Master's Thesis

Computer-based Classification of Running Characteristics of an Internal Combustion Engine with Airborne Sound Measurements in End-Of-Line Quality Control

Andreas Fuchs
9773819

Graz, June 10, 2016

Master's programme Electrical Engineering and Audio Engineering
066 413

University of Music and Performing Arts Graz
Graz University of Technology

Advisor:
Dipl.-Ing. Dr.techn. Alois Sontacchi

Assessor:
O.Univ.Prof. Mag.art. DI Dr.techn. Robert Höldrich

Institute of Electronic Music and Acoustics

Abstract

The engine isn't running smoothly, it's knocking! – A motorcar mechanic is intuitively capable of assessing the running characteristics of an internal combustion engine by listening to its sounds. In the course of evolution the human sense of hearing has evolved to a measuring instrument with exceptional adaptability, fault tolerance and hierarchical structuring ability. After training, the ear can thus be used for quality assessment.

However, for quality assurance at the end of the production line, automation and integrity are of central importance, which is why the condition on that specific day and suchlike of a specially trained individual must not interfere with the assessment. Thus, it is tempting to use the human hearing expertise as a template for automatic classification based on acoustic measurements of the running engine in order to test and attest the mechanical quality. By using modern hardware and efficient algorithms, an easily used system is developed, which is - in a best-case scenario - sufficiently robust to deliver reliable results in the production process.

From existing measurement data from an internal combustion engine, characteristic sound emissions in the duty cycle of the engine are specified in different operation conditions. With the appropriate algorithms, the alignment on the duty cycle can be performed solely with the microphone signal itself. It is therefore possible to define automatically computable characteristics, which contain the desired information and form the input of the computer-based classification. By algorithms of supervised machine learning, the statistically optimal processing of features is guaranteed for classification. As a result, there is an automatic, cost-effective and precise discrimination between different running-characteristics of the engine.

Zusammenfassung

Der Motor läuft nicht rund, er klopft! - Ein Mechanikermeister weiß intuitiv auf sein geschultes Gehör zu vertrauen, um die Laufeigenschaften eines Verbrennungsmotors zu beurteilen. Der menschliche Gehörsinn hat sich im Laufe der Evolution zu einem Messinstrument mit außergewöhnlicher Adaptionsfähigkeit, Fehlertoleranz und hierarchischer Strukturierungsfähigkeit entwickelt. Mithilfe von Training kann somit das Ohr als verlässlicher Qualitätswächter fungieren.

Für die Qualitätssicherung am Ende der Produktionslinie sind allerdings Automatisierbarkeit und Integrität von zentraler Bedeutung, weshalb die Beurteilung unabhängig zur Tagesverfassung und dergleichen einer speziell geschulten Einzelperson sein muss. Somit liegt es nahe, die menschliche Hörexpertise als Vorlage für die maschinelle Klassifikation zu benützen, um anhand akustischer Messungen des laufenden Motors die mechanische Qualität zu prüfen und zu attestieren. Durch den Einsatz moderner Hardware und effizienter Algorithmen wird ein möglichst einfach zu bedienendes System entwickelt, das im Idealfall ausreichend robust ist, um im Produktionsalltag zuverlässige Ergebnisse zu liefern.

Aus vorhandenen Messdaten eines Verbrennungsmotors werden die in unterschiedlichen Betriebszuständen charakteristischen Schallemissionen innerhalb eines Arbeitsspiel des Motors spezifiziert. Mit den entsprechenden Algorithmen kann die Einteilung in Arbeitsspiele ausschließlich anhand des Mikrofonsignals erfolgen. Damit ist es möglich, maschinell berechenbare Merkmale zu definieren, die die gewünschte Information enthalten und den Eingang der computergestützten Klassifizierung bilden. Durch Algorithmen des maschinellen, überwachten Lernens wird die aus statistischer Sicht optimale Verarbeitung der Merkmale zur Klassifizierung garantiert. Daraus ergibt sich eine automatisierte, kostensparende und präzise Unterscheidung zwischen unterschiedlichen Laufeigenschaften des Motors.

Acknowledgements

I would like to express my gratitude to Dipl.-Ing. Dr.techn. Alois Sontacchi for all his insight and constructive ideas during our instructive discussions about this thesis. His lectures in Music Information Retrieval inspired me strongly to focus in digital signal processing and the extra-ordinary special care he took of the students in these lectures should not be unmentioned. My thanks go also to O.Univ.Prof. Mag.art. DI Dr.techn. Robert Höldrich, who always had the overall image with the finest details in his mind. DI Manuel Brandner provided me with useful insights of possible algorithms and properties of the measurement data.

I like to thank my parents, who supported me my whole life in every aspect and made my studies in Electrical Engineering and Audio Engineering in Graz possible. I would like to thank my sister Niki, whose success and versatility is always an inspiration.

Special thanks go to my partner Amina, who showed endless patience with my studies and gives me the strength to reach my goals.

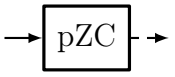
Contents

Block Diagram Legend

Signals

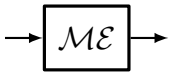
⟶	Signal or envelope	⟹	Signal in cycles or frames
- - - -►	Tick or ignition locations	= = = =►	Tick locations in cycles
·····►	Parameter or scalar	⁝⁝⁝⁝►	Scalar per cycle (features)

Commonly Used Blocks

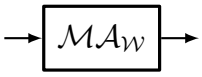


Positive zero-crossing detection

Peak-picking (location of peaks)

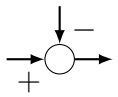


Moving median filter

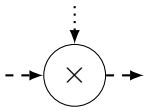


Weighted moving average filter

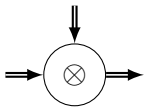
Signal Operations



Addition/subtraction



Scalar multiplication



Element-wise matrix multiplication

Scalar division

1 Introduction

In the field of Music Information Retrieval (MIR), for a computer usable data is extracted from a music signal. The most common example is the identification of the song titles and artists, which was already predicted to be possible by Downie in 2003 [8], and is now a common app for smart phones. Although this thesis does not work with music signals, the objective is similar. There is no doubt that an internal combustion engine has a characteristic rhythmic sound and, as the human ear is very capable of pattern recognition, changes of a rhythmic structure are easily detectable. If the engine does not run smoothly, the periodic structure in the sound emission is disturbed, which changes the rhythm. A faulty operation of an internal combustion engine, which is directly associated with sound emission, is knocking, which is caused by an uncontrolled combustion. From this it follows that for different running characteristics of an engine, the relevant information can be manifested in the airborne sound emission. Therefore, the main objective of this thesis is to extract this information into useful features, which values depend on the running characteristics and can be used to classify the engine. A feature reduces the sound pressure time series to a single value and is often calculated for a sliding time-window to allow for changes.

The final goal of this research is to develop a new approach to cost-efficiently detect manufacturing defects of an internal combustion engine in end-of-line quality control with a microphone signal. The method should work as simple and cheap as possible, so in the ideal case only one microphone at a fixed location is enough. The main advantages of this approach is that no physical connection to the engine must be made, neither for an electronic signal to the engine control unit nor to a vibro-acoustic sensor. The microphone can be installed at a fixed point in the measurement room and the connected computing unit can immediately return an assessment. This thesis is a preliminary study, if this goal is reachable, and should provide a useful set of tools for this task. A data set with 16 microphones at different test-runs is available, where not actual defect engines were measured, but the engine control unit of one good engine was used to create faulty conditions, like the missing combustion of one cylinder or a missing pilot injection.

Due to the practical focus and innovative field, this thesis is not structured into a theoretical and practical part, but follows the functional structure of the developed procedures. The required theory from mechanics, technical acoustics and signal processing are explained by the problem at hand.

Chapter 2 gives an overview of the recorded signals and taken measurements. A preliminary evaluation of the spatial distribution of the sound emission is presented and the common concept of crankshaft angle analysis is evaluated for different signals. In a crankshaft angle analysis a signal is sliced into the duty cycles of the engine and equi-angular resampled. Therefore, events can be connected to a certain point in the duty cycle.

Crankshaft angle analysis needs additional signals from the engine, so chapter 3 generalizes the concept to be performed solely with the microphone signal. Although, only the alignment on the whole duty cycle is possible, this is still an important pre-processing step for the feature calculation. As a by-product some of the faulty conditions of the measured engine can also be detected, as the alignment works with the ignitions of the four cylinders and is therefore sensitive to a missing cylinder ignition.

Chapter 4 defines the so-called envelope shape features on the duty cycles, which are mainly used for the classification of the different simulated running-characteristics. Special characteristics of the average shape of the envelope of the sound emission are used for the distinction. Chapter 5 finally evaluates the performance of the features with a classifier under different framework conditions. Due to the thorough manual feature definition and optimization, at most five features are necessary for a given classification task. Thus, a simple k-nearest neighbor classifier performs well and also a further feature reduction is due to the low dimensionality easily performed. Although many of these steps right up to the definition of features can also be performed by advanced machine learning algorithms, the size of the measured data set limits their practicalness.

The structure of the thesis also has another motivation considering the amount of signals and measurements, which are analyzed in the respective chapters. In chapter 2 all measured signals are introduced and the capabilities, if all signals are used for cycle alignment, is shown. Furthermore, one microphone is selected for further analysis. In chapter 3 the main focus is only on this microphone signal. Here, some of the measurements with higher revolutions speeds must be put out of the focus, as they do not hold the required characteristics of distinct transients. Chapter 4 then concentrates on the differences between the measured conditions for the engine only at two operation speeds and for one microphone for the envelope shape features. At last, in the evaluation of the classifier in chapter 5, the view is widened again to other microphones and general real-word application parameters.

2 Signal Analysis

In this chapter the data set used in this master thesis is presented and closely examined. Section 2.1 gives an overview over the measurement set-up, the recorded signals and how different running-characteristics were produced. In section 2.2 the results of an analysis by listening to the microphones signals is presented to find the most apparent differences of the running-characteristics. With this information, two measures are defined in section 2.3 to examine the spatial directivity for relevant characteristics. As one of the design goals is to use as less microphones as possible and one microphone shows a distinct good performance in the spatial distribution, the development of algorithms in chapter 3 can later focus on only one microphone.

Section 2.4 explores the possibilities of crankshaft angle analysis by using equi-angular resampling. Also, the concept of signal slicing into frames corresponding to one duty cycle is presented, which is used throughout this thesis.

2.1 Recorded Signals and Measurements

The data set to be analyzed in this thesis contains 16 distributed sound pressure measurements around an internal combustion engine for different operating conditions. Before the beginning of this thesis, the measurement was performed by members of the Institute of Electronic Music and Acoustics of the University of Music and Performing Arts Graz. Because of the preliminary nature of the project, the measurements were performed in an anechoic chamber to suppress interfering effects due to reflections or exterior noise. The used Müller-BBM PAK system for recording the signals yields a high-precision measurement at a sampling rate of $f_s = 65536\,\mathrm{Hz}$. [1]

Apart from the 16 microphone signals seven additional signals were recorded, which need physical contact to the engine or are directly delivered from the engine control unit itself:

Rail Pressure (3) The pressure-volume diagram for a 4-stroke diesel engine is shown in figure 1(a). The ignition/combustion is triggered by the injection of fuel into the cylinder by the injector and happens shortly before the cylinder pressure reaches its maximum. A four-stroke engine cycle consists of two full

1. This sampling rate is rather common in technical acoustic measurement systems, as also the Brüel&Kjaer Pulse platform uses it. The reason is that with $65536 = 2^{16}$ samples per second, a short-time Fourier transform with a suitable frame length of a power of two will result in a whole number of frames per second.

rounds with each round containing two ignitions [26]. Figure 1(b) shows the recorded rail pressure signal for one cycle for idle run, which is the pressure in the injector. Only three sensors were used, so only the rail pressure of three of the four cylinders are recorded.

Injections (1) This signals shows the time instants of the opening and closing of the injectors, whereby a negative pulse is associated with the closing. In a Common-Rail system the injection pressure and the injection timing can be adapted by the engine control unit to optimize fuel consumption and exhaust and sound emissions [25].

Cogwheel (2) The enginge control unit is connected to crankshaft position sensors in order to measure the rotational speed and position of the crankshaft. A common principle is the measurement by inductive sensors in proximity to a cogwheel, which is connected to the crankshaft [29]. Each cog on the cogwheel produces an impulse. In the measurement two signals are recorded:

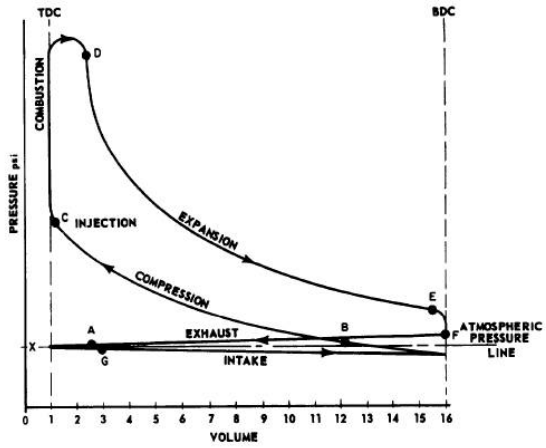
cogwheel 1 One impulse per round.

cogwheel 122 Signal with 122 impulses per round.

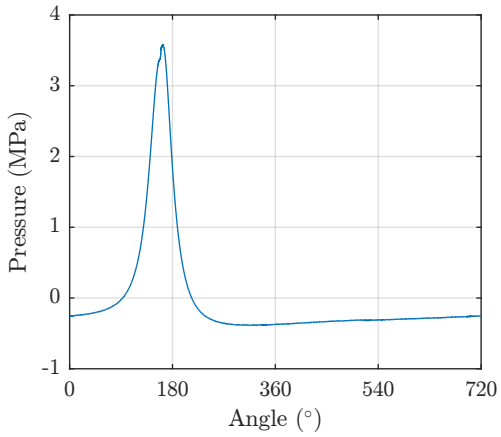
Figure 2(a) shows the two signals for one half round in an idle run.

Belt (1) The belt signal consists of one impulse per round, which is triggered by the belt, superimposed with a calculated rpm signal by the engine control unit. A lowpass filtered version or simply the mean of this signal is therefore a good approximation of the current rounds per minute (rpm), as can be seen in figure 2(b).

Due to its complexity a car engine can not be modeled by a monopole sound source and shows a distinctive directivity pattern. Therefore, acoustic cameras are often used to locally identify sound sources of the engine [23], [13]. For a precise localization the demand for a large number of microphones of a spherical microphone array [19] is contrary to the objective of a simple and cheap measurement technique. As a compromise a total number of 16 microphones were placed around the engine to capture spatial sound emission properties, as shown in figure 3. The numbering starts with microphone 1 at the top above the engine. Then the six microphones $2-7$ are placed in a horizontal plane, which is at about $2/3$ height of the whole array, and form the *upper ring*. A second horizontal plane at about $1/3$ height (*lower ring*) contains the microphones $8-13$, which are placed in the vertical centers of the upper ring, so that the triplets of neighbors form triangles of approximately equal area. The microphones $14-16$ are in a horizontal plane at the bottom of the array

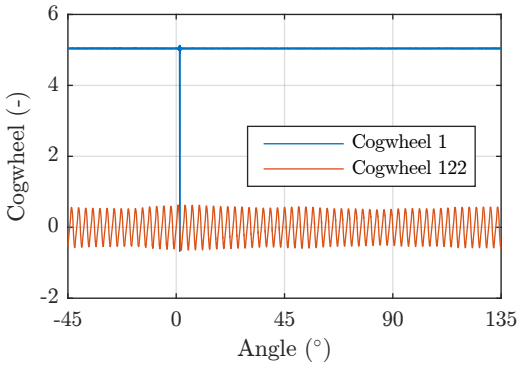


(a) Cylinder Pressure-volume diagram [24].

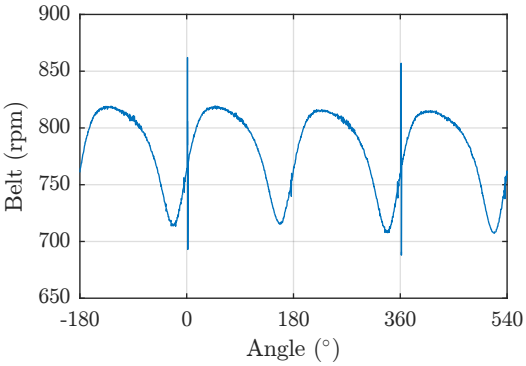


(b) Rail pressure signal 1 for one cycle (idle run). 0 ° is not top dead centre.

Figure 1 – Pressure in a diesel 4-stroke engine.



(a) Cogwheel signals.

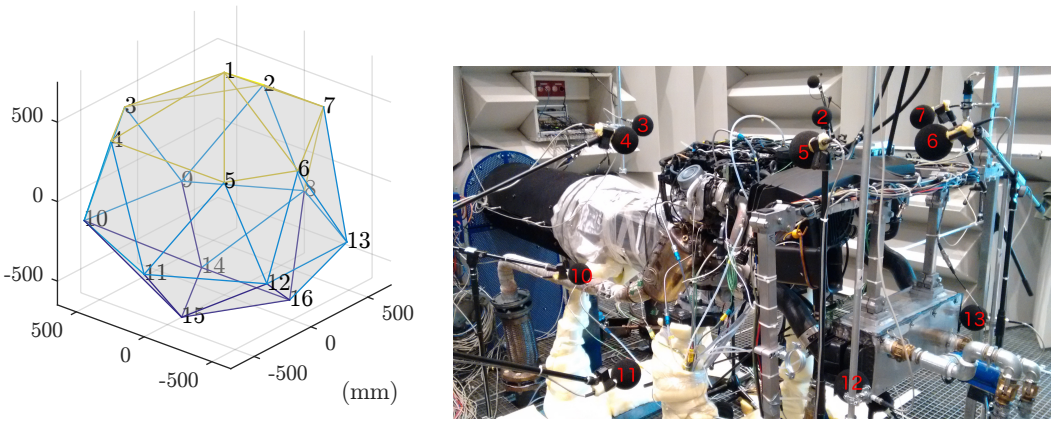


(b) Belt signal.

Figure 2 – Recorded signals for position and rotation speed analysis of the crankshaft. 0 ° is not top dead center.

and form an equilateral triangle, which is again rotated in a way that the neighbors with the lower ring form triangles of equal area. The diameter of the array is approximately 1.2 m.

Various measurement runs were performed to capture the engine sound emission in different situations. Each measurement is identified by two categories. The first category is called **condition**. As the goal is to analyze the running-characteristics of the engine, the engine control unit was programmed to produce out of tune parameters or erroneous control. Therefore, a faulty condition of the engine is simulated. A detailed description is found in table 1. The conditions in the raw data can also be identified by a numeric id or an unique string identifier, which are



(a) (b)

Figure 3 – Illustration (a) and picture (b) of the microphone positions around the engine.

also shown in the table. Pilot injections are injections with less fuel shortly before and possibly after the main injection to reduce sound emissions [25]. As only the pilot injection before the main injection could be turned off, injections after the main injection are always present.

The second category is called **operation** and describes the current operation mode of the engine, e.g. idle or run-ups. Also each different load on the engine is considered a different operation. Depending on the condition, not all operations were measured, because a high load on an engine with faulty parameters may damage the engine permanently. All operations with 100 % or 300 Nm load were not measured with a disabled cylinder (id 9-12) or misaligned injections (id 13,14). The different operations are listed in table 2, which also have numeric and string identifiers.

As the measurements were done without much intermission, the engine is considered to be warm for all measurements. Therefore, a stationary state for the stationary runs can be assumed and the results of this thesis are only applicable to a warmed-up engine.

Id	String	Description
1-5	v001_0x	**base condition 1-5** The standard configuration is measured 5 times for reproducibility.
6	v002	**rail pressure increase**
7	v003	**rail pressure decrease**
8	v004	**no pilot injection**
9	v005	**no cylinder 1** The combustion of the cylinder is disabled.
10	v006	**no cylinder 2** see above
11	v007	**no cylinder 3** see above
12	v008	**no cylinder 4** see above
13	v009	**5° later injection** The main and pilot injections are 5° after the top dead center.
14	v010	**5° earlier injection** The main and pilot injections are 5° before the top dead center.

Table 1 – Overview of the measured conditions.

Id	String	Description
1	nidle	**idle**
2	n1500_l000	**run @ 1500 rpm, 0% load**
3	n1500_l050	**run @ 1500 rpm, 50% load**
4	n1500_l100	**run @ 1500 rpm, 100% load**
5	n2000_l000	**run @ 2000 rpm, 0% load**
6	n2000_l050	**run @ 2000 rpm, 50% load**
7	n2000_l100	**run @ 2000 rpm, 100% load**
8	n4400_l000	**run @ 4400 rpm, 0% load**
9	n4400_l050	**run @ 4400 rpm, 50% load**
10	n4400_l100	**run @ 4400 rpm, 100% load**
11	nru_l100	**run up @ 100% load**
12	nru_300Nm	**run up @ 300 Nm load**
13	nru_100Nm	**run up @ 100 Nm load**

Table 2 – Overview of the measured operations relevant for this thesis.

2.2 Analysis by Ear

The first step in the analysis of the recorded audio signals is a listening comparison. Therefore, the hearable differences between the conditions were recorded in writing. This procedure was repeated for all operations, which were measured for all conditions. Many differences can only be assessed in the direct comparison between the faulty and base conditions and are difficult to hear without a reference. Because the engine is not mounted in a car body, the usual transfer paths for the engine sound are not present [31] and the emitted sound is therefore unusual with many high-frequency components. The following statements are generalizations, derived from listening to the recordings of microphone 1 above the engine:

— At slower revolutions also temporal information can be derived, at higher revolutions mainly spectral comparisons can be made and, as a result, there is less information content. In the fastest stationary runs at 4400 rpm sometimes no difference can be heard. At 780 rpm = 13 rounds/second (idle) 26 ignitions occur per second and can not be separated individually by ear, although the transient character of the ignitions is recognizable.

— More load on the engine also provides more information in separating the conditions.

— The difference between different operations for one condition is much higher than the difference between conditions for one operation. Hence, the revolution speed and the load have larger impact on the sound than the faulty parameters. Therefore, run-ups are also hard to classify by ear, because the sound changes permanently and small, varying characteristics between to run-ups can not be remembered.

— The strongest effect between the conditions for slow revolutions lies between

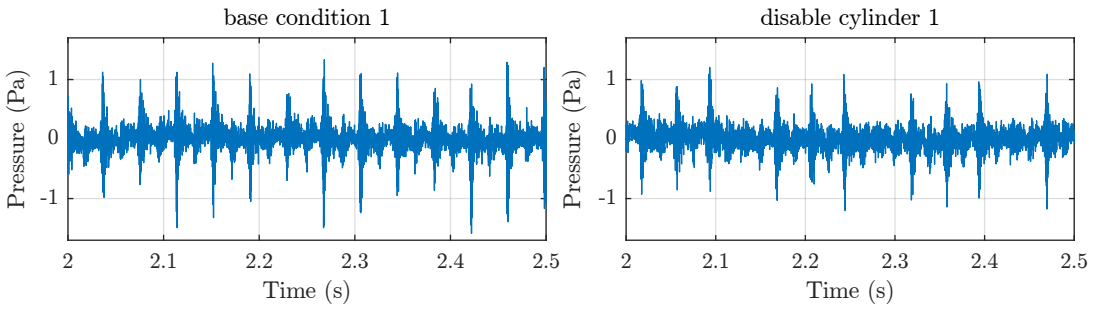


Figure 4 – Comparison of the sound pressure for base condition 1 and the disabling of cylinder 1 in idle run.

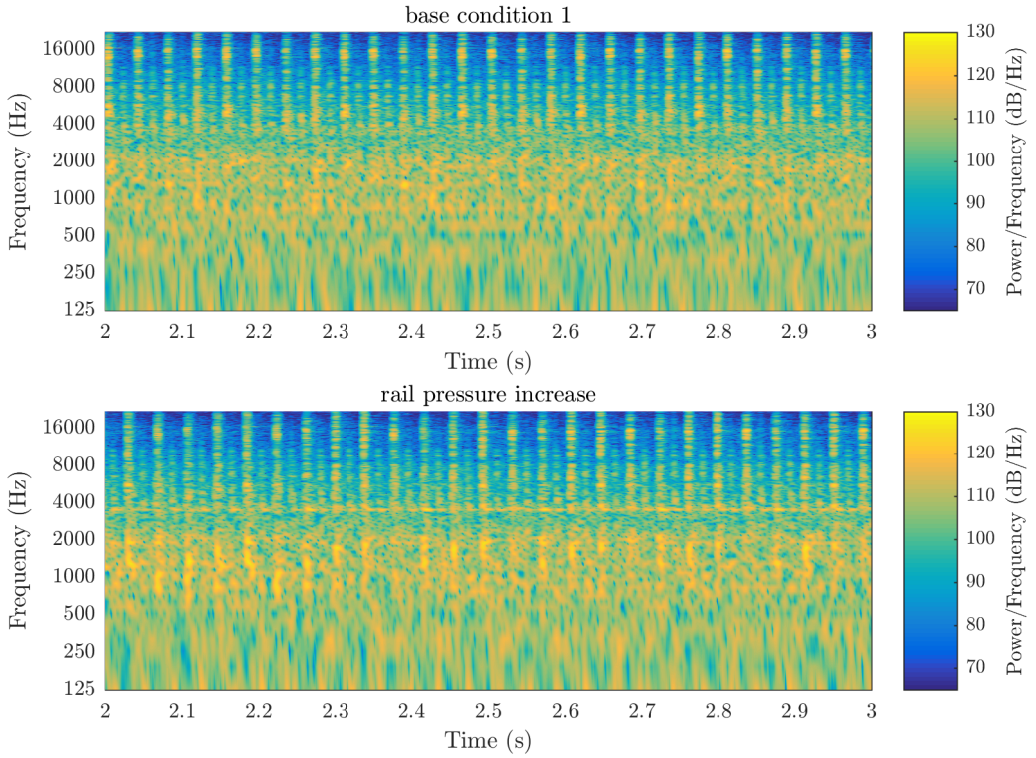


Figure 5 – Spectrograms (1024 points, 87.5 % overlap, Hann-window) for base condition 1 and increased rail pressure in idle run. The associated hearing impression *stronger* may be caused by the stronger components in the frequency range $1-2\,\mathrm{kHz}$ during the ignitions. The strong component at $4\,\mathrm{kHz}$ between the ignitions for the increased rail pressure is most likely caused by moments of second order and is exploited in a specialized feature in section 4.1.3.

the ones with one disabled cylinder and the rest. The missing cylinder produces a distinct rhythm from the ignitions, shown in figure 4. If the signal is played at the quarter sampling rate, the expected rhythmic structure with two short and one long note is clearly audible.

— The conditions with a time-shifted injection (13, 14) have a perceived faster rhythm.

— Other characteristics are tonal or narrow-band noise components, which may be temporally modulated. These do not follow any particular structure. Also the terms *stronger* and *harsher* are used to describe differences to the base conditions, e.g. for the increased rail pressure condition in idle. A spectrogram compares these two conditions in figure 5.

2.3 Spatial Distribution

Although 16 microphones were recorded, the goal is to use as less microphones as possible. To select a suitable microphone for the algorithm development, the spatial directivity was examined. For a comprehensible comparison, only a frequency-dependent single-value measure for one stationary run for each microphone is desirable. A 1/3-octaves frequency resolution, which is realised with twelth-order butterworth bandpass filters, should be sufficient. For the assessment of the varying sound pressure, the time-weighted sound pressure level according to DIN EN 61672-1 [1] from the sound pressure $p(t)$ is calculated with

$$L_\tau = 20 \log_{10} \left[\frac{\sqrt{1/\tau \int_{-\infty}^{t} p^2(\xi) \mathrm{e}^{-(t-\xi)/\tau} \, \mathrm{d}\xi}}{p_0} \right] . \tag{1}$$

τ is the time-constant and is 0.125 s for time-weighting *fast* (L_F) and 1 s for time-weighting *slow* (L_S), p_0 is the sound pressure reference level 20 μPa. In the standard no down-sampling mechanism is mentioned, so the levels are actually defined for every sample. The sometimes mentioned *impulse* weighting is discouraged by the standard itself, as it is not suitable to assess the loudness of transient sounds or the risk of hearing damage. The short attack time of 35 ms might be useful, the release time of 1.5 s is totally inappropriate in this scenario. Nevertheless, the L_S level should be a reliable measure for stationary characteristics of the microphone signal, because it will only have minimal fluctuations due to the long weighting period. The median level of the last six seconds of the recording, which is twelve seconds long, is taken as final value.

A second measure was constructed, which should consider the impulse characteristics of the signals. To preserve the peaks in the signal, but create a level in decibel, equation (1) is used with a very short time-constant of $\tau = 50\,\mu\mathrm{s}$ [2], as commonly used for a *peak* measurement [19]. From this highly fluctuating level, the basis level and average peak level [15] are calculated, but without the A frequency-weighting. The third-bands are analyzed separately. The basis level is the sound pressure level, which is exceeded for 95 % of the observation period. The average peak level is exceeded at 1 % of the observation period. Because of this inverse definition of percentiles, the basis level is identified by its correct percentile, denoted as $L_{\mathrm{p}5}$.

2. So τ equals approximately 3.3 samples.

Similarly the average peak level is defined as L_{p99}.[3] Within this thesis we are only interested in the relative impulse characteristics, therefore the impulse measure is defined as:

$$L_{\mathrm{Imp}} = L_{\mathrm{p99}} - L_{\mathrm{p5}}. \tag{2}$$

The two measures L_{S} and L_{Imp} are presented in a color-plot with the third-bands and microphones as axes. To sharpen the color values, up to two normalizations are applied:

1. As the L_{S} is still an absolute value, the broad-band level is subtracted from each third-band. Such a normalization is sometimes called *intensity ratio* and is only applied to the L_{S}.
2. The values in each frequency band are normalized by subtracting the mean level of all microphones in this band. The mean is not calculated from decibel values but the equivalent power values. Therefore, the omni-directional frequency response is suppressed, as would have happened to any frequency-weighting. This normalization is applied to L_{S} and L_{Imp}.

The normalized impulse directivity L'_{Imp} is shown in figure 6 for the base condition 1 with both normalizations. As mentioned in section 2.2, the heard differences between the conditions are small. The L'_{Imp} is also similar between the conditions, and therefore only one condition is presented here. As one can see in the spectrograms in figure 5 above, the ignitions create transients in the high frequency range. Figure 6 shows that these transients are most prominent in descending order in the microphones 1, 6, 5 and 4 above 4 kHz. These microphones are all placed close together (top and 1/3-third of the upper-ring) and the corresponding direction from the engine was already shown in the picture in figure 3(b). For higher revolutions per minute or more load on the engine, this directivity is lost. In the mid-frequency range between 500 Hz and 4 kHz, no dominant directivity can be found. In the lower frequency range below 500 Hz some strong, single peaks are visible, which change significantly with the revolution speed and the load. Nevertheless, the first microphone shows a good impulse capture ability for the high frequency-range for nearly all shown operations and should be the first choice for algorithms depending on high-frequency transients.

The normalized stationary directivity L'_{S} is shown in figure 7. Because the values

3. In [21] also the term percentile level is used, although they are defined as the basis and average peak level.

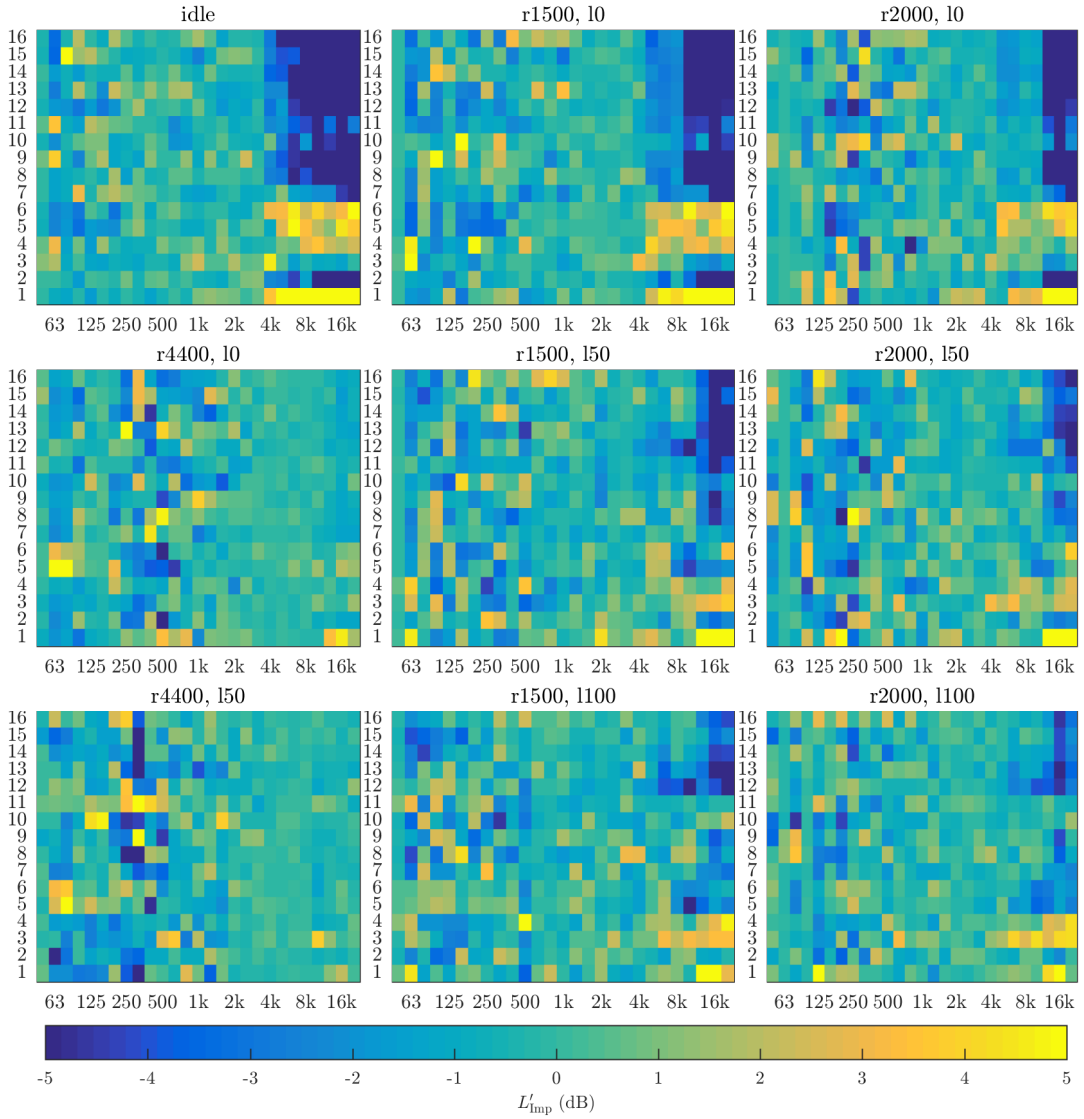


Figure 6 – Normalized impulse directivity L'_{Imp} for base condition 1. On the horizontal axis are the third-bands with center frequencies 50 Hz – 20 kHz. The microphone number corresponds to the vertical axis.

are also normalized to the broadband energy, it is possible to see which microphone has the most amount of stationary energy in a third-band. Again three blocks can be distinguished. For the very low frequency-range below 160 Hz microphones 5, 6 and 12 show the most amount over all operations. These microphones are all close together in front of the engine. In the lower mid-range ($160 - 500$ Hz), microphones 8, 9 and 10 show the most reliable properties. These are located in the lower ring in the back of the engine and are therefore quite the opposite direction as microphones 5, 6 and 12. Frequencies above 500 Hz are nearly equally strong in the microphones 1-4, with the strongest components in microphone 1 for the top end. So unless the focus is particularly below 500 Hz, microphone 1 is definitely a good choice for further analysis, because of the efficient combination of lots of broadband stationary energy and strong transients.

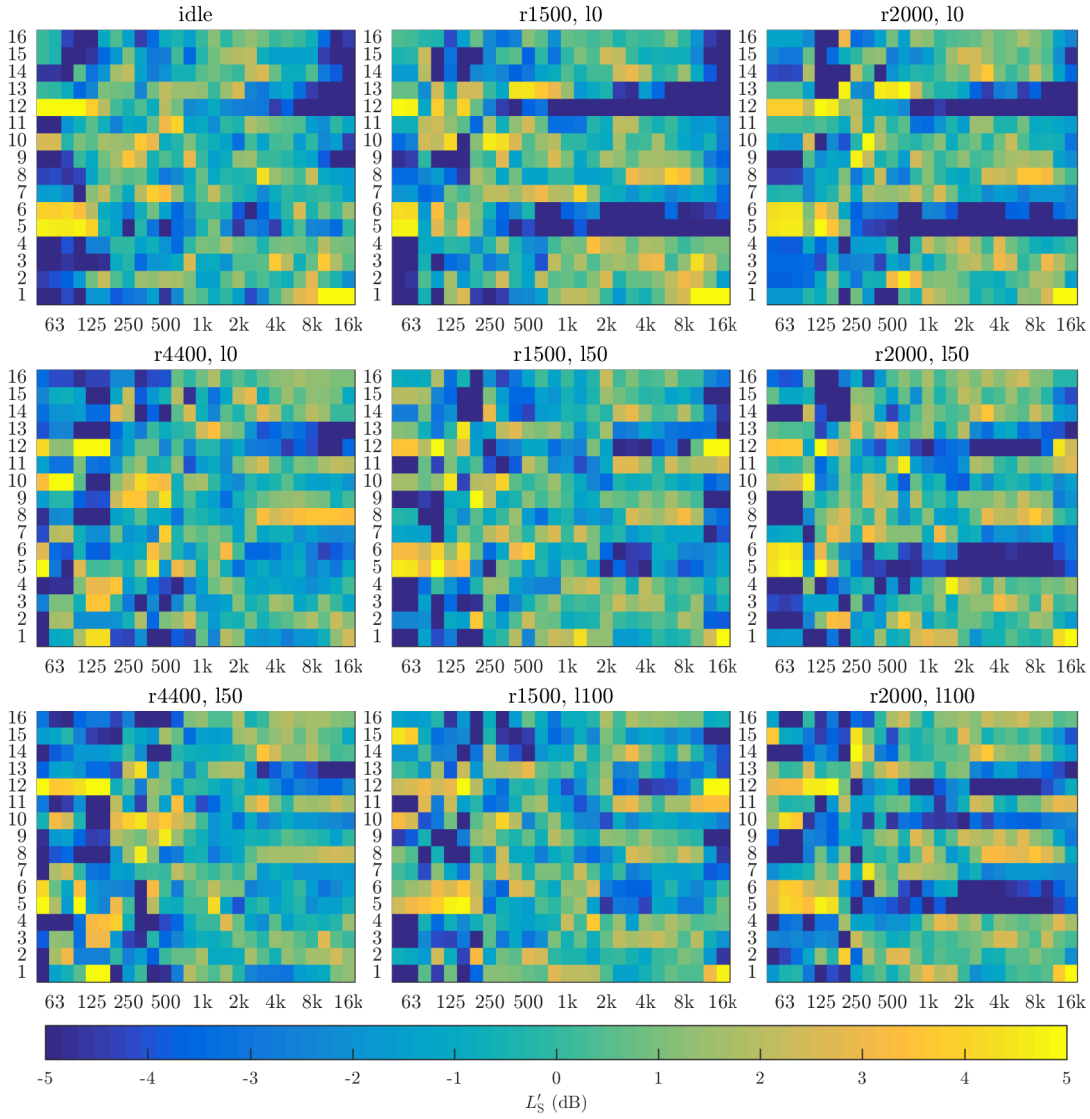


Figure 7 – Normalized stationary directivity L'_S for base condition 1. On the horizontal axis are the third-bands with center frequencies 50 Hz – 20 kHz. The microphone number corresponds to the vertical axis.

2.4 Cycle and Angle Alignment

Although it is possible to define a suitable microphone in the last section, a crankshaft angle analysis is a good way to determine the characteristics of a duty cycle and therefore the running characteristics [31]. In a crankshaft angle analysis, the sound pressure signal is sliced into frames, which are one duty cycle long. Inside one frame the signal is equi-angular resampled, i.d. instead of a constant time interval between two samples, a constant crankshaft angle is used. The alignment in whole cycles can be done by the cogwheel 1 signal and is described in section 2.4.1. The interpolating algorithm for the angle alignment is discussed in section 2.4.2. The verification of the algorithms and the general effects of the alignments are presented in chapter 2.4.3.

Because the cogwheel 1 signal has one tick per round and not per cycle, it is not possible to distinguish between the first or second round of a cycle. Also the exact timing of the top dead center is unknown. Therefore, in this master thesis the angle is never aligned to the top dead center or the cycle. Since the main objective is to measure microphone signals, which have no physical connection to the engine and no intrinsic defined delay, this is not a drawback.

2.4.1 Whole Cycle Alignment

The used recording system guarantees perfect synchronization between all signals for one measurement. Therefore, it is possible to extract the exact time instant of a whole round from the cogwheel 1 signal and use it to slice any other signal into frames. Afterwards the signal is resampled, so that all frames have the same number of samples.

The cogwheel 1 signal has a constant offset of about 5 V with negative peaks (sinc-impulses) marking one round (figure 2(a)). The lowness of the peak is dependent on the revolution speed and decreases for higher revolutions, which may disturb robust peak picking in run-ups. Therefore, the signal is pre-processed in the following way:

- Apply oversampling on demand
- Subtract the mean and then invert the signal
- Slice the signal into frames of 10 000 samples [4]
- Normalize each frame by its maximum value so every frame has a maximum

4. As this corresponds to approximately 1/6 second, there are always at least two rounds in every frame. With this short frame-size the normalization quickly adapts to changes.

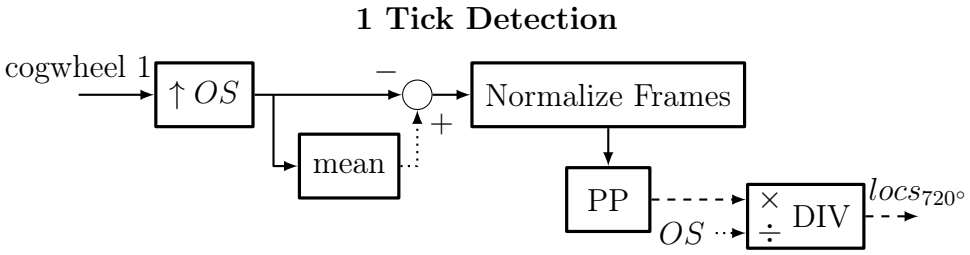
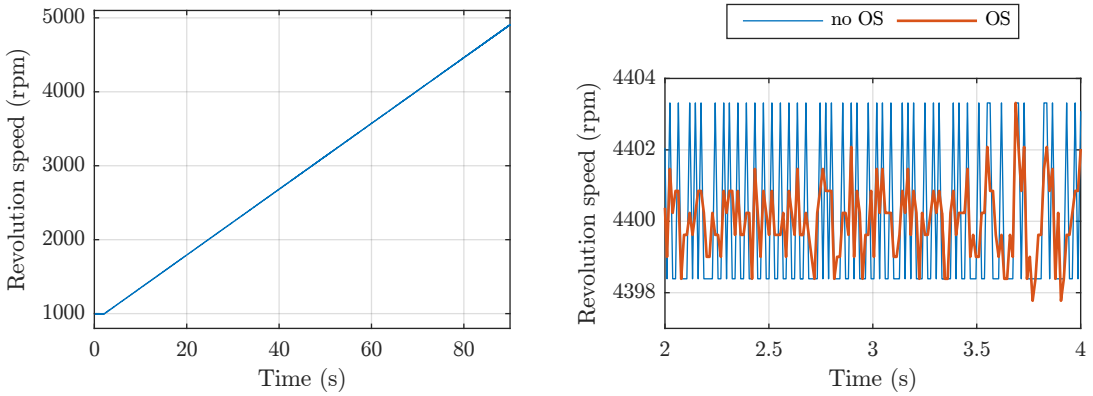


Figure 8 – Retrieval of the tick locations from the cogwheel 1 signal. In the block PP the peak-picking is performed and the locations of the peaks is returned. For the frame-wise normalization, frames of 10 000 samples without overlap are normalized to their respective maximum.



(a) Run-up.

(b) 4400 rpm with and without oversampling.

Figure 9 – Rpm calculated from the cogwheel 1 signal for base condition 1.

of 1; the last frame is normalized to the second-to-last frame, because the last frame may not contain a peak

— Find the location of the peaks, which are higher than 0.5

This algorithm is called *1 Tick Detection* and is shown in a block diagram in figure 8. The output of this algorithm is referred to as $locs_{720°}$, as the locations of two rounds of the engine are calculated. The frame-wise normalization can be efficiently done in MATLAB without any noticeable speed loss and provides a very robust detection, as shown in figure 9(a). The number of samples between two peaks can easily be converted to rounds per minute (rpm). But without oversampling, the detectable rpm are quite quantizied, because the number of samples is an integer. At 4400 rpm one round is approximately 891 samples long, which gives an uncertainty of approximately $5\,\mathrm{Hz} > 1\,‰$ without oversampling, see figure 9(b). If the signal is eight times oversampled, an even more accurate detection of the rpm is possible.

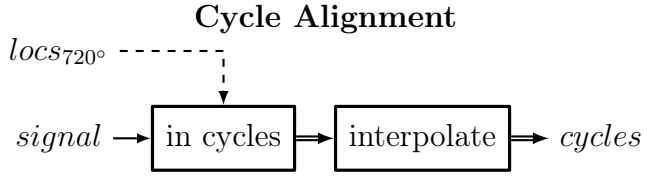


Figure 10 – Cycle alignment for a signal with the corresponding locations of two rounds in $locs_{720°}$ and resampling to 12288 samples per cycle.

After the localization of the peaks of the cogwheel 1 signal, any other signal of this measurement can be aligned in whole cycles, simply by slicing into frames at the found peaks. For better comparison all frames are resampled to the same length, regardless of the revolution speed, to get a fixed angle sampling interval. To avoid aliasing, we always must resample to more samples. The most samples are needed for idle ($\approx 780\,\mathrm{rpm}$), as strict lower bound 700 rpm are used:

$$700\,\mathrm{rpm} = 11.\dot{6}\,\mathrm{rounds/s} = 5.8\dot{3}\,\mathrm{cycles/s} \Rightarrow \lceil N_{\mathrm{cycle}} \rceil > 11235\,\mathrm{samples/cycle}.$$

A practical choice is $N_{\mathrm{cycle}} = 4096 \cdot 3 = 12288\,\mathrm{samples}$. It is very close to the minimum demand and has very small prime-factors (all 2 except one 3), which will guarantee a quick Fast-Fourier transform [16]. It is also dividable into useful parts such as half rounds. Therefore, one sample equals $\frac{720\,°}{12288} \approx 0.06\,°$. The block diagram for the alignment of any signal into cycles on locations describing two rounds is shown in figure 10.

2.4.2 Angle Alignment

If the revolutional speed is perfectly constant over time, the time sampling interval ($T_s = 15.25\,\mu\mathrm{s}$) would equal the angle sampling interval ($0.06\,°$). But as we could already see in figure 2(b), the revolution speed has a variation of more than 10 % in idle. In the desired domain of equi-angular sampling, the varying time for a constant angle can be seen as a very extreme form of jitter. Jitter is defined as usually small and often stochastic phase variations of the clock of an analog-to-digital converter (ADC), which creates phase-modulations in the time-discrete domain resulting in a lower signal-to-noise ratio [6]. For an accurate crankshaft analysis, it is therefore necessary to align the signal to the angles in the cycle.

The cogwheel 122 signal is used as reference signal for the angle. The signal is sliced into frames and resampled as described in section 2.4.1. The signal is very similar

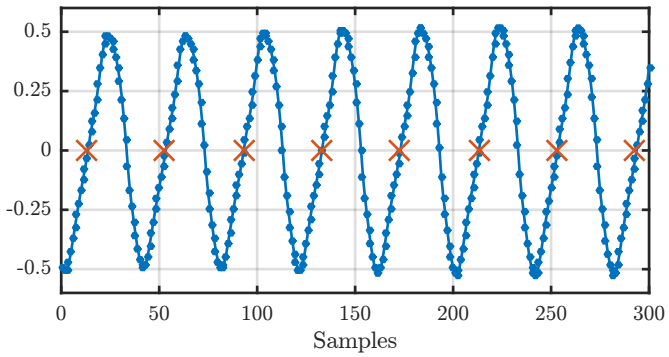


Figure 11 – Cogwheel 122 signal with the actual samples marked as blue dots. The orange crosses mark the *true* samples of $\Gamma[n]$ (*ticks*).

to a triangular signal with 122 periods per round. Instead of peak-picking on the maximum, we use the positive zero-crossing to avoid a glancing intersection at the peak. They can easily be found by using the sign and diff function, which will result in an 1 for a positive, -1 for a negative zero-crossing and will be zero otherwise.

$$\operatorname{sgn}(x) = \begin{cases} -1 & x < 0 \\ 0 & x = 0 \\ 1 & x > 0 \end{cases} \tag{3}$$

$$\operatorname{diff}(x[n]) = x[n] - x[n-1] \quad n = 2, \ldots, N \tag{4}$$

$$\Gamma[n] = \operatorname{diff}(\operatorname{sgn}(x[n])) > 0 \tag{5}$$

So the signal $\Gamma[n]$ will only be *true* at the positive zero-crossings, as shown in figure 11. The indexes where $\Gamma[n]$ is true, are the sample locations of a $\frac{360\,^\circ}{122} \approx 3\,^\circ$ rotation per *tick*. If the diff function is applied again to these tick locations, it is possible to see the time-variation for a constant angle (blue line in figure 12). As one can already see in figure 11, the timing between two ticks is quantizied by the sampling rate. By oversampling the signal, the locations of the zero-crossings can be determined with more accuracy. With an oversampling factor of eight, a clear periodic structure is visible along the angle. The four periods in one cycle in figure 12 correspond to the four ignitions. Each ignition accelerates the crankshaft and results in a fast decline of the time between ticks. Afterwards the crankshaft slows down until the next ignition $180\,^\circ$ later. For performance reasons the locations of the zero-crossings have been pre-calculated for the whole data set for each cycle.

From the locations of the $3\,^\circ$ ticks from the cogwheel 122 signal, it is possible to

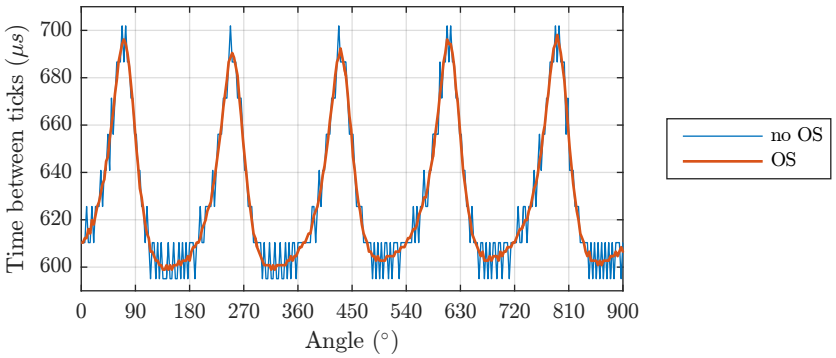


Figure 12 – Time variation between the found ticks. In blue without oversampling, in orange the cogwheel 122 signal was oversampled 8 times.

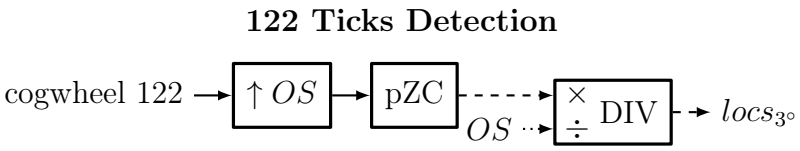


Figure 13 – Retrieval of the tick locations from the cogwheel 122 signal. The block pZC contains the finding of the locations of the positive Zero-Crossings.

perform the equi-angular resampling. Again, these locations correspond to the *true* values of $\Gamma[k]$ (eq. (5)), which describe the positive zero-crossing. The procedure for calculating the locations of the *true* values of $\Gamma[n]$ is referred to as positive Zero-Crossing (pZC) in the block diagrams. The block diagram for this *122 Tick Detection* with oversampling is shown in figure 13. The output $locs_{3°}$ contains the locations of the 3 ° ticks.

From the cogwheel 122 signal we only get 244 locations per cycle in $locs_{3°}$, so the first step in the equi-angular resampling is to interpolate the location indices to 12288 locations per cycle, resulting in the full angle resolution of 0.06 °. Now the desired signal to be equi-angular resampled is interpolated at exact these locations. This *Angle Alignment* procedure is shown in figure 14. The alignment, i.d. the interpolation, is performed for every cycle separately. For both interpolation steps a cubic-spline interpolation is used with a continuos second derivative [28]. It is important to notice that a cubic spline may produce extreme values while extrapolating outside end-knots. Therefore, it is necessary to append one round before and after the actual cycle.

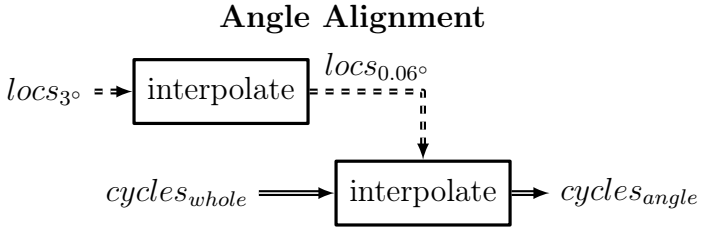


Figure 14 – Angle alignment with the corresponding locations for the 3° ticks from the cogwheel 122 signal. Both the locations and the signal ($cycles_{whole}$) must be cycle-aligned.

The whole angle alignment algorithm can be summarized as following:

Part 1 - Preparation

1. Find the location of the peaks of the pre-processed and over-sampled cogwheel 1 signal (1 peak/round).
2. Slice the over-sampled cogwheel 122 signal on these locations into frames of the length of two cycles, whereby the two rounds in the center are the actual cycle.
3. Resample all frames to $2 \cdot OS \cdot 12288$ samples with the oversampling factor $OS = 8$.
4. Find the locations of the ticks of the cogwheel 122 signals in all frames.

Part 2 - Signal Alignment

1. Slice the signal into frames of the length of two cycles, analogous to step 2 of Part 1.
2. Resample all frames to $2 \cdot 12288$ samples.
3. Perform the equi-angular resampling with the locations found in step 4 of Part 1.
4. Use the two rounds in the center of the two cycles only.

This algorithm is also shown in the block diagram in figure 15. The lower-third part of the figure correspond to part 2 of the algorithm. Also note, that the right-third of the figure is only needed for angle alignment. As mentioned before, the first part of the algorithm has been calculated for the whole data set. The second-part is done on-the-fly, because much interpolating is done and the complete alignment of the whole data set would need more than 100 GB of space. Also with the pre-calculated data a stationary run is aligned in a few seconds. Oversampling shows a significant improvement in the retrieval of the locations of the cogwheel 1 and cogwheel 122 signal on the alignment. However, the alignment of the signal can be calculated

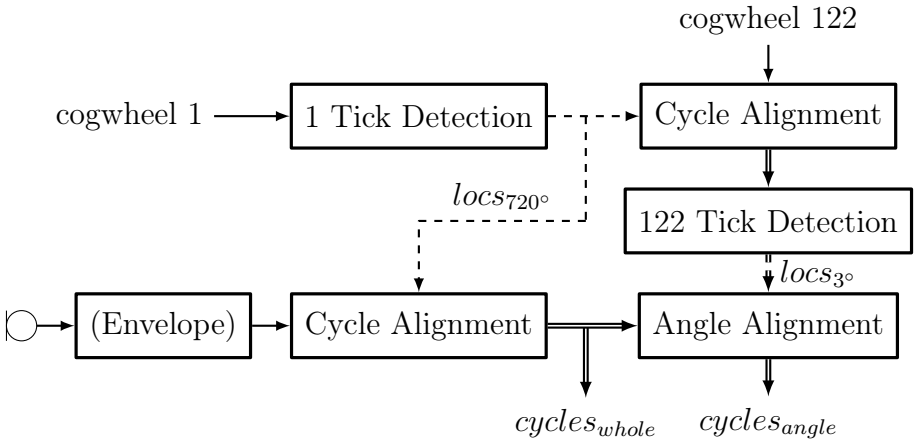


Figure 15 – Whole cycle and angle alignment of a microphone signal with cogwheel signals.

without oversampling, as oversampling the signal itself does not significantly improve the accuracy of the alignment.

2.4.3 Evaluation

Figure 16 shows the envelope of the microphone signal 1 in equi-angular resampled frames. The four ignitions are clearly visible, as well as a peak between the ignitions. The second-order fictitious forces of piston and connection rod should be the cause of this peak, as they run at double frequency of the engine [31]. Also the non-uniform gas forces create moments of second order in a four-stroke engine [26].

The accuracy of the alignment can be shown with the injections signal. This signal controls the opening and closing of the injectors and is therefore created by the engine control unit itself and triggered by the crankshaft angle. Figure 17 shows the difference between the alignment of the injections signal on only the whole cycles and angle alignment. In the first case, the variations between the cycles are visible, because the angle axis is derived as (time-)linear interpolation between the cycle ends. If equi-angular resampling is used instead, the three discrete injections are nearly perfectly aligned between the cycles.

As a measure of this alignment, the cross-correlations with zero-lag between all

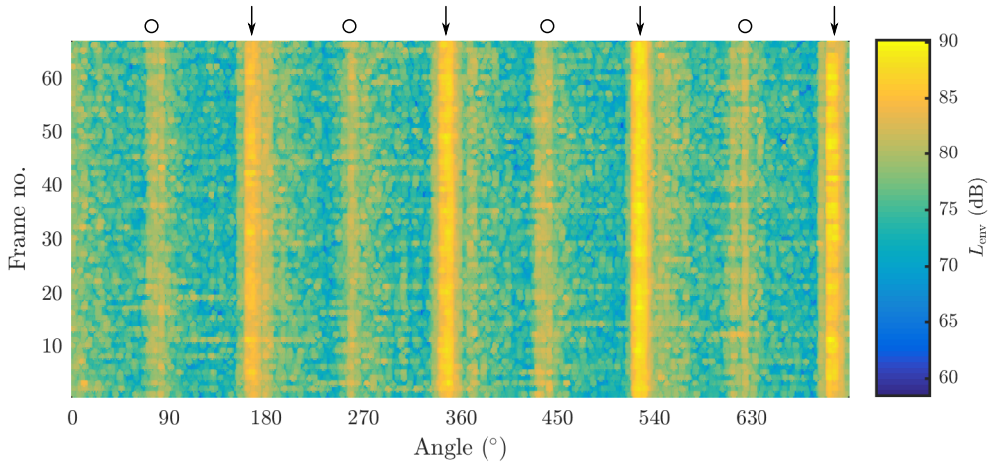


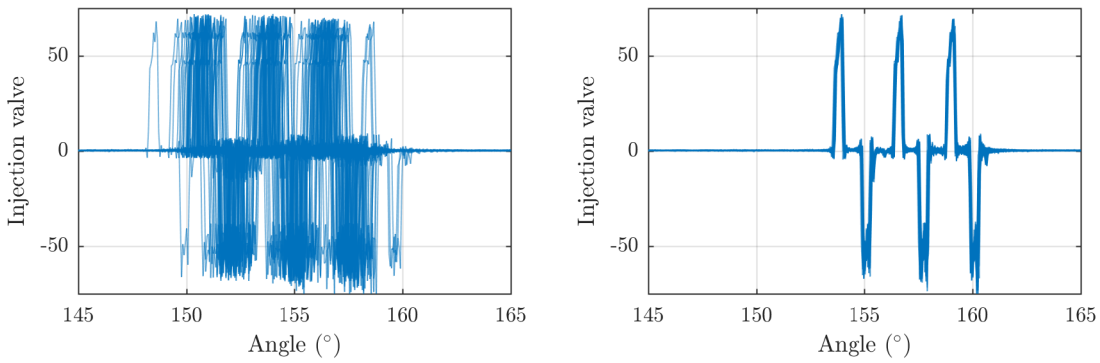
Figure 16 – Equi-angular resampled and cycle aligned envelope of microphone signal 1 for an idle run. The arrows mark the locations of the ignitions, second order moments are marked with a circle.

cycles, normalized by their respective L2-norms, is defined as

$$r_{ij} = \frac{c_i^{\mathbf{T}} c_j}{||c_i||_2 ||c_j||_2}, \qquad i \neq j. \tag{6}$$

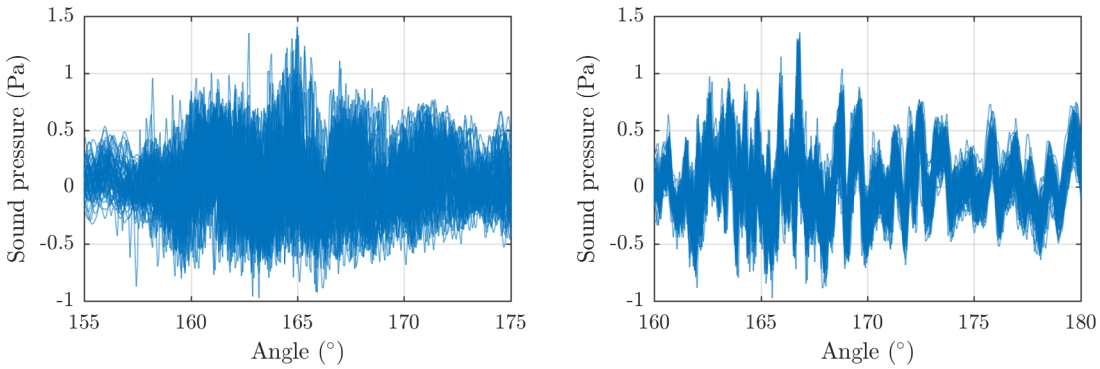
Where c_n is a vector containing all samples of the n-th cycle. The correlation coefficients of the whole cycle alignment of the injections signal are not necessarily normally distributed with a median value of 0.106 and have the levels of -0.06 and 0.38 for the 25- and 75-percentiles for this measurement (base condition 1, idle). With angle alignment, the median value is 0.938 with 0.92 and 0.95 as the same percentiles. For a microphone signal this high-accuracy for the angle alignment can not be reached, as can be seen in figure 18. The figure also shows the ignition, which is delayed in angle due to sound propagation to the microphone signal. A possible explanation of this poorer performance is the low frequency vibration of the sound-emitting surfaces of the engine. But without angle alignment, no reliable high-frequency phase information can be extracted. The median value for the whole cycle alignment of 0.12 and 0.05 and 0.2 as 25 and 75 percentiles is improved to 0.47 with 0.42 and 0.51 as percentiles by using angle alignment.

Figure 19 shows a box plot of the correlation coefficients for the five base conditions and the stationary runs for the microphone signal 1. Especially for idle runs there is apparently much variation between the cycles, resulting in a low correlation (≈ 0.1) between the cycles, if only cycle alignment is used. With the alignment on the angle, this correlation is significantly increased up to 0.5. Although the shown base



(a) Whole Cycle Alignment (b) Angle Alignment

Figure 17 – Alignment comparison for the injections signal during the ignition. Each cycle of this measurement is plotted as a single line over the angle. The 20 ° range corresponds to a time duration of 4.3 ms.



(a) Whole Cycle Alignment (b) Angle Alignment

Figure 18 – Alignment comparison for a microphone signal during the ignition. Each cycle of this measurement is plotted as a single line over the angle.

conditions are repetitions without any parameter change, significant differences are observable. At high revolutions the variation between cycles is less, as even without equi-angular resampling high correlation coefficients of 0.8 are reached, which even can not be improved by angle alignment. A possible explanation may again be the engine vibrations, as their influence can not be corrected by equi-angular resampling. This characteristic will later by shown in section 4.2, when a feature is based on the variation of the revolution speed.

The tendency of a smaller difference between the correlation coefficients at higher revolutions between the methods can also be found for other microphones. Nevertheless, the behavior of the influence of the load for one revolution speed may change

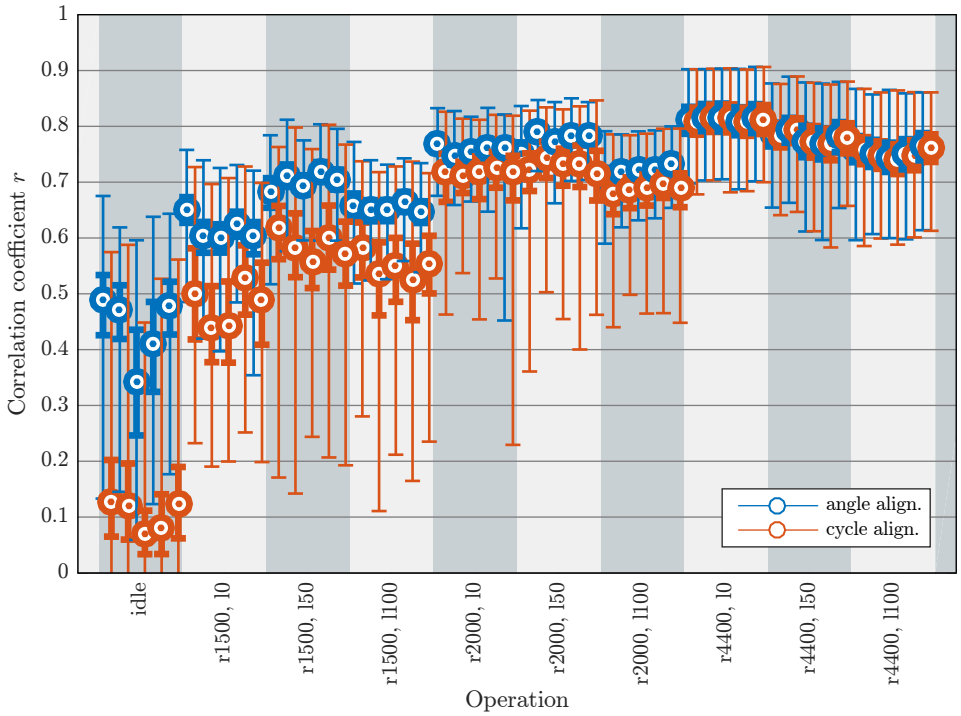


Figure 19 – Box plot of the correlation coefficients with minimum and maximum values (thin lines), the 25 and 75 percentiles (thick lines) and the median value (dots) for the five base conditions and the stationary runs for the microphone signal 1. Angle alignment (orange) with equi-angular resampling shows the biggest improvement at lower revolutions. At high revolutions nearly no difference can be found, if the signal is only sliced into frames by cycle alignment (blue).

completely for other microphones and should therefore not be analyzed into detail by using only one microphone channel. Because these correlations are out of the scope of this thesis, they are not further analyzed.

2.5 Summary

This chapter focused on providing an overview of the available data set. The signals taken directly from the engine contain much and very exact information. Nevertheless, in the analysis by ear, only small differences could be heard in the microphone signals between the different conditions. The most dominant sounds are transient peaks at lower revolution speed, which are caused by the ignitions in each cylinder. One of two crude measures is defined to evaluate this quality, as it will likely be the basis for features we use in the classification. The other defined measure is based on the stationary energy. The two measures are used to evaluate the spatial

distribution, so it is possible to focus on only one microphone in the algorithm development. The microphone on top of the engine shows the most promising result and is therefore used in the next chapters.

With crankshaft angle analysis it is possible to divide the signal into meaningful, characteristic parts, but we need at least two additional signals from the engine, which may contain the information we want to extract from the audio signal anyway. The focus on the next chapter is using only the microphone signal for cycle alignment to extract as much information from the microphone signal as possible.

3 Cycle Alignment with a Microphone Signal

This chapter is focused on the reduction of required signals without loosing relevant information. As discussed in the previous chapter, crankshaft angle analysis will show engine characteristics reliably, nevertheless two additional signals from the engine are necessary. In this chapter algorithms are presented, which are able to find the occurrence of the ignitions in the microphone signal and are used to perform whole cycle alignment, similar to section 2.4.1. Although it is not possible to do a finer angle alignment, the cycle alignment will then be used to define robust features in chapter 4, which will only require the microphone signal as input. As by-product of this development, it will be possible to already detect a missing cylinder combustion.

Another motivation for a closer look at the ignitions can be found by studying the vibro-acoustical mechanisms in an internal combustion engine. The most relevant aspects in connection to this thesis are presented in the next two paragraphs out of the book *Fahrzeugakustik* by Martin Pflüger et al. [23].

The excitation mechanisms in an internal combustion engine can be divided into combustion excitations, mechanical excitations and other excitations like airflow. All of this mechanisms result in vibrations of the engine structure and are therefore influenced by the respective transfer paths. Together with the properties of the sound emitting surfaces of the engine and the coupling to airborne sound transmission, the final sound is determined. Because for the measurements discussed in this thesis the engine was not mounted in a car body, we only focus on the sound sources on the engine block. Inside a duty cycle various pressure and collision impulses create together with liquid or gas flow processes vibrations of the structure. Additionally belt drives and other aggregates like injection pump, generator, oil and water pumps cause excitation.

Nevertheless, the combustion chamber pressure is one of the most dominant excitation mechanism for a diesel engine, which matches with the analysis by ear in section 2.2. The transient peaks correspond to the single ignitions. The combustion chamber pressure is dependent on the revolution speed and shows a decrease in amplitude for higher frequencies (low-pass characteristic). But the band-pass characteristic with a very steep lower slope of the transfer path results in a more high-pass characteristic of the ignition induced airborne sound. This could already be seen in the spectrograms in figure 5, as during the ignition energy is mainly added

at higher frequencies. As the mechanical excitations of the crankshaft are in-phase with the combustions, they also add to the final ignition sound. On the other hand, second order moments, which occur with the double frequency, may happen exactly between two ignitions. As already discussed, the ignition sound is dependent on the revolution speed, as well on the load and the parameters of the engine control unit (e.g. injection timing).

All of the measured faulty conditions (table 1) are closely linked to the ignition. The most obvious one is, of course, the lack of the ignition in one cylinder (condition id 9-12). The other conditions influence the timing (13, 14) or the combustion process itself (6, 7, 8). Therefore, we will focus on finding parameters of the characteristic sound of an ignition. The goal in this chapter is mainly to extract time information of the ignition by defining an ignition-based envelope (section 3.2) to perform cycle alignment with only the microphone signal. The characteristic shape of cycle-aligned ignitions is then used for features for the classification in the next chapter.

After the definition of the ignition-based envelope, a rpm-detection is implemented on this envelope in section 3.3, to evaluate the timing precision. This algorithm separates itself from other rpm-detection algorithms, as it calculates the revolution speed between the ignitions. Because the performance of the up to this point defined algorithms depends on ten non-independent parameters, a Monte-Carlo simulation is used to optimize the parameter set in section 3.4.

If one ignition is missing in the signal, we will get a characteristic pattern in the calculated revolution speed. As we want to use this pattern to find the four faulty conditions of a missing cylinder, a specific performance measure for the Monte-Carlo simulation is created, which simultaneously is a reliable feature for the detection of a missing cylinder, which is later discussed in section 3.6.

The parameters of the algorithms are based on simplicity of their role in the algorithms and are therefore also dependent on the average revolution speed. Because the average revolution speed should be known at an engine test bench, this is not a big drawback. Also, the retrieval of the revolution speed from engine sound is already studied in literature, e.g. [11]. Nevertheless, another proposal for the rpm-detection over a time period is presented in section 3.5. With this algorithm it is possible to detect a missing cylinder solely with the microphone signal. This procedure is summarized and evaluated in section 3.6. Finally, in section 3.7 a cycle alignment algorithm, which now only requires a microphone input signal, is compared to the already discussed whole cycle and angle alignment of chapter 2.

Because all algorithms depend on moving average and moving median filters, the specific characteristics are highlighted beforehand in section 3.1.

As the ignition sound is most dominant for lower revolutions and engine tests are typically started with no load, this chapter focuses on this operations: idle, 1500 rpm @ 0 % load, 1500 rpm @ 50 % load and 2000 rpm @ 0 % load. The algorithms will be developed using microphone 1, which is placed directly above the engine.

3.1 Zero-phase Moving Average and Median Filtering

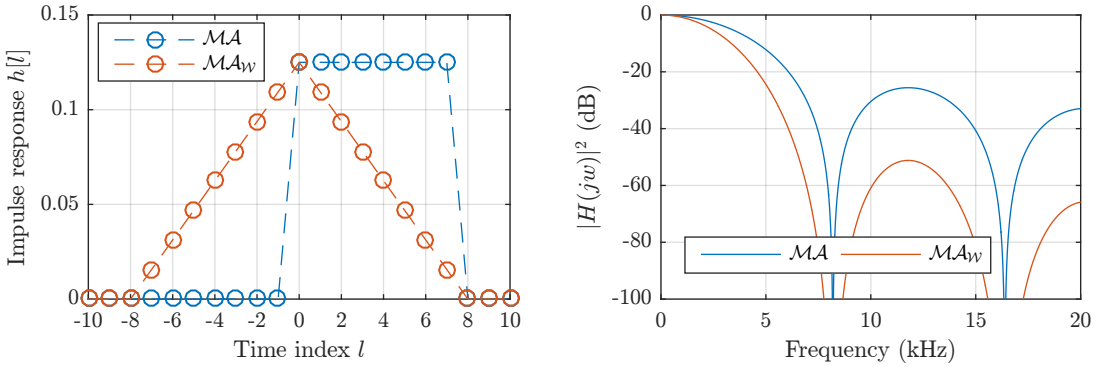
A moving average filter (MA) is a FIR-filter with the impulse response [7]

$$\mathcal{MA}\{x[n], L\} = \sum_{k=-\infty}^{\infty} x[n-k]h[k], \quad h[l] = \begin{cases} \frac{1}{L} & 0 \leq l \leq L-1 \\ 0 & \text{else.} \end{cases} \tag{7}$$

As only one constant factor for multiplication is needed, the filter is one of the simplest digital filters. As the name suggests, the filter calculates the average or mean for a sliding window along the data. In the frequency domain the filter has a low-pass characteristic and suppresses high-frequency components. As in many measurement scenarios, the desired signal is superimposed with high-frequency noise, moving average filtering is a common technique for improving the signal-to-noise ratio [27]. But the implementation of the moving average filter in equation (7) has a major drawback, as it introduces a phase-shift. If, for example, a moving average filter is applied to the squared microphone signal to calculate an envelope, no angle alignment with a cogwheel signal is possible. Therefore, to compensate for this shift, MATLAB's zero-phase filtering function *filtfilt()* was used, which uses forward and backward filtering. If the filter coefficients of the moving-average are used for filtfilt, the resulting impulse response is the triangular function

$$\mathcal{MA}_{\mathcal{W}}\{x[n], L\} = \sum_{k=-\infty}^{\infty} x[n-k]h[k], \quad h[l] = \begin{cases} \frac{L-|l|}{L} & -L \leq l \leq L-1 \\ 0 & \text{else.} \end{cases} \tag{8}$$

Thus, a zero-phase weighted moving average is created. Figure 20 shows the impulse and the magnitude frequency response for the two moving average filter implementations. The weighted moving average results in a squared magnitude response with a higher suppression of high frequency components due to the doubled filter order. The area below both impulse responses equals to one, which yields the 0 dB gain



(a) Impulse Response. (b) Frequency Response.

Figure 20 – Comparison of the impulse and frequency response for the standard moving average $\mathcal{MA}$ (blue) and the zero-phase weighted moving average filter $\mathcal{MA}_\mathcal{W}$ (red).

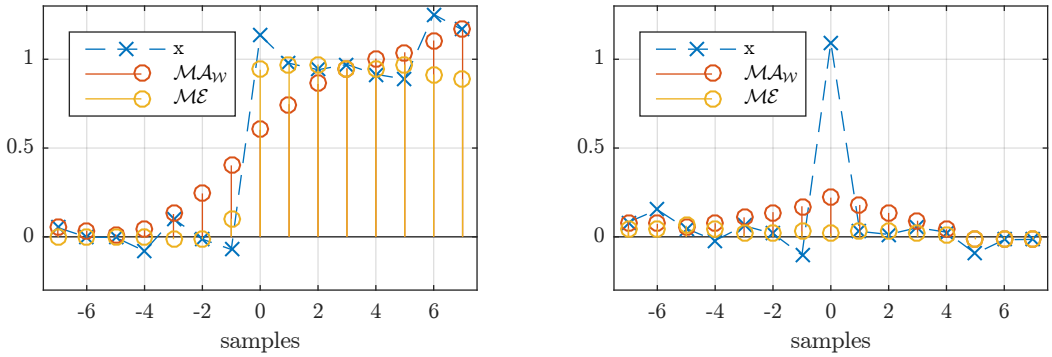
at 0 Hz. The introduced loss of causality by forward and backward filtering will not significantly restrain the real-time capabilities in this application, as even a few seconds of buffering will only introduce an acceptable amount of latency. Therefore, the zero-phase weighted moving average, denoted as $\mathcal{MA}_\mathcal{W}$, will be used for nearly all cases in this thesis.

From a more statistical perspective the median of a sample is more robust against outliers in the data, as one extreme value will not affect the median, contrary to the mean. The median value is defined as the value in the centre of a sorted dataset, which is denoted with $x_{(l)}$. If the number of samples L in the dataset is even and the sample consists of metric data than the mean of the two values in the centre is used as median [17]

$$\tilde{x} = \begin{cases} x_{(\frac{L+1}{2})} & \text{for } L \text{ odd} \\ \frac{1}{2}[x_{(\frac{L}{2})} + x_{(\frac{L}{2}+1)}] & \text{for } L \text{ even.} \end{cases} \tag{9}$$

The median $\tilde{x}$ minimizes the sum of the absolute errors $|x_i - \tilde{x}|$, whereas the mean minimizes the sum of the squared errors $(x_i - \bar{x})^2$ [17]. In the digital domain it is possible to define a moving median filter of length L by calculating the median for each time sample with the closest L neighbors

$$\mathcal{ME}\{x[n], L\} = \text{median}\{x[n - l_l : n + l_u]\} \qquad \begin{matrix} L \text{ even}: & l_l = \frac{L}{2},\ l_u = \frac{L}{2} - 1 \\ L \text{ odd}: & l_l = l_u = \frac{L-1}{2}. \end{matrix} \tag{10}$$



(a) Step function.

(b) Dirac function.

Figure 21 – Effect of filtering with a moving average (MA) and moving median (ME) filter for the step function and a Dirac impulse. Both functions are superimposed with Gaussian noise with a standard deviation of 0.1. Both filters are 5 samples long ($L = 5$).

Due to the symmetric definition around the current sample, this filter does not include any time delay. Moving median filters are non-linear filters and can effectively be used to suppress impulse intereferences with good retention of the desired signal. They also show robust noise smoothing properties for gaussian, uniform and laplacian distributed noise [3]. As with the moving average, the filter length defines the trade-off between noise suppression and signal retention.

Mathworks significantly improved the speed and RAM usage of the *medfilt1* function in the MATLAB 2015b release, which does exactly the zero-phase moving median filtering described in equation (10). Figure 21 compares the effect of median and weighted moving average filtering on two particular functions superimposed with Gaussian noise with a standard deviation of 0.1 to show the noise suppressing capabilities. In the left figure 21(a) the step function is used to show the behavior at edges. The median filter preserves the steep edge, whereas the moving-average slurs the edge. One has to notice that the odd length of the filter is especially good for preserving the edge with a moving median, as with an even length, the median filter would have produced a value of 0.5 at sample 0 (without noise). In the right figure 21(b) the single impulse is completely canceled out by the median filter. A moving-average smears the impulse.

In short, a moving average can be seen more as a smoothing filter, whereas the moving median is a noise suppression technique for impulse interference and retention of sharp edges.

3.2 Ignition-based Envelope

As mentioned in this chapter's introduction, we want to calculate an envelope function of the microphone signal, which will have robust and relevant information about the ignitions of the engine, most importantly the timing. The envelope of a signal is a more or less smooth curve connecting the outer positions of a signal [14]. There are different ways to calculate the envelope, MATLAB's new *envelope*[5] function contains three distinctions:

Analytic The magnitude of the analytic signal, where the imaginary part is the Hilbert-transform of the real signal, is the envelope of the signal. The time-domain Hilbert transform correponds to a phase shift of positive frequency components by $+90\,^\circ$ and negative components by $-90\,^\circ$ [12].

RMS The root-mean-squared envelope is calculated by using a sliding window on the root-mean-square [16]. This is equivalent to $\sqrt{\mathcal{MA}\{x^2[n], L\}}$ with a zero-phase (but not weighted) moving average. Of course, the window length L can be specified.

Peak MATLAB finds peaks with a minimum distance and uses spline interpolation for the samples in-between [16]. The minimum distance must be specified and can also be seen as a window length.

Figure 22 shows a comparison of the different envelope functions. The Hilbert transform shows the envelope for the two superimposed sines in the right figure especially well, but flutters nearly as much as the original signal for the measurement signal in the left figure. This could be removed with an additional low-pass filter. The rms envelope gives a good estimation of the energy for the measurement signal, but follows the signal for superimposed sines. The peak signal has a very similar envelope as the analytic envelope for the two sines; its drawback is the spline interpolation, which creates oscillation for the measured microphone signal, as can be seen in the left figure.

Although the rms and peak envelopes can sharpen the localization of the ignitions, another approach from the field of music information retrieval is used. In [9] Fitzgerald proposes a very efficient and simple method for the separation of harmonic and percussive parts of a monaural audio signal by using median filtering. Admittedly, the purpose of separating the harmonic parts of an audio recording for remixing has nothing to do with this thesis, still parts of the algorithm can be used here

5. The function was introduced in the 2015b release.

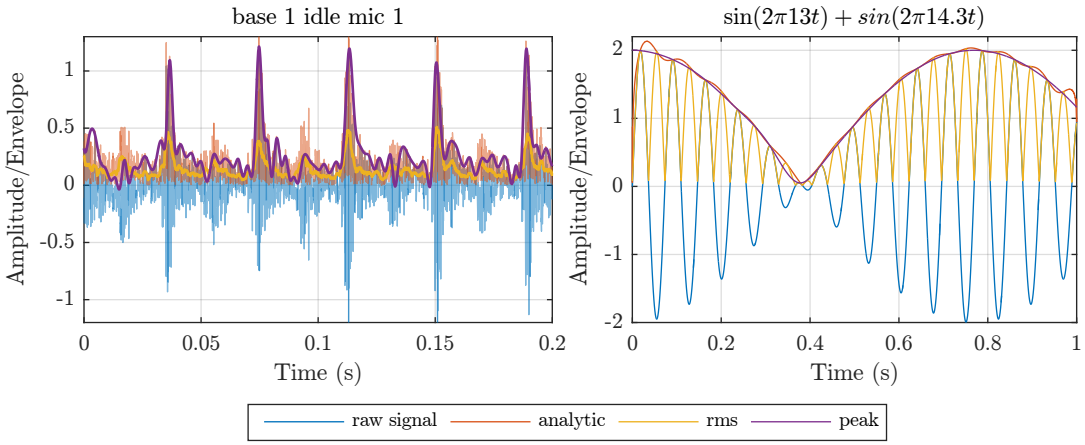


Figure 22 – Comparison of different envelope calculation techniques. In the left figure the microphone signal 1 for the base 1 condition in idle is plotted, on the right two sines with a frequency of $\frac{780}{60}\,\mathrm{rpm} = 13\,\mathrm{Hz}$ and $13\,\mathrm{Hz} + 10\,\% = 14.3\,\mathrm{Hz}$. The *rms* and *peak* envelope are calculated by using a window length of 100 samples.

effectively, as the percussive separation is used to extract the ignition sound and an information-loaded envelope of only the ignitions can be extracted directly. An overview of the algorithm is shown in figure 23.

The median-filtering for the harmonic/percussive seperation is done on the short-term Fourier transform magnitude spectrogram. The short-term Fourier transform of a signal $x[n]$ is defined as [2]

$$S[w,m] = \sum_{n=0}^{N-1} x[n]w[n-mD]\mathrm{e}^{-j2\pi nw/N} \tag{11}$$

with the frequency bin index w and the time frame index m. The order of the bin indexes has been chosen to correspond to the dimensions in MATLAB's FORTRAN memory-order. $w[n]$ is the window function to reduce leakage, typically a Hamming window is used. N is the total number of samples, D is the inverse of the frame rate or the number of samples between the beginning of two neighboring frames. It is also possible to zero-pad the input sequence $x[n]$ to a greater number of samples to increase the frequency resolution of the Fourier transform.

Fitzgerald now creates a *percussive enhanced* spectrogram $P[w,m]$ and a *harmonic enhanced* spectrogram $H[w,m]$. For each time frame $S[:,m]$ a moving median filter is applied to the frequency bins w, thus suppressing all impulse interference, i.d. all

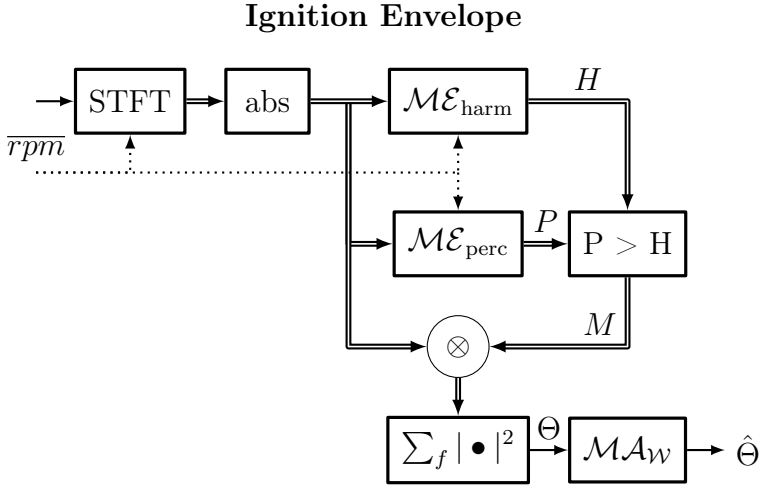


Figure 23 – Calculation of the Ignition-based Envelope with median filtering.

narrow-band frequency components, and therefore enhancing the percussive part [9]

$$P[:,m] = \mathcal{ME}\{|S[:,m]|, L_{\text{perc}}\}. \tag{12}$$

On the other hand, if the moving median filter is applied to every frequency bin $S[w,:]$ over all time frames m, all impulses in each frequency band are cancelled out, thus enhancing the harmonic structure [9]

$$H[w,:] = \mathcal{ME}\{|S[w,:]|, L_{\text{harm}}\}. \tag{13}$$

This spectrograms are now used to generate a mask, which is then applied element-wise to the the original spectrogram for percussive enhancment [9]

$$M[w,m] = \begin{cases} 1 & P[w,m] > H[w,m] \\ 0 & \text{otherwise,} \end{cases} \tag{14}$$

$$\hat{P}[w,m] = S[w,m] \otimes M[w,m]. \tag{15}$$

Figure 24 compares $S[w,m]$ and $\hat{P}[w,m]$ for the same time slot. The percussive enhancement mostly uses the high frequency energy above 7 kHz. Under this frequency the mask $M[w,n]$ is mostly zero, resulting in the white areas for $-\infty$ dB.

Fitzgerald also creates a harmonic enhancement mask, by exchanging P and H in equation (14), and also the synthesis with the inverse Fourier transform. As we are neither interested in a harmonic enhancement nor in audio-signals, we directly use

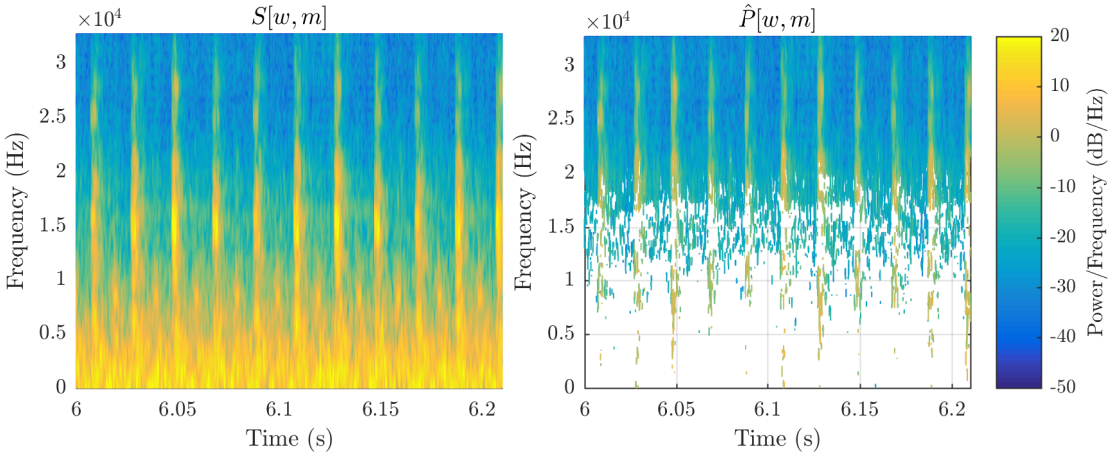


Figure 24 – Spectrogram before (left) and after (right) percussive enhancment after [9] for base run 1 with 1500 rpm and 0 % load.

$\hat{P}[w, m]$. As our goal is to find a ignition-based envelope, we simply define

$$\Theta[m] = \sum_{w=0}^{N_{\mathrm{FT}}-1} |\hat{P}[w, m]|^2 \tag{16}$$

as envelope. As the time index m suggests, this envelope has a lower sampling rate than the original signal, as it is only defined for every time frame of the short-term Fourier transform. Finally, a weighted moving average filter is applied, to further smooth the envelope

$$\hat{\Theta}[m] = \mathcal{MA}_{\mathcal{W}}\{\Theta[m], N_{\mathrm{MA}}\}. \tag{17}$$

Figure 25 shows the envelope of the original spectrum, which is calculated similar to equation (16). A more suitable term than percussive separation in this field is *transient separation*, as the transient ignitions are extracted. With the separation we get an envelope, which is nearly zero between the ignitions plus forms a distinct peak at the ignition time, even though it is hard to identify the ignitions in the envelope of the original spectrum by eye. As the ignitions were very distinctive in the spectrogram in Figure 24, the correct detection of the ignitions can be verified. Because the ignitions show the most energy in the upper frequency range, the transient separation mainly works as a high-pass filter. Nevertheless, the suppression tunes itself to the signal and therefore useful components even below 500 Hz can be considered, as can bee seen in figure 24.

In the implementation of the algorithm six parameters can be used to tune the

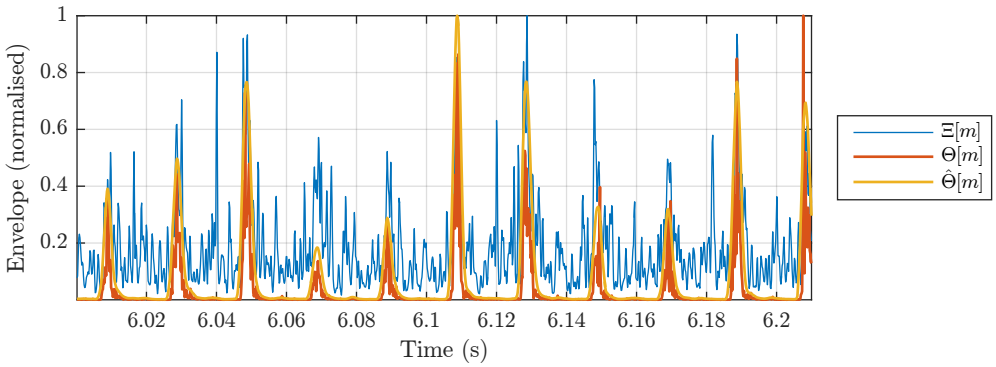


Figure 25 – Envelope of the original spectrum ($\Xi[m] = \sum_{w=0}^{N_{\mathrm{FT}}-1} |\hat{S}[w,m]|^2$), the percussive (transient) enhanced spectrum $\Theta[m]$ and the smoothed enhanced spectrum $\hat{\Theta}[m]$ ($N_{\mathrm{MA}} = 10$) for base run 1 with 1500 rpm and 0 % load. Because the energy of harmonic components is suppressed, the envelopes are normalized for better visualization.

Parameter	Description
N	Length of the time window for idle operation.
overlap	Overlap of the time frames in percent.
fft factor	The time frame is zero padded to the length N times the given value to increase frequency resolution. The number of frequency bins is not affected by the revolution speed.
perc factor	Is the ratio of the number of frequency bins to use as L_{perc} to the total number of frequency bins.
harm factor	Is the ratio of the length to use as L_{harm} to the length of one round of the engine.
N_{MA}	Weighted moving average filter length for smoothing $\Theta[m]$.

Table 3 – Parameters of the Ignition Envelope algorithm.

performance, which are shown in table 3. As the rate of ignitions per second is determined by the revolution speed, the STFT is ideally scaled to the revolution speed, so the number of time frames between two ignitions does not depended heavily on the revolution speed. As we do not want another signal from the engine, the parameters for the algorithm are scaled by an integer, which is a multiple of the approximate idle revolution speed (780 rpm) for the average rpm in the signal to analyze. With this approach the number of time samples for the envelope $\hat{\Theta}[m]$ between two rounds can not change more than a factor two for the same parameter set for different revolution speeds. Furthermore, if an engine is on the test bench, normally the approximate revolution speed is known.

To tune the parameters to a optimal set, a performance measure is needed. One goal

is to get the correct timing, so the next section focuses on calculating the revolution speed out of the smoothed envelope $\hat{\Theta}$. Afterwards, it is possible to optimize the parameter set in section 3.4.

3.3 Temporal RPM Detection

In this section the time interval between two ignitions is used to calculate the rpm per half-round. The detection of ignitions is closely related to the onset detection in music signals, where transient regions in a complex signal are located. The detection of onsets in a music signal consists usually of the following steps: The original audio is pre-processed and reduced to a detection function, which shows the occurrence of transients in the signal. Before the peak-picking to find the locations of the onsets, some post-processing steps may be performed to improve detection. Some common techniques for post-processing are normalization and adaptive thresholds [4].

The ignition-based envelope function calculated in the last section already contains every desired feature of a detection function, as it is sub-sampled and shows distinctive peaks during an ignition. As a result, we only need to perform robust peak-picking on the locations of the ignitions to calculate the revolution speed. As for every round exactly two ignitions happen in normal operation, it is possible to calculate the current revolution speed $\rho[j]$ from the time interval between two ignitions in rpm with

$$\rho[j] = \frac{60}{2} \frac{1}{t[j] - t[j-1]} \,\text{rpm}, \qquad j = 2, \ldots, J \tag{18}$$

$t[j]$ is the absolute time of the j-th detected ignition. J is the total number of detected ignitions.

Although the peaks of the ignition-based envelope are very distinct, their height changes significantly even in one measurement. Therefore, two algorithms were created, where each of them combines post-processing of the averaged envelope, peak-picking and rpm calculation and is designed to be robust to the varying envelope shapes caused by different revolution speeds and loads on the engine.

Adaptive Threshold Algorithm

As proposed in [4], normalization and an adaptive threshold with a (nonlinear) smoothing filter is used as pre-processing. The envelope is normalized to a maximum

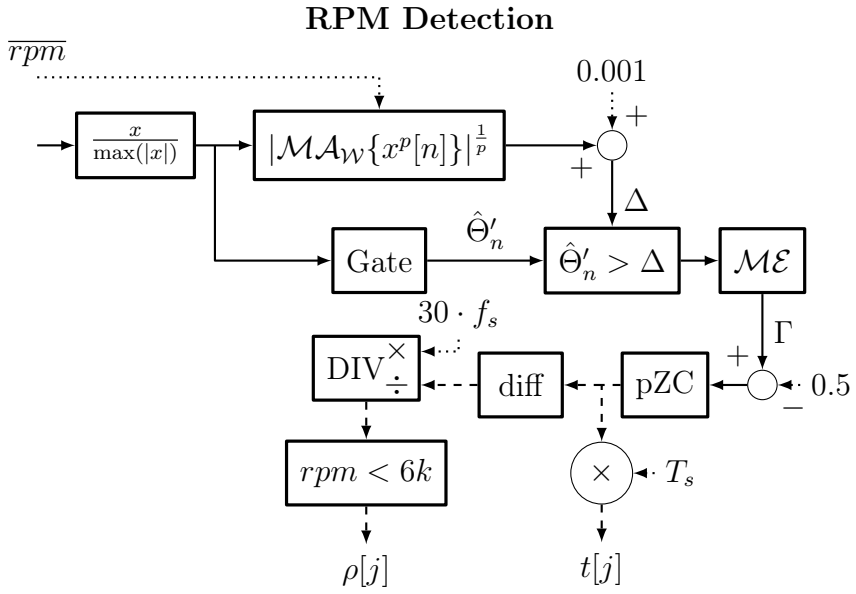


Figure 26 – Temporal RPM detection through onset detection on the ignition-based envelope with the Adaptive Threshold Algorithm. The block pZC finds the locations of the positive zero-crossings.

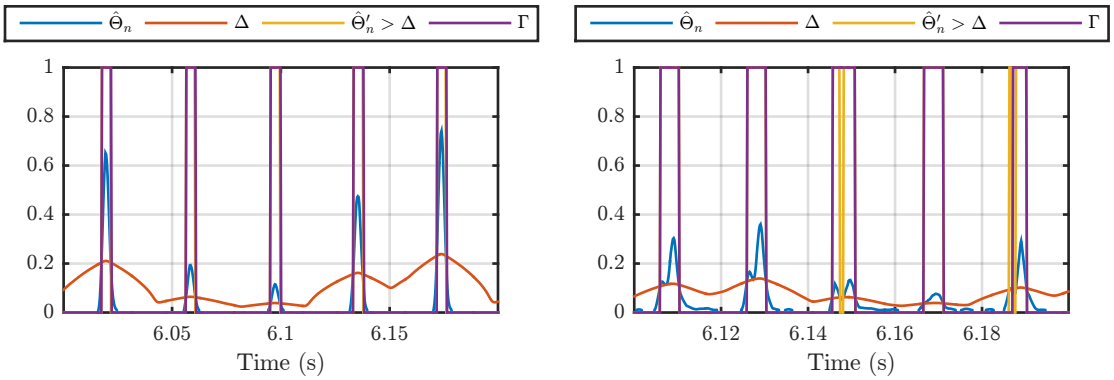
value of 1, denoted as $\hat{\Theta}_n[m]$. The adaptive threshold is defined as

$$\Delta[m] = 0.001 + \left|\mathcal{MA}_{\mathcal{W}}\left\{\hat{\Theta}_n[m]^p, N_{\mathrm{MA}}\right\}\right|^{\frac{1}{p}}. \tag{19}$$

The power p gives the ability to control how much the adaptive threshold is influenced by high values. The first addend of 0.001 ensures a higher adaptive threshold for very low envelope values. Additionally, all values of the normalized envelope $\hat{\Theta}_n[m]$ below 0.01 are set to 0, denoted as $\hat{\Theta}'_n[m]$ (gate function). A location of an ignition is then found, where $\hat{\Theta}'_n[m] > \Delta[m]$. As $\hat{\Theta}'_n[m]$ can fluctuate around the threshold, thus making a distinct detection difficult, a moving median filter is used to only identify these areas as ignitions, where the threshold is definitely exceeded. The filter length is set to 15 % of the expected time between two ignitions ($L_{\mathrm{med}} = \frac{60}{2rpm} \cdot 0.15$):

$$\Gamma[m] = \mathcal{ME}\{\hat{\Theta}'_n[m] > \Delta[m], L_{\mathrm{med}}\}. \tag{20}$$

Finally, the locations of the positive flanks of the boolean signal $\Gamma[m]$ are used as the time locations $t[j]$ in equation (18). As a sanity check, all rpm values above 6000 rpm are discarded by removing the corresponding $t[j]$ values and recalculating the rpm signal $\rho[j]$. A graphic overview of this algorithm is presented in figure 26.



(a) Base 1 idle.

(b) Base 1 1500 rpm @ 50 % load.

Figure 27 – Detection of the ignitions with an adaptive threshold and median filtering. For an idle run, the ignitions form very distinct peaks and no median filtering is necessary (left figure). Nevertheless, for higher revolution speeds and load on the engine, median filtering can improve the detection (right figure).

Parameter	Description
MA factor	Length of the weighted moving average filter of equation (19) as a factor of the length of approximate one round, i.d. if the given value is 1, N_{MA} is as long as one round.
p	Power of equation (19).
marks median	The median filtering defined in equation (20) can be used optionally.

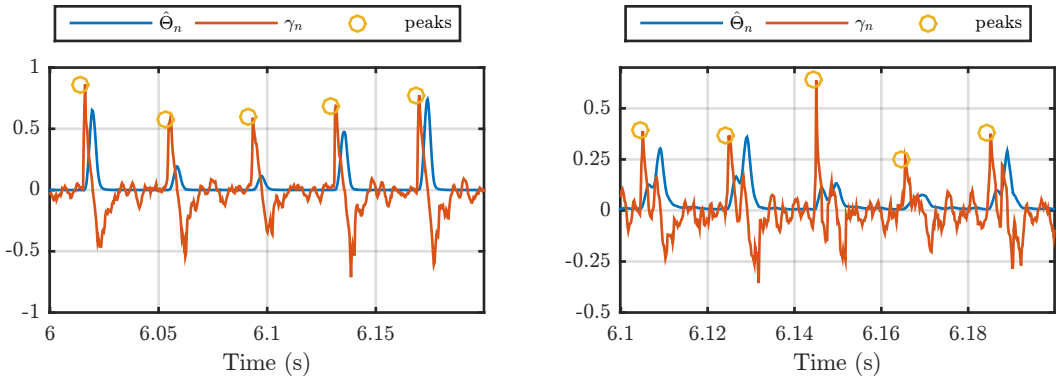
Table 4 – Parameters of the RPM Detection Adaptive Threshold algorithm.

All relevant signals for the detection are shown in figure 27. The adaptive threshold Δ ensures a robust detection for the distinct peaks for the idle run in the left figure. For higher revolutions with more load on the engine, the ignition-based envelope still shows the ignitions, but the shape is distorted. Nevertheless, the ignitions can be identified and median filtering equalizes small errors by performing a majority vote in each sliding window to decide, if an ignition is present or not.

In the implementation of the algorithm, three parameters can be adjusted, which are shown in table 4. Note, that as for the ignition-based algorithm, the average rpm of the signal must be known.

Divided Gradient Algorithm

A second algorithm, which detects high ascents for small values of the averaged envelope, was developed. As the Adaptive Threshold Algorithm may detect four



(a) Base 1 idle.

(b) Base 1 1500 rpm @ 50 % load.

Figure 28 – Detection of the ignitions with the Divided Gradient Algorithm.

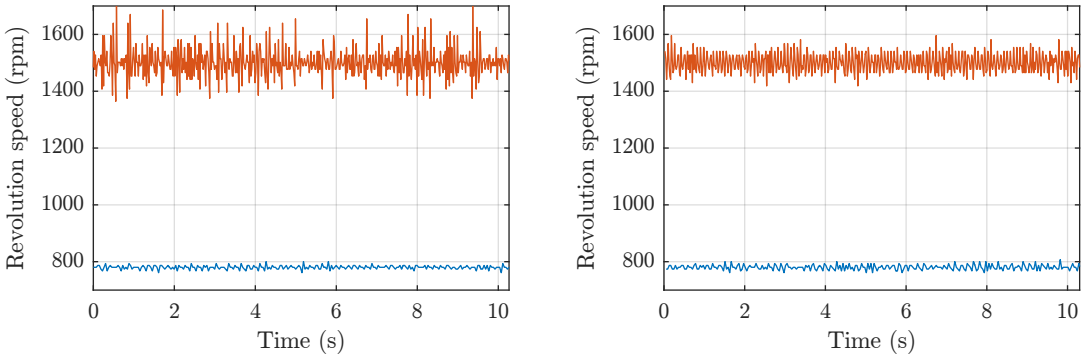
ignitions even if one cylinder is turned off, this algorithm was originally developed to detect only the really happened ignitions. The algorithm defines a detection function $\gamma[m]$ as

$$\gamma[m] = \frac{\mathrm{diff}(\hat{\Theta}[m])}{\hat{\Theta}[m]}, \tag{21}$$

which shows a distinct peak at the beginning of an ignition. The peaks are now detected with the powerful MATLAB function *findpeaks*. As a fixed threshold would not work for different revolution speeds and loads on the engine, the detection function is normalized ($\gamma_n[m]$) to a maximum value of 1 and the root-mean-square of $\gamma[m]$ is used as *Minimum Peak Prominence* for the function *findpeaks*. Instead of the absolute height, the peak prominence measures the height of the peak to the higher of the two neighboring valleys in the intervals to the neighboring peaks. Therefore, a peak may have a lower absolute value, but is more prominent because of its surroundings [16].

The scaling with the rms-value shows a robust ignition detection, even for detection functions with various peaks, which do not correspond to an ignition, as shown in figure 28. Even though for idle a simple thresholding would suffice for the detection function in the left figure, the right figure shows the robust detection with the peak prominence for signals with less prominent peaks. As sanity check, also the input parameter *Minimum Peak Distance* of the *findpeaks* function is also used, which is set to 20 % of the average time for one round.

One particular advantage of this algorithm is, that it has no external parameters which may influence the performance. But as with all algorithms presented in this



(a) Adaptive Threshold Algorithm.

(b) Divided Gradient Algorithm.

Figure 29 – Detection of the revolution speed from the ignition-based envelope by the two developed algorithms for idle (blue line) and 1500 rpm with 50 % load (orange line).

chapter so far, the average revolution speed must be known.

Figure 29 shows the difference of the rpm-detection of the two algorithms for idle and a run at 1500 rpm with 50 % load for base condition 1. Excerpts have already been shown in figure 27 and 28. As one can easily see, the algorithms are more challenged at higher revolution speeds and the performance is dependent on the parameters of the rpm detection as well as the ignition-based envelope calculation. Therefore, the next section will focus on the optimization of the parameter set.

3.4 Parameter Optimization

With the algorithms from the last two sections it is possible to calculate the revolution speed between two ignitions given a microphone signal and the average revolution speed. The revolution speed between two ignitions may not be the main interest in this thesis, but will give a good measure of the time precision of the ignition detection and therefore the cycle alignment.

If the algorithms presented in section 3.2 and 3.3 are combined, up to ten adjustable parameters influence the performance, if we count the selection between the two RPM-detection algorithms as extra parameter. Additionally, some parameters depend on each other. As the performance of one specific parameter set also varies for different operations and conditions, it is not possible to find an optimal parameter set by hand. Therefore, a stochastic optimization was approached, also known as Monte-Carlo Method [22]. First, useful ranges and sampling points for each range

are defined for each parameter in the *parameter pool*. Then, a certain amount of parameter sets is randomly chosen and the locations of the ignitions for each parameter set are calculated. Finally, depending on performance measures, the best parameter set is chosen, which should be close to the optimal solution, if enough parameter combinations are calculated.

As the calculation of the spectrogram and median filtering in the calculation of the ignition-based envelope $\Theta[m]$ takes more time than the rest, a mixture of Monte-Carlo and a Brute-Force search was used. Therefore, the whole pool of the length of the moving average filter to calculate $\hat{\Theta}[m]$ as well as all parameters for the RPM detection is calculated for one envelope.

Section 3.4.1 will discuss the parameter space and the influence of some parameters. Then, suitable performance measures are presented in section 3.4.2. As these algorithms try to measure the timing between two ignitions, they are very sensitive to a missing one. But this sensitivity can be exploited to define a robust feature for missing cylinder detection, which is also used as a performance measure. For the optimization the influence of various conditions on the algorithm is neglected on purpose, except for a missing cylinder. As these algorithms are pre-processing steps for some of the features in chapter 4, less data in the optimization is favorable, so over-fitting to the data set is avoided. Therefore, only the first base condition and the missing cylinder 1 condition is used in the optimization. The result of the optimization and performance for other conditions is presented in section 3.4.3.

3.4.1 Parameter Space

The Monte-Carlo method is used to find a suitable parameter set for the application. Each parameter has its own pool of predefined values, from which the parameter value is randomly chosen for each run of the optimization. The parameter set of the run with the best performance measures is then assumed to be the optimal solution, as long as a significant amount of parameter sets have been calculated. The total number of possible combinations N_{tot} for P parameters is

$$N_{\text{tot}} = \prod_{i=1}^{P} N_i, \tag{22}$$

where N_i is the number of values in the i-th pool. To minimize simulation time it is important to use only necessary values in each pool. As already mentioned, not a

Parameter	Pool	#	Method	Algorithm
N	$\{64, 128, 196, 256, 384, 512, 768, 1024\}$	8	MC	
overlap	$\{2/3, 3/4, 5/6, 9/10\}$	4	MC	
fft factor	$\{1, 2, 4\}$	3	MC	
perc factor	$\{0.1, 0.3, 0.5, 0.7, 0.9, 1.1\}$	6	MC	
harm factor	$\{0.01, 0.05, 0.1, 0.5, 1, 5, 10\}$	7	MC	
N_{MA}	$\{1, 2, 4, 8, 12, 16, 20, 24, 28, 32, 36\}$	11	BF	
algorithm	{'threshold', 'gradient'}	2	BF	
MA factor	$\{0.05, 0.1, 0.15, 0.2, 0.3, 0.4, 0.6, 0.8, 1, 1.2, 1.5\}$	11	BF	'threshold'
p	$\{1/\sqrt{3}, 1/\sqrt{2}, 1, \sqrt{2}, \sqrt{3}, 2, 3\}$	7	BF	'threshold'
marks median	{'yes', 'no'}	2	BF	'threshold'

Table 5 – Parameter pools for the optimization. Parameters with the Monte-Carlo method (MC) are randomly chosen, whereas all values from the pool are calculated for the Brute-Force (BF) parameters.

pure Monte-Carlo optimization was used, as a significant performance boost could be gained by splitting the optimization in a Monte-Carlo and Brute-Force part. The calculation of the ignition-based envelope $\Theta[m]$ takes more time than the rest of the algorithms, especially if the lengths of the median filters in the transient separation are long.[6] Therefore, if we have calculated $\Theta[m]$, all subsequent calculations are done with all the values from the pools of the respective parameters.

In table 5 all parameters with their respective pools, number of values in the pools and their method are given. The MC parameters are needed to calculate $\Theta[m]$ and are chosen randomly for each run of the optimization. With equation (22) this results in 4032 different MC parameter sets for the ignition envelope. All pool values for the BF parameters are calculated, but the last three parameters in the table are only valid for the Adaptive Threshold algorithm, as the Divided Gradient algorithm does not have any parameters (see section 3.3 for the algorithms). The three parameters for the *threshold* algorithm have 154 combinations, the *gradient* algorithm needs to be computed once. Together with the 11 combinations of the N_{MA} parameter for the smoothing of the envelope, this yields in $(154 + 1) \cdot 11 = 1705$ combinations for the Brute-Force part, which are always calculated for each parameter set of the MC parameters. So the whole parameter pool has about $4032 \cdot 1705 \approx 6.9$ million combinations.

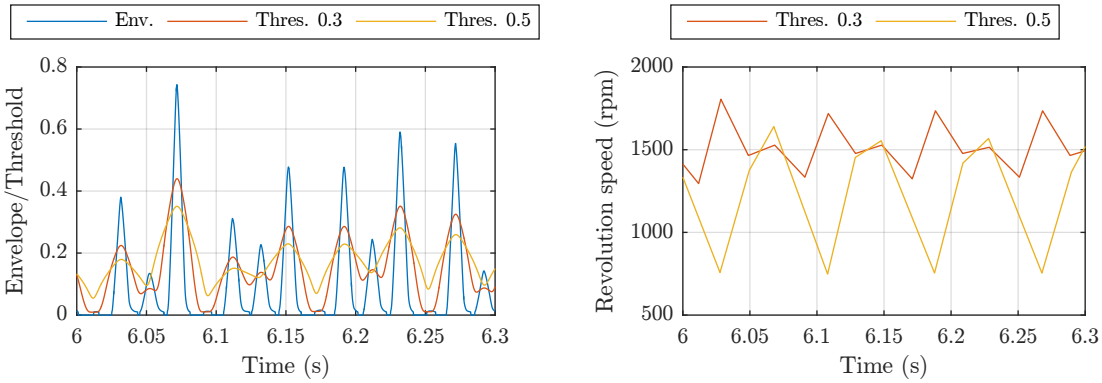
6. This speed difference is even more significant, if the old implementation of the *medfilt1* function in a MATLAB release prior to 2015b is used. The development of the Monte-Carlo optimization started on this old implementation, which led to the split.

N is the length of the time window for the spectrogram, and together with the *overlap* parameter, the samplingrate is defined in regard to the revolution speed of the final envelope, i.d. how many samples are between two ignitions. This is done by scaling N in regard to the revolution speed. The sparsest sampling is for a big window with minimum *overlap* (N = 1024 with $^2/_3$ as *overlap*) with approximate 7.4 samples between two ignitions. The finest sampling with $N = 64$ with $^9/_{10}$ *overlap* has approximately 394 samples between two ignitions. The parameter *harm factor* is the ratio of the length of the moving median filter used for harmonic enhancement L_{harm} (eq. (13)) to the duration of one ignition. A *harm factor* of ten will cancel out the single ignitions, as the filter will return the median over ten ignitions. A very short filter length of 0.01 will only cancel out very short transient events or the median filter may be discarded, if we have less than 100 sample points between the ignitions.

The parameter *perc factor* is the ratio of L_{perc} (eq. (12)) to the total available frequency bins. The number of frequency bins is affected by the *fft factor* parameter and is $(N \cdot fft\_factor)/2 + 1$, as N is always even. Therefore, for a *perc factor* of 1.1 no moving median filter is used, as the median for this time-frame is put in every frequency-bin for this time-frame. Again, if we have a very small *perc factor*, only very sharp harmonic components are filtered. The *fft factor* parameter is added for an additional degree of freedom between time and frequency resolution.

The influence of the parameter N_{MA} is intuitive, as it just smooths the envelope. High values for this parameter will slur the sharp edge of the ignition envelope, whereas small values may not smooth the envelope enough for a reliable detection of ignitions.

The parameter *MA factor* has a great influence on the detection, depending on the ignition of all four or only three cylinders. Figure 30 shows this for a missing cylinder condition for a stationary run with 1500 rpm with 0 % load. The envelope contains three high and one small peak. The small peak is approximately at the location of the fourth cylinder and is probably caused by mechanical excitation of the crankshaft or the pistons. For a *MA factor* of 0.3 this small peak is also detected, as the adaptive threshold falls below the envelope. If the *MA factor* is increased, the adaptive threshold is more stationary and only three ignitions per cycle are detected. In this situation the algorithm might be considered failing regarding the calculation of the rpm. But the outcome of the algorithm for this scenario is perfectly predictable, as the revolution speed will be half, if one ignition is missing. This characteristic



(a) Threshold detection for two different values for the *MA factor*. With a shorter value, the threshold is more adapted to the envelope, thus a small peak is detected at the position of the missing ignition (orange line). A *MA factor* of 0.5 recognizes the missing ignition (yellow line).

(b) If one ignition is missing, the time interval between two ignitions is doubled, resulting in a half rpm value and a characteristic pattern (yellow line). Although the threshold with a *MA factor* of 0.5 finds a fourth peak, its location is inaccurate (orange line).

Figure 30 – Comparison of the influence of different values of the *MA factor* for the no cylinder 1 condition and 1500 rpm with 0 % load.

pattern can be seen in the right figure for the yellow line. Two *correct* revolution speeds are followed by one half as big. As one of the objectives of this thesis is to detect a missing cylinder, a performance measure will be defined in the next section, which will be sensible to this particular pattern.

The effect of the median filter in the Adaptive Threshold Algorithm is tested by doing the calculation with ('yes') and without ('no') the filtering. So, the *marks median* parameter controls, if the median filter is active, and a positive or a negative effect can be observed. The pool for parameter p is created from geometric considerations.

3.4.2 Performance Measures

Considering the characteristic pattern for the missing cylinder, the optimization has two goals to trim the algorithms. First, we want to have a precise localization of the ignitions for performing cycle alignment solely with the microphone signal. Secondly, we want to detect a missing cylinder. Therefore, two performance measures are defined, one for each goal.

Root-Mean-Squared Error A precise localization of the ignitions will have an accurate calculation of the revolution speed $\rho[j]$ after equation (18). As shown in section 2.4.1, the revolution speed can be calculated very accurately from the

cogwheel 1 signal and is used as reference. But, as the cogwheel 1 signal yields only one revolution speed value per round, the values are interpolated at the time locations $t[j]$ of the found ignitions. This signal is referred to as $\rho_{\mathrm{cg1}}[j]$. As strong deviations should be avoided for the cycle alignment, a quadratic measure with the root-mean-square (RMS) is used:

$$\Psi = \frac{\left[\frac{1}{N}\sum_{j=1}^{J}\left(\rho[j]-\rho_{\mathrm{cg1}}[j]\right)^2\right]^{\frac{1}{2}}}{\bar{\rho}_{\mathrm{cg1}}}. \tag{23}$$

$\bar{\rho}_{\mathrm{cg1}}$ is the mean of the revolutions speeds calculated from the cogwheel 1 signal and normalizes Ψ to a relative RMS error, which values can be better compared across different revolution speeds.

Ignition Count Feature The second measure is defined to detect the characteristic pattern of the revolution speed, if one cylinder is missing. If all cylinders are combusting, the revolution speed values for each ignition $\rho[j]$ should be close to the average revolution speed $\overline{rpm}$ of the signal. This value must be known anyway, as the Ignition Envelope and also the Temporal RPM Detection algorithms need an average revolution speed for parameter scaling. It is easily derived, that if one cylinder is missing, $\rho[j]$ only consists of three values per cycle, whereas two are close to $\overline{rpm}$ and one should be close to $\overline{rpm}/2$, as was already shown in figure 30(b). Therefore, to check for the ignition of 3 or 4 cylinders in a signal block, the number of values of $\rho[j]$, which are close to $\overline{rpm}$, is counted, as well as the number of values, which are close to $\overline{rpm}/2$. In each case a tolerance band of $\pm 20\,\%$ is used. Afterwards, the two values are each divided by the total number of samples N of $\rho[j]$, see left part of figure 31. These two values can be seen as coordinates $[x_1, x_2]$ and are now transformed to create one single feature, which will be 1, if all cylinders are combusting, and 0, if only three cylinders are combusting. Figure 32 shows the derivation of this coordinate transform. If only three cylinders are combusting, then x_1 will have a value of $2/3$ and x_2 a value of $1/3$ (point A). On the other hand, if all cylinders work fine, all values of $\rho[j]$ will be close to $\overline{rpm}$ and x_1 will be 1 and x_2 zero (point B). Therefore, we need a shift of the coordinate system to a new origin at $(2/3,\ 1/3)$, a rotation by $-45\,^\circ$ and a scaling by $3/\sqrt{2}$ [30]:

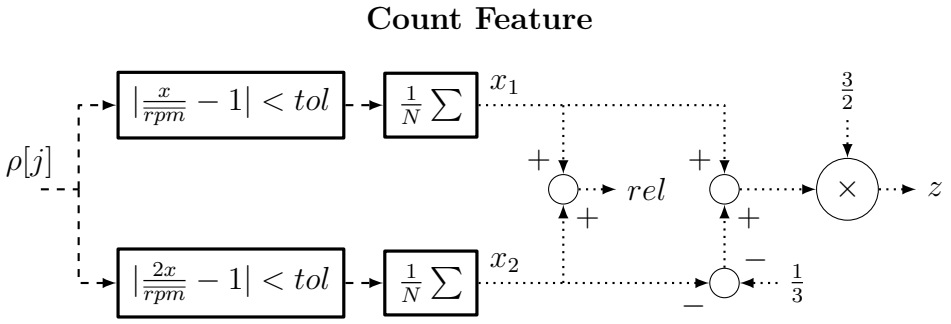


Figure 31 – Calculation of the Count feature with counting rpm-values in tolerance bands and the coordinate transform for z.

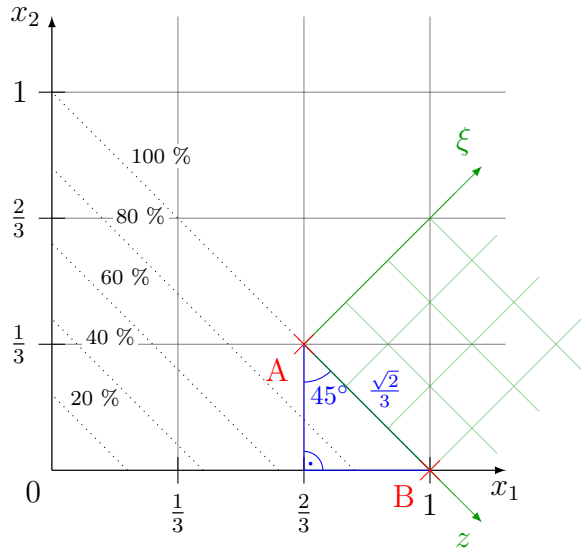


Figure 32 – Derivation of the coordinate transform of the count feature as well as the reliability boundaries (dotted lines).

$$\begin{bmatrix} z \\ \xi \end{bmatrix} = \begin{bmatrix} \cos(-\frac{\pi}{4}) & \sin(-\frac{\pi}{4}) \\ -\sin(-\frac{\pi}{4}) & \cos(-\frac{\pi}{4}) \end{bmatrix} \begin{bmatrix} x_1 - 2/3 \\ x_2 - 1/3 \end{bmatrix} \frac{3}{\sqrt{2}}$$
$$z = \left(x_1 - x_2 - \frac{1}{3} \right) \frac{3}{2}. \tag{24}$$

The first coordinate z of the new coordinate system is by design a feature for detecting a missing cylinder (condition ids 9-12) and is called the *Ignition Count feature*. As we want to find an additional optimal parameter set for this task, the feature is also used as performance measure in the optimization. The performance of the feature is later evaluated in section 3.6 with the optimized parameter set, which was

found in the Monte-Carlo simulation. Figure 31 also shows the variable *rel*, which is the reliability of the detection and corresponds to the dashed lines in figure 32. This sum of x_1 and x_2 defines how many percent of the ignitions were in either of the two tolerance bands and is ideally 100 %. Therefore, the reliability of z can also be assessed.

3.4.3 Optimization Results

The Monte-Carlo simulation was stopped after 4.272.730 parameter sets were calculated, which is approximately 62 % of the parameter pool. The first goal of the optimization is a precise localization of the ignition. Because the localization should be independent from the load on the engine and the revolution speed in a certain range, each parameter set was calculated for the first seven stationary operations, i.d. idle, 1500 rpm and 2000 rpm with 0, 50 and 100 % load, for base condition 1. The normalized RPM RMS-Error Ψ is calculated. In a first reduction, only parameter sets with $\Psi < 0.03$ for all calculated operations are considered valid. This reduces the pool of parameter sets to 932 combinations. Although all these parameter sets may produce good results, the parameter set with the lowest sum of Ψ over the seven operations is considered optimal. After this metric the best parameter set of the 932 combinations scores with 0.15, the worst with 0.1858, which is a difference of approximately -1.9 dB. Nevertheless, the optimal set for the localization of the ignitions is shown in the middle column in table 6.

The parameter *marks median* has the smallest effect, as in the ordered parameter sets of the sum of Ψ, parameter sets, which only differ in the *marks median* parameter, are next to each other. The first ten parameter sets have all the same parameter N, *perc factor* and *harm factor*. The *overlap*, N_{MA} and *MA factor* parameters show minor variations, which is not surprising, as they are all closely related to the amount of smoothing applied in the time domain. Of special interest are the lengths of the median filters for the transient separation, which are at the edges of the respective parameter pools. In practice, these values correspond to an effective $L_{\mathrm{harm}} = 2$ and $L_{\mathrm{perc}} = 283$. Therefore, the moving median for the harmonic enhancement is at its minimum length, whereas for the transient enhancement a static median can be used, as we only have 257 (positive) frequency bins. No parameter set of the 932 combinations uses the Divided Gradient Algorithm. All values from the *overlap* and *fft factor* pools occur in the 932 combinations. For the parameter p only values above $\sqrt{2}$ produce good results.

Parameter	Ignition Locations	Ignition Count
N	128	1024
overlap	5/6	3/4
fft factor	4	1
perc factor	1.1	0.3
harm factor	0.01	10
N_{MA}	36	1
algorithm	'threshold'	'threshold'
MA factor	0.3	0.6
p	2	2
marks median	'yes'	'yes'

Table 6 – Optimized parameter sets for the two different optimization tasks.

The performance of the optimized parameter set is now evaluated on all conditions and microphones for the four basic operations in figure 33. For idle, only the missing cylinder conditions show a deviation above $-35\,\mathrm{dB}$. For all other conditions the calculation of the revolution speed from the ignition locations works for all microphones. For higher revolution speeds and more load on the engine, the impulse sound caused by the ignition is less dominant, which results in a poorer performance, especially on other microphones than microphone 1, which was in the optimization.

Although the parameter set for the ignition locations also indicates the missing cylinder, the second performance measure is used for the optimization of ignition counting. For this measure, also the idle and 1500 rpm with 0 % load operations for missing cylinder 1 were included in the Monte-Carlo optimization. To reduce the parameter sets, the following limits were defined: The deviation of the ignition count feature z must be less than 0.01 for base condition 1 and 0.05 for the missing cylinder. All found ignitions must be inside the tolerance bands ($rel = 1$) and Ψ must be smaller than 0.05. More than 600.000 parameter sets fulfill these rather strict requirements. The best 100.000 sets show no deviation to the optimal values for z. The finally chosen parameter set in table 6 was double checked by hand for speed and a robust detection by choosing a parameter set with a long *MA factor*, compare figure 30(a). A more detailed evaluation is presented in section 3.6, where the whole detection of a missing cylinder with only the microphone signal is presented.

Although the algorithms can now be tuned to the optimal parameter set for the respective task, the last remaining external input to the algorithms, beside the microphone signal, is superseded in the next section, before the ignition detection and the microphone-based cycle alignment are evaluated.

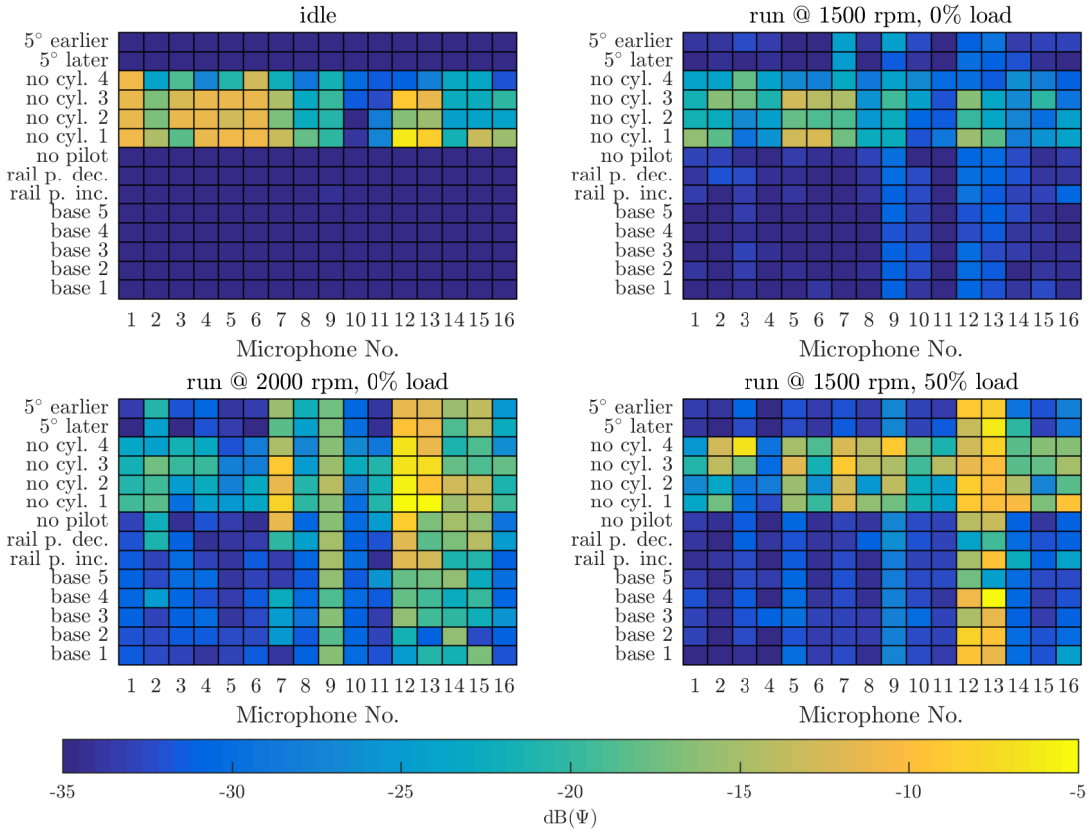


Figure 33 – RPM-RMS Error in decibel for all conditions and microphones with the optimized parameter set for ignition localization. If one cylinder is missing, a higher deviation is expected.

3.5 Average RPM Detection

The algorithms for the Ignition-based Envelope and the Temporal RPM Detection need the average revolution speed $\overline{rpm}$ as additional input to scale the parameters with the revolution speed. The algorithms do not need the exact rpm, as only an integer, which describes how many times the average revolution speed is faster than in idle mode, is used for scaling. Therefore, in a practical application, the algorithm could be set by hand to the current test procedure. The Ignition Count feature also needs the target revolution speed to perform the counting in the tolerance bands. For the correct location of the tolerance bands, a more precise target revolution speed is necessary. Therefore, another rpm detection algorithm was developed, which calculates the average revolution speed in a signal block. With this algorithm, all other algorithms developed in this chapter only need the microphone signal as input.

The algorithm uses the auto-correlation of the input signal, as the four ignitions show a similar waveform. The average time between the peaks is then used to

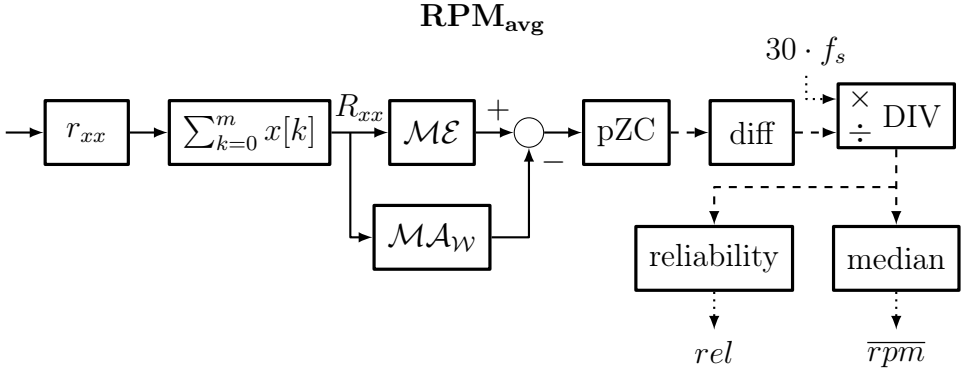


Figure 34 – Block diagram for the average RPM calculation.

calculate $\overline{rpm}$. With some pre-processing, no threshold must be defined for peak-picking. Instead, the simple localization of the positive zero-crossing can be used. The block diagram is shown in figure 34.

The auto-correlation is calculated using the unbiased cross-correlation

$$r_{xx}[m] = \begin{cases} \frac{1}{N-|m|} \sum_{n=0}^{N-m-1} x[n+m]x^*[n] & m \geq 0 \\ r_{xx}^*[-m] & m < 0. \end{cases} \tag{25}$$

The auto-correlation is shown as a purple line in figure 35. In the algorithm only the half for positive m is used. Also, at the most $^3/_4$ of the auto-correlation or 0.75 s are used, so the auto-correlation shows distinct peaks. Therefore, the algorithm to calculate the average revolution speed $\overline{rpm}$ works best, if it is applied to a single short block of 1 second length of the microphone signal for a stationary run.

The shape and height of the peaks differ for different revolution speeds and loads on the engine, therefore peak-picking would have introduced new parameters. As the derivative of a function is the slope, the integral of a function must have the steepest slope, if the function itself has a peak. In the time-discrete domain the derivative is the diff-function (eq. (4)), the integral can be expressed by the cumulative sum. For the auto-correlation this can be defined with

$$R_{xx}[m] = \sum_{k=0}^{m} r_{xx}[k]. \tag{26}$$

The cumulative sum is calculated for the whole signal block. In electronic circuits, an integrator is also considered a low-pass filter, which amplifies slow changes compared to high frequency components. Nevertheless, high-frequency fluctuations, which may

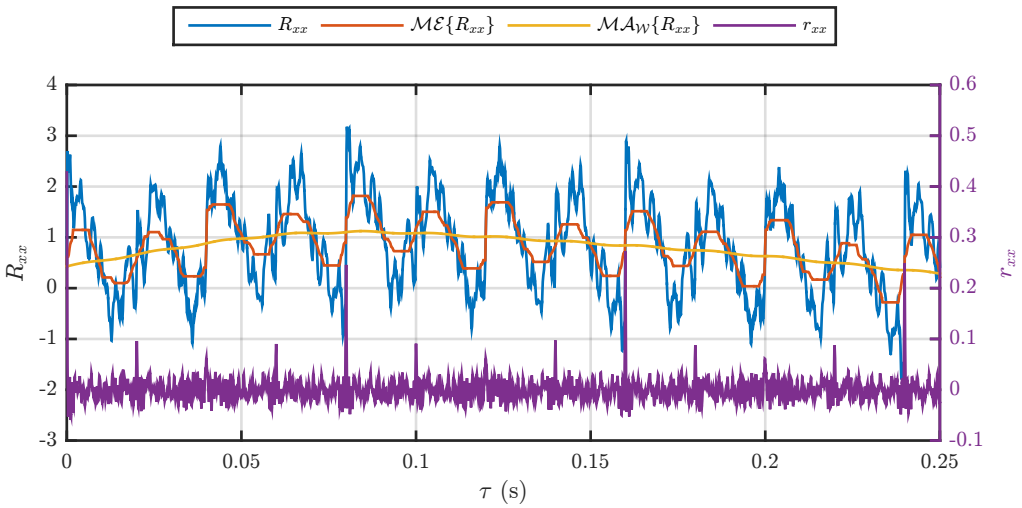


Figure 35 – Auto-correlation r_{xx}, integrated auto-correlation R_{xx} and the filtered versions for R_{xx} for base condition 1 with 1500 rpm with 100 % load.

disturb a reliable detection, are still visible in the blue curve in figure 35. Therefore, a moving median filter is applied. The length of this moving median filter is 819 samples, which is the quarter of the time between two ignitions at an absolute minimum revolution speed of 600 rpm. The disturbing low frequency components are canceled by subtracting the output of the moving median filter with the output of a weighted moving average filter with a filter length of 3277 samples, which is the time between two ignitions at 600 rpm. Afterwards, simply the locations of the positive zero-crossings must be calculated, as they correspond to the positive crossing for the orange curve ($\mathcal{ME}$) of the yellow curve ($\mathcal{MA}_W$) in figure 35. The difference between the locations of the positive zero-crossings is actually one more the time between the different ignitions. As we are interested in the average revolution speed, only the median of the revolution speeds is used as $\overline{rpm}$ for the scaling of the parameters. In the block diagram in figure 34, also a reliability value *rel* is given as well. One possible calculation is to check, if the 25 and 75 % percentiles are more than 15 % apart from the median value.

Again, the performance of this algorithm is evaluated at four basic operations in figure 36. For all microphones and all conditions the average revolution speed $\overline{rpm}$ is calculated for a time block of one second duration. The relative error is the difference between the calculated $\overline{rpm}$ and the actual revolution speed of the measurement, divided by the the actual revolution speed. Surprisingly, the algorithm works best for 1500 rpm with 0 % load, where it works for all microphones and all conditions. The

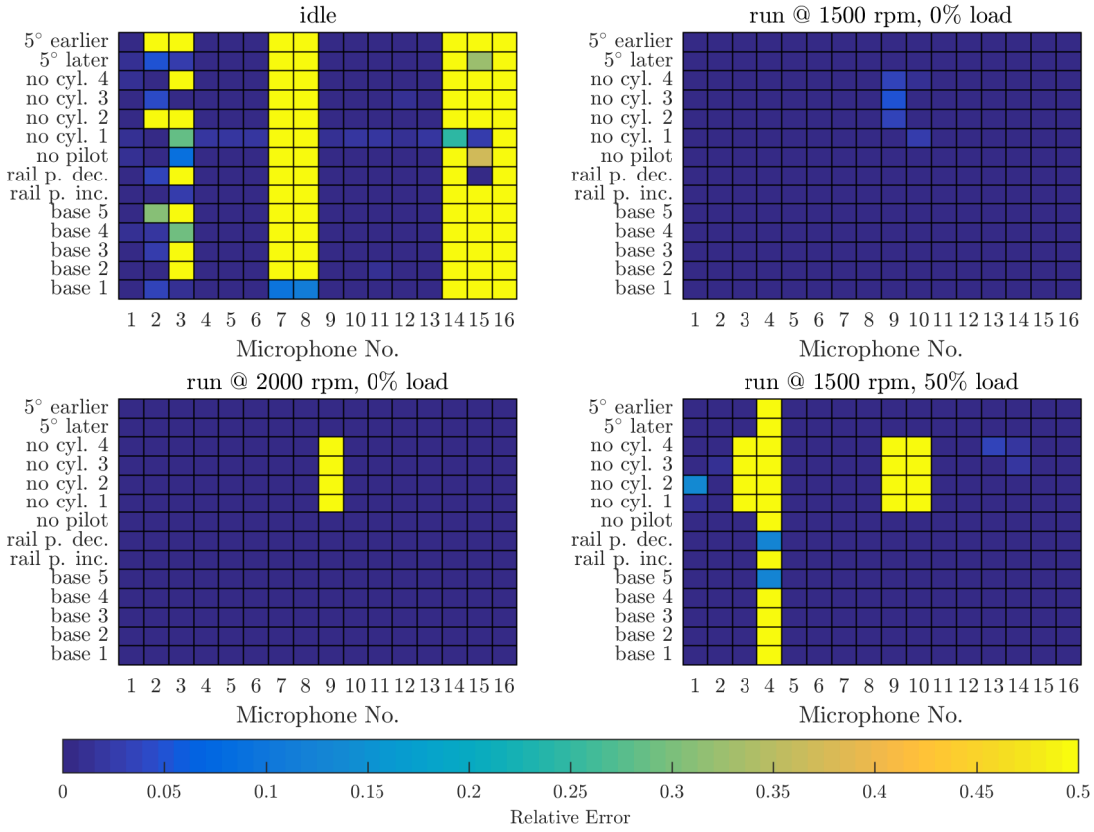


Figure 36 – Relative error for the $\mathrm{RPM}_{\mathrm{AVG}}$ algorithm. If the relative error is above 0.5, the Ignition-based Envelope Algorithm will have incorrect scaling. The Ignition Count feature needs an even more reliable estimation.

second aspect is remarkable, as for a missing cylinder the correct revolution speed is also detected. Therefore, the algorithms needed for the Ignition Count feature will get the correct scaling. For other operations not all microphones perform equally well, especially in the case of the missing cylinder conditions. Nevertheless, the algorithm produces a reliable detection for microphone 1 in these operations. If the deviations in figure 36 are compared to the reliability in figure 37, the validity of the *rel* measure with the deviation of the 25 and 75 percentiles to the median can be assessed.

3.6 Ignition Check

The average revolution speed $\overline{rpm}$ is necessary for the scaling of the parameters for the Ignition-based Envelope algorithm, the Temporal RPM Detection algorithm and defines the tolerance bands in the calculation of the Ignition Count feature, which was introduced in section 3.4.2. With the algorithm of the last section, it is

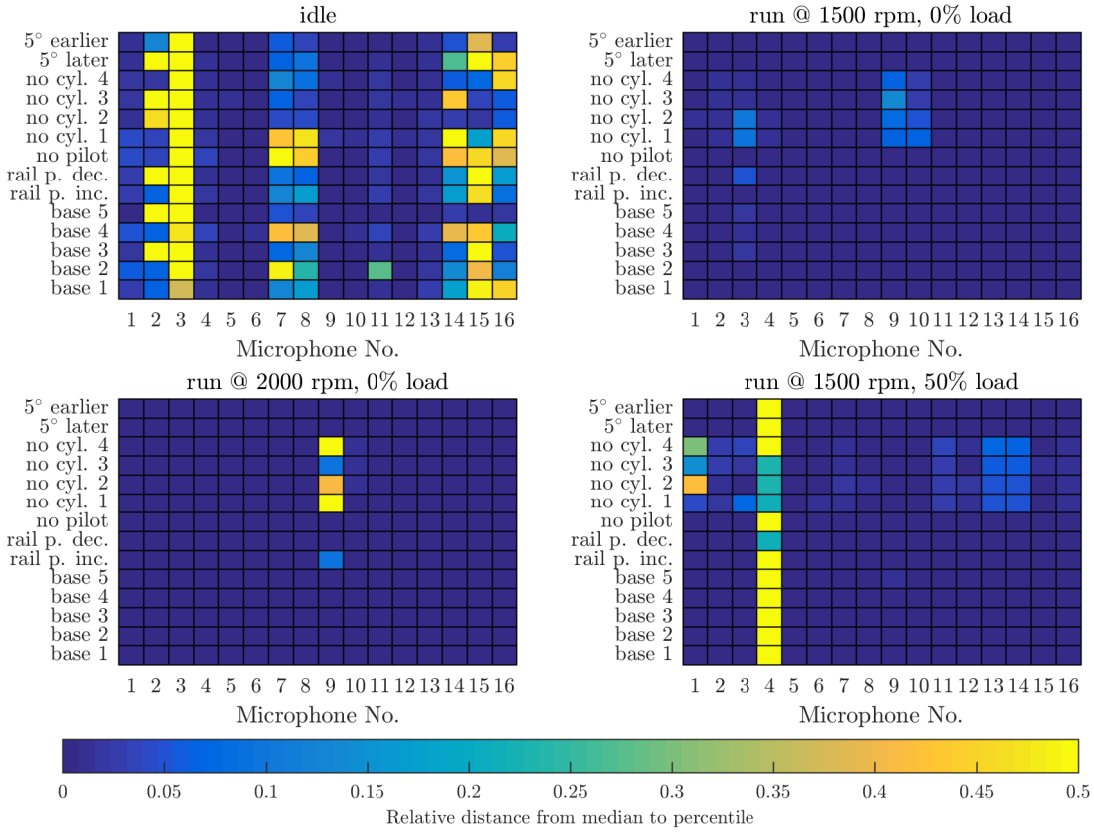


Figure 37 – The relative distance between the larger of the 25 and 75-percentiles to the median is used as measure for the reliability for the $\mathrm{RPM_{AVG}}$ algorithm. Therefore, if the percentiles are close to the median, the estimation is reliable.

now possible to combine these algorithms to the whole procedure for the Ignition Check, which is shown in figure 38, and where only the microphone signal is needed as input. The desired values of the feature z are 1 for four ignitions and zero, if only three cylinders are combusting. The ignition-based envelope and the rpm detection use the optimized parameter set for this purpose from table 6.

The feature z is now evaluated for all conditions and all operations with a revolution speed below 2500 rpm and 0 %, 50 % and 100 Nm load. These selection includes 5 stationary runs and the only run-up, which was measured for all conditions. Therefore the evaluation is based on 84 measurements, where only 4 of them were in the optimization process, which can be seen as the training of the classification process. Each measurement is sliced into blocks of one second length. For each block the average revolution speed $\overline{rpm}$ and the values z and rel are calculated according to the block diagram in figure 38. Each value-pair is then plotted in the upper left part of figure 39. As one can see, a perfect detection is reached. The theoretical

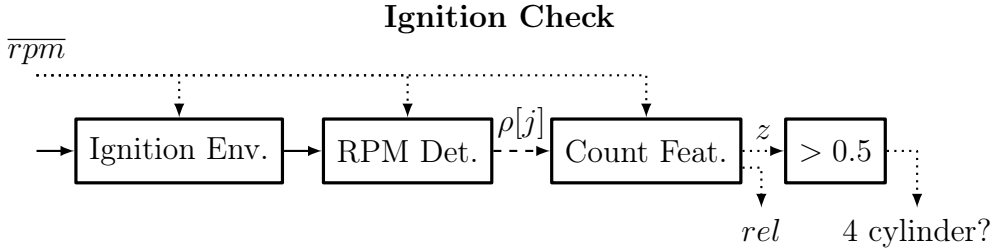


Figure 38 – Block diagram for the Ignition Check algorithm, which tests, if all four cylinders are combusting. The whole detection solely depends on the microphone signal, as $\overline{rpm}$ can also be calculated from the microphone signal.

threshold of 0.5 is also a reasonable value for the most simple classifier of a scalar threshold decision, as was already shown in figure 38 in the most right block with a boolean output. The robustness can be assessed by comparing the inter-quantile range between the 0.005 and 0.995 quantiles for each group to the respective median, because 99 % of the values lie in this interval. The median for the three cylinder group is 0 with an inter-quantile range of 0.423. Especially for lower revolutions with few ignitions in the signal block of one second, some deviation to the perfect value of zero is expected, as the slicing is not synchronized to the duty cycle. Therefore, there might not be exactly two times more ignitions with the full revolution speed than with the half revolution speed. For the four cylinder group the median is 1 with an inter-quantile range of 0.227. Of course, z can never be greater than one, which results in a skewed distribution, see lower left part of figure 39. The dashed lines in the left figures show the neighboring extreme values, which have a distance of 0.317 to each other.

One extra value with $z = 0.04$ and $rel = 0.36$ is outside the limits of the picture in order to make the distribution more visible. A high reliability rel may not correlate to a closer value of z to 1 respectively 0, nevertheless a lower bound of $rel \geq 0.85$ will provide a robust classification.

The correct detection and its reliability is also evaluated for all microphones and compared in figure 40. For microphone 1 the detection works for all shown operations flawlessly. As the engine has a distinct directivity, the other directions (microphones) proof to be more challenging. Nevertheless, for idle, where the ignitions produce the most prominent peaks in the envelope, only the microphones 14 and 15 produce an error and a decreased reliability can be found in the microphone 2, 3 and 7. To show a wrongful detection, the reliability is multiplied by -1 in the figure resulting in the yellow fields. As all parameters in the algorithms were only optimized with the signal

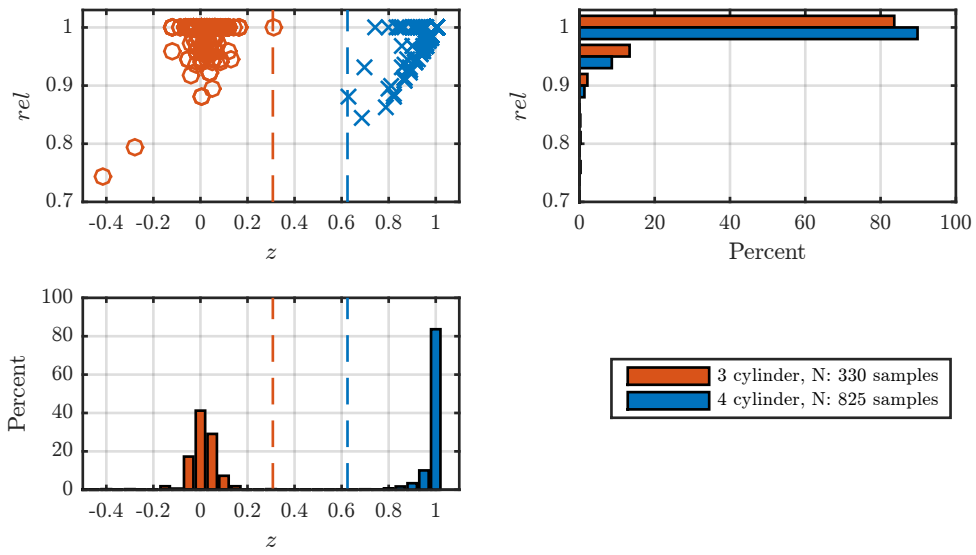


Figure 39 – Evaluation of the Ignition Count feature for all measurements with less than 2500 rpm and maximum 50 % respectively 100 Nm load.

from microphone 1, this nearly omni-directional detection for an idle run shows the robustness. At higher revolution speeds this robustness is lost. Surprisingly, the stationary operation with 1500 rpm with 50 % load shows also a good detection for many microphones. For the other two operations in the figure, the microphones 4, 5 and 6 can be considered the best microphones, although the detection is not flawless. Together with the best performing microphone 1, these are the exact microphones, which showed the most impulse capture ability in the spatial distribution analysis in section 2.3.

In conclusion, the detection of a missing cylinder can be considered solved. The detection works for microphone 1 for different operations, or in idle for 11 different microphone positions. The poorer performance of the other microphones for higher revolution speeds may be explained by the fit of the parameters to microphone 1 in the optimization. Nevertheless, for the following part of the thesis only the conditions with four ignitions are considered. This is especially important in the next section for cycle alignment, as it assumes to align on every fourth ignition.

3.7 Microphone-based Cycle Alignment

Finally, the microphone-based cycle alignment can be performed. The Monte-Carlo simulation yielded the optimized parameter set for the finding of the locations of the

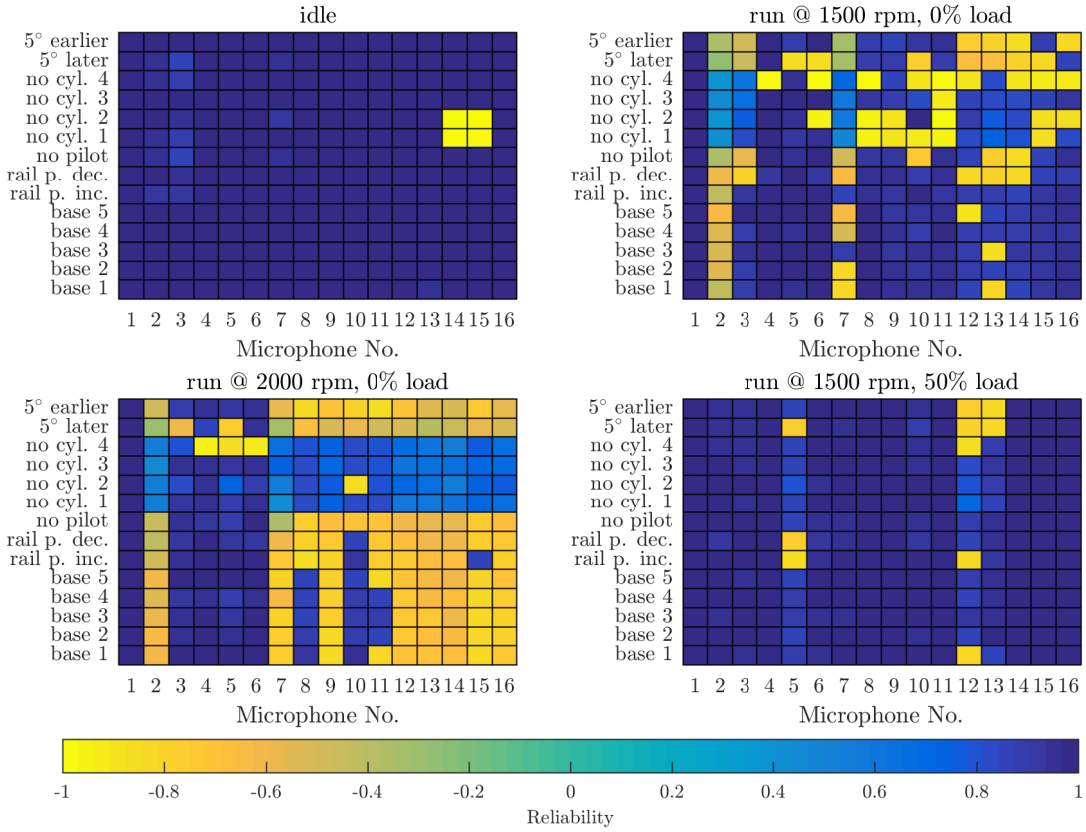


Figure 40 – Reliability for the Ignition Count feature for all microphones. A value of 1 means perfect reliability, whereas zero means, that no detected rpm value could be associated to a tolerance band. If the detection is wrong, than the reliability is multiplied by -1. Therefore, a bright yellow field means, that the algorithm would trust a wrong detection.

single ignitions. Together with the average revolution speed $\overline{rpm}$ and the assurance of four ignitions, all relevant parts for the *Ignition Locations* procedure in figure 41 are present. The locations of the ignitions $t[j]$ in seconds are multiplied by the sampling rate and rounded to samples. Afterwards every fourth value is taken to get the locations in $locs_{720°}$. Therefore, the output is the same format as the *1 Tick Detection* algorithm for the cogwheel 1 signal in section 2.4.1. As a result, the same cycle alignment from figure 10 can be used and all algorithms for the microphone-based cycle alignment procedure are available and shown in figure 42.

In the next chapter envelopes will be used instead of raw microphone signals for the classification of faulty conditions, which doesn't change the alignment procedure itself, hence the optional envelope block in figure 42. Therefore, figure 43 shows the alignment of an ignition-based envelope with the cogwheel 1 signal and the microphone signal itself. 0° is neither dead center, nor will the two algorithms have

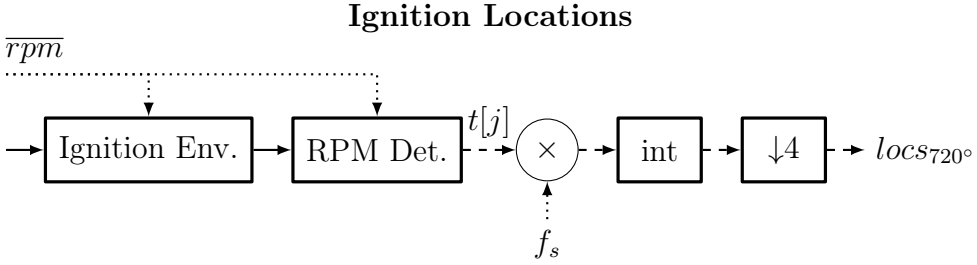


Figure 41 – Calculation of the sample locations of the ignitions for the cycle alignment from a microphone signal.

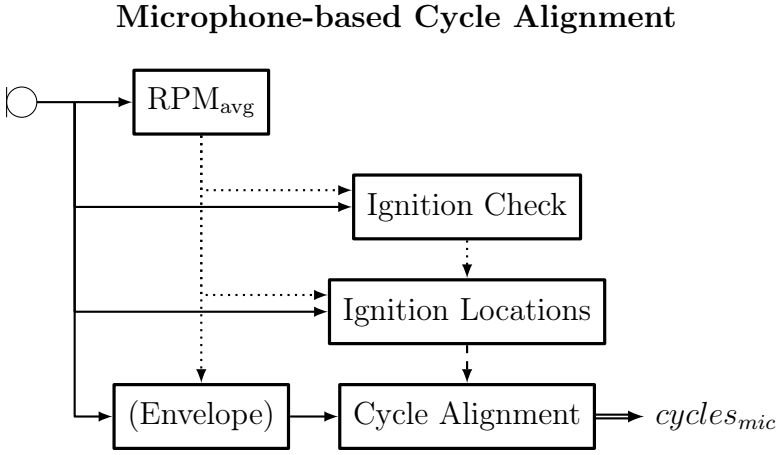
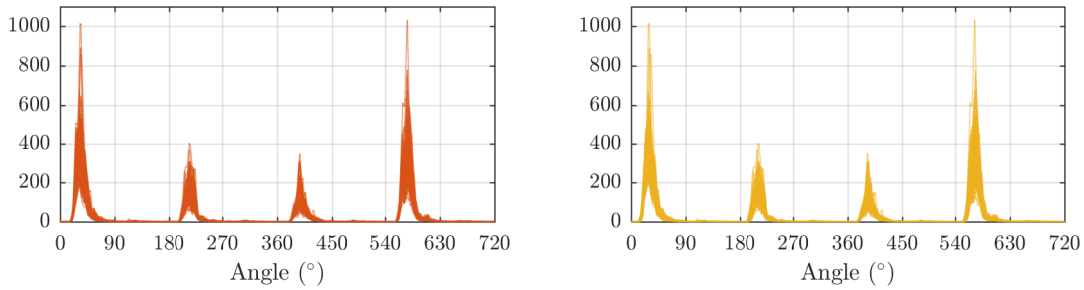


Figure 42 – Microphone-based Cycle Alignment.

the same reference point of the crankshaft angle. For the figure this angle shift has been corrected, so that the same ignition happens at almost the same angle. The envelope is calculated with the Ignition-based Envelope algorithm with the parameter set for the Ignition Locations of table 6. Only the parameter N_{MA} was changed to 3, otherwise the single ignitions would be smoothed to a nearly triangle form.

The variations between the cycles are now subject to differences between the cycles and can not be further improved even with a finer angle alignment. This is shown in figure 44 with a box plot comparison of four alignment methods. The correlation coefficients are calculated according to eq. (6) in section 2.4.3 for the ignition-based envelope described above with microphone-based cycle alignment and cogwheel-based cycle and angle alignment of section 2.4. For a full comparison, a new cycle alignment technique is also shown. The signal (or envelope) is just sliced into frames of the length of two rounds calculated from the average revolution speed. This *rpm-based* cycle alignment is of course the most imprecise, but also the most simple form



(a) Cogwheel-Based. (b) Microphone-Based.

Figure 43 – Cycle Alignment for an ignition-based envelope of all cycles calculated for base condition 1 with 1500 rpm with 0 % load.

of cycle alignment.

The correlation between the cycles is for the envelope in general much higher, and also the differences between angle alignment and cogwheel-based cycle alignment are much less. The alignment with the microphone signal shows in most of the cases no significant difference to the cycle alignment with the cogwheel 1 signal, if no load is on the engine. For the operations with 100 % load at 1500 rpm the correlation is generally less, but the alignment with only the microphone signal is not as exact as the alignment with cogwheel signals. The rpm-based cycle alignment shows a much poorer performance in idle, as the revolution speed fluctuates more. It is obvious that we observe different medians for the five base measurements and also the inter-quartile range is much higher. For higher revolutions this difference is even smaller, especially for 1500 rpm with 0 % load nearly no difference can be seen between the alignment techniques, as the engine control unit apparently stabilizes the revolution speed rather well. Nevertheless, rpm-based cycle alignment holds two disadvantages, which are not covered by the correlation coefficient. First, the cycle alignment is completely sensitive to revolution speed changes. The cycle alignment of a run-up will completely fail and this technique can only be used for stationary runs therefore. Second, no connection to the angle of the crankshaft is made. The cogwheel-based alignment techniques are based on the angle and also for the microphone-based cycle alignment the ignitions will be aligned to a fix angle. This means that, although there is no reference point to the crankshaft angle, the ignitions for two measurements will happen at approximately the same angle, if they are evaluated with the same alignment technique. For rpm-based cycle alignment the angle at which the ignitions occur can not be predicted. Although there might be rare cases where even the rpm-based cycle alignment is sufficient, the certainly

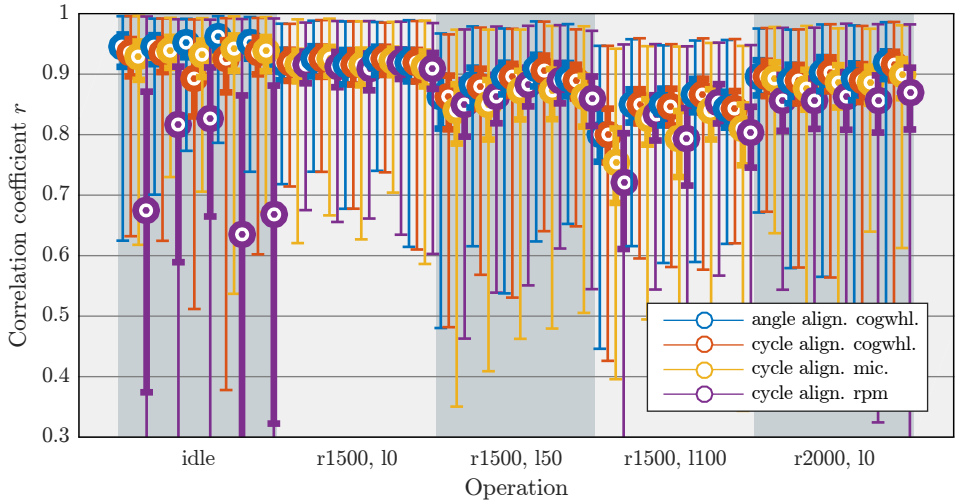


Figure 44 – Box plot of the correlation coefficients with minimum and maximum values (thin lines), the 25 and 75 percentiles (thick lines) and the median value (dots) for the five base conditions and five stationary runs for the ignition-based envelope of microphone signal 1.

more generalized and stable approaches are the microphone-based cycle alignment with the detection of the ignitions and the cogwheel based techniques. As for low revolution speeds microphone-based cycle alignment works quite well in comparison to an alignment based on the cogwheel signals, it can also be used for the calculation of cycle-based features in the next chapter.

3.8 Summary

The main goal of this chapter is to perform cycle alignment with only the microphone signal. During that process various algorithms are developed, which can be exploited to detect a missing cylinder.

The ignition-based envelope enhances reliably the ignitions, which are then located with an onset detection. From these locations two important aspects can be derived. First, the revolution speed between two ignitions can be calculated and used to detect a missing ignition. This already yields in the reliable classification of four of the faulty conditions considered in this thesis. Second, as will be shown in the next chapter, robust features can be defined on a cycle aligned envelope, where the locations of the ignitions are used to perform cycle alignment, where no other signal than the microphone signal is necessary.

Although, the algorithms for both aspects are partly identical, they work best if an optimized parameter set is used for the task at hand, thus resulting in different calculations. Performance considerations for further optimization are out of the scope of this thesis, as for meaningful considerations at least some specifications of the hard- and software system should be known.

4 Features

In a classification task, features, which are later fed into a classifier, are calculated from the source signal. This basic concept is already used in section 3.6, where the *Ignition Count feature* z is calculated from a signal block. This single value can be easily compared to a threshold afterwards. In general a classification doesn't work with only one feature and a single comparison, but an arbitrary number of features F and a more sophisticated classifier in the F-dimensional space spanned by the features.

This chapter focuses on the definition of features, which are able to indicate a faulty condition. The periodicity of an internal combustion engine is used to define features based on the duty cycle. In section 4.1 the so-called *Envelope Shape features* are presented, which describe specific characteristics of a cycle aligned envelope for the classification. Seven of these features are defined in total, but in most cases four of them are enough for classifying all the measured faulty conditions except the missing cylinder, which is already covered by the Ignition Count feature.

This features can all be calculated using any form of cycle or angle alignment already discussed. If the feature is sensitive to the form of the alignment, the differences will be shown. In section 4.2 another feature is defined, which is not based on the microphone signal, but on the cogwheel 122 signal. It analyses the speed droop in a duty cycle of the crankshaft and is also able to detect a missing cylinder combustion with much less computational complexity and effort. This should be seen as an example that signals coming directly from the engine contain much information as well. Therefore, it doesn't make sense to do cycle or angle alignment with cogwheel signals of a microphone signal to detect a missing cylinder combustion, because the relevant information is already derivable from the cogwheel signal itself.

This chapter focuses on the differences between the conditions and only analyses microphone 1 for the idle and 1500 rpm @ 0 % load operations. The microphone-based cycle alignment works best at these operations and the envelope shape features are also very sensitive to a different revolution speed or load, as the envelope of the microphone signal changes significantly. The features are based only on cycle alignment, i.d. without equi-angular resampling, but for the cogwheel-based and the microphone-based approach. As we have already seen, the difference between the cogwheel-based cycle alignment and angle alignment is marginal for the envelope and, therefore, doesn't affect the classification.

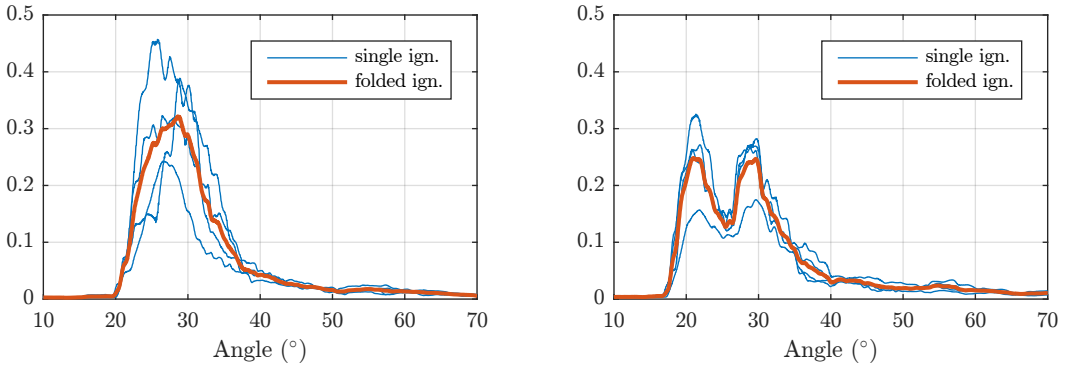
In this chapter only the derivation and the discriminating abilities of the features are shown for the mentioned operations and microphones. The overall performance of the features in the classification task is presented in the next chapter with additional considerations of the robustness.

4.1 Envelope Shape Features

The envelope shape features try to extract as much precise information from a cycle aligned envelope as possible. A RMS-based envelope is used, which is included in MATLAB's *envelope()* function and was already discussed in the beginning of section 3.2. The length of the moving window is set to 20 % of the time between two ignitions. Therefore, either the average revolution speed $\overline{rpm}$ is needed again or the time of the whole cycle, which is already known from the obligatory cycle alignment, is used as reference. The envelope shows the most discriminating properties, if the envelope is not calculated on the broadband signal, but on bandpass filtered signals in the upper frequency range. Sixth-order Butterworth octave-bandpass filters with zero-phase forward and backward filtering, resulting in steep 12-th order filters, are used. The reasons were already discussed in section 3.1. The standard octave bands from 125 Hz to 16 kHz were analyzed as center frequencies, but for the classification the 1, 4, 8 and 16 kHz octave bands are sufficient. Of course, also the ignition-based envelope could be used as envelope function, but despite the increased computational effort, it showed no direct improvement to the bandpass-filtered RMS envelopes for the classification.

Although, the envelopes of the four ignitions may differ between each other and the cycles (see figure 43), a characteristic shape of the ignition can be found for various conditions, if the four half-rounds of a cycle are added together, which can be seen as a fold of the four half-rounds. For the microphone-based cycle alignment, the ignitions in the microphone signal can not be assigned to the single cylinders and even for the angle alignment it is not possible to know, if one round is the first or the second half of the duty cycle, which starts with a certain cylinder. Therefore, it is not useful to define features, which describe the shape of the envelope of the ignition of a certain cylinder. Of course, it is mandatory that all cylinders combust, which we can assume is always the case, since the *Ignition Count feature* is already defined and evaluated. Also the microphone-based cycle alignment needs four ignitions to work properly.

Figure 45 shows the envelope of the single ignitions for one cycle as blue lines



(a) Base condition 1.

(b) 5 ° earlier injection.

Figure 45 – The 16 kHz octave-band filtered envelope is angle aligned and the four half rounds, with one ignition each, of one cycle are shown (blue lines). The folded ignitions (orange line) are divided by four to represent the mean of the single ignitions. The measurements are an idle-run on microphone 1.

for two different conditions as well as the folded envelope (orange line) divided by four for scaling. Although the folding can be seen as loss of information for the single cylinders, there is still enough information for the distinction between the two shown conditions. One of the most noticeable differences is the different shape of the folded envelope, which shows one distinct peak in the base condition and two for the shifted injection timing. These shape differences are the basis for the envelope shape features defined in this section, where not only the shape itself, but the angle position of the shape can be evaluated, although this angle position depends heavily on the used alignment technique. In figure 45 the fine angle-alignment was used, which is based on the cogwheel signal. Therefore, the envelope of the 5 ° earlier injection condition really shows the most energy 5 ° earlier, because the difference between the combustion timing to the crankshaft angle is deterministic. If the alignment is done only with the microphone signal, no information about the crankshaft angle during the ignition is available. Therefore, the angle difference for these two conditions will be lost, as the microphone-based cycle alignment calculates the angle index from the ignitions itself and the most energy of the ignition will always be in a similar region. As the shape of the envelope is not affected and still shows the two peaks for the 5 ° earlier injection, it is still possible to distinguish the two conditions with only the microphone signal.

The whole pre-processing chain for the calculation of the envelope shape features is shown in figure 46. As already discussed, different alignment techniques are combined with different center frequencies for the octave-band filter before the envelope

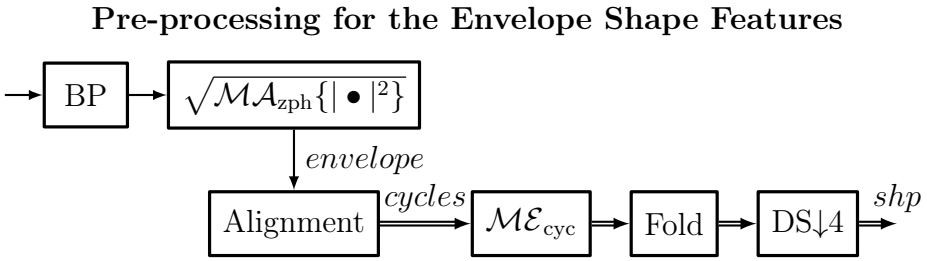


Figure 46 – Complete pre-processing chain for all the envelope shape features with an octave-band filtered RMS envelope.

calculation. For a more stable shape of the envelope, a moving median filter is applied. This filter calculates the median of the envelope over five cycles for each angle. Afterwards the four ignitions are folded into one half-round and downsampled by a factor 4, as the envelopes do not contain high-frequency components. Therefore, the full cycle of 12288 samples for two rounds is reduced to 768 samples for one-half round, thus resulting in an angle resolution of approximately 0.2 °/sample.

In the following subsections, the envelopes for different center frequencies are compared for microphone-based and cogwheel-based cycle alignment for all conditions with four ignitions. In each section the features, which are derived from the respective envelope, are also presented.

4.1.1 Bandpass filtered Envelope @ 16 kHz

As already seen in figure 45, a bandpass filtered envelope at 16 kHz changes its shapes for different conditions. The shape with two distinct peaks actually occurs for all faulty conditions, except the decreased rail pressure, see figure 47. For the cogwheel-based cycle alignment also the shifted injections conditions are each distinguishable from the rest.

For this case, the first envelope shape feature is defined, which will only work for the cogwheel-based cycle alignment. The *Envelope Onset 16k* is the angle of the onset of the ignition. The calculation follows the block diagram in figure 48. A moving median filter ensures a zero-phase noise suppression and retention of the sharp edge of the onset for the pre-processed envelope shape, which is calculated according to figure 46. Afterwards the onset is calculated by subtracting the mean and finding the location of the first positive zero-crossing, which is converted to degrees (°). This yields a single value for each ignition shape per cycle. It is stated again that each of this shapes is calculated as the median of five cycles in the envelope shape

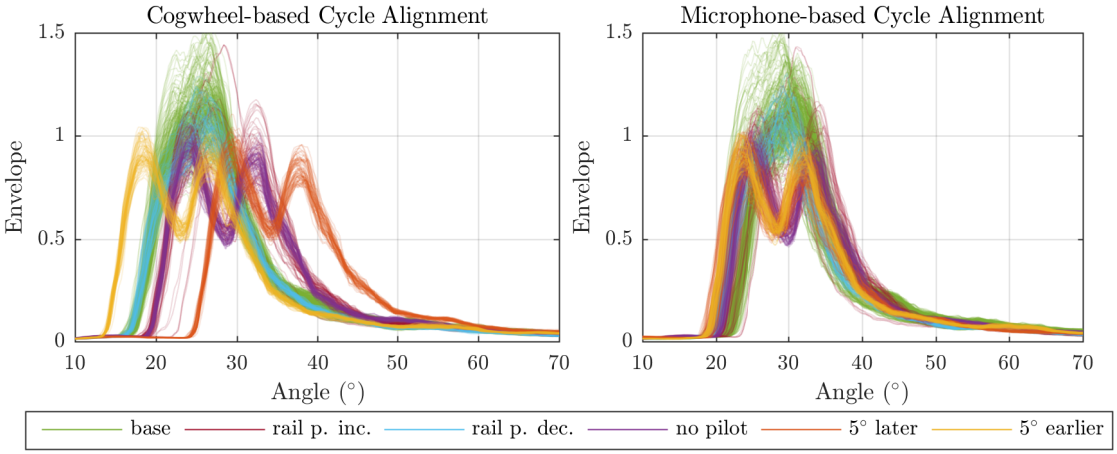


Figure 47 – Folded shape for a 16 kHz octave-band filtered envelope for idle.

Envelope Onset Feature

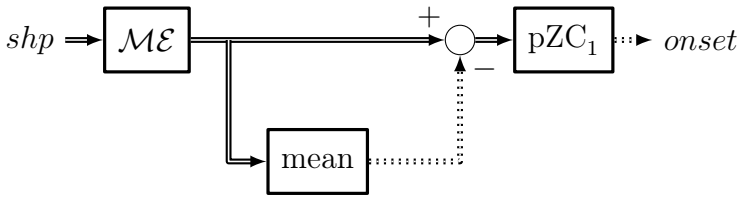
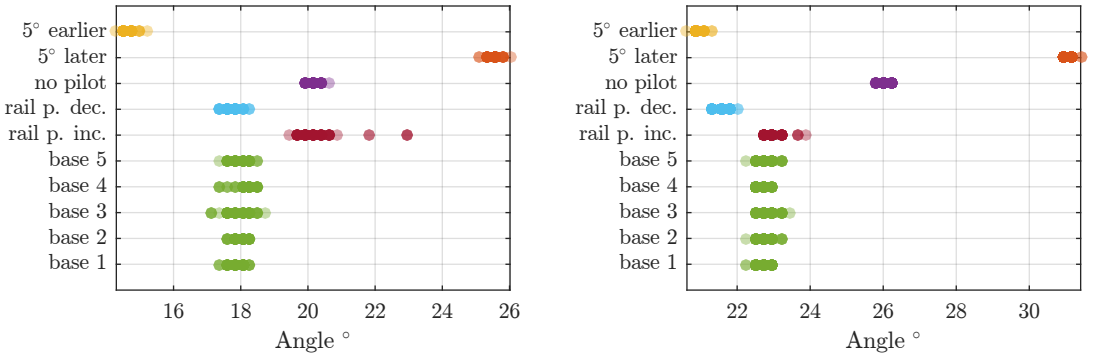


Figure 48 – Calculation of an Envelope Onset feature. A typical length for the moving median filter is 1/8-th of the half-round. The block pZC_1 returns the angle of the first zero-crossing.

pre-processing.

Figure 49 shows the discriminating properties of the Envelope Onset 16k feature. For idle the shifted injected conditions are each individually separable. Also, the no pilot injection and the increased rail pressure conditions can be separated from the rest, although not from each other. The 5 base conditions show the repeatability and can not be distinguished from the decreased rail pressure. At 1500 rpm the difference between the conditions is less, although some classification is still possible. Although we need the cogwheel 1 signal for the cycle alignment, it should be possible to detect four of the remaining five faulty conditions with this single feature.

Nevertheless, without the cogwheel 1 signal and microphone-based cycle alignment there is practically no difference in the Envelope Onset 16k feature for the different conditions and the shape difference must be evaluated. Also, this will provide a better distinction between the no pilot injection and increased rail pressure conditions for the cogwheel-based cycle alignment. The second peak of the envelope for



(a) Idle. (b) 1500 rpm @ 0 % load.

Figure 49 – Values of the *Envelope Onset 16k* feature for two operations and all conditions with cogwheel-based cycle alignment.

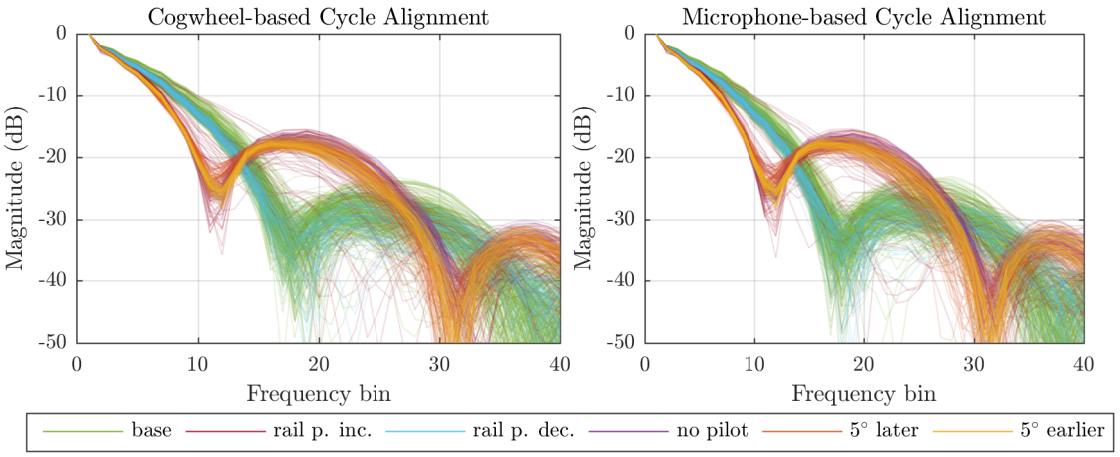


Figure 50 – Magnitude of the frequency spectrum of the folded shape of the 16 kHz octave-filtered envelope for idle.

the respective faulty conditions can be seen as a delayed version, thus resulting in a comb-filter frequency pattern, as shown in figure 50. As the angle-shift is only visible in the phase of the Fourier transform, there is no difference between the two alignment techniques shown in the figure. The feature which exploits the difference between the conditions is called *Spectral Shape 16k* and calculates the sum of the damping in decibels in certain frequency ranges.

The magnitude of the Fourier transform is shifted to have unit gain at 0 Hz, see block diagram in figure 51. Therefore, the feature is not dependent on the absolute values of the envelope and no calibration of the recording system will be necessary, if all features fulfill this requirement. Afterwards, three ranges are defined, which contain characteristic information. For idle the first range with index j reaches from bin 8

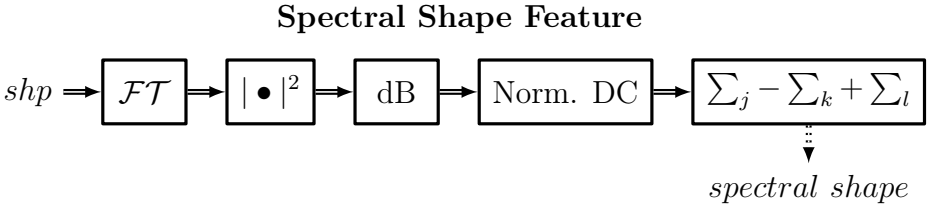


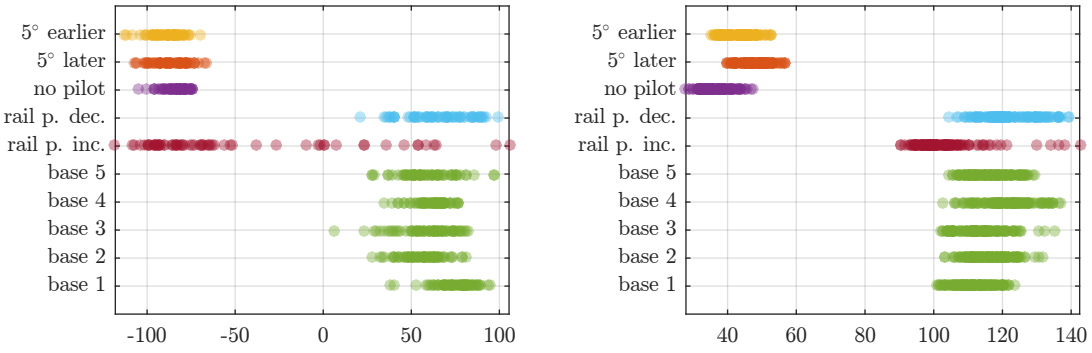
Figure 51 – Spectral Shape feature calculation. The DC component of the Fourier transform is used for normalization. The ranges j, k and l depend on the center frequency of the band-pass filtered envelope and the revolution speed.

to 12, the second range k from bin 16 to 22 and the third range l from frequency bin 29 to 31. The decibel values in each band are summed up and the first and third ranges are added together, whereas the second range is subtracted. Thus, the base conditions and decreased rail pressure condition will have a greater value than the other 4 conditions. Although, the summation of decibel values has no meaningful physical interpretation, it provides a strong discrimination between the conditions.

As for 1500 rpm the shape of the ignition will cover are larger angle range due to the higher rotational speed, the indexes for the ranges must be adapted and shifted to smaller values, as the comb-filter frequency decreases. The discrimination ability of this feature is shown in figure 52. The condition of increased rail pressure stretches over the whole range for idle and will result in wrong classifications. Nevertheless, the angle shifted injection and the no pilot conditions are definitely separated from the other conditions, even with only microphone-based cycle alignment, if the increased rail pressure is ignored for the moment. But to be able to distinguish between them, a cogwheel signal is necessary for the Envelope Onset 16k feature, if only this bandpass-filtered envelope at 16 kHz is used.

4.1.2 Bandpass filtered Envelope @ 8 kHz

The envelope in the octave-band around 8 kHz shows well the difference in the alignment techniques, see figure 53. For the cogwheel-based cycle alignment the ignition is shifted by the different timing of the engine control unit, but an apparently mechanical impulse 90 ° later is always at the same angle position for all conditions. In the microphone-based cycle alignment, the ignition is the reference for the angle and happens therefore always at the approximate same angle. But the angle difference to the mechanical impulse stays naturally the same, therefore the mechanical impulse is shifted in the angle for the last two conditions. Therefore, it is at least possible to



(a) Idle. (b) 1500 rpm @ 0 % load.

Figure 52 – Values of the *Spectral Shape 16k* feature for two operations and all conditions with microphone-based cycle alignment.

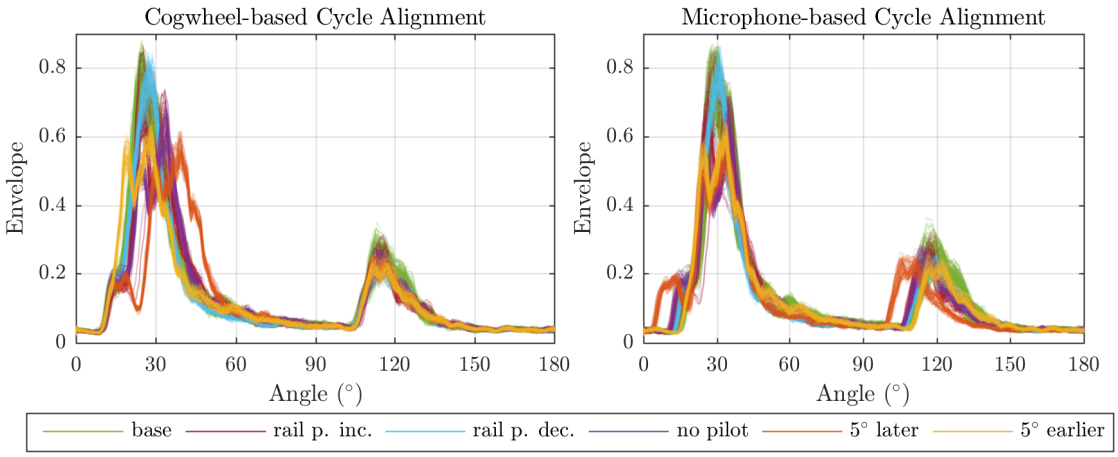
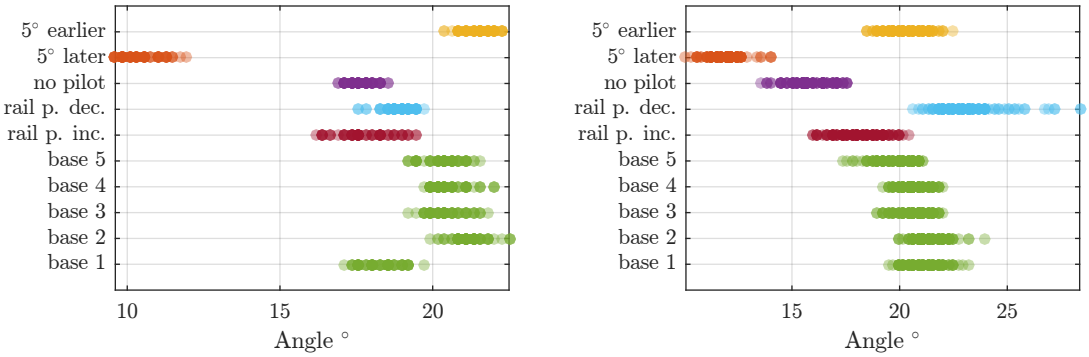


Figure 53 – Folded shape for a 8 kHz octave-band filtered envelope for idle.

distinguish these two from each other, even with microphone-based cycle alignment, and plus, the difference in the timing to the no pilot condition is also significant.

As again an onset is analyzed in the folded shape, another envelope onset feature according to the block diagram in figure 48 is defined. The *Envelope Onset 8k* feature detects the first steep slope in the second quarter-round (90 °−180 °) and will therefore find no difference between the conditions for the cogwheel-based alignment. In case of a microphone-based alignment, a discrimination is possible, as can be seen in figure 54. The Envelope Onset 8k feature compensates the not applicable Envelope Onset 16k feature for microphone-based cycle alignment, thus enabling to not only detect, but to discriminate between the angle shifted injection and the no pilot conditions. Only for 1500 rpm the separation between the 5 ° later and the no pilot conditions may not work flawlessly. The Envelope Onset 8k feature is defined



(a) Idle. (b) 1500 rpm @ 0 % load.

Figure 54 – Values of the *Envelope Onset 8k* feature for two operations and all conditions with microphone-based cycle alignment.

for the onset in the second quarter-round, as here the discrimination works better for both revolution speeds.

4.1.3 Bandpass filtered Envelope @ 4 kHz

The next sub-band can be used to detect and discriminate the conditions with the modified rail pressure. Figure 55 shows the folded envelope shape for an idle run for the two cycle alignment techniques. It is easy to see that in the second quarter-round more energy is present at certain angles for the increased rail pressure conditions than for all other conditions. An increased energy in this frequency region was already visible in the spectrogram in figure 5. For cogwheel-based cycle alignment this difference is even bigger, as the mechanical impulse is aligned to the same angle. If only microphone-based alignment is used, this difference is less. As measure the ratio between the energy from 138 ° to 169 ° and the whole half-round is used and called *Energy Ratio 4k 2*. Again, by using the ratio, the feature is independent from absolute values; the block diagram is shown in figure 56. An additional number is used in the name, because the feature correlates with the energy in the second quarter-round and another energy ratio will be defined in this sub-band for the first quarter-round for the detection of the decreased rail pressure condition.

The Energy Ratio 4k 2 feature effectively discriminates the increased rail pressure conditions from all the other conditions, regardless of revolution speed and alignment technique. An especially remarkable aspect can be found in the right half of figure 57. The decreased rail pressure condition results in a lower value than all

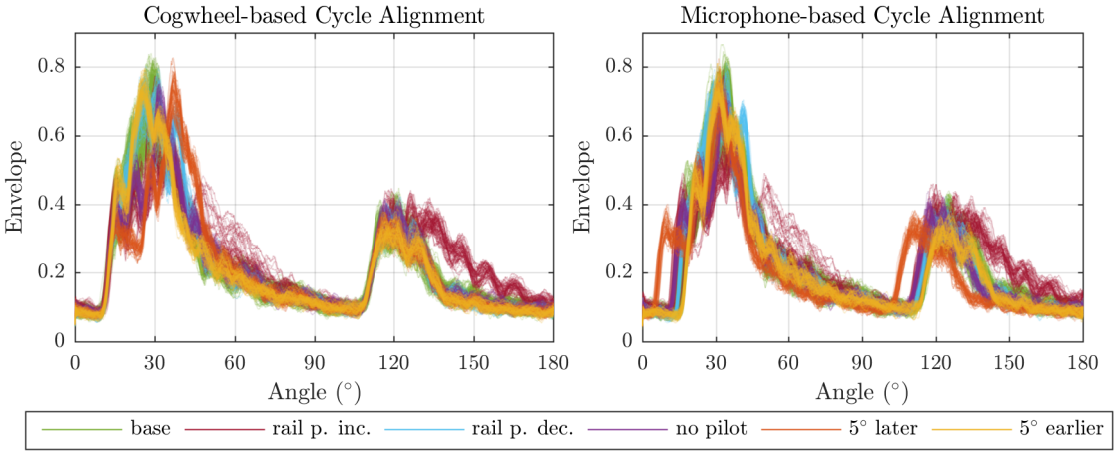


Figure 55 – Folded shape for a 4 kHz octave-band filtered envelope for idle.

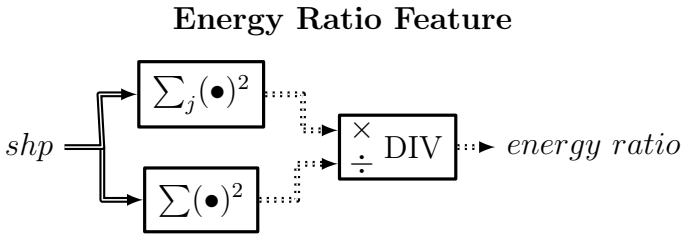
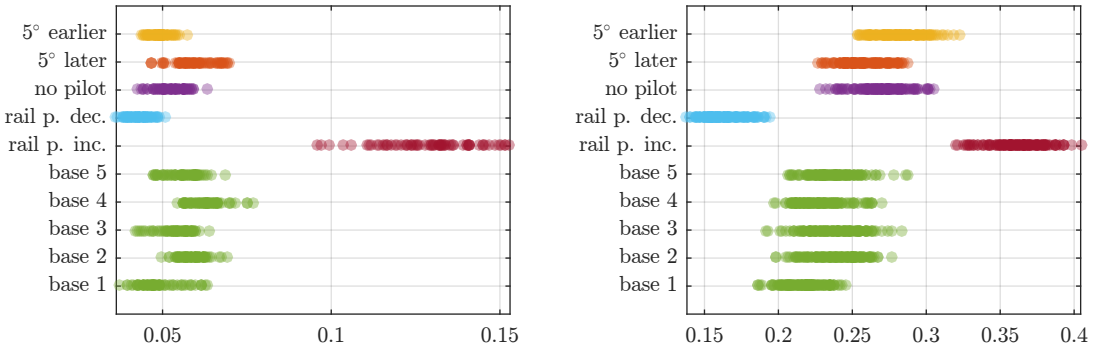


Figure 56 – Calculation of an Energy Ratio feature, which divides the energy in a certain range j through the full energy in the shape.

other conditions, whereas the increased rail pressure condition yields higher values as all the other conditions. Therefore, a linear correlation between the rail pressure and this feature at this operation can be assumed.

Nevertheless, the small distance between the base conditions and the decreased rail pressure can cause a false classification in certain cases. Also, for the idle run this distance is less. As all other faulty conditions can more or less be discriminated between each other and the base condition, only the shapes of the decreased rail pressure condition and the base conditions are shown in figure 58. The most significant difference for all sub-bands down to 125 Hz can be found in the 4 kHz sub-band during the ignition. Unfortunately the envelope for the base conditions changes shape for the different revolution speeds. Therefore, for each revolution speed a specialized feature is defined, so the discrimination between the conditions is maximized.

As already mentioned, another energy ratio is defined in this sub-band. The *Energy Ratio 4k 1* feature also uses the procedure shown in the block diagram in figure 56. Nevertheless, the energy ratio is calculated from the energy from 28 ° to 46 ° to the



(a) Idle. (b) 1500 rpm @ 0 % load.

Figure 57 – Values of the *Energy Ratio 4k 2* feature for two operations and all conditions with microphone-based cycle alignment.

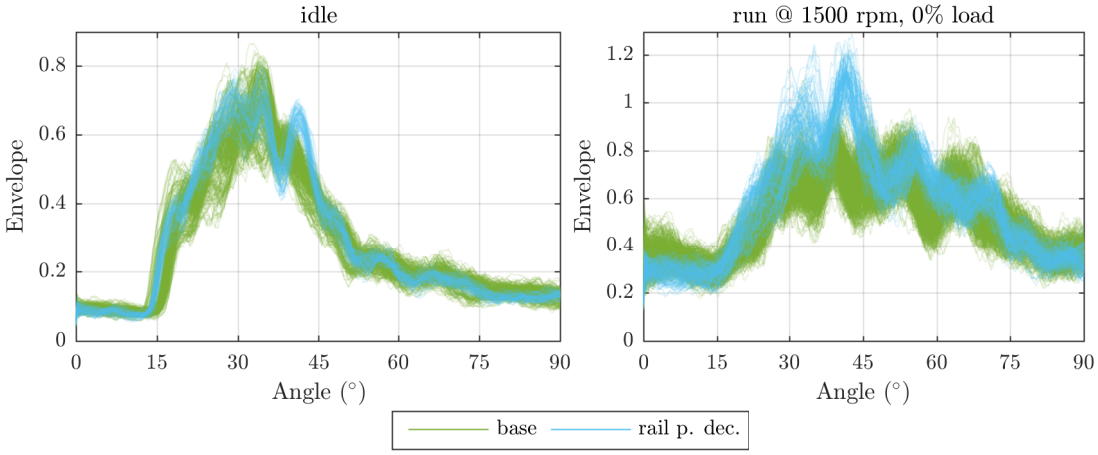
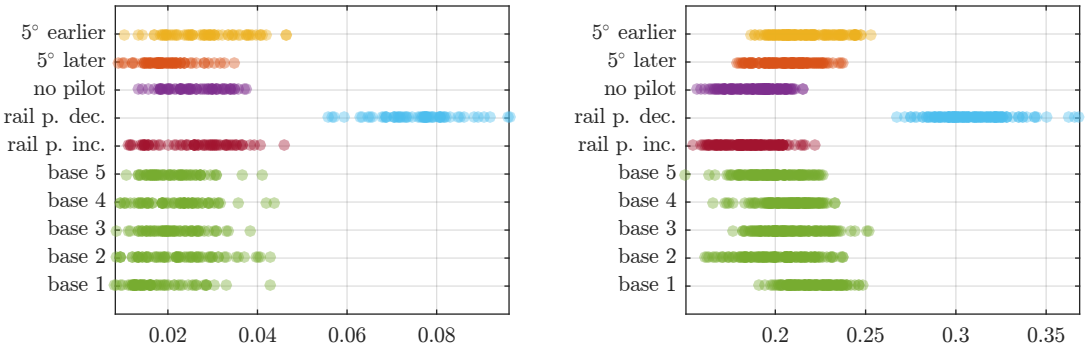


Figure 58 – Difference of the filtered envelope shape at 4 kHz between the base and reduced rail pressure conditions for idle and 1500 rpm with microphone-based cycle-alignment. Only the first quarter-round is shown.

energy in the full half-round. Of course, this feature is used for the operation at 1500 rpm with 0 % load. This highly specialized feature only shows a significant difference for the reduced rail pressure condition, as can be seen in figure 59(b).

For idle this energy ratio is not the optimal solution, as can be seen in the left half of figure 58. But, the decreased rail pressure shows a very distinct peak around 42 °. Therefore, a slice is taken from the envelope shape spanning from 37 ° to 46 °. In this slice a linear polynomial fit is calculated and evaluated at the corresponding sampling points, see figure 60. This regression line is now subtracted from the actual shape in this slice and the root-mean square is calculated, thus, resulting in the RMS error of the linear fit. As the envelope of the reduced rail pressure



(a) *Linear Fit Error 4k* feature for idle.

(b) *Energy Ratio 4k 1* feature for 1500 rpm with 0 % load.

Figure 59 – Features for the detection of the increased rail pressure conditions at their respective useful revolution speeds, calculated with microphone-based cycle-alignment.

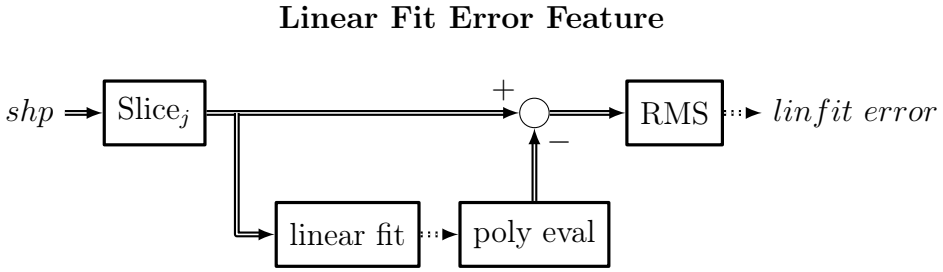
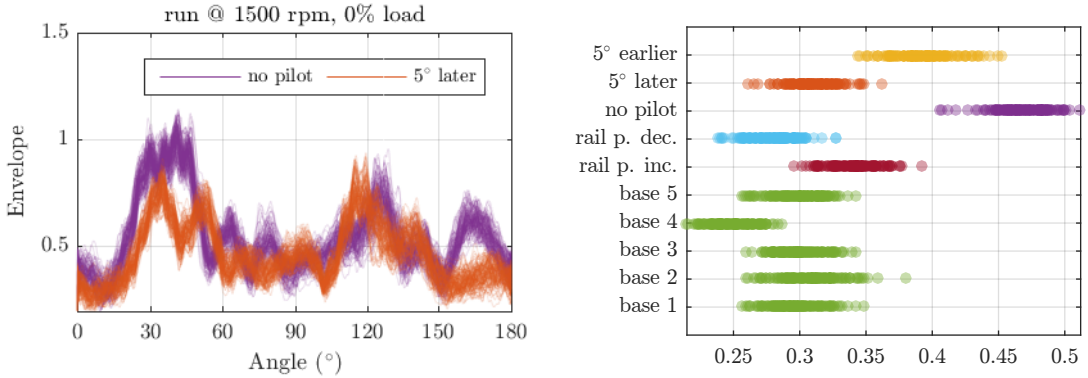


Figure 60 – Calculation of the Linear Fit Error feature. For a slice j of the shape, a linear fit is calculated and evaluated at the appropriate sampling points. The RMS-error to this linear model is used as feature.

condition oscillates in the slice, the linear fit produces a significant higher error than the fit to the nearly constant slope for the base conditions. The respective feature is called *Linear Fit Error 4k* and its values are shown in figure 59(a) for microphone-based cycle alignment. The feature also works for the cogwheel-based alignment techniques, but as the slice is very narrow, the angles must be adapted accordingly to the changed reference point of the crankshaft angle.

4.1.4 Bandpass filtered Envelope @ 1 kHz

With all features defined up to this point, all conditions can be represented with significant different feature values for the cogwheel-based cycle alignment for idle and the run with 1500 rpm and 0 % load. A good classifier should therefore be able to distinguish between these features. For microphone-based alignment all conditions



(a) Folded Envelope shape.

(b) Feature values.

Figure 61 – Derivation and values for the *Energy Ratio 1k* feature to discriminate the no pilot from the 5 ° later injection condition for microphone-based cycle alignment at 1500 rpm.

are also separated by the already defined features for idle. What remains is the slight overlap for the Envelope Onset 8k feature in figure 54 between the no pilot injection condition and the 5 ° later injection condition for 1500 rpm. As all the tools for a new feature, which is able to widen this distance, are already present, another feature is defined for this particular case.

In the sub-band of 1 kHz the folded envelope shows significantly more energy during the ignition for the no pilot condition than for the 5 ° later injection condition. Therefore, the *Energy Ratio 1k* feature is defined, which calculates the energy in actually two angle spans and normalizes the sum of these two to the energy in the whole half-round according to the block diagram already shown in figure 56. The two spans reach from 22 ° to 48 ° and 155 ° to 180 ° degree for maximal discrimination, see figure 61(a).

The feature is only useful for the desired task and separates in this particular case the wanted conditions, see figure 61(b). Although, the no pilot injection condition is also separated from other conditions, this is of minor relevance, as other features showed already significant differences here. The same goes for the idle condition for the microphone-based cycle alignment or both operations for the cogwheel-based cycle alignment. Here, the Energy Ratio 1k feature yields no new information.

As now all conditions with four ignitions show significant different values for idle and the run at 1500 rpm with 0 % load for the defined envelope shape features, no further features based on the microphone signal are defined, as these conditions should be separable.

4.2 Speed Droop Feature

One further feature is defined, which solely works with the cogwheel 122 signal and stresses the relevance of the cycle alignment with only the microphone signal, if the running-characteristics should be classified on the airborne sound emission. Because in this case, no additional signal from the engine is necessary. On the other hand, if we have for example both cogwheel signals, it is indeed possible to perform angle alignment, but the information, if, for example, one cylinder is not combusting, can already be derived from the cogwheel signal itself. Therefore, it is not useful to specify features in this case, to detect the missing cylinder in an additional microphone signal.

As this thesis focuses on airborne sound emission, only one feature is presented here, which only relies on the cogwheel 122 signal and detects a missing cylinder. This feature is called *Speed Droop* and measures the maximum time variation difference between two 3 ° ticks in one duty cycle of the cogwheel 122 signal. This time variation was already shown in the angle alignment in section 2.4.2 in figure 12. If one cylinder is missing, the difference between the longest time between two ticks to the shortest time is greater than for conditions with four ignitions. The block diagram for this feature is shown in figure 62. This features needs no extra input but the cogwheel 122 signal. In the upper part of the block diagram, the average rpm are calculated from the mean of the difference between the positive zero-crossings (pZC) of the triangular shaped cogwheel 122 signal. This scalar rpm value is used to slice the signal into frames with the length of the run-time of two average rounds (rpm-based cycle alignment). Although this is the most imprecise form of cycle alignment, it is still sufficient in this application. In each frame the 122 Ticks Location procedure (compare figure 13) is used to find the exact locations of the cogwheel 122 ticks. Afterwards the difference function is applied and the difference between the maximum and minimum value for each frame is used as feature value.

Figure 63 shows the values of the feature for idle and 1500 rpm with 50 % load. Although a classification should be possible for idle, the operation at 1500 rpm with 50 % load shows a very stable discrimination. This discrimination is nearly lost without load and is therefore not shown. The feature was calculated by using an oversampling factor $OS = 8$, which results in the eight discretization steps between the integer sample numbers. It can also be seen that the speed droop is less for higher revolution speeds, which is the reason for the higher correlation between the cycles at higher revolution speeds in figure 19.

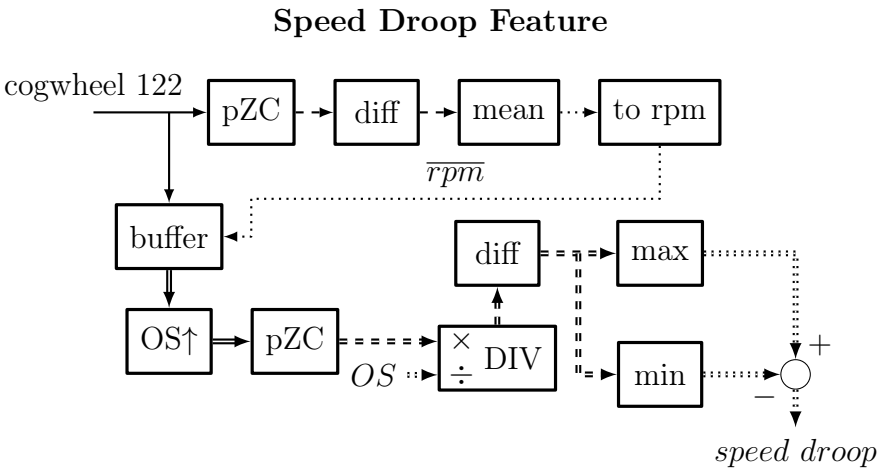
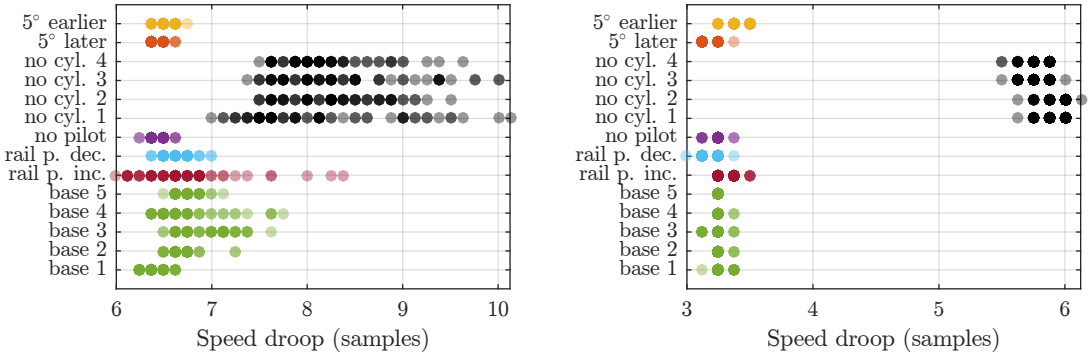


Figure 62 – Calculation of the Speed Droop feature from solely the cogwheel 122 signal for detecting a missing cylinder.



(a) Idle. (b) 1500 rpm @ 50 % load.

Figure 63 – Values of the *Speed Droop* feature for two operations and all conditions.

Although a more reliable discrimination in idle is desirable, the speed droop feature shows the possible information content in signals coming directly from the engine control unit. Of course, these signal contain much information about the running-characteristics, as they are derived from the internal sensors. But, as in this chapter, features, which are solely based on the microphone signal, showed already a good discrimination between the conditions, the physical connection to the engine for additional signals is not necessary, which may be advantageous in the test procedure.

4.3 Summary

With the pre-processing steps for the cycle alignment of the previous chapters, it is possible to define simple yet powerful features for the classification of the faulty conditions. The shape of the folded ignitions per cycle is very distinctive for different conditions in certain frequency-bands. Therefore, as often necessary in acoustics, the analysis must be split into several frequency bands to extract sufficient information. The envelope shape features do not rely on an absolute calibration of the microphone signal, as they work on energy ratios, spectral shapes and onsets. Nevertheless, the angle indexes are absolute values, which means that the measurement procedure and analysis must provide a fix angle shift between the crankshaft angle and the angle of the alignment. Therefore, the simple rpm-based cycle alignment is not suitable for the envelope shape features. Interestingly, the microphone-based cycle alignment is even more robust in this regard than the cogwheel-based techniques. If the microphone is placed closer or further away from the engine, the angle ranges presented in this chapter need to be adapted only for the cogwheel-based techniques. But, if the ignitions are delayed in the microphone signals, as it is in this case, no adjustment must be made for the microphone-based cycle alignment.

In the next chapter the whole classification procedure is summarized and the classification results for different framework conditions are presented for the envelope shape features.

5 Classification Results

In this chapter the classification abilities of the envelope shape features are closer examined under different framework conditions. As the Ignition Count feature showed satisfactory results in section 3.6 due to the reliable detection on different microphones or revolution speeds and loads (even with run-ups), the task of detecting a missing cylinder is considered solved. Therefore, according to table 1, six different conditions remain: The five times measured base condition, the two conditions with modified rail pressure (increased and decreased) and the three conditions with modified injection (no pilot injection, 5 ° later and 5 ° earlier injection). The classification can be done for two different objectives: First, the task is to detect any faulty condition, which results in a binary classification base condition versus the others. Second, the different conditions shall be identified, which is in the most difficult case a six-group classification. Here, the classifier assumes, that each condition can only happen individually. As not all of the envelope shape features are equally robust, the classification will also be tested on a reduced set of conditions, namely without the modified rail pressure conditions.

The classification result analysis is split in two parts, in section 5.1 the classification abilities, if cogwheel signals are available, are presented. In section 5.2 the classification with only the microphone signal is shown.

As one of the most intuitive algorithms, a k-nearest neighbor classifier is used for all the classifications presented in this chapter, which actually does not try to extract a model, but works memory-based [10]. If the class or group of a new sample should be predicted, the distance to all the training samples is calculated. The mode of the groups of the k closest samples is assigned to the sample. To guarantee robustness for possible outliers, k is chosen to be ten. A higher value of k produces greater regions for the group with more smoothing. For smaller values, also single sample points have an effect [5]. Of course, k should be smaller than the number of training samples in each group. Otherwise, if one group has more samples in the training data (base condition), the decision boundaries of this group will expand. As for idle runs, only about 60 samples are available per measurement, the choice of $k = 10$ gives enough room for splits in training and test sets. As distance metric the most intuitive euclidean distance (L2-norm) between the query sample and the training samples is used.

The significant distances between the different groups, i.d. conditions, were already

clearly visible in the last chapter for each individual feature. Therefore, it is not surprising, that even this simple classifier will provide good results. The features span a F-dimensional space, depending on how many features are used, and a high distance in one dimension, i.d. feature, is sufficient for groups to be separated, even though in all other dimensions no discrimination may be possible. For the distance metric to work properly, the features need to be standardized to zero mean and unit variance, so that each feature is weighted equally.

An overfitted classifier works well on the training data, but will produce lots of errors on new data. If the degrees of freedom of a classifier is high, the classifier is fitted to nearly every sample in the training set. The degrees of freedom of a k-nearest neighbor classifier is N/k, where N is the total number of samples in the training set. This results in a total maximum of N/k non-overlapping neighborhoods [10]. Nevertheless, the feature values shown in the last chapter suggest that each condition will have just one neighborhood, where all samples - with small exceptions at the edges - will belong to one condition.

Nevertheless, overfitting might still occur and can not be ruled out. As only one engine was measured in the same room, another model of the engine might produce different envelope shapes. As mentioned in chapter 2, the measurement room was anechoic. The influence of a room can roughly be tested by convoluting a room impulse on the measurement signal and performing the analysis. As already mentioned, the revolution speed and the load on the engine have a significant impact on the sound of the engine and therefore also on the envelope. Which means, it is not possible to use a measurement from a different operation as validation of a trained classifier. As a result, stratified 10-fold cross-validation will be used to assess classifier performance [10]. Also, the repeatability of the measurements is considered due to the five base measurements, which only shows minor deviations.

To assess the quality of a classification task, two common measures are used [10], [16]:

Classification Loss As loss function the fraction of misclassified test data samples is used. Of course, this should be zero.

Edge For a k-nearest neighbor classifier the score of a classification is the posterior probability that a new data point is part of a certain group, i.d. the number of the neighbors with that group label divided by the total number of neighbors. The difference for the posterior probability of the true group to the maximal probability of the false groups is a measure of certainty, which

is ideally one. In this case, all neighbors are from one group, whereas for a value of zero, the neighbors are exactly split into two groups. Here, different strategies exist for the k-nearest neighbor classifier. In literature, ties are often broken randomly [5], in this thesis the group of the closest sample of the nearest neighbors is used. Nevertheless, it is a guess. The average of all these differences for the test data set is called the edge. It should be noted that also for a value smaller than - but close to - one a perfect classification is possible. Therefore, even though two classification tasks may have a classification loss of zero, one can still be more stable, as the edge is higher.

The two measures are calculated for every fold of the cross-validation and the median and inter-quartile range are used to evaluate the performance of a classification task. Of course, the inter-quartile range should be small, otherwise the performance of the classifier depends on the selection of the samples.

Another important factor for the envelope of the signal is the microphone position. Here, two cases can be considered. First, the whole classification can be done on a different microphone with the same features. Second, another classification task can be done, which is definitely not a requirement for the classification, but shows the robustness. In this case, the classifier learns on one microphone and tries to classify from the signal of another microphone. Even in this scenario, it is at least possible to detect the no pilot condition and the conditions with the shifted injection timing, although they might not be separable from each other.

Table 7 gives an overview of the considered envelope shape features and their relevance for the alignment techniques. Also each feature is assigned an index, which will be used as short-hand in the next sections. With the alignment techniques proposed in this thesis the Envelope Onset 16k (2) and the Envelope Onset 8k (3) features are only useful in one alignment scenario.

5.1 Cogwheel-based Classification

In this section it is assumed that at least the cogwheel 1 signal is available for the classification task. The whole procedure for this framework is shown in figure 64. The cogwheel 1 signal has one impulse per round at a fixed angle. On the one hand, this impulse can be used to perform cogwheel-based cycle alignment for the calculation of the envelope shape features. On the other hand, an average revolution speed $\overline{rpm}$ can be easily derived from the *1 Tick Detection* procedure, as it is just

	Feature Name	Envelope	Cogwheel-based Alignment	Microphone-based Alignment
1	Spectral Shape 16k	RMS 16 kHz	useful	useful
2	Envelope Onset 16k	RMS 16 kHz	useful	not useful
3	Envelope Onset 8k	RMS 8 kHz	not useful	useful
4	Linear Fit Error 4k	RMS 4 kHz	useful	useful
5	Energy Ratio 4k 1	RMS 4 kHz	useful	useful
6	Energy Ratio 4k 2	RMS 4 kHz	useful	useful
7	Energy Ratio 1k	RMS 1 kHz	not necessary	useful

Table 7 – List of all envelope shape features and their validity for classification for different alignment techniques.

Classification with Cogwheel-based Cycle Alignment

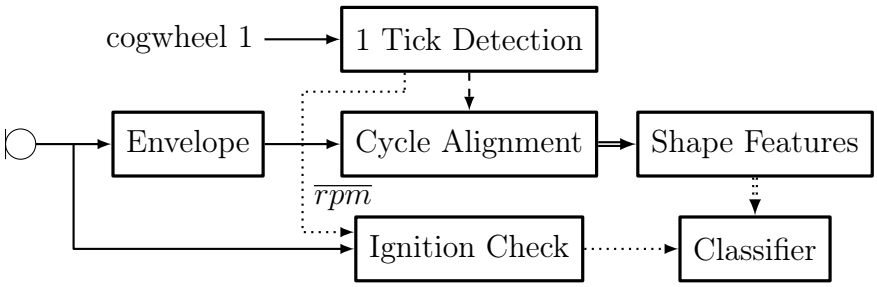


Figure 64 – Overview of the classification procedure with cycle alignment with the cogwheel 1 signal.

the mean time between the ticks and converted to rounds per minute. Therefore, the Ignition Count feature z can be calculated from the raw microphone signal and fed into the classifier. As a simple threshold comparison is sufficient, the classifier in the block diagram consists actually of this comparison and the k-nearest neighbor classifier for the envelope shape features. Of course, the k-nearest neighbor classifier could also be used for the Ignition Count feature, but it can not improve the already perfect classification with its increased computational complexity.

Figure 65 shows the block diagram for the classification task, if also the cogwheel 122 signal is available. As this signal contains 122 ticks per round, the more precise angle alignment can be performed. Nevertheless, the difference between the two alignment techniques is marginal for the (smoothed) envelope, as was already shown in section 3.7. The structure of the procedure is similar to the cycle alignment, only the upper right part in the block diagram is added, where the cogwheel 122 signal is evaluated and used for the angle alignment of the cycles. The detection of a missing

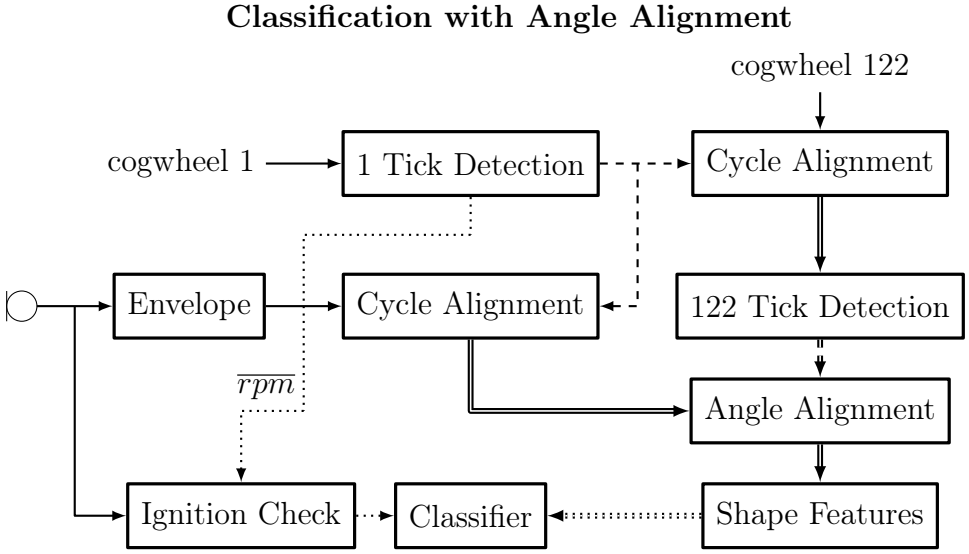


Figure 65 – Overview of the classification procedure with angle alignment with the two cogwheel signals.

cylinder is not affected by the additional signal and leads to the same combination of the classification of a threshold detection for the Ignition Count feature and the k-nearest neighbor classifier for the envelope shape features. It is obvious that angle alignment can only decrease the variation of some of the envelope shape features. Therefore, the results for cogwheel-based cycle alignment are primarily presented here. Although, in the most challenging task, where only two envelope shape features will be used for a full classification, the difference to angle alignment will be presented.

Again, Table 7 shows the seven envelope shape features and their validity for the alignment techniques. For the cogwheel-based alignment only five of them are useful or necessary. Nevertheless, as already discussed in the definition of the features in section 4.1, not all features are useful for both revolution speeds (idle and 1500 rpm). Also, the discriminating abilities for each feature could be already discussed from the feature values. Table 8 summarizes this discrimination abilities for the cogwheel-based cycle alignment in the upper right part for idle operation and in the lower left part for the operation with 1500 rpm and 0 % load. If a feature is able to discriminate between two conditions, the respective index from table 7 is entered into the cell. If the discrimination is not certain, as the feature values are close together, the index is put into braces. As two features (Linear Fit Error 4k (4) and Energy Ratio 4k 1 (5)) only work at one revolution speed, at maximum four features of the five useful features for this alignment technique are in the respective halves of table 8.

1500 rpm \ idle	base	rail inc.	rail dec.	no pilot	5 ° later	5 ° earlier
base	\	2, 6	4	1, 2	1, 2, 4	1, 2
rail inc.	6	\	2, 6	6	2, 6	2, 6
rail dec.	(2), 5, (6)	2, 5, 6	\	1, 2, 4	1, 2	1, 2, (4)
no pilot	1, 2	1, 2, 6	1, 2, 5, 6	\	2, (4)	2
5 ° later	1, 2	1, 2, 6	1, 2, 5, 6	2	\	2
5 ° earlier	1, 2	1, 2, 6	1, 5 ,6	2	2	\

Table 8 – Discrimination abilities of the envelope shape features from table 7 for cogwheel-based cycle alignment for idle (upper right part) and 1500 rpm with 0 % load (lower left part).

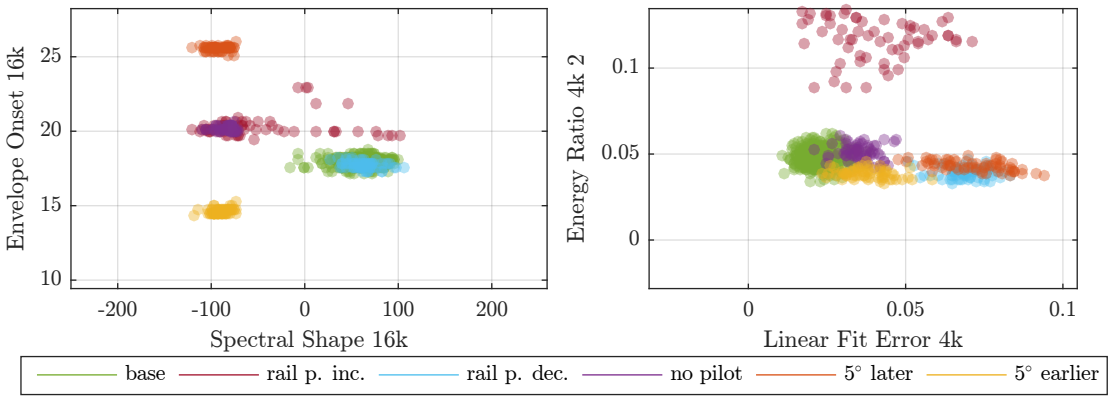


Figure 66 – Scatter plot for the features used for classification for idle with cogwheel-based cycle alignment on microphone 1.

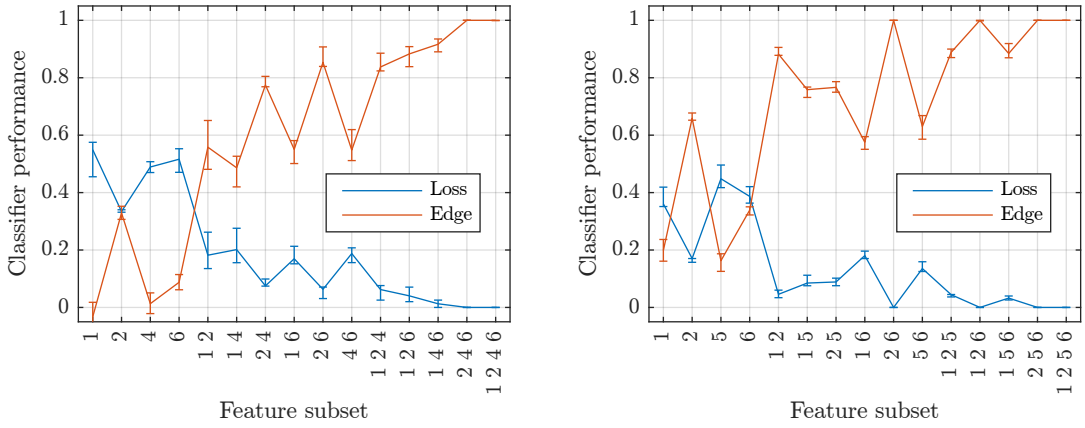
The four features for idle operation are shown in figure 66. The axis limits in the figure correspond to the interval $[-3.5, 3.5]$ times the standard deviation around the mean of the standardization of each feature. At first sight it is evident that in the left part of the figure the most difficult distinction is between the rail pressure increase and the no pilot condition plus between the rail pressure decrease and base conditions. With the additional features of the right part, six distinct neighborhoods without any overlap are present in the four dimensional feature space, although some of them are closer together than others. As already seen in table 8, the Spectral Shape 16k (1) and Envelope Onset 16k (2) features are redundant, as the Envelope Onset 16k feature has all the discrimination features of the Spectral Shape 16k feature.

Not surprisingly, the classification works rather well with this four features. For the stratified 10-fold cross-validation, the median of the missclassifcation loss is

zero with an inter-quartile range of zero. The median of the edge is also a perfect one, but the inter-quartile range is for various runs approximately around 0.001, which is still extra-ordinary. The closest occurring value to one for the edge for not perfect cross-validation runs is 0.9989. One interesting aspect can be shown, if the redundant Spectral Shape 16k feature is not used. In this case, the statistics of the classification loss does not change and the median of the edge is also one. But, the inter-quartile range is for repeated runs of the cross-validation definitely smaller. The reason of this marginal better performance with less features is the so-called curse of dimensionality, which has many manifestations. One possible cause is that the prediction is more difficult at the edges for higher dimensions [10]. Even though the median and the inter-quartile range is used on the performance measures of the individual cross-validation runs, they slightly change for repeated runs in the range below one percent.

Figure 67(a) shows all subsets of these four features for idle operation. Due to the low number of features, it is actually possible and useful to calculate all possible subsets and find the subset with the best performance with as less features as possible. The concept of using the actual classifier for feature selection and an useful performance measure is called the wrapper approach [18] and its general main disadvantage is its complexity, as the number of subsets for F features is $2^F - 1$. Also, for a good performance measure, the whole classification and test procedure must be run a number of times (cross-validation). Nevertheless, for four features this results in 15 subsets, where the performance of each subsets can easily be calculated and no specialized search strategy, such as sequential forward/backward selection or random mutation [18] is necessary. The figure confirms the choice of the subset of three features without the Spectral Shape 16k (1) feature, as it is the subset with the lowest number of features by providing the best reachable performance.

For the operation with 1500 rpm with 0 % load in table 8, also four features are present, as only the Linear Fit Error 4k (4) feature is exchanged for the Energy Ratio 4k 1 (5) feature. In comparison to idle operation, the no pilot injection condition and the increased rail pressure condition are now discriminated by three of the four features, and also the distinction between the decreased rail pressure and the base condition might be fulfilled by two features beside the specialized Energy Ratio 4k 1 (5) feature. Therefore, the whole classification is possible with just two features, the Envelope Onset 16k (2) feature and the Energy Ratio 4k 2 (6) feature. In the performance measures classification loss and edge no difference is visible, as for two and four feature respectively, the median of the classification loss for the



(a) Idle.

(b) 1500 rpm with 0 % load.

Figure 67 – Classifier performance with the median and 25 and 75 percentiles as errorbars of the misclassification loss and edge for possible feature subsets for cogwheel-based cycle alignment on microphone 1.

ten folds of the cross-validation is zero and the median for the edge is one, with an inter-quartile range of zero for both measures. Therefore, no information is lost by using just two features for successfully identifying six different conditions. This is also visible in figure 67(b).

As two dimensions can easily be plotted, the feature values and the decision boundaries for the k-nearest neighbor classifier are shown in figure 68. In the left part, the classification with cogwheel-based cycle alignment is shown in comparison to the classification with angle alignment in the right half. No qualitative difference can be found, also the performance measures for the classification are equal. The only difference lies in the marginal higher variation in each group for the Envelope Onset 16k (2) feature for cycle alignment, which has no effect on the classification. Therefore, angle alignment yields no improvement in the classification on the microphone signal envelope. All five base conditions are shown in the figure, and show together approximately the same variation in each feature as the single measurements for the faulty conditions. If we assume that the variation of the feature values is the same for a faulty condition as for the base conditions, the classification is robust. It can also be seen that the base condition is surrounded by the faulty conditions, which is not surprising, as the faulty conditions are symmetrical in their respective aspect and each of the features is optimized for this one aspect, i.d. the modified rail pressure or the modified injection timing.

All classifications up to this point were done by using the microphone 1 signal. But,

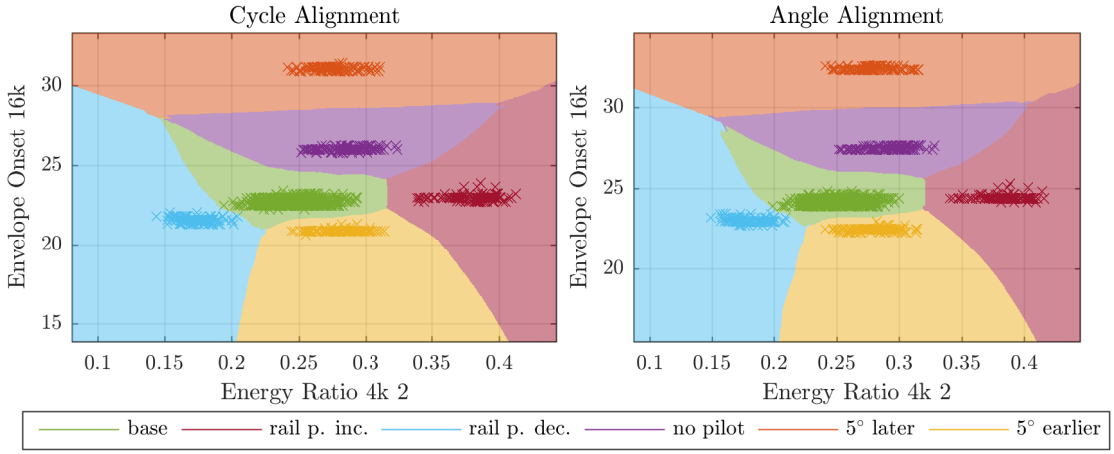


Figure 68 – Scatter plot of the two features used for classification with cycle and angle alignment with the neighborhoods for the k-nearest neighbor classifier for 1500 rpm with 0 % load on microphone 1.

as an internal combustion engine has a distinct directivity and the envelope shape features are defined on band-limited frequency ranges, the performance of the envelope shape features depend on the position of the microphone. As they were defined to characteristic properties in the envelope of the microphone 1 signal, the envelope shape features work obviously best for this microphone. Nevertheless, a reliable detection of a faulty condition is still possible, for example, with the envelope of the signal from microphone 4. In this scenario, detection means that the individual conditions are only separated between the base conditions and the faulty ones. With the same three features for idle, this results in a median classification loss for the cross-validation of 0.015 with an inter-quartile range of 0.015. The median for the edge is 0.95 with an inter-quartile range of 0.015. Both the training and the validation is done on microphone 4. Although, some data points are misclassified, by building a statistics over a few seconds, still a reliable detection is possible. For the separation of the individual classes, the median of the classification loss is still only 0.03 with an inter-quartile range of 0.02, the median of the edge is 0.85 with an inter-quartile range of 0.06. Although these values suggest a good classification in most of the cases, the confusion matrix in table 9 shows a more detailed analysis of a classification training and test with 50 % holdout. The discrimination between the base condition and the increased rail pressure condition is prone to most of the mistakes, although, the whole classifier can be considered to be too sensitive, as most of the errors are false alarms. Therefore, the conclusion can be drawn that, although some features may be sensible to the microphone position, the envelope shape features should be robust to minor changes in the placement of the microphone, which

actual \ predicted	base	rail p. inc.	rail p. dec.	no pilot	5 ° later	5 ° earlier
base	143	11	5	6	0	0
rail p. inc.	0	27	0	6	0	0
rail p. dec.	0	0	32	0	0	0
no pilot	1	3	0	28	0	0
5 ° later	0	0	0	0	34	0
5 ° earlier	0	0	0	0	0	33

Table 9 – Confusion matrix for classification on the microphone 4 signal with cogwheel-based cycle alignment for idle and the Envelope Onset 16k (2), Linear Fit Error 4k (4) and Energy Ratio 4k 2 (6) features. The actual classes are the rows and the predicted classes the columns.

is the relevant factor here. It is not necessary for the application to define features, which work perfectly over an angle of 60 °, which is the angle between microphone 1 and 4.

The two most robust features of the envelope shape features are the Spectral Shape 16k (1) and the Envelope Onset 16k (2) features, which separate the no pilot condition and the two conditions with modified injection timing from each other and the base conditions. If only the two features are used, it is possible to train the k-nearest neighbor classifier on microphone 1 and separately identify the three faulty and the base condition in idle operation, if the conditions with the modified rail pressure are omitted. This results in a classification error below 5 ‰. This confirms the robustness of at least some of the envelope shape features for cogwheel-based alignment techniques.

5.2 Microphone-based Classification

In this section the classification with only microphone-based cycle alignment is examined. As no other input than the microphone signal is necessary, a measurement device using this technique would be independent from the engine itself. Of course, the classification must be adapted (i.d. trained) to the special characteristics. Figure 69 shows the whole procedure for this classification task. The microphone signal is used as input to the average rpm calculation to further decrease the number of required input parameters. Afterwards, the Ignition Check is performed to test if all cylinders are combusting. As for the cogwheel-based aligned classification, a simple threshold detection on the Ignition Count feature z is part of the classifier. If all

Classification with Microphone-based Cycle Alignment

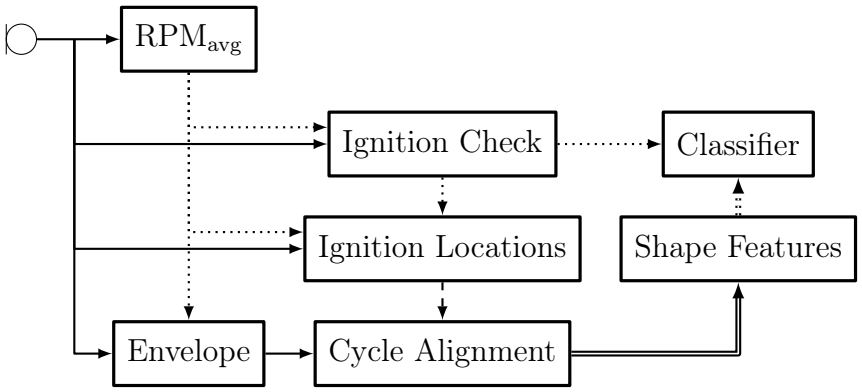


Figure 69 – Overview of the microphone-only classification procedure.

cylinders are combusting, the time locations of the ignitions can be used to perform cycle alignment on the (bandpass-filtered) envelope. Afterwards, the envelope shape features are calculated and fed into the second part of the classifier, also a k-nearest neighbor algorithm.

Table 10 summarizes the discrimination abilities of the envelope shape features for microphone-based cycle alignment. As derived in chapter 4, four features are necessary for idle operation (upper right part of the table), whereas five features should be needed for the operation with 1500 rpm and 0 % load. The for cogwheel-based cycle alignment powerful Envelope Onset 16k (2) feature has no discriminating abilities with microphone-based cycle alignment and is replaced by the Envelope Onset 8k (3) feature for the discrimination between the last three conditions. Also, the Spectral Shape 16k (1) feature is not redundant anymore. At 1500 rpm with 0 % load, again the Linear Fit Error 4k (4) is replaced by the Energy Ratio 4k 1 (5) feature for the discrimination of the decreased rail pressure condition. The only reason for the definition of the Energy Ratio 1k (7) feature was the discrimination between the no pilot injection condition and the 5 ° later injection condition. As other features may also be able to discriminate between these two, the feature may hold redundant information.

Figure 70 shows the four features needed for idle in a scatter plot. Again the features in the left part focus on the discrimination of the no pilot injection condition and the conditions with modified injection timing. The two features in the right part of the figure are each specialized on the modified rail pressure conditions. The cross-validation yields the optimal result with a median misclassification loss of zero

1500 rpm \ idle	base	rail inc.	rail dec.	no pilot	5 ° later	5 ° earlier
base	\	6	4	1	1, 3	1
rail inc.	6	\	4, 6	6	3, 6	3, 6
rail dec.	5, (6)	(3), 5, 6, (7)	\	1, 4	1, 3, 4	1, 3, 4
no pilot	1, (3), 7	1, 6, 7	1, 3, 5, 6, 7	\	3	3
5 ° later	1, 3	1, 3, 6	1, 3, 5, 6	(3), 7	\	3
5 ° earlier	1, (7)	1, (6)	1, 5, 6, 7	3	3, (7)	\

Table 10 – Discrimination abilities of the envelope shape features from table 7 for microphone-based cycle alignment for idle (upper right part) and 1500 rpm with 0 % load (lower left part).

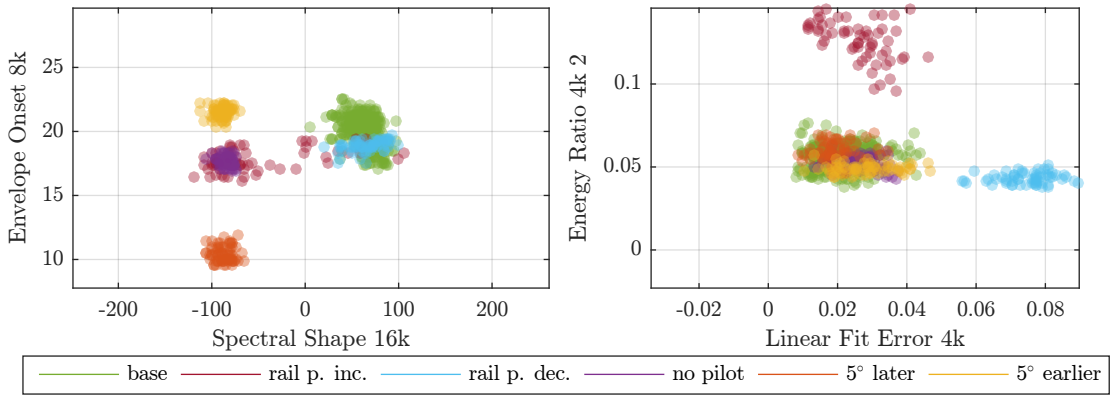
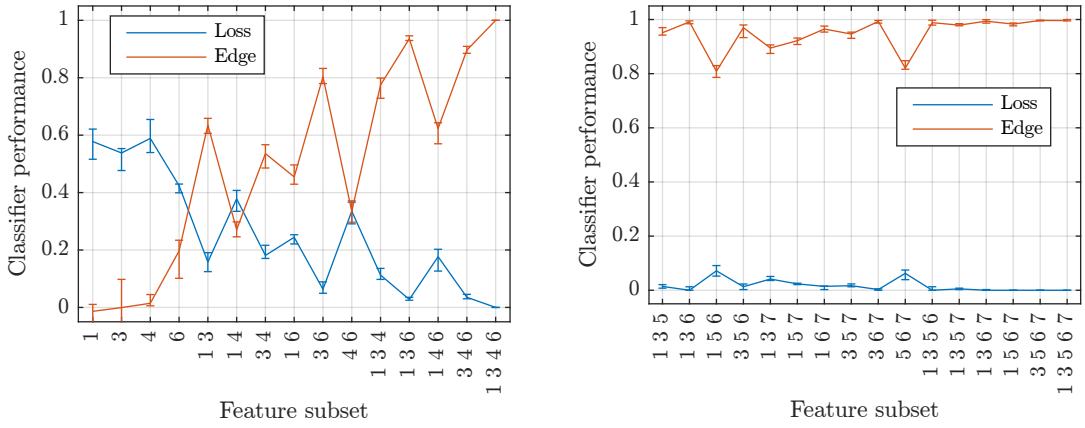


Figure 70 – Scatter plot of the features used for classification for idle with microphone-based cycle alignment on microphone 1.

and a median edge of one with respective zero inter-quartile ranges. In figure 71(a) the performance measures for all possible subsets of the four features are shown. As already suspected from table 10, for a good classification all four features are necessary. Together with the evaluation of the Ignition Count feature, all measured conditions can not only be detected, but individually classified with only the microphone signal, which is the main objective of this thesis.

For the operation with 1500 rpm and 0 % load the feature values change naturally. From the lower left half in table 10 it is possible to define multiple subsets from the five features, which should be able to distinguish between all conditions. With all five features, the classifier makes no mistake in the classification in any fold of the cross-validation. The median of the edge is 0.996 with an inter-quartile range of 0.003. The performance of all subsets including more than three features is shown in figure 71(b), as less features will always provide a poorer performance in this



(a) Idle.

(b) 1500 rpm with 0 % load.

Figure 71 – Classifier performance with the median and 25 and 75 percentiles as errorbars of the misclassification loss and edge for possible feature subsets for microphone-based cycle alignment on microphone 1.

task. Various subsets with an acceptable performance exist. If only three features are used, two subsets perform equally well: The subset consisting of the Spectral Shape 16k (1), the Envelope Onset 8k (3) and the Energy Ratio 4k 2 (6) features and the subset consisting of the Envelope Onset 8k (3), the Energy Ratio 4k 2 (6) and the Energy Ratio 1k (7) feature. Due to the additional discrimination of the decreased rail pressure by the Energy Ratio 4k 2 (6) feature, the Energy Ratio 4k 1 (5) feature is apparently not necessary in this situation. The four features of these two subsets are shown in figure 72 and the selection of the two subsets with their three respective features is comprehensible. The difference between the two subsets is below the variation of the performance measures for multiple runs and can therefore be considered equal. Although the classifier performs worse with any of these two subsets with three features than for all five features, this reduction is below 1 % in the median of the performance measures. As these values assume to classify each sample point individually, this minor deviations can safely be ignored, if the classification is run over a signal block with a couple of cycles. Therefore, for this classification task three features are considered to be enough.

The whole classification procedure also works with the signal from another microphone. As in this scenario of microphone-based cycle alignment various steps for the feature calculation are dependent on the microphone signal, not only the features must provide useful values, but also the average revolution speed calculation and the cycle alignment itself must work with the new signal. As shown in section 3.5

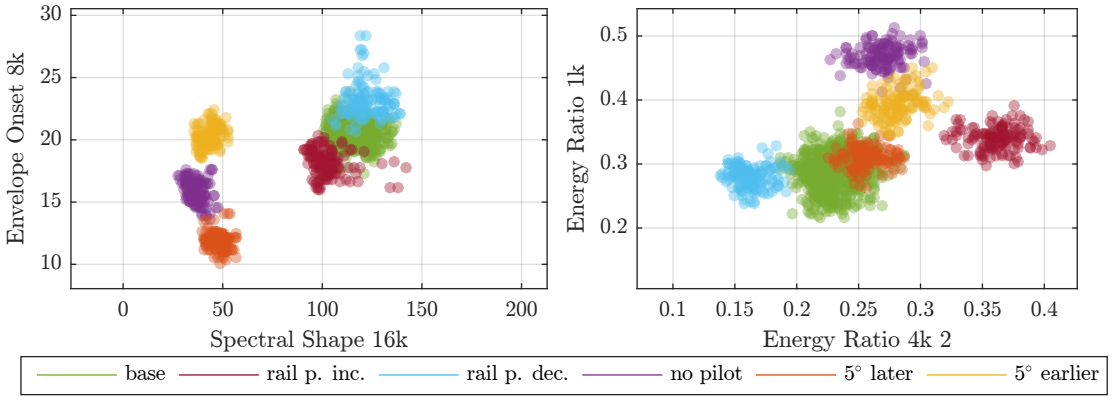


Figure 72 – Scatter plot of the features used for classification for 1500 rpm with 0 % load with microphone-based cycle alignment on microphone 1.

and 3.7, the two latter perform as good with microphone 4 as with microphone 1. Therefore, the envelope shape features can also be calculated for this microphone signal with microphone-based cycle alignment and used for classification. If the tasks is only to detect a faulty condition in idle operation, the classifier still works well with a median classification loss of zero with an inter-quartile range of 0.015, if the same four features are used as for microphone 1. The median edge for the cross-validation is 0.96 with an inter-quartile range of 0.04, if the classifier is trained and tested on microphone 4. If the individual conditions should also be classified, the median classification loss is 0.04 with an inter-quartile range of 0.024 and the median of the edge is 0.87 with an inter-quartile range of 0.024. The confusion matrix for this classification task with a hold-out of 50 % is shown in table 11. Again, it can be seen that some features do not work equally robust for the different microphone positions. Nevertheless, in this scenario the envelope shape features also provide a reasonable classifier, although the whole derivation was done on another microphone.

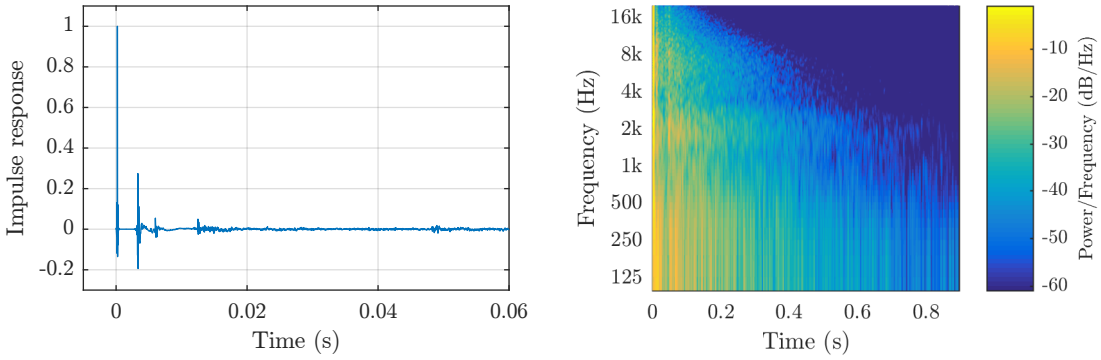
Again, the Spectral Shape 16k feature together with an Envelope Onset feature are the most robust features, which is not surprising, as an energy ratio is more likely to change for another microphone position than the shape of the envelope or timing differences. Therefore, the Spectral Shape 16k (1) and the Envelope Onset 8k (3) feature are used to individually identify the no pilot injection condition, the conditions with modified injection timing and the base condition, where the conditions with modified rail pressure are left out, by training the classifier on microphone 1 and testing on microphone 4 for idle operation. Although in this scenario, the detection of the three faulty conditions works with only these two features, the no pilot

actual \ predicted	base	rail p. inc.	rail p. dec.	no pilot	5 ° later	5 ° earlier
base	163	0	3	0	0	0
rail p. inc.	0	23	0	7	3	0
rail p. dec.	0	0	33	0	0	0
no pilot	0	1	0	32	0	0
5 ° later	0	0	0	0	34	0
5 ° earlier	0	0	0	0	0	33

Table 11 – Confusion matrix for classification on the microphone 4 signal with microphone-based cycle alignment for idle and the Spectral Shape 16k (1), Envelope Onset 8k (3), Linear Fit Error 4k (4) and Energy Ratio 4k 2 (6) features. The actual classes are the rows and the predicted classes the columns.

injection condition and the 5 ° later injection condition can not be separated safely. If the Energy Ratio 1k feature (7) is additionally used, which was designed for this purpose, although not for idle but for 1500 rpm, the three features can perfectly identify the three faulty conditions in this task.

Due to the preliminary type of this research, the engine was measured in an anechoic chamber. As outlook for the real-world application, the developed algorithms are tested, if a typical room will have a significant influence. The room impulse response of a small room with a reverberation time of 0.9 s is convoluted with the raw microphone signal and is shown in figure 73. Of course, the shape of the room impulse response will affect the performance of the algorithms. All algorithms work with transient events, therefore the shape of the early reflections is especially important. On the other hand, if the direct component is much stronger and isolated in time, the influence of the room should be manageable. With a close-up microphone positioning, as was done in the measurement, the relationship between the direct component and the first reflections can be calculated with mirror sources [20]. In the simulated room impulse response, the first reflection arrives 3 ms after the direct component, the mirror source is therefore approximately only 1 m further away. For the microphone 1 above the engine this could be seen as a reflection from the ceiling, which is only 50 cm above the microphone. Also, the damping of the first reflection of the simulated impulse response should be easily achievable, as a factor of approximately 1/4 can be interpreted that the first-order mirror sources must be at least four times further away then the direct component, even with a 100 % reflecting surface. As the distance to the engine was 0.6 m, this is also not an unrealistic scenario. The figure shows the spectrogram as well, where the typical faster decay in



(a) Early Reflections (b) Spectrogram

Figure 73 – Room impulse response for testing the algorithms with early reflections and reverberation.

the higher frequency range is seen. Therefore, the influence will also be less for the high-frequency envelope shape features or the ignition-based envelope, which focus on the high frequency range.

The first part of the classification with microphone-based cycle alignment must be the Ignition Count feature to check if all four ignitions can be detected. Figure 74 shows the Ignition Count feature z for idle operation and 1500 rpm with 0 % load, if the signal of microphone 1 is convoluted with the above shown impulse response. Each sample is a slice of one second duration, for which the average revolution speed $\overline{rpm}$ is also calculated. 1 % of the samples with four cylinders have a value of -1 for z with a perfect reliability and are outside of the limits of the figure. Nevertheless, the separation of the conditions works in this scenario.

As the used impulse response is just one example, a detailed analysis of the influence on the envelope shape features is not useful. Nevertheless, the more challenging situation is visible in figure 75 in comparison to figure 72 for the operation with 1500 rpm and 0 % load, if the microphone signal with the virtual room impulse response is used for training and testing in the classification. Although the feature values are similar, the variance in each group is increased, thus resulting in more overlap between the groups. In the calculation of the features, the algorithm for the detection of the average revolution speed $\overline{rpm}$ sometimes produced an unreliable result. In this case, the average revolution speed was retrieved from the belt signal, as this is not the main focus of the thesis. For the classification with microphone-based cycle alignment, two subsets of each three features perform equally well on the raw microphone 1 signal. With the additional virtual room impulse response, the first

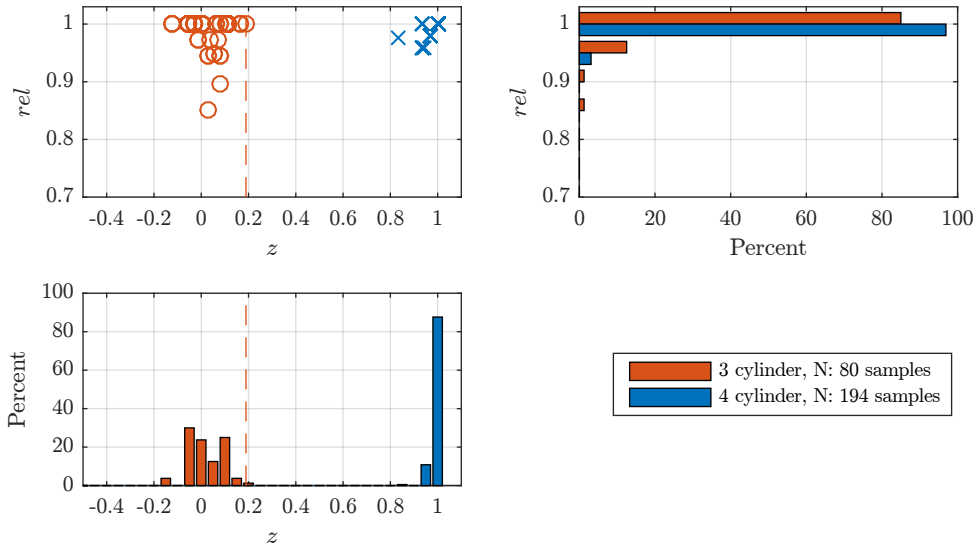


Figure 74 – Ignition Count feature z and reliability for all conditions with a virtual room impulse response on microphone 1 for idle and 1500 rpm with 0 % load.

subset consisting of the Spectral Shape 16k (1), the Envelope Onset 8k (3) and the Energy Ratio 4k 2 (6) features performs quite well with a median misclassification loss of 0.02 and an inter-quartile range of 0.03 in the cross-validation. The median edge for this subset is 0.94 with an inter-quartile range of 0.018. The other subset scores comparably with the Energy Ratio 1k (7) feature instead of the Spectral Shape 16k (1) feature. Again, the features, which are designed to discriminate the rail pressure, are not as robust as the others, as they do not work properly in the, therefore omitted, idle operation.

Finally, another classification task is performed. The classifier is trained under the laboratory conditions with the raw microphone signal 1 (anechoic) and tested with the feature values from the virtual room for 1500 rpm with 0 % load with one of the two subsets for this classification task. The confusion matrix is shown in table 12. Although some misclassification between the faulty conditions occur, the classifier still works reliably, even though the only used signal is the microphone signal, which even changed significantly due to the added room impulse response between training and test.

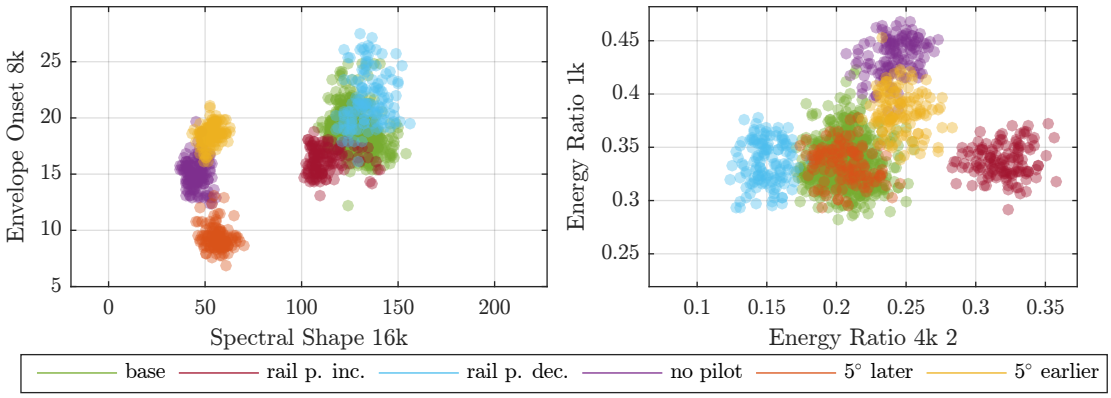


Figure 75 – Scatter plot of the features used for classification for 1500 rpm with 0 % load with microphone-based cycle alignment on microphone 1 with a virtual room impulse response.

actual \ predicted	base	rail p. inc.	rail p. dec.	no pilot	5 ° later	5 ° earlier
base	618	0	24	0	1	0
rail p. inc.	3	125	0	0	0	0
rail p. dec.	6	0	130	0	0	0
no pilot	0	0	0	91	35	2
5 ° later	0	0	0	0	128	0
5 ° earlier	0	0	0	27	0	101

Table 12 – Confusion matrix for classification on the microphone 1 signal with a virtual impulse response with microphone-based cycle alignment for 1500 rpm with 0 % load and the Spectral Shape 16k (1), the Envelope Onset 8k (3) and the Energy Ratio 4k 2 (6) features. The classifier was trained on the raw microphone signal. The actual classes are the rows and the predicted classes the columns.

5.3 Summary

In this chapter the discrimination abilities of the envelope shape feature are examined with a k-nearest neighbor classifier under different framework parameters. Although the features are not robust against different revolution speeds and loads on the engine, it is possible to separate the different conditions for one operation. Therefore, the classifier needs to be trained for every operation and needs to know the current operation. As the revolution speed can be easily detected and tests of a new engine start usually with no load, the classification can still run automatically.

The envelope shape features need to use cycle alignment, so the single ignitions can be folded to one shape. A fixed reference point to the crankshaft angle is used,

if cogwheel (or belt) signals are used for the alignment. Then it is possible to easily detect a modified injection timing, which results in the very robust Envelope Onset 16k feature. Nevertheless, with microphone-based cycle alignment a reliable classification is also possible, although other features need to be defined.

The robustness of the microphone-based cycle alignment and the respective features show good results, if a virtual room is added to the measurement signals. This should be seen as an outlook that the developed algorithms can also work under more realistic environment parameters such as a more reverberant room.

6 Conclusion

The main objective of this thesis to reliably detect even small changes in the engine sound can be considered solved. Although the differences can be heard in a direct comparison, it would require much training for a human to be able to classify the engine sound emission into the seven categories without a reference. Of course, a computer-based classification also requires training and adaption, but then the detection runs automatically.

An internal combustion engine has by-design a natural periodicity. Of course, more information can be extracted from any time series signal from the engine, if this periodicity is used to perform cycle or angle alignment. This is typically done with additional signals (cogwheel or belt signals), which require a physical connection to the engine itself. As this thesis focuses on microphone signals, chapter 3 concentrates on cycle alignment with the microphone signal by detecting the time locations of the ignitions. As in normal operation four ignitions correspond to one duty cycle, cycle alignment can be performed. As four ignitions are not always present in the measurement data, which is of course considered a faulty condition, the sensitivity to a missing cylinder ignition is exploited in order to actually detect the missing ignition. The used concept of pattern recognition can easily be generalized for two or more missing ignitions, although for two missing ignitions two characteristic patterns must be considered, as the two missing cylinders may be adjacent or separated in the duty cycle. It is also likely that the parameter set needs to be adjusted so that the long time during two adjacent missing cylinder ignitions is detected correctly.

As most engines will produce transient peaks during the ignition, if they are not mounted inside a car-body, both procedures of the microphone-based cycle alignment and the ignition detection should work with different engine models. Therefore, the whole concept is a powerful set of tools for pre-processing microphone recordings of the engine sound emissions for a detailed analysis, which is not limited to the task of classification in this thesis. Nevertheless, a parameter adjustment, as it is presented here with the Monte-Carlo optimization, is highly recommended. Apart from the fine tuning of the procedure for end-of-line quality control, further research could establish the performance of the algorithms in a live monitoring set-up during the operation of a mounted engine with microphones inside the engine compartment.

In chapter 4 the envelope shape features are not designed to be general indicators of a faulty engine, although the shape difference for the 16k filtered envelope may

occur for other engines as well, as the reason is a not precise injection, which leads to a non-uniform acceleration of the piston. The presented envelope shape features are examples of possible features on the average half-round of a cycle-aligned envelope of a microphone signal. The classification shows that onsets and spectral shape features are more robust than energy ratios, which are more likely to change for other microphone directions. Nevertheless, for other classification tasks, environmental parameters or engines, it is most likely to be necessary to define different envelope shape features.

Chapter 5 finally shows the classification abilities of the defined envelope shape features for this classification task, where among other things a wrongly timed injection with a shift of only 5 ° to the crankshaft angle is reliably detected. Overall, the out-of tune conditions of the engine are not only detected, but can be individually classified. Although less features are needed and the variation in some features is marginally less, if additional cogwheel signals are used for the cycle alignment, the same performance of the classifier can be reached only with the microphone signal. The robustness of the features against microphone displacement is double-checked by using another microphone for the classification, which is 75 cm away from the microphone used for the derivation of the algorithms and features. The robustness of the features to a room impulse response is only superficially analyzed, but nevertheless shows promising results for the application of the algorithms in more reverberating environments.

The analysis shows that only a few seconds of audio material is enough for a robust detection. Without any detailed performance optimization the feature calculation lasts for about ten seconds for ten seconds of audio-material on a personal computer. As some parts can easily be parallelized for optimal use of a multi-core processor, a real-time implementation is definitely possible. Depending on the finally used platform (DSP, MATLAB, ...), special care must be taken in the filter design. Either the phase delay must be considered in the feature definition or zero-phase filtering should be considered by introducing some latency, due to the required buffering for the acausal procedures. As no immediate reaction to the classification is necessary in most cases, even a latency of one second seems to be acceptable.

Out of the scope of this thesis is the evaluation, if the sound emission of the artificially created faulty conditions is comparable to the sound emission caused by manufacturing defects. Here, additional measurements are inevitable, which are firmly motivated by the findings in this thesis.

References

[1] DIN EN 61672-1. Elektroakustik - Schallpegelmesser - Teil 1: Anforderungen (IEC 616702-1:2002), 2005.

[2] J. Allen. Short term spectral analysis, synthesis, and modification by discrete Fourier transform. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 25(3):235–238, June 1977.

[3] B. Audone and M. Borsero. Median filters to suppress interference. In *2015 IEEE International Symposium on Electromagnetic Compatibility (EMC)*, pages 366–371, August 2015.

[4] J. P. Bello, L. Daudet, S. Abdallah, C. Duxbury, M. Davies, and M. B. Sandler. A Tutorial on Onset Detection in Music Signals. *IEEE Transactions on Speech and Audio Processing*, 13(5):1035–1047, September 2005.

[5] Christopher M. Bishop. *Pattern recognition and machine learning*. Information science and statistics. Springer, New York, 2006.

[6] Michael Dickreiter. *Analoge Schallspeicherung, analoge Tonregieanlagen Hörfunk-Betriebstechnik, digitale Tontechnik, Tonmeßtechnik*. Handbuch der Tonstudiotechnik. Saur, München, 7th edition, 2008.

[7] Gerhard Doblinger. *Zeitdiskrete Signale und Systeme: eine Einführung in die grundlegenden Methoden der digitalen Signalverarbeitung*. Schlembach Fachverlag, Wilburgstetten, 2007.

[8] J. Stephen Downie. Music information retrieval. *Annual review of information science and technology*, 37(1):295–340, 2003.

[9] Derry Fitzgerald. Harmonic/percussive separation using median filtering. In *Proc. of the 13th Int. Conference on Digital Audio Effects (DAFx-10) 2010*, Graz, 2010.

[10] Jerome Friedman, Trevor Hastie, and Robert Tibshirani. *The elements of statistical learning*, volume 1. Springer, Berlin, 2001.

[11] Svend Gade, Thorkild Find Pedersen, Henrik Herlufsen, and Hans Konstantin-Hansen. Practical experience with RPM estimation using the Autotracker algorithm. In *IMAC-XXV: Conference & Exposition on Structural Dynamics*, 2007.

[12] Ian Glover and Peter M. Grant. *Digital Communications*. Pearson Education, Essex, 2010.

[13] Philipp Grams, Dejan Arsic, and Hugo Fastl. Schallortung an Motorprüfständen über Grad-Kurbelwinkel. In *DAGA 2015 Nürnberg*, March 2015.

[14] C. Richard Johnson Jr, William A. Sethares, and Andrew G. Klein. *Software Receiver Design: Build your Own Digital Communication System in Five Easy Steps*. Cambridge University Press, August 2011.

[15] Christoph Lechner. *Ringversuch Messung der Schallimmission 2005 - Vergleich von Messmethoden am Beispiel Straßenverkehrslärm*. Umweltbundesamt GmbH, Wien, 2006.

[16] MathWorks. MATLAB 2015b Documentation, 2016.

[17] Hans-Joachim Mittag. *Statistik*. Springer-Lehrbuch. Springer, Berlin, Heidelberg, 2016.

[18] Dunja Mladenić. *Feature Selection for Dimensionality Reduction*, volume Subspace, Latent Structure and Feature Selection of *Lecture Notes in Computer Science*. Springer, Berlin Heidelberg, 2006.

[19] Michael Möser. *Messtechnik der Akustik*. Springer, Berlin, Heidelberg, 2010.

[20] Michael Möser. *Technische Akustik*. Springer, Berlin, Heidelberg, 2015.

[21] Gerhard Müller and Michael Möser, editors. *Taschenbuch der technischen Akustik: mit 119 Tabellen*. Springer, Berlin, 3rd edition, 2004.

[22] Thomas Müller-Gronbach, Erich Novak, and Klaus Ritter. *Monte Carlo-Algorithmen*. Springer-Lehrbuch. Springer, Berlin, Heidelberg, 2012.

[23] Martin Pflüger, Franz Brandl, Ulrich Bernhard, and Karl Feitzelmayer. *Fahrzeugakustik*. Springer, Wien, 2010.

[24] Integrated Publishing. Actual Combustion Cycles. `http://www.tpub.com/engine3/en3-15.htm`, April 2016.

[25] Konrad Reif and Robert Bosch GmbH, editors. *Moderne Diesel-Einspritzsysteme: Common Rail und Einzelzylindersysteme*. Bosch Fachinformation Automobil. Vieweg + Teubner, Wiesbaden, 1st edition, 2010.

[26] M. Ristau. Verbrennungsmotoren. In *Vieweg Handbuch Maschinenbau*, pages 912–977. Springer, 2007.

[27] David C. Stone. Application of median filtering to noisy data. *Canadian Journal of chemistry*, 73(10):1573–1581, 1995.

[28] Eric W. Weisstein. Cubic Spline. `http://mathworld.wolfram.com/CubicSpline.html`, January 2016.

[29] Wikipedia. Crankshaft position sensor. https://en.wikipedia.org/w/index.php?title=Crankshaft_position_sensor&oldid=711874503, March 2016. Page Version ID: 711874503.

[30] Eberhard Zeidler, editor. *Springer-Taschenbuch der Mathematik.* Springer Fachmedien, Wiesbaden, 2013.

[31] Peter Zeller, editor. *Handbuch Fahrzeugakustik: Grundlagen, Auslegung, Berechnung, Versuch ; mit 85 Tabellen.* ATZ/MTZ-Fachbuch. Vieweg + Teubner, Wiesbaden, 2nd edition, 2012.

Andreas FUCHS
(Name in Blockbuchstaben)

9773819
(Matrikelnummer)

Erklärung

Hiermit bestätige ich, dass mir der *Leitfaden für schriftliche Arbeiten an der KUG* bekannt ist und ich diese Richtlinien eingehalten habe.

Graz, den 31.05.2016

.....
Unterschrift der Verfasserin/des Verfassers