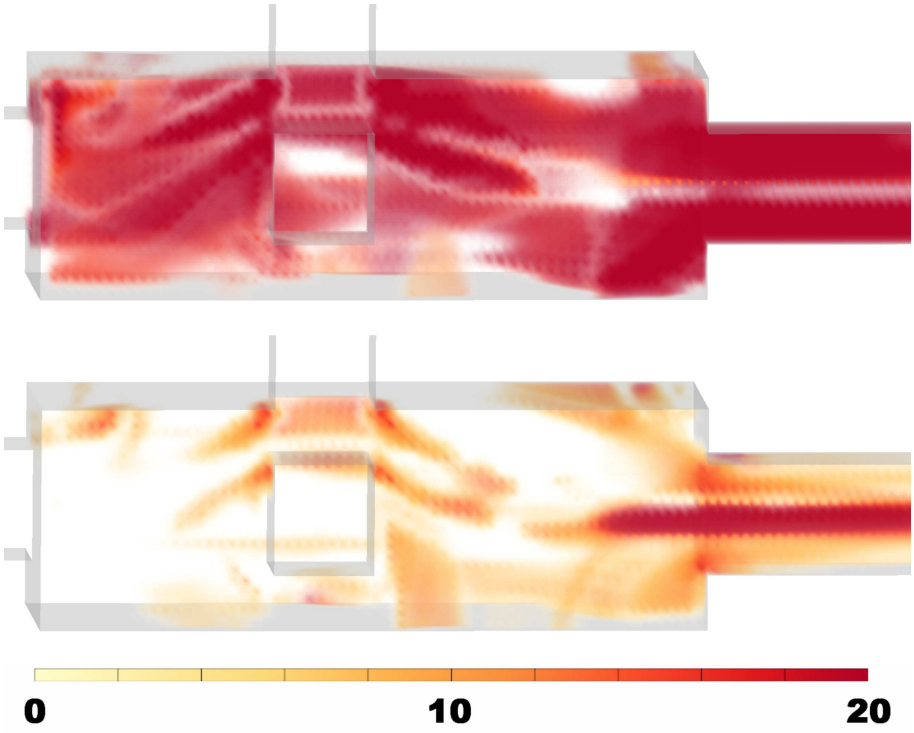


Figure 4: Vorticity method: Both figures show the same threshold, but the scalar vorticity field has been computed from the original field (top) and the 5th approximation (bottom), respectively. We see a strong reduction of the structures in the back of the top picture when taking the most dominant modes only. In the outflow, the vorticity field based on the most dominant modes reveals one instead of two vortices. This effect is called vortex braiding. For further discussion we refer to the text.

the relative energy gain in adding the i^{th} mode to the approximation based on $i-1$ modes. The blue curve in teh same figure shows the relative precision gain with increasing number of modes used in the approximation. The precision is computed as the time-averaged relative L^2 error of the approximation and is plotted against the number of modes used for the approximation. The bumps in the function graph indicate modes with the same order of energy. Including just one of them will increase the error since it means splitting an energy scale. The gaps in the graph indicate zero gain, which can not be plotted due to the logarithmic scale on the y-axis. The rather strong fluctuations indicate that we are touching upon the boundaries of numerical accuracy for this dataset.

Dette er en postprint-versjon / This is a postprint version.
DOI til publisert versjon / DOI to published version: 10.1111/j.1467-8659.2011.01926.x

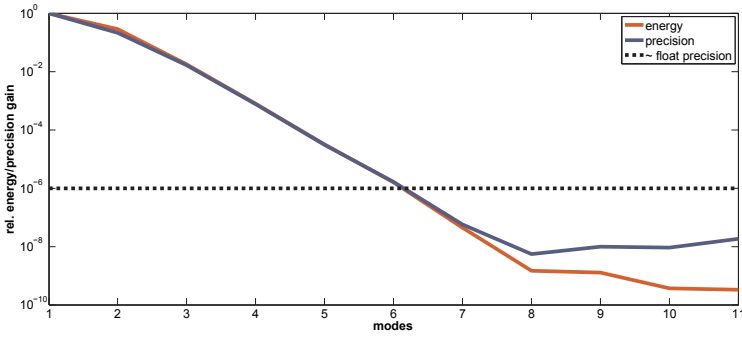
Vorticity thresholding: First, we examine the relative L^2-error of the reconstruction per time steps. Fig. C.3(b) shows the error curves for approximations based on the i most dominant modes.

The vorticity field is derived from the approximations using the first 5 modes (accounting for 88 percent of the motion energy), in addition to the original field. We confine ourselves to the demonstration of the impact of the usage of POD-based approximations. Therefore, we choose a threshold for the original field and apply it to the field derived from the approximation (The choice of a suitable threshold value is a subject by itself and not covered herein). Fig. 4 shows the respective fields. Not surprisingly, rather strong vortex regions are present in the vicinity of the vertical inlet. Even though these features are present in both fields, we see that we have a reduced response for the approximation. On of the most interesting observations is that the two vortices in the outflow, present in the original field, become one, when focusing only on the largest energy-scales. The center of the vortex is in the middle of the two previously detected ones. This can be attributed to vortex braiding, which is due to spiraling of these two vorticies around each other [17]. This also shows that our proposed method captures features which cannot be obtained by post filtering. This is simply due to the fact that this structure was not detectable in the original output at the first place.

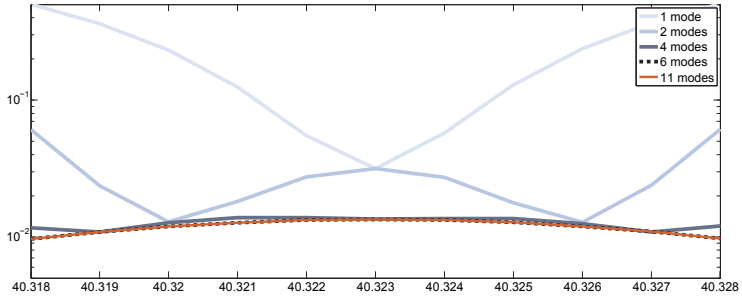
5.2 Turbulent channel flow

This data set is a *direct numerical simulation* (DNS) of a fully developed turbulent channel flow at frictional Reynolds number Re_τ of 180. The flow domain is bounded by two infinitely large parallel solid walls, and the flow is driven by constant mean pressure gradient in the in the streamwise (x) direction. . The boundary conditions are non-slip on the solid walls and periodic else. The data are produced by a *Spectral Element Method* (SEM) solver developed at *The Norwegian Defence Research Establishment* (FFI) [219]. The dataset consists of 2146689 vertices arranged in a rectilinear grid. We decompose the field using POD with 11 snapshots (all time steps), the red curve in Fig. C.5(a) shows the energy gain adding the i-th mode to the approximation using of order $i-1$. The blue curve shows the precision gain, showing by how much the error decreases adding the i-th mode. The error is the relative error between original flow field and the approximation with respect to the L^2-norm, averaged over all time steps.

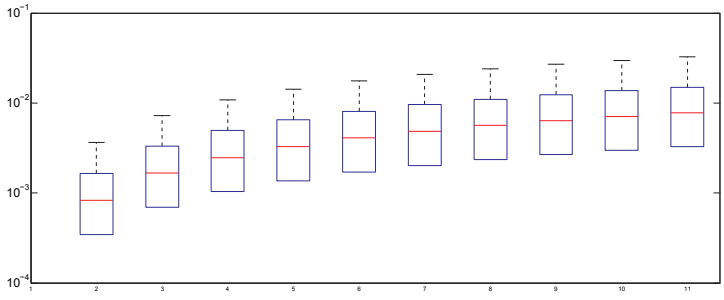
Vorticity: As in the previous data set, we examine the relative L^2-error of the reconstruction per time step first. Fig. C.5(b) shows the error curves for approximations based on the i most dominant modes. We see that for a higher



(a) Energy and error statistics for n-th order approximation



(b) Temporal error evolution for selected approximations



(c) integrative error per integration length

Figure 5: (a) The relative energy and precision gain of the ith order approximation compared to the $(i-1)$st approximation. The precision is calculated using the time averaged relative L^2-error of the respective approximations. (b) The function graph of the L^2-error plotted against time for several approximations. (c) Box plot of the integrative error of the 8-th order approximation. The groups represent particles seeded at all cell centers and advected for $n \cdot dt$, dt being the constant time-sampling distance of the data set.

Dette er en postprint-versjon / This is a postprint version.
DOI til publisert versjon / DOI to published version: 10.1111/j.1467-8659.2011.01926.x

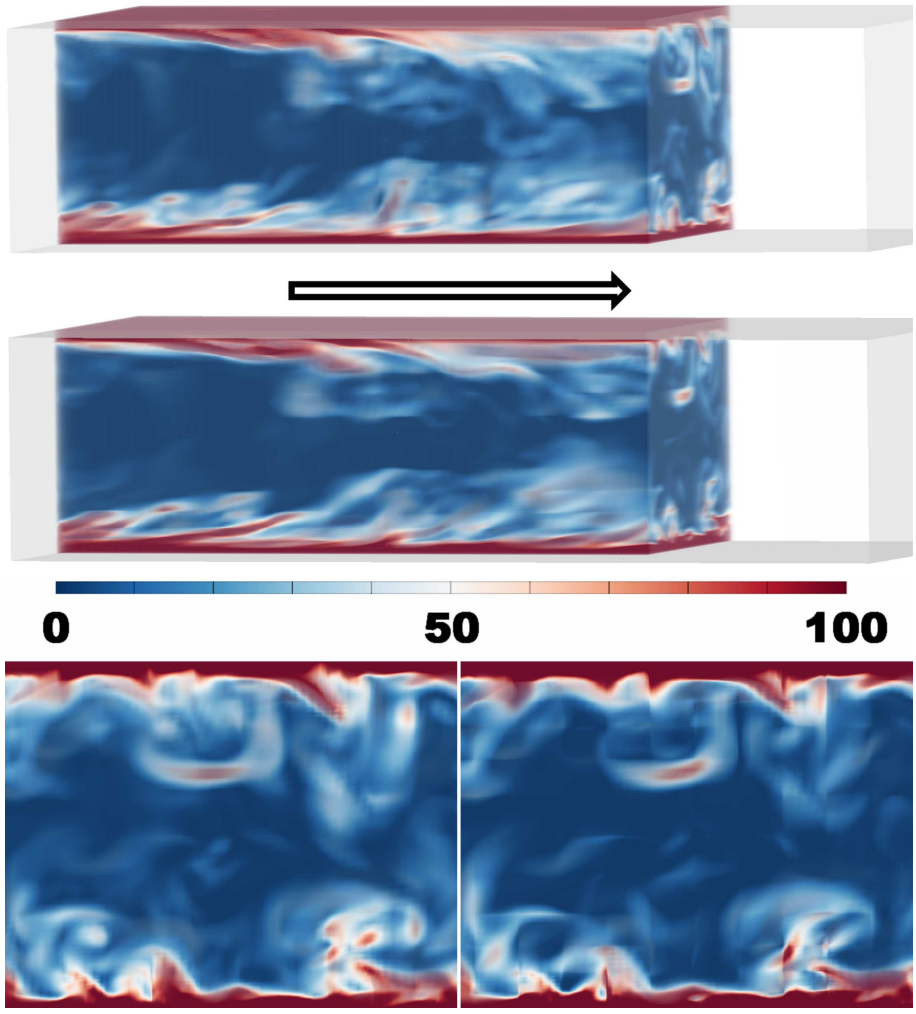


Figure 6: Vorticity: Both figures show the scalar vorticity field, using the same color scale, but the fields has been computed from the original field (top) and the 4-th approximation (middle), respectively. The bottom row shows the original and the approximated field on the back-facing clipping plane, to the left and right, respectively. On of the most interesting observations is the behavior of the three features in the bottom right corner. For discussion of this and further features we refer to the text.

Dette er en postprint-versjon / This is a postprint version.
DOI til publisert versjon / DOI to published version: 10.1111/j.1467-8659.2011.01926.x

number of modes the approximation quickly converges to the total reconstruction using all modes. The relative L^2-error is already for an approximation using only the first 4 modes of the order of 10^{-2}, i.e., 1%

We obtain the vorticity from the original field and the approximations using 4 modes, accounting for approximately 93 percent of the motion energy. Fig. 6 displays a subsection of the flow domain bounded by two clipping planes, one orthogonal to the mean flow (yz-plane) and one normal to this first plane and the solid walls (xz-plane). In order to assess the impact of the use of different approximations, we applied the same color coding to the scalar vorticity fields derived from original field and the approximation. On the stream-aligned clipping plane we see a clear reduction of structures of high scalar vorticity. Vortex tubes emerging from the wall have better spacial coherence. The two planes in the bottom row are the fields on the first clipping plane for the original field (left) and the approximation (right). On these planes we observe how single features disappear (left lower corner, upper right corner). In the bottom right corner only one of three apparently equally strong features of the original field turns out to be kinetically important. Other structures, as the horizontal feature in the middle upper part of the cross section, consist unchanged.

Finite-Time Lyapunov Exponent: We examine the integration error, seeding a particle per vertex and advecting them over the whole time span the data set is defined. Every time the integration time is a multiple of the time interval corresponding to the distance between two time steps we compare the particle position from advection in the original field with the particle position obtained from advection in the approximated field. Every multiple of this time interval is represented as one group in the box plot in Fig. C.5(c). The lines in the middle of the single boxes represents the median of the group, where the box itself is the interquartile range. The whiskers give information about the tail-shape in the distribution, showing the maximum/minimum sample value that is not considered an outlier, i.e., within median $\pm 2.7 \times$ (standard deviation). For normal distributed data this corresponds to approximately 99% of the sample points [132]. We see that the median is of order 10^{-3} to 10^{-2}, and the upper whisker reaches at most to a height roughly corresponding to 0.05. Given that the local error is on average of order 10^{-2} as seen in Fig. C.5(b), we can conclude that the loss of precision induced by integration is sufficiently small. This certainly ensures an accurate and effective usage of POD for integration based feature detectors as well. This is an interesting result, since local errors usually grow exponentially when integrated.

Application of the proposed approach to feature extraction based on *finite-time Lyapunov exponents* (FTLE), introduced by Haller [59] is also studied as follows. Fig. 7 shows the FTLE field for the original flow on top and the FTLE

Dette er en postprint-versjon / This is a postprint version.
DOI til publisert versjon / DOI to published version: 10.1111/j.1467-8659.2011.01926.x

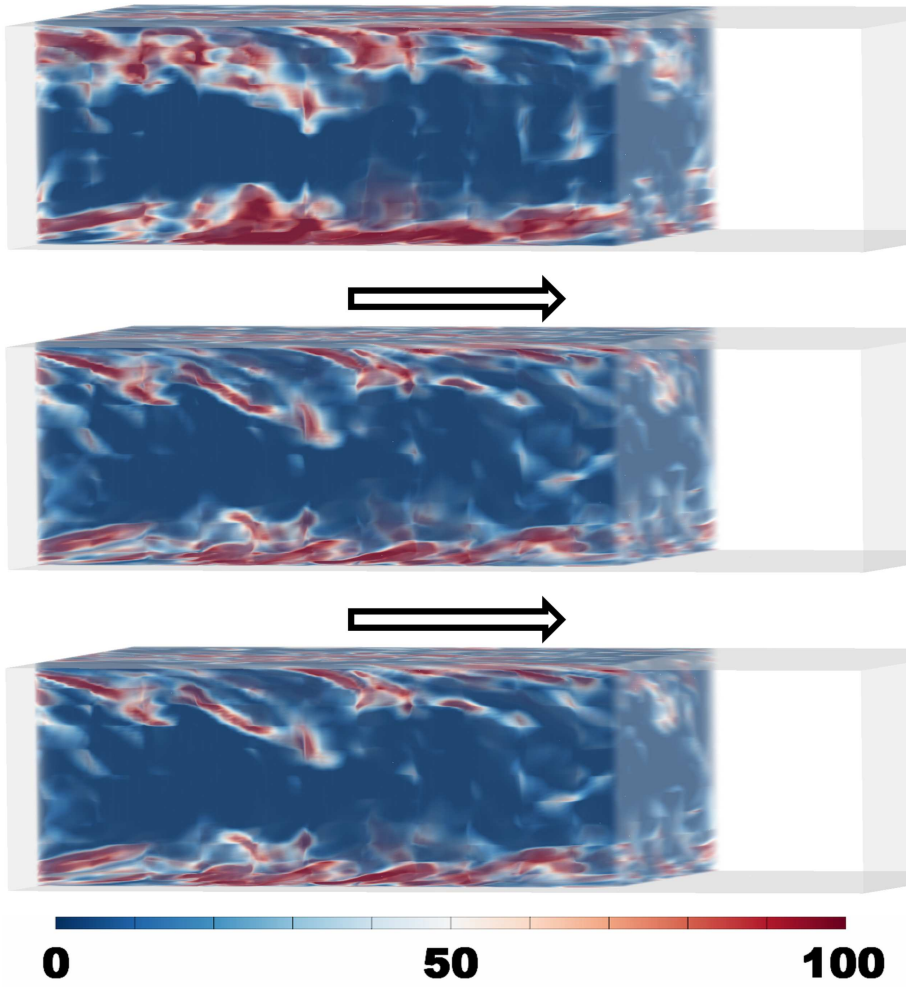


Figure 7: FTLE: As in for the local methods, we keep the color map unvaried for all three figures, but the FTLE field visualized has been computed from the original field (top), the $2nd$ (middle) and the $4th$ approximation (bottom). Observe that the focusing on the energetically most dominant scales of motion energy yields a more crispy and detailed output with finer lobes. Adding two more modes does not chance the output, even though just two modes where used for the first approximation. This indicates that the dissipative scales have to be interpreted as "noise" in the context of integration-based feature extraction. We refer to the text for further discussion.

Dette er en postprint-versjon / This is a postprint version.
DOI til publisert versjon / DOI to published version: 10.1111/j.1467-8659.2011.01926.x

field for the approximation based on 2 modes in the middle. The first 2 modes capture approximately 79% of the energy. We see that certain features increase its visibility and are more crisp. Furthermore the focusing on the energetically most relevant scales reveals finer lobes that appear to be a single one in the FTLE field based on the original flow. An interesting observation is that taking the approximation from 2 to 4 modes does not change the output significantly. This can heuristically be explained by the fact that FTLE is due to its definition strongly dependent on the transport properties of the underlying flow. These properties are in turn dominated by the large motion energy-scales. We see that structures related to high FTLE values are highly energetic per se. Here, the reduction achieved through POD is a reduction of noise.

6 Discussion and future work

This paper presents a new methodology for energy-scale aware feature extraction, which is based on the POD of the flow field. We show that the proposed approach is applicable for both local and integration-based methods. The application to the integration-based FTLE method verifies the fact that particle motion is mainly due to large energy-scales. This indicates that low energy-scales should be considered as noise in this context. Indeed, FTLE of the approximated field shows finer detail than FTLE of the original field.

It is worthwhile to note that POD of the velocity field is essentially not compatible with the widely used λ_2 criterion by Jeong and Hussain [81]. The main reason for this is that POD extracts the large scales, containing significant amount of energy. In terms of the wavenumber space, these scales are in the low-wavenumber region. On the other hand, λ_2 is mostly related to small scale vortical structures. They are, in turn, related to the high wavenumber region [217]. This implies that, computing λ_2 over the field reconstructed using just first few POD modes, corresponds to computing λ_2 on a field which does not have small scale vortical structures. Therefore, λ_2 may just show numerical noise, if only the first few modes are used in the approximation.

On the other hand, POD can also be applied to the vorticity field instead of the velocity field. The modes, or structures, of the flow will be ordered by POD according to their enstrophy, which is directly linked to the energy dissipated at the smallest scales of motions [137]. As suggested by Kostas et al. [101], dominant vortical structures in a flow can be more effectively extracted by constructing the POD using vorticity. Therefore, the next step in this work will be to implement vorticity into our POD solver and proposed methodology described herein.

In this paper, we studied turbulent channel flow driven by constant pressure

gradient. The turbulence kinetic energy and energetic scales were the main interests here and the problem was isolated accordingly. Even though different flows, such as flows under the influence of adverse pressure gradient, spatially developing boundary layers, thermal boundary layer or free shear flows, show different characteristics, the proposed method is in general applicable to any situation, because it has already been shown that application of POD on these different flows is possible and effective. Impact of these different conditions and different geometries on the proposed method will certainly be subject of future studies.

7 Acknowledgments

The authors want to thank Carl Erik Wasberg from the Norwegian Defence Research Establishment (FFI) for providing the DNS data set of the turbulent channel flow. The CFD simulation of a flow in a T-junction is courtesy of AVL List GmbH, Graz, Austria. The project SemSeg acknowledges the financial support of the Future and Emerging Technologies (FET) programme within the Seventh Framework Programme for Research of the European Commission, under FET-Open grant number 226042.