



# Statistical analysis of flow field variations via independent component analysis

J. Laichter<sup>1</sup> · P. Kranz<sup>1</sup> · S. A. Kaiser<sup>1</sup>

Received: 30 October 2023 / Accepted: 22 January 2024 / Published online: 23 February 2024  
© The Author(s) 2024

## Abstract

The link between in-cylinder flow and subsequent combustion in a single-cylinder gasoline spark-ignition engine is analyzed via independent component analysis (ICA). Experimentally, the two in-plane components of the velocity are measured in the central cylinder plane by high-speed particle image velocimetry (PIV) with the engine running slightly lean at 1500 rpm in skip-fired mode. In 213 cycles, measurements are made during the late compression stroke before ignition with approximately 1° crank-angle temporal resolution. ICA then decomposes the set of 213 flow fields at each time step, yielding a set of “source” flow patterns—the independent components (IC). The temporal coherence between the ICs is then examined in a persistence analysis, comparing each IC with the one from the previous time step starting at ignition timing and going backwards in time. The results show which ICs persist how long throughout the compression stroke. To investigate the link between the ICs and combustion, the crank angle at which 10% of the fuel are burned (CA10) in each cycle is correlated with the extent to which a given IC can be found in each flow field. The most persistent IC can be traced over more than half of the 70 degrees crank angle over which images were acquired. The IC that correlates best with CA10 visually more resembles some of the flow features found in conditional averaging of fast-burning versus slow-burning cycles.

## 1 Introduction

Cycle-to-cycle combustion variations (CCV) in internal combustion engines lead to variations in output torque and emissions. This is more pronounced in spark-ignited (SI) engines, and even more so in lean operation, the extremes being knocking or misfiring cycles. The causal chain that explains the nature of CCV is still ongoing research, even though CCV have been investigated for decades. The further development of optical diagnostics made it possible to obtain data like instantaneous flow fields and flame propagation with high spatiotemporal resolution. This has been utilized to better understand the underlying reasons for CCV and to improve predictive models (Fontanesi et al. 2015; Granet et al. 2012; Zeng et al. 2017).

Since the turbulent flow in the cylinder of an engine is far from isotropic homogeneous turbulence, but instead is structured on a large scale (e.g., swirl or tumble), it is reasonable

to assume that CCV in part is driven by variations of large or intermediate spatiotemporal scales (Reuss 2000). Therefore, one might look for “typical” flow features, which can be linked to combustion metrics such as peak pressure, combustion phasing, and indicated mean effective pressure (IMEP), such that ultimately engine design can be informed by this insight. For this, a combination of in-cylinder pressure-based combustion analysis and time-resolved imaging diagnostics is needed. High-speed imaging systems with frame rates that are adequate to resolve the turbulent time scales in combination with large storage can collect time-resolved data to study combustion-related events on a statistically significant basis (Fajardo et al. 2006; Peterson et al. 2011). Next, objective and automated processing tools are needed to analyze the large amount of imaging data. Conditional averaging is a very simple yet useful method. In the context of CCV, typically images of the flow, fuel/air mixing, or flame propagation are averaged conditionally to stemming from a subset of cycles with a specific combustion outcome, for example, being among the fastest or slowest-burning ones (Bode et al. 2019; Laichter and Kaiser 2022). Correlation maps (Laichter et al. 2023) are used to investigate how well a single-valued quantity correlates with a scalar field quantity at any point in a two-dimensional field. Stiehl et al. (Stiehl et al. 2016)

✉ J. Laichter  
Judith.laichter@uni-due.de

<sup>1</sup> Institute for Energy and Materials Processes, University Duisburg-Essen, Carl-Benz-Straße 199, 47057 Duisburg, Germany

investigated the in-cylinder flow field and fuel spray in a direct-injection spark-ignition engine via such correlation maps. The results indicated that CCV of the large-scale tumble vortex significantly influenced the spray shape of the second injection. When data from a large number of cycles (on the order of thousands) is available, machine learning can offer valuable insights. Such machine-learning approaches were employed by Hanuschkin et al. (2021). They used the two-dimensional cross sections of the flame at a single crank-angle degree (CAD) shortly after ignition to predict high- or low-energy combustion cycles.

In addition, proper orthogonal decomposition (POD) has found a range of applications in combustion research, including internal combustion engines (Borée 2003; Chen et al. 2011; Liu and Haworth 2011). POD decomposes data in terms of dominant structures by energy and frequency (Chen et al. 2013, 2012; Druault et al. 2005; Graftieux et al. 2001). However, in the resulting decomposition, even the leading-order modes do not necessarily resemble “typical” flow features, i.e., features that one might recognize in a given single cycle (Chen et al. 2013). POD is a good tool to compare data sets (Abraham et al. 2015) or to identify subsets of data that have common features (Chen et al. 2014), but the connection between what is found in the POD mode structures and their coefficients on the one hand and physical processes on the other is not a priori guaranteed or obvious.

It appears that an alternative method, or at least further processing, is required to extract features that can be directly connected with observed flow and combustion behavior. This work explores if independent component analysis (ICA) of vector fields can extract flow features that are potentially responsible for the difference between late and early burns in lean spark-ignited premixed iso-octane/air combustion in an engine. The example studied here utilizes high-speed 2D-particle image velocimetry (PIV) to investigate cycle-to-cycle combustion variation in an optically accessible research engine.

Independent component analysis is a statistical tool used in various fields. Like POD, it is also based on single value decomposition (Hyvärinen 1999). Hyvärinen et al. (2001) investigated ICA in terms of mathematics and statistics and their research group provides an open-source code for MatLab®, FastICA (Hyvärinen n.d.).

The classic example for ICA is the so-called “cocktail party problem.” Voice signals are randomly mixed, and the task is to extract the individual speakers’ voices. ICA or POD can be used to separate the single-source signals. While POD just gives a set of dominant mixtures, ICA is able to extract the independent sound sources. The difference between ICA and POD is that ICA extracts statistically independent and non-Gaussian components from large data sets. This means that underlying statistically independent structures (blind sources) in signal mixtures are separated

by ICA, while POD modes are always mixtures of dominant patterns (Chen et al. 2012). ICA is used without knowledge of the physically properties of the sources, and therefore, the number of independent components in a data set is usually unknown (Bizon et al. 2013a; Hyvärinen and Oja 2000). However, in contrast to POD, ICA does not sort dominant structures by level of energy.

The application of ICA in engine research is quite rare. Bizon et al. (2013a, 2016b) showed that ICA with two components applied to images of flame luminosity in a diesel engine could identify two separate combustion events, the main combustion zone along on the central spray axes and the later combustion near the piston bowl. Bizon et al. (2016) also studied how ICA could be used to study the evolution of combustion luminosity for individual cycles and cycle sequences. They showed that time-dependent coefficients carry information about the CCV morphology. ICA could extract and separate moving sources from the background using a nonlinear mixing model. This was further elaborated on using artificial examples to show the limitations of applying linear ICA to nonlinear mixtures. A study based on direct numerical simulation (DNS) used ICA to analyze the streamwise velocity fluctuations occurring in turbulent channel flows. The results indicate that ICA may be suitable to connect statistical with structural descriptions of turbulence (Wu and He 2022).

Since ICA appears to fulfill the intended task of identifying “typical” large-scale structures in scalar images, it should also be applicable to velocity fields, which is explored in this work. It is demonstrated that ICA of vector fields works and provides a plausible physical picture of the processes, leading to either early or late burning cycles.

## 2 Experimental setup and PIV measurements

### 2.1 Optical engine and operating conditions

Flow imaging was performed by high-speed PIV in an optically accessible four-stroke single-cylinder SI engine with four valves and a pent roof. Figure 1 illustrates the relevant part of the engine, and Table 1 summarizes the operating conditions. The engine was operated at 1500 rpm, and flow fields were captured from  $-90$  to  $-20^\circ\text{CA}$  (i.e., for 70 crank angles) in 213 fired cycles. Running the engine in “skip-fired” mode, each fired cycle was followed by two motored cycles, which decreases the thermal load and the potential influence of the residual gas on CCV. The fuel was injected via port-fuel injection (PFI), which yields a near-homogeneous mixture at the timing of ignition. From each cycle’s record of the in-cylinder pressure trace, the heat-release rate (HRR) was calculated, and the crank angle at which 10% of

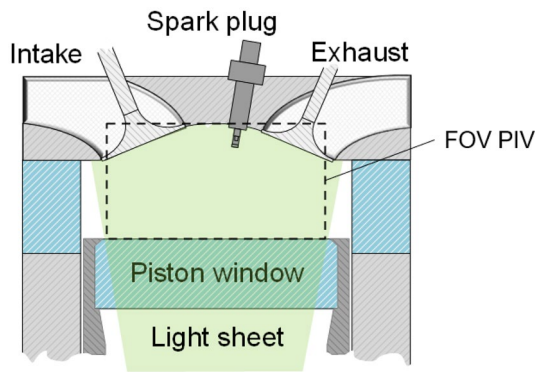


Fig. 1 Sketch of the optical engine with the field of view

Table 1 Engine properties and operating conditions

Engine type	4-stroke single cylinder
Fuel	Iso-octane
Air–fuel ratio $\lambda$	1.1
Bore/stroke	84 /90 mm
Geom. comp. ratio	9:1
Engine speed	1500 RPM
Intake pressure	1 bar
Spark timing	$-20^{\circ}\text{CA}^a$
IMEP	7.7 bar
COV of IMEP	1%

<sup>a</sup>This paper assigns  $0^{\circ}\text{CA}$  to compression top-dead center, i.e., crank angles during intake and compression are negative

the fuel is burned (CA10) was used as an indication of the speed of the initial phase of combustion.

### 2.2 Optical diagnostics and image processing

A fused-silica cylinder liner and a flat piston window provided optical access. Silicone-oil droplets were introduced  $\sim 50$  cm upstream in the intake manifold. Two laser pulses from a frequency-doubled Nd:YVO<sub>4</sub> (edgewave IS335, 532 nm, 1 mJ/ pulse, 9 ns pulse duration), temporally separated by  $\Delta t = 20 \mu\text{s}$  illuminated the tumble plane in a 0.6-mm-thick light sheet. The light scattered by the oil droplets was imaged by a CMOS camera, with 10,000 image pairs acquired per second, corresponding to about one image pair per crank angle. From each image pair, velocity vectors were calculated using LaVision’s DaVis 8.4. Multi-pass cross-correlation with an interrogation window size decreasing from  $64 \times 64$  pixels to  $32 \times 32$  pixels with 50% overlap resulted in a vector spacing of  $914 \mu\text{m}$ . The moving average of five vector fields was calculated before ICA was applied. A more detailed description of the diagnostics and the image analysis can be found in Laichter and Kaiser (2022).

### 2.3 Pressure-based analysis and PIV measurements

In Fig. 2, the pressure traces and the pressure-based quantity CA10 are plotted. With ignition at a  $-20^{\circ}\text{CA}$  and a relatively lean mixture of  $\lambda = 1.1$ , the IMEP at this operating point is 7.7 bar with a COV of 1%, which would be considered perfectly acceptable operating stability in a production engine. Nevertheless, as is typical for spark-ignited engines, the pressure traces show significant variability. A histogram

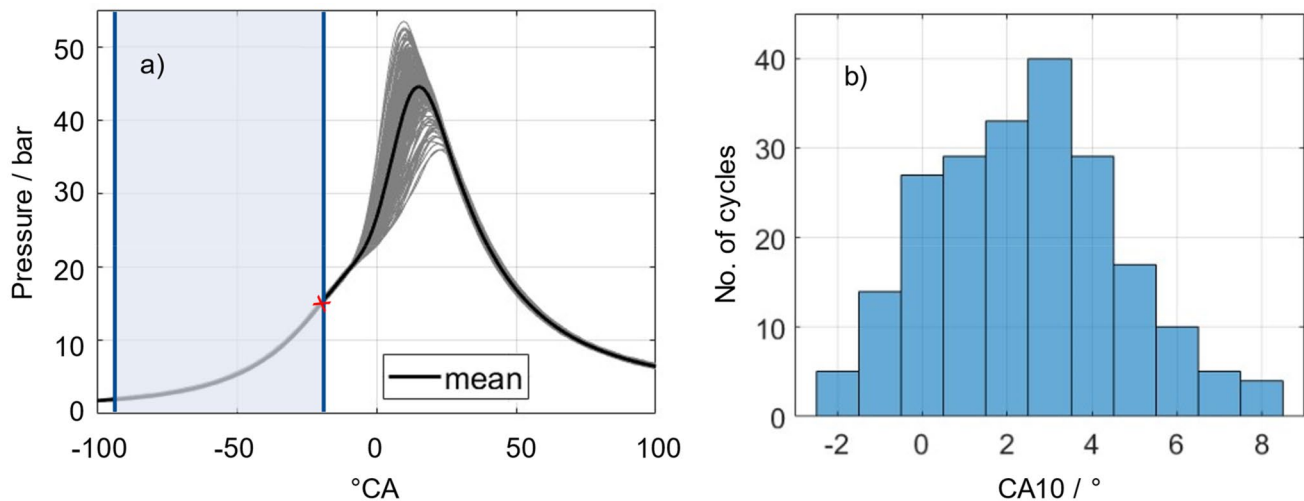


Fig. 2 (a) Pressure traces for all cycles in gray and their mean value in black. The red x indicates ignition, and the investigated crank-angle (=time) range is marked in light blue. (b) Histogram of CA10

of CA10 is shown in Fig. 2b. The distribution is near-Gaussian, slightly skewed to earlier crank angles and with the peak bin centered at 3°CA after combustion TDC. The mean is at 2.4°CA. 74% of the cycles reach CA10 between 0 and 4°CA.

To give an overview of the flow during compression, Fig. 3 shows flow fields for selected crank angles. The first three columns contain flow fields from different single cycles and the last column shows the flow fields averaged over the whole data set. At -90°CA, the typical tumble motion can be found in each single cycle and in the average. The tumble vortex center is in a different position for each single cycle, and on average located somewhat to the right of center in the field of view. The peak velocity magnitude also differs from cycle to cycle. At -60°CA, the vortex center of the average flow has moved to the exhaust (i.e., right) side of the combustion chamber, and on a single-cycle basis, the position of the vortex center still varies. This behavior continues

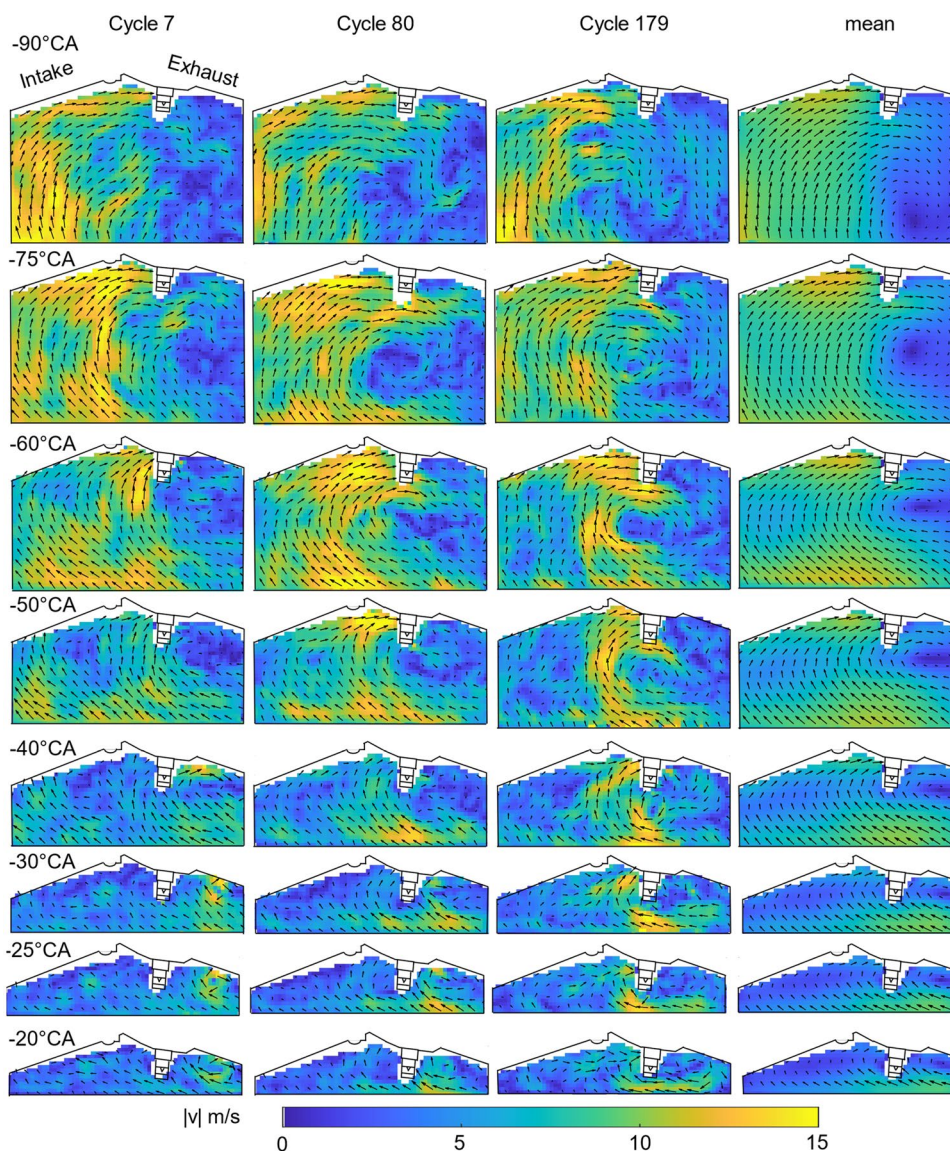
until ignition timing at -20°CA. In Cycle #7 and Cycle #80 a vortex can still be seen close to the spark plug, but in cycle #179 and also in the average the tumble vortex is now much less pronounced, reflecting the expected tumble breakdown toward TDC.

### 3 Methodology

#### 3.1 Snapshot ICA

ICA is a statistical tool for finding underlying independent source flow structures  $s_1, s_2, \dots, s_n$  from a set of flow fields  $x_1, x_2, \dots, x_m$  at a given crank angle. Each flow field  $x$  contains different proportions of the sources  $s$ . To what extent a flow field contains each of these sources is described in the mixture matrix  $A$ . This mixture matrix provides two pieces

**Fig. 3** First three columns: flow fields from three cycles at selected crank angles. Fourth column: mean flow at these crank angles. Only every 3rd vector in each dimension is shown



of information, the absolute value and the polarity. However, this information is determined through a scaled version of the source signal; thus, it is difficult to recover the length and orientation of the vector. The model can be written as

$$x = As \leftrightarrow s = Wx, \tag{1}$$

with  $W=A^{-1}$ . While  $x$  is known from the measurements (here the flow fields),  $A$  and  $s$  are unknown and must be determined iteratively, resulting in the best estimate  $y = Ws$  that maximizes the statistical independence of these estimated components  $y$  (Hyvärinen et al. 2001). By assuming that the sources  $s$  are statistically independent, where each source is characterized by a non-Gaussian distribution, the maximization of the statistical independence of the estimate  $y$  solves the basic ICA problem from Eq. (1) (Hyvärinen and Oja 2000).

As the number  $n$  of the independent sources  $s$  is often unknown and much smaller than the number  $m$  of the flow fields  $x$ , the plain application of the ICA model would result in a  $m \times n$  mixing matrix  $A$ . Thus, the rank of the data needs to be decreased. This is usually done by pre-processing based on principal component analysis (PCA), for example, by POD, so that only a few eigenvalues and vectors remain as input to ICA (Hyvärinen 1999; Bizon et al. 2013b).

### 3.2 Synthetic example

To obtain a first understanding what applying ICA to flow fields does, a synthetic example was exercised, as summarized in Fig. 4a. Three sources are randomly mixed to create 500 mixture samples. Each mixture contains all three sources, but to a different extent. POD creates as many modes as mixture samples are given to the algorithm. Here, only the first three are shown. The first POD mode  $\varphi_1$  represents the mean value of all mixture samples. The remaining

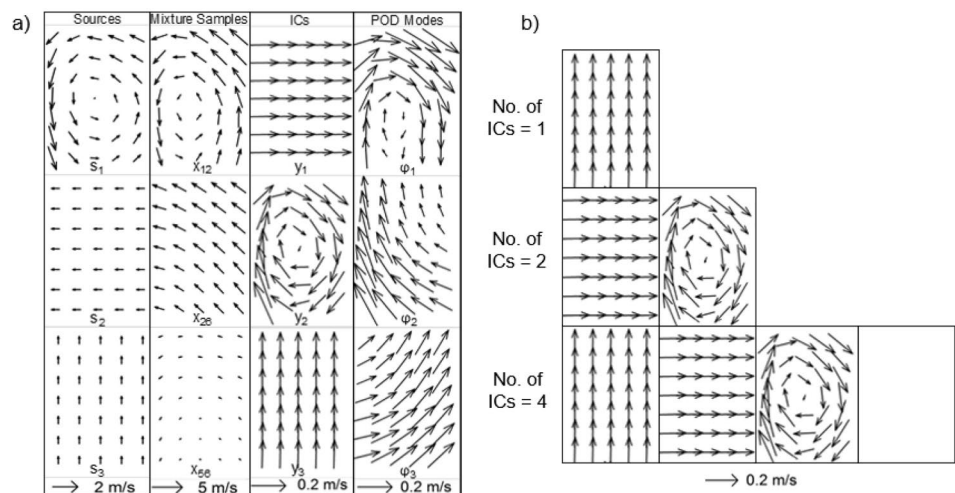
POD modes are sorted by energy in descending order. None of them resembles any original source. In comparison with that ICA only provides as many ICs as demanded from the algorithm in the first place. In this case, we know that exactly three sources exist, and the results show very similar flow fields compared to the original sources. Only the polarity and the absolute value are different, as expected.

More realistically, the number of relevant ICs is not known a priori. A parameter study of the number of ICs is shown in Fig. 4b. We see that if the number of ICs is smaller than the number of sources, ICA only finds one or two out of three sources, but not a mixture of the three. Repeating the calculation may yield one or two different ICs, but each is always a scaled version of one of the original sources. When the FastICA (Hyvärinen n.d.) algorithm searches for more than three ICs (here, 4), it simply returns only three matrices as a result.

### 3.3 Determination of the number of ICs

The synthetic example presented in Fig. 4 showed that ICA works in principle on flow fields and that the algorithm is able to extract flow fields that are very similar to the original sources. However, in real engine data two inter-connected challenges exist: first, it is not known a priori how many sources exist, and second, the data contains noise. The latter is minimized—at the cost of some temporal resolution—by calculating the moving average of five images for each single cycle before ICA is applied and by only using the lower-order POD modes. The former—not knowing the “relevant” number of ICs—is addressed here by varying the number of ICs (1, 2, 3, 5, 8, 10, and 20) and comparing the results. For completeness, Table 2 shows the portion of the total energy within the POD modes used for the ICA with various numbers of ICs.

**Fig. 4** (a) ICA and POD applied to synthetic flow fields with the number of ICs set to the known number of sources (three), (b) the ICs found by FastICA with different numbers of requested ICs



**Table 2** Cumulative energy contained in POD modes for different numbers of ICs

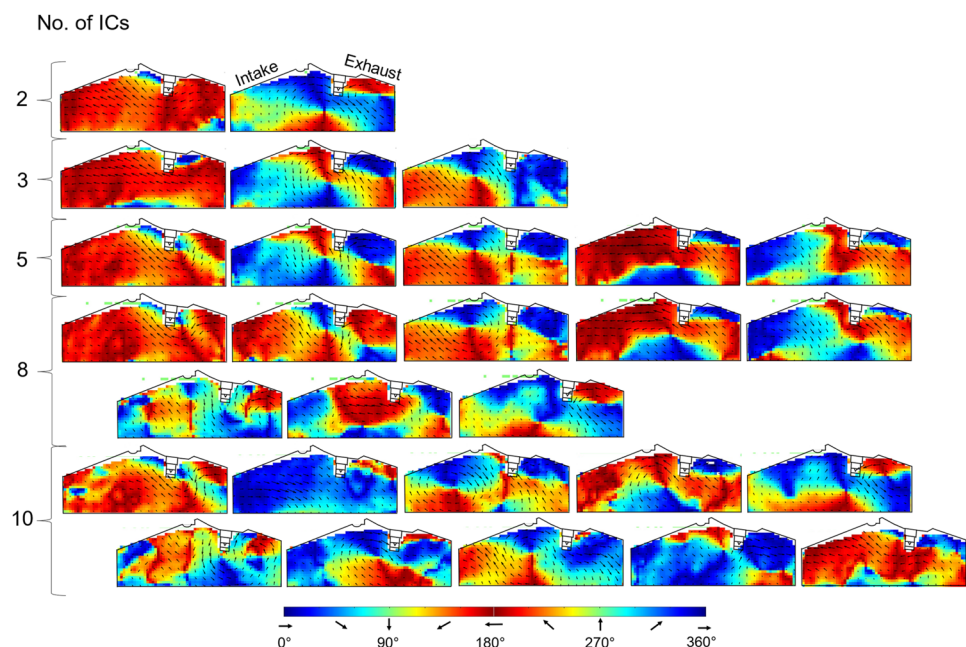
Number of ICs	% of energy in POD modes
2	78.8
3	85.1
5	91.6
8	95.3
10	96.3
20	98.7

Figure 5 shows some of the results of this parameter study. Because it is not possible to compare the absolute values of the ICs, the background color represents not the magnitude but the orientation of each vector. This simplifies an initial visual inspection, which was used to manually sort the ICs for the purposes of this figure. For example, the first column of the first row (2 ICs requested) shows a flow directed to the intake side. The three ICs in the second row are arranged such that the first one shows a similar flow pattern, etc. With 10 calculated ICs the pattern is similar, but some smaller-scale structures appear. In this column, the flow in each field is predominantly directed toward the intake side. The ICs in the second column feature a large central vortex. As previously noted, the sign (polarity) of the ICs is arbitrary. Reviewing the second column provides a clear illustration of this. While in the first row (2 ICs) the vortex rotates clockwise, in the second row (3 ICs) it is counter-clockwise.

If 8 or 10 (or 20, not shown) ICs are calculated, some of the ICs show relatively small flow patterns. One reason

might be turbulence. The ICA algorithm first filters the data set through a POD, and if the number of ICs is set to, e.g., 10, in a next step only the first 10 POD modes are used to approximate the original sources. The more POD modes are used for the ICA, the more turbulence is contained by these higher-order POD modes. Furthermore, the repeatability of the ICA becomes less robust. Calculating 8, 10, or 20 ICs several times results in more variations between runs than returned for a lower requested number of ICs. By calculating five or less ICs, larger flow patterns are visible. Some of the ICs are visually similar to ICs found with eight requested ICs, which we take to be an indication of convergence. Requesting three ICs yields close to a subset of the five ICs below, and two yields a subset of three. Since we are mostly interested in large-scale patterns and but want as much information as possible, we set the number of ICs to 5, noting that this choice remains somewhat arbitrary.

Applying ICA at a specific crank angle for the decreased data set of five POD modes from all 213 recorded cycles results in 70 sets (from the 70 crank angles) of five ICs each. As FastICA requires a scalar input, the two-dimensional velocity field was rearranged. Each  $u$ - and  $v$ -component of the velocity field with its  $49 \times 64$  subregions was transformed into a  $1 \times 3136$  vector. These two vectors were then combined into a single  $1 \times 6272$  vector which was used for ICA. Reversing the re-arrangement, the output ICs were then separated to obtain the two velocity components, resulting in two matrixes with the original dimensions of  $49 \times 64$  subregions, i.e., two-component vector fields.

**Fig. 5** Parameter study with different numbers of requested ICs at  $-45^\circ\text{CA}$ 

### 3.4 Post-processing

This snapshot ICA finds a set (here, 5) of independent components that are “typical” for the ensemble of flow fields at a given crank angle. Since the ICs are neither connected in time, nor sorted like, for example, modes of a POD, further post-processing is indispensable. In this study, two different methods are applied to the resulting data set: a persistence analysis and an analysis of correlation with combustion. The flowchart in Fig. 6 illustrates the post-processing. The basic idea behind sorting the ICs is to presume temporal coherence or persistence. That is, for a flow feature occurring during the compression stroke to be relevant for the later combustion event it must persist for an extended period of time.

To identify whether an IC persists throughout the compression stroke, the similarity of pairs of flow fields across time is quantified by the relevance index  $RI$ . Liu and Haworth (2011) presented this metric:

$$RI = \frac{\mathbf{x}_A \cdot \mathbf{x}_B}{|\mathbf{x}_A| |\mathbf{x}_B|}, \tag{2}$$

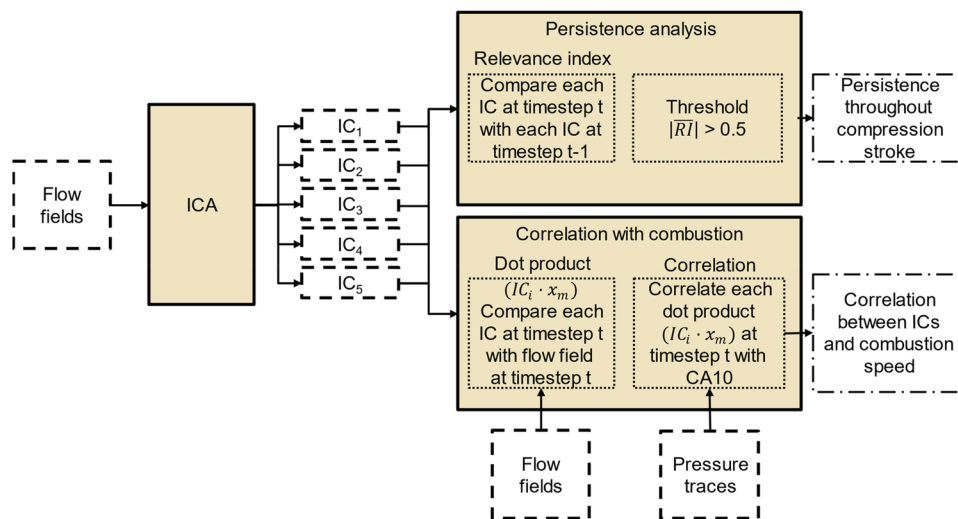
where at each point in the flow field  $(\mathbf{x}_A, \mathbf{x}_B)$  represents the dot product of the two velocity vectors  $\mathbf{x}_A$  and  $\mathbf{x}_B$ , and  $|\mathbf{x}|$  is the magnitude of  $\mathbf{x}$ . The resulting relevance index varies between  $-1$  and  $1$ , with  $1$  corresponding to perfect alignment and  $-1$  to anti-polar alignment. Starting at ignition timing, we compare each  $IC_{i,t}$  at timestep  $t$  with each  $IC_{i,t-1}$  at timestep  $t-1$ , the metric for similarity being the spatial average of  $\overline{RI}$ . However, since the polarity of each ICs is random,  $\overline{RI} = -1$  indicates similarity as much  $\overline{RI} = 1$  does. Hence, the absolute value of  $|\overline{RI}|$  is used to identify persistence. An example of such a  $|\overline{RI}|$  combination matrix is given in Fig. 7. As shown in the top-left table in that figure, the

two ICs with the highest  $|\overline{RI}|$  are first linked. Here, these are  $IC_3$  at  $-20^\circ CA$  ( $IC_{3,-20^\circ CA}$ ) and  $IC_1$  at  $-21^\circ CA$  ( $IC_{1,-21^\circ CA}$ ). These ICs are now marked as already paired when looking for the best correlation among the remaining pairs of ICs. The algorithm continues with descending  $|\overline{RI}|$  (gray, large arrows between tables in Fig. 7) until every  $IC_i$  has either a predecessor or there is no remaining  $|\overline{RI}| > 0.5$ . Below this threshold, the flow fields are considered not similar.

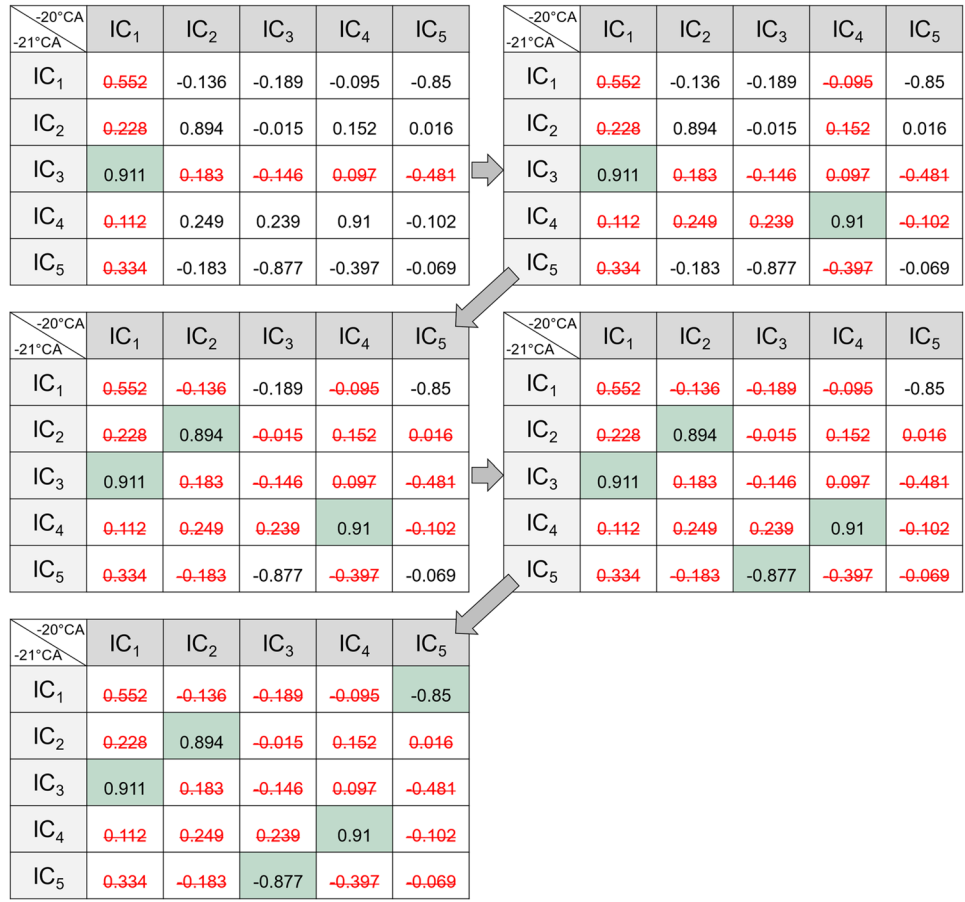
The pairing procedure is subsequently executed for the preceding time step (comparing  $-21^\circ CA$  with  $-22^\circ CA$ , then  $-22^\circ CA$  with  $-23^\circ CA$ , and so on) until no further pairs are detected. This results in five chains of ICs starting at ignition timing back to the crank angle at which no combination of ICs can be found anymore. The table in Fig. 8 shows the chain starting back from  $IC_2$  for the first few crank angles. An example is shown at the bottom of the figure (here, later crank angles are used for a wider FOV).

Now that we have identified persistent, typical “components” of the flow occurring in the compression stroke, how can we find out if any of them are “responsible” for a certain outcome of the (cyclically variable) combustion? For that purpose, the link between combustion and flow is analyzed by a correlation between CA10 and the degree to which each flow field contains a specific IC. In the first step, each IC is compared with each flow field for each timestep—by calculating a pair’s dot product ( $IC_i \cdot x_m$ ). Since only one IC is being compared here with a set of flow fields, the magnitude of the respective IC can also be considered. Thereby, regions in the IC with strong flow are given more importance than those with weaker flow. The resulting scalar quantity for each IC and each timestep is then checked for its correlation with CA10. The correlation coefficient  $R_{a,b}$  is a measure to what extent two parameters are linearly correlated (Holický 2013):

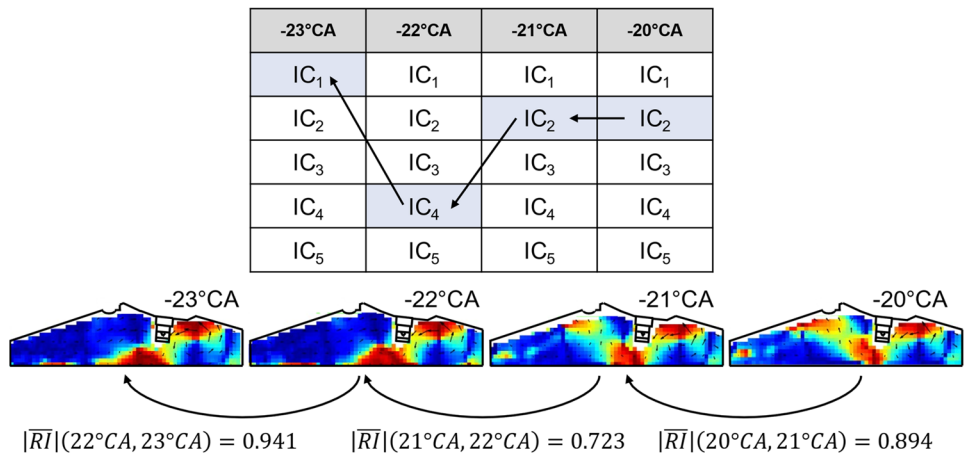
Fig. 6 Flowchart of the post-processing of the flow fields



**Fig. 7** Relevance index for all combinations of IC<sub>i</sub>s at -20°C CA and -21°C CA



**Fig. 8** Example of the resulting persistence chain



$$R_{a,b} = \frac{\sum_{i=1}^n (a_i - \bar{a})(b_i - \bar{b})}{\sqrt{\sum_{i=1}^n (a_i - \bar{a})^2 \sum_{i=1}^n (b_i - \bar{b})^2}}, \quad (3)$$

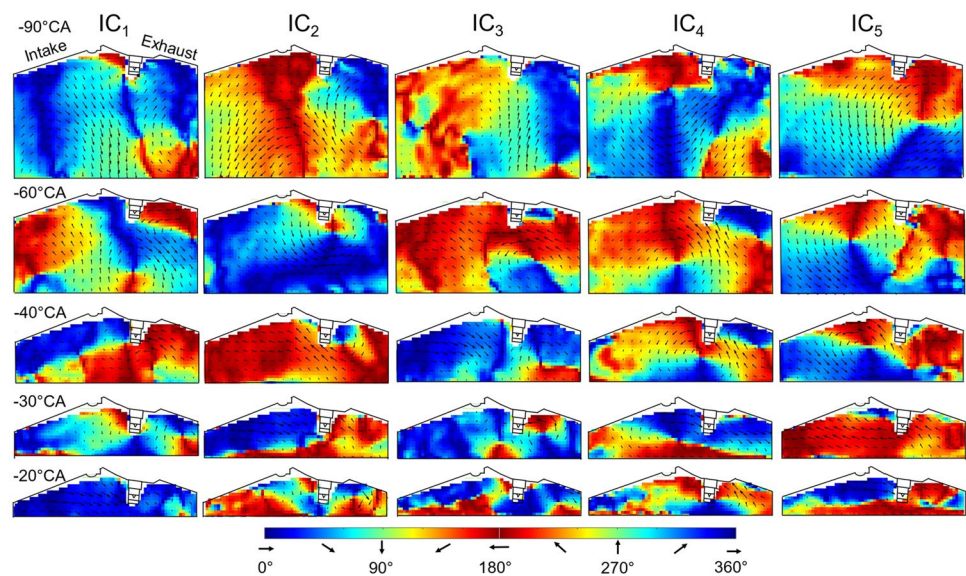
A potential physical correlation between the dot product (i.e., flow) and CA10 (i.e., combustion) may not necessarily be linear, but we will see that the scatter in the data does not warrant looking for a higher-order correlation.

### 4 Results

Figure 9 shows the five ICs at selected crank angles. Since ICs are inherently unsorted, there is not necessarily (in fact, more likely not) a connection in time between the ICs in a given column. At -90°C CA, IC<sub>5</sub> shows the tumble vortex with its center on the exhaust side. IC<sub>4</sub> shows a flow



**Fig. 9** ICs at selected crank angles (not sorted)



field with two vortices, and  $IC_1$  to  $IC_3$  show a flow pattern without dominant vortex. Here, the flow is mostly directed to the piston or the pent roof. At  $-60^\circ CA$ , the field of view covers the whole combustion chamber, and in each IC, a single vortex is found, but located at different positions. At  $-40^\circ CA$ , the vortices in each IC remain, but they are decreasing in size due to the compression. Furthermore, the vortex centers are located at different positions. Closer to ignition, at  $-30^\circ CA$ , only three ICs contain a vortex.  $IC_2$  contains two flow regions at the spark plug with flow directed against each other, and  $IC_5$  shows a larger-scale flow pattern. At ignition timing at  $-20^\circ CA$ ,  $IC_2$  to  $IC_4$  have relatively small flow structures and a simple pattern is not visible. However,  $IC_1$  shows a strong horizontal flow, and  $IC_5$  includes a vortex next to the spark plug.

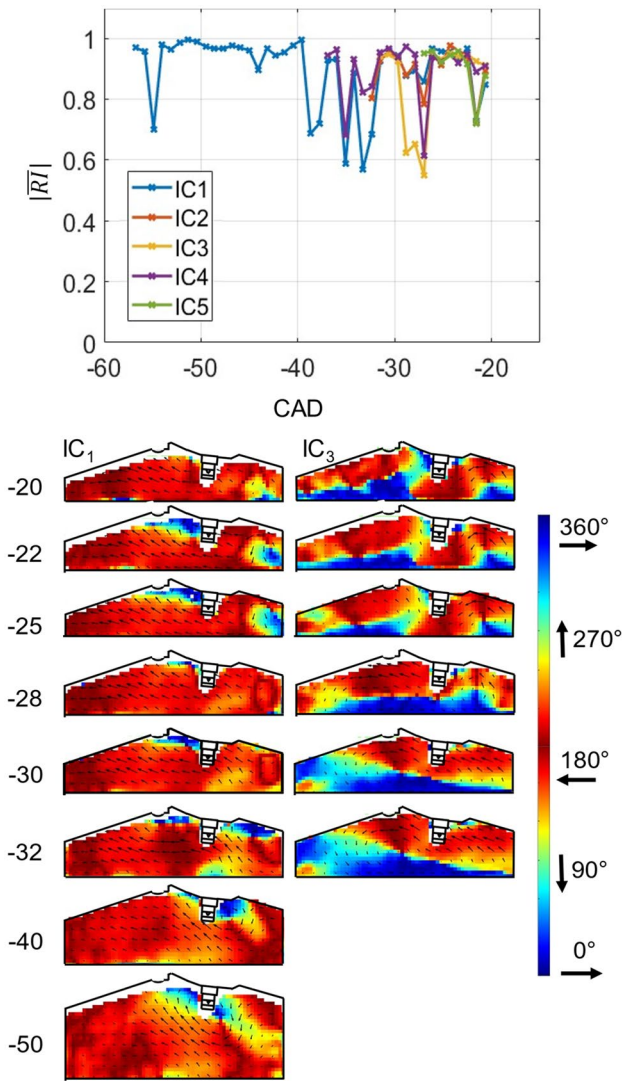
To connect the ICs with each other and find flow patterns that persist during the compression stroke, similar flow fields are linked based on their relevance index as described in Sect. 3.4. Figure 10 shows the outcome of this persistence analysis, as well as two example ICs. Note that the global polarity of the ICs does not carry any meaning; therefore, each temporal sequence is fixed to one polarity. The most persistent IC ( $IC_1$ ) has high relevance indices not only shortly after ignition and but also between  $-40^\circ CA$  and  $-55^\circ CA$ . However, between  $-32^\circ CA$  and  $-40^\circ CA$  this IC shows significant changes in its  $|\overline{RI}|$ . The other ICs only persist from ignition to  $-27^\circ CA$  and  $-37^\circ CA$ . Even though each IC has a different persistence in time, all of the connected ICs show similar drops in their  $|\overline{RI}|$ , e.g., at  $-27^\circ CA$  and  $-36^\circ CA$ . One reason for these drops may occur due to the formation of new flow patterns. The algorithm compares all ICs from two consecutive time steps and then arranges them based on their highest RI values. If two ICs become more similar over time, for example, due to the movement

of vortex centers toward the center, and simultaneously new flow patterns occur, only one IC can be associated with the old pattern. The RI will then decrease.

The corresponding flow pattern of  $IC_1$  and  $IC_3$  is given on the bottom in Fig. 10. The vortex near the spark plug in  $IC_1$  at earlier crank angles gradually disappears and leaves a nearly unidirectional flow from right to left at ignition. The other example IC in Fig. 10,  $IC_3$ , is less persistent. At  $-32^\circ CA$ , it shows a vortex with its center on the intake side next to the spark plug. At  $-28^\circ CA$ , a second smaller vortex appears on the exhaust side, and toward TDC, this one becomes more prominent.

Next, to gain information about the relation between the speed of combustion and ICs, each IC is compared to each (moving-average time-filtered) snapshot at the corresponding crank angle. The resulting dot product indicates how much of an IC is contained within a snapshot. This scalar quantity is then correlated with the pressure-based combustion quantity  $CA_{10}$ .

Figure 11a shows examples of the procedure. Since the magnitude is being taken into account by the dot product, color coding and vector length represent that magnitude. The flow field in cycle 7 has a vortex center on the exhaust side close to the pent roof, while the vortex in  $IC_1$  is more centered in the horizontal direction. The dot product of  $IC_1$  and  $x_7$  is only 0.0037, i.e., these two flow fields have very little in common. In cycle 80, the vortex center is more centered horizontally, which better matches the IC with a dot product of  $-0.1533$ . However, the faster flow in the IC below the spark plug does not fit with the low velocities in the same region in cycle 80. Among the three example cycles here, the best fit between IC and snapshot has cycle 179. The vortex center is vertically more centered than it is in cycle 80 and that leads to higher velocities below the spark plug. The



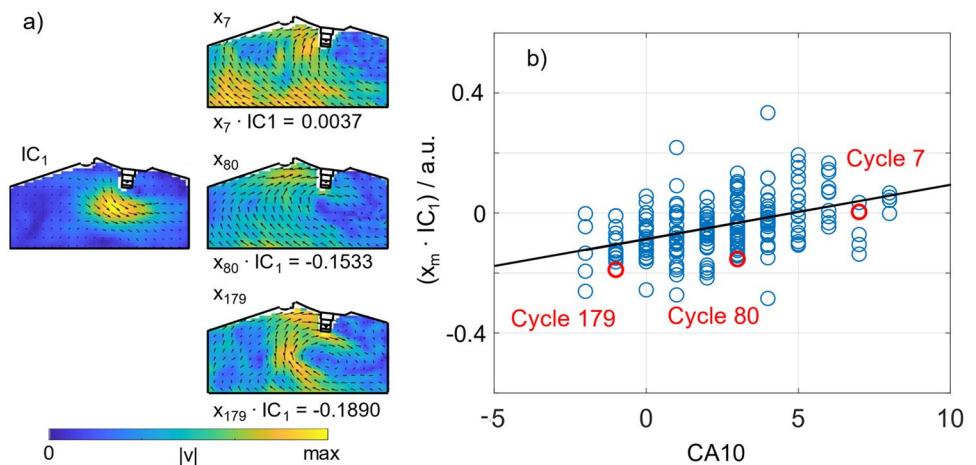
**Fig. 10** (top) Persistence analysis with absolute value of the average  $R|$  vs. CAD, (bottom)  $IC_1$  and  $IC_3$ . The global polarity of the ICs is matched to the initial IC for a better comparison

corresponding dot product is  $-0.189$ . Figure 11b shows the correlation of combustion speed and dot product as a scatter plot for the whole data set at  $-55^\circ\text{CA}$ . In this case, indeed, if much of  $IC_1$  is contained in a snapshot, combustion tends to be faster (earlier CA10) and vice versa. Here, the overall correlation coefficient is 0.438.

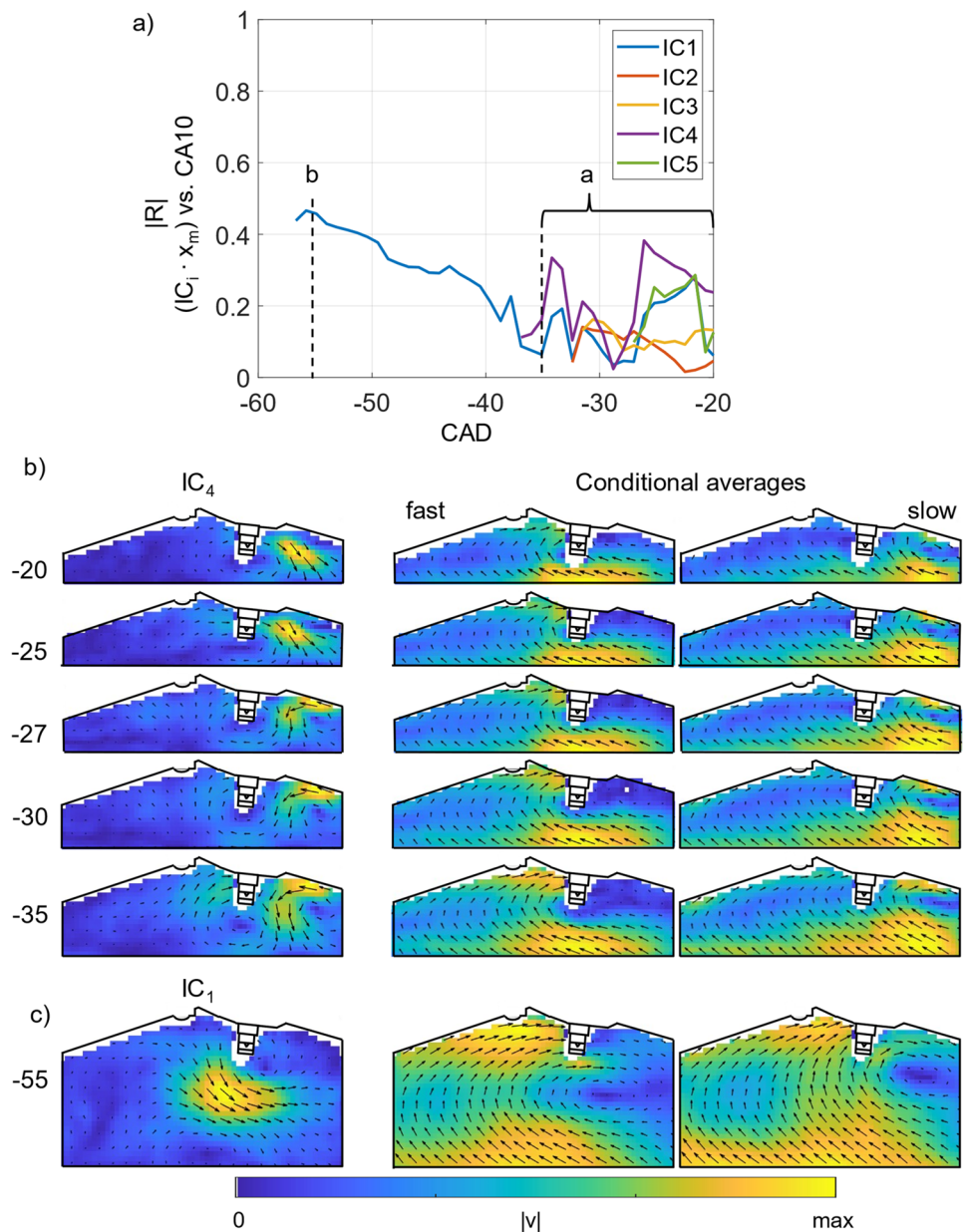
Figure 12a now combines the persistence analysis from Fig. 10 and the correlation analysis from Fig. 11. Here, the absolute value of the correlation coefficient  $|R|$ , representing the degree of correlation between the dot product ( $IC_i \cdot x_m$ ) and CA10, is plotted vs. CAD. The length of each trace indicates the persistence of an IC, while  $R$  indicates how relevant the IC is for the speed of the subsequent combustion. Tracing the correlation from its individual starting point to ignition, the most persistent IC ( $IC_1$ ) has a relatively high  $R$  value of 0.48 at  $-55^\circ\text{CA}$ , slowly decreasing to 0.05 at  $-30^\circ\text{CA}$ . At ignition ( $-20^\circ\text{CA}$ ), this IC has the second lowest correlation coefficient. The correlation coefficients of  $IC_2$  and  $IC_3$  are more constant but also low with  $\sim 0.05$  and  $\sim 0.1$ , respectively.  $IC_5$  starts at  $-28^\circ\text{CA}$  with an  $R$  value of 0.1, increasing to about 0.25 between  $-25^\circ\text{CA}$  and  $-22^\circ\text{CA}$ , then decreasing toward ignition to 0.1.  $IC_4$  has the highest  $R$  value and by this metric seems would be an important flow pattern for the combustion speed. This IC is shown in the first column of Fig. 12b for different CADs. At  $-35^\circ\text{CA}$ , two vortices dominate and there is just a relatively weak flow on the intake side of the combustion chamber. This flow pattern remains until  $-27^\circ\text{CA}$ . The vortex center on the exhaust side disappears at  $-25^\circ\text{CA}$ , such that at ignition, there is a single vortex around the spark plug with a strong flow toward the piston on the exhaust side.

Our prior analysis of the same data set, which relied on conditional averaging as outlined in reference (Bode et al. 2019), showed that in the 35 fastest-burning cycles, the pre-combustion flow exhibited a vortex encircling the spark plug, whereas on average, a vortex on the exhaust side was linked to slower combustion. The corresponding flow fields

**Fig. 11** (a) Example of an IC and selected snapshots at  $-55^\circ\text{CA}$ . The resulting dot product is given below each snapshot. (b) Corresponding correlation with CA10 over all 213 cycles at  $-55^\circ\text{CA}$ . The results from the three example snapshots are marked in red



**Fig. 12** (a) Correlation coefficient of  $(IC_i \cdot x_m)$  and CA10 vs. CAD, (b)  $IC_4$ —the IC with the highest correlation coefficient near ignition—and conditionally averaged flow fields at the same crank angles, (c)  $IC_1$ —the most persistent IC—and conditionally averaged flow fields at  $-55^\circ CA$



are shown in the second (fast) and third (slow) column of Fig. 12b. In the present study,  $IC_4$  has the strongest correlation with combustion speed. This IC contains the two vortices, one centered around the spark plug and another on the exhaust side between  $-35^\circ CA$  and  $-27^\circ CA$ . Toward TDC, the IC shows only one vortex centered at the spark plug, which aligns only with the conditional average of the 35 fastest cycles. This suggests that the more important flow pattern influencing combustion involves the vortex center around the spark plug and a strong flow on the exhaust side.

In the first column of Fig. 12c, the flow pattern of  $IC_1$  at  $-55^\circ CA$ , which also correlates well with the combustion speed, is shown. The main feature of  $IC_1$  is a region of strongest lateral flow below the spark plug. In

the corresponding conditional averages, the differences between fast and slow cycles are less obvious than in the later stages of flow development. For the slow cycles, the vortex center is on the exhaust side, while for the fast cycles, the vortex center is more horizontally centered. Visually,  $IC_1$  has very little in common with conditional averages or the difference between them. This and the then following steady decline of this ICs correlation with combustion may indicate that  $IC_1$  contains a flow component that is important early on but is then gradually replaced by more than one other component. Some of this energy redistribution may also occur outside of the PIV plane, such that it is not detected in the experimental data.

## 5 Conclusions

This work focused on identifying links between the flow and the speed of combustion in an optical internal combustion engine operated on a slightly lean homogeneous iso-octane/air mixture. In particular, the goal was to examine if ICA based on vector fields measured with high-speed PIV might provide an objective and quantitative analysis of flow patterns to find such a link. While techniques based on POD had previously been used widely in flow analysis and combustion research, so far ICA had only been applied for a few examples of scalar combustion images (Bizon et al. 2013a, 2016b). The expected key difference between POD and ICA was that ICA would produce “source components,” i.e., flow structures that could be linked to other observables, while POD would do this only under exceptional cases, if at all (Chen et al. 2013). However, the independent components (ICs) identified by ICA are not inherently sorted by, e.g., energy (as in POD), requiring further post-processing to re-establish some of the temporal coherence that the original PIV data had.

To test the basic procedure, we initially applied ICA to a synthetic example in which the number of underlying sources was known a priori. ICA was then employed to extract flow patterns from high-speed PIV data gathered from 213 engine cycles. In such empirical data, the number of sources is not known. A parameter study suggested that for the current data set five ICs are sufficient. A persistence analysis found that one independent component could be traced from  $-60^\circ\text{CA}$  to ignition at  $-20^\circ\text{CA}$ , while the other four ICs persisted at the most from  $-38^\circ\text{CA}$  to ignition. As a metric of the speed of combustion progress, the pressure-based cumulative heat release, in particular CA10, was utilized. By quantifying the correlation between the similarity of independent components and flow fields and CA10, we investigated the relationship between independent components and the speed of combustion. This examination showed that the combustion relevance of the most persistent IC mostly declines throughout the upper compression stroke, while the IC that correlates best with combustion speed is one of the less persistent ones. The latter also visually more resembles some of the flow features found in conditional averaging of fast-burning vs. slow-burning cycles (Laihter and Kaiser 2022). The combination of ICA and the current implementation of a persistence analysis does not give a direct indication as to how any presumed underlying re-organization of the flow occurs.

This work presented an example of how ICA of velocity fields could be used to identify differences in the in-cylinder flow that impact cycle-to-cycle combustion performance in an internal combustion engine. A broader

range of conditions and engine configurations needs to be studied to comprehensively characterize how ICA may be used for this purpose and to develop guidelines for best practices, similar to what has been done for POD (Chen et al. 2012, 2013). Furthermore, employing ICA on a three-dimensional numerical simulation could expand the restricted field of view offered by experimental data and provide deeper insights into the underlying flow patterns in an engine associated with CCV.

**Author contributions** JL and SAK contributed to study conception. JL set up the experiment, collected, analyzed the data, and wrote the manuscript draft. PK created the synthetic example. All authors jointly discussed and revised the manuscript, and they read and approved the final submitted version.

**Funding** Open Access funding enabled and organized by Projekt DEAL. Financial support of this work by the Deutsche Forschungsgemeinschaft (DFG) within the framework of the research unit FOR 2687, project number 423271400 and 349537577, is gratefully acknowledged.

**Availability of data and materials** The data supporting the findings reported in this paper are camera images, both snapshots and image sequences. They are available from the authors upon request.

## Declarations

**Conflict of interest** The authors declare that they have no known conflict of interest or personal relationships that could have appeared to influence the work reported in this paper.

**Consent for publication** The authors confirm that this work is original and has not been published elsewhere, nor is it under consideration for publication elsewhere, and consent for the publication in experiments in fluids.

**Ethical Approval** Not applicable.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Abraham PS, Yang X, Gupta S, Kuo T-W, Reuss DL, Sick V (2015) Flow-pattern switching in a motored spark ignition engine. *Int J Engine Res* 16(3):323–339
- Bizon K, Continillo G, Lombardi S, Mancaruso E, Sementa P, Vaglieco BM (2013a) Independent component analysis of combustion

- images in optically accessible gasoline and diesel engines, SAE Technical Paper 2013-09-08
- Bizon K, Lombardi S, Continillo G, Mancaruso E, Vaglieco BM (2013b) Analysis of Diesel engine combustion using imaging and independent component analysis. *Proc Combust Inst* 34(2):2921–2931
- Bizon K, Continillo G, Lombardi S, Sementa P, Vaglieco BM (2016) Independent component analysis of cycle resolved combustion images from a spark ignition optical engine. *Combust Flame* 163:258–269
- Bode J, Schorr J, Krüger C, Dreizler A, Böhm B (2019) Influence of the in-cylinder flow on cycle-to-cycle variations in lean combustion DISI engines measured by high-speed scanning-PIV. *Proc Combust Inst* 37(4):4929–4936
- Borée J (2003) Extended proper orthogonal decomposition: a tool to analyse correlated events in turbulent flows. *Exp Fluids* 35(2):188–192
- Chen H, Reuss DL, Sick V (2011) Analysis of misfires in a direct injection engine using proper orthogonal decomposition. *Exp Fluids* 51(4):1139–1151
- Chen H, Reuss DL, Sick V (2012) On the use and interpretation of proper orthogonal decomposition of in-cylinder engine flows. *Meas Sci Technol* 23(8):85302
- Chen H, Reuss DL, Hung DLS, Sick V (2013) A practical guide for using proper orthogonal decomposition in engine research. *Int J Engine Res* 14(4):307–319
- Chen H, Xu M, Hung DLS, Zhuang H (2014) Cycle-to-cycle variation analysis of early flame propagation in engine cylinder using proper orthogonal decomposition. *Exp Therm Fluid Sci* 58:48–55
- Druault P, Guibert P, Alizon F (2005) Use of proper orthogonal decomposition for time interpolation from PIV data. *Exp Fluids* 39(6):1009–1023
- Fajardo CM, Smith JD, Sick V (2006) PIV, high-speed PLIF and chemiluminescence imaging for near-spark-plug investigations in IC engines. *J Phys: Conf Ser* 45:19–26
- Fontanesi S, d'Adamo A, Rutland CJ (2015) Large-Eddy simulation analysis of spark configuration effect on cycle-to-cycle variability of combustion and knock. *Int J Engine Res* 16(3):403–418
- Graftieaux L, Michard M, Grosjean N (2001) Combining PIV, POD and vortex identification algorithms for the study of unsteady turbulent swirling flows. *Meas Sci Technol* 12(9):1422–1429
- Granet V, Vermorel O, Lacour C, Enaux B, Dugué V, Poinot T (2012) Large-Eddy Simulation and experimental study of cycle-to-cycle variations of stable and unstable operating points in a spark ignition engine. *Combust Flame* 159(4):1562–1575
- Hanuschkin A, Zündorf S, Schmidt M, Welch C, Schorr J, Peters S, Dreizler A, Böhm B (2021) Investigation of cycle-to-cycle variations in a spark-ignition engine based on a machine learning analysis of the early flame kernel. *Proc Combust Inst* 38(4):5751–5759
- Holický M (2013) Introduction to probability and statistics for engineers. Springer, Berlin Heidelberg, Berlin, Heidelberg
- Hyvärinen A. Available at <<https://research.ics.aalto.fi/ica/fastica/>>
- Hyvärinen A (1999) Fast and robust fixed-point algorithms for independent component analysis. *IEEE Trans Neural Netw* 10(3):626–634
- Hyvärinen A, Oja E (2000) Independent component analysis: algorithms and applications. *Neural Netw off J Int Neural Netw Soc* 13(4–5):411–430
- Hyvärinen A, Karhunen J, Oja E, Haykin S (2001) Independent component analysis. John Wiley & Sons Inc, New York, USA
- Laichter J, Kaiser SA (2022) Optical investigation of the influence of in-cylinder flow and mixture inhomogeneity on cyclic variability in a direct-injection spark ignition engine. *Flow Turbul Combust*
- Laichter J, Kaiser S, Rajasegar R, Srna A (2023) Impact of mixture inhomogeneity and ignition location on early flame kernel evolution in a direct-injection hydrogen-fueled heavy-duty optical engine, SAE Technical Paper 2023-32-0044
- Liu K, Haworth DC (2011) Development and assessment of POD for analysis of turbulent flow in piston engines, SAE Technical Paper 2011-01-0830
- Peterson B, Reuss DL, Sick V (2011) High-speed imaging analysis of misfires in a spray-guided direct injection engine. *Proc Combust Inst* 33(2):3089–3096
- Reuss DL (2000) Cyclic variability of large-scale turbulent structures in directed and undirected IC engine flows, SAE Technical Paper 2000-01-0246
- Stiehl R, Bode J, Schorr J, Krüger C, Dreizler A, Böhm B (2016) Influence of intake geometry variations on in-cylinder flow and flow-spray interactions in a stratified direct-injection spark-ignition engine captured by time-resolved particle image velocimetry. *Int J Engine Res* 17(9):983–997
- Wu T, He G (2022) Independent component analysis of streamwise velocity fluctuations in turbulent channel flows. *Theor App Mech Lett* 12(4):100349
- Zeng W, Sjöberg M, Reuss DL, Hu Z (2017) High-speed PIV, spray, combustion luminosity, and infrared fuel-vapor imaging for probing tumble-flow-induced asymmetry of gasoline distribution in a spray-guided stratified-charge DISI engine. *Proc Combust Inst* 36(3):3459–3466

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.