Mechanical Systems and Signal Processing 235 (2025) 112871

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

journal homepage: www.elsevier.com/locate/ymssp

Full Length Article

Adaptive polynomial filter applied to in-cylinder pressure signal in internal combustion engines

Pedro Gabana [a], Blanca Giménez [a,*], Andrés Melgar [a], Alfonso Horrillo [b]

[a] Department of Energy and Fluid Mechanics Engineering, University of Valladolid, Paseo del Cauce 59, E-47011 Valladolid, Spain
[b] CIDAUT Foundation, Plaza Vicente Aleixandre Campos 2, Parque Tecnológico, E-47151 Boecillo-Valladolid, Spain

ARTICLE INFO

Communicated by Didier Remond

Keywords:
Adaptative filter
Polynomial filter
Savitzky-Golay filter
ICE pressure diagnostic
Cycle averaging

ABSTRACT

Pressure record processing is a powerful tool for research and diagnosis of internal combustion engines. From the pressure record, using the energy conservation equation, it is possible to calculate the rate of heat release (*RoHR*) and thus quantify the evolution of the combustion process. In order to obtain accurate *RoHR* results, it is necessary to remove as much noise as possible from the pressure signal, without removing relevant information from the studied phenomenon. In reciprocating internal combustion engines, pressure data are recorded discretely and synchronized with the crankshaft angle. Both pressure signal and its derivative are used in *RoHR* calculation. For the integration of differential equations using data as a boundary condition, it is necessary to have no discontinuities in the function and its derivative.

This work presents a novel methodology for filtering discrete data adaptively and converting it into a continuous and derivable function. For this, polynomial fits are used in each interval between two experimental data. The polynomial order and the number of points used to make the fit are chosen depending on the value taken by a convolution of the signal. With the adaptive filter it is possible to reduce noise significantly in parts of the cycle where signal-to-noise ratio is low without affecting the parts where signal-to-noise ratio is high. The novelty and main advantage of this filtering methodology is that it is configured with only one parameter independent of the operating conditions of the engine while preserving the information of cycle-to-cycle variations.

1. Introduction

The evolution of the pressure on the piston in a cycle of an Internal Combustion Engine (*ICE*) determines the operation of the engine. Knowing the value of the pressure in the combustion chamber during the time interval that the cycle lasts (one cycle of the engine 40 ms at 3000 rpm) allows to calculate the work transmitted to the piston, a large part of which is transmitted to the engine crankshaft. It is also possible to process the pressure records to obtain the speed at which heat has been released (rate of heat release *RoHR*) when the piston is near the top dead center. This heat release is a consequence of the thermochemical processes that take place during the combustion process. The phenomenology of these processes is the subject of numerous studies due to their implications for the performance and emissions of the engine.

The variables characterizing the state of the thermodynamic system evolving in the combustion chamber of an *ICE* are pressure,

* Corresponding author.
E-mail addresses: pedro.gabana@uva.es (P. Gabana), blanca.gimenez@uva.es (B. Giménez), andres.melgar@uva.es (A. Melgar), alfhor@cidaut.es (A. Horrillo).

https://doi.org/10.1016/j.ymssp.2025.112871
Received 28 July 2023; Received in revised form 11 April 2025; Accepted 15 May 2025
Available online 23 May 2025

Nomenclature

Variables

a	coeficients used to calculate weighted average
Atten	attenuation (bar)
c	convolution
i	pressure data index
Im	imaginary part.
j	complex unit
k	harmonic number multiple of fundamental frequency
m	number of regression points on each side of the polynomial
n	polynomial order
N	number of data points
Ncycles	cycles number
p	pressure (bar)
Re	real part
V	dimensionless frequency
x	angular data position (°)
X	duration of engine cycle
z	interval index

Acronyms

APF	Adaptative Polynomial Filter
CA	Crank Angle (°)
CCV	Cycle to Cycle Variations
DFT	Discrete Fourier Transform
FF	Frequency decomposition Filter
H2	Hydrogen
ICE	Internal Combustion Engine
MCD	Mean Cycle Difference (bar)
NG	Natural Gas
RMSE	Root Mean Square Error
RoHR	Rate of Heat Release (J/°)
SGF	Savitzky-Golay Filter
TF	Transfer Function

Greek letters

α	crank angle (°)
ω	angular frequency (rad/s)
ω_F	fundamental angular frequency (rad/s)
ω_N	data sampling angular frequency (rad/s)

Subscripts

c	cutoff
exp	experimental
filt	filtered
fin	final
i	position of a value
ini	initial
k	relative to a multiple of fundamental frequency
l	index around position i
s	stop
0	relative to central point

temperature, volume and mass. The fuel and air masses evolving in the cycle are measured. The crank angle is also used to both calculate the volume (by means of a kinematic angle-volume relationship) and synchronize the recording of pressure data [1,2].

For pressure measurement, piezoelectric sensors are used [3,4] since they have good frequency response characteristics and low sensitivity to temperature variations that take place during the combustion process [5]. However, temperature measurement is difficult due to the low frequency response of temperature sensors, therefore it is usually calculated from pressure, volume and mass using the equation of state of the system [6].

The pressure signal is used to determine relevant parameters in the engine performance analysis. For example, the maximum pressure, which is a mechanical restriction [7], the mean indicated pressure [8], the indicated torque [9,10], mechanical losses, etc. Further pressure signal analysis allows to study knock [11,12] or correct the reference of the crankshaft sensor position [13]. *RoHR* obtained from pressure records is useful to study quantitatively the combustion process. It is complemented with visualization techniques that are more expensive and restricted to undemanding operating conditions from a mechanical or thermal point of view.

The *RoHR* is calculated using diagnostic thermodynamic models based on the integration of a system of differential equations in which the combustion chamber pressure and its derivative are boundary conditions [1,14]. The Runge-Kutta method is usually used for the numerical integration of equations. The assumptions of this integration method require the functions being integrated to be continuous and derivable and can be evaluated at any point chosen by the method. Therefore, the *RoHR* calculation needs high quality pressure records regarding continuity and derivability. Pressure records consist of a series of data equispaced angularly according to the angular resolution of the sampling. In order to evaluate the pressure at any point indicated by the numerical method, it is usual to interpolate linearly between the data of two consecutive points, which implies that at the points the functions are not derivable.

Pressure records are also used in the study of cycle-to-cycle variations (CCV) [2,3]. CCV consists of the fact that, under stationary engine operating conditions (engine speed, torque, consumption…), in each cycle, for the same crankshaft rotation angle, the pressure takes different values. These pressure fluctuations are due to the random nature of the fluid movement due to turbulence, which causes the combustion process to evolve differently every cycle. As a consequence, heat release (*RoHR*) occurs differently in each cycle. CCV has negative consequences both on performance (power and efficiency) and on pollutant emissions.

1.1. Filtering techniques for in-cylinder pressure in ICE

In the literature it is possible to find different techniques for in-cylinder pressure records filtering such as cycle averaging, frequency decomposition filtering and weighted-average based filters.

A way to filter *ICE* pressure records is cycle averaging the signals recorded from different cycles while the engine is stabilized at certain operating conditions. Thus, a mean cycle is obtained which is representative of the average combustion process [15,16]. If *CCV* in *ICE* combustion process is analyzed, cycle averaging is not a valid processing technique since it eliminates the unique information of each cycle [15–18].

Another common technique is pressure records filtering based on their decomposition into harmonic series. Each harmonic is a complex number corresponding to one frequency and characterized by their amplitude and phase. The last step of the filtering process consists of reconstructing the signal from the harmonics, once certain harmonics have been removed or modified [15–18]. This type of filters will be called frequency filters (*FF*) in this work.

The filtering assumes that the phenomenon to be analyzed is manifested in the pressure record in certain harmonics. These harmonics will be called characteristic harmonics of the process in this work. The noise is present in all the harmonics of the signal. The harmonics where the phenomenon is not present (not characteristic harmonic of the process) can be eliminated without any modification of the information of the phenomenon, since, in this case, the harmonic is composed only of noise.

However, in the characteristic harmonics of the process it is not possible to eliminate the noise; we can only hope that the component of the phenomenon in these harmonics is much higher than the noise one, and then the noise contribution is not significant. To attenuate or completely eliminate unwanted harmonics, the harmonics of the original signal are multiplied by a series of harmonics that correspond to those of the transfer function of the filter. The transfer function is usually zero for harmonics that are considered not to be characteristic of the process and one for those that are.

Identifying the characteristic harmonics of the phenomenon is the basis for successful filtering using this technique. In general, it is assumed that all the characteristic harmonic frequencies are below a certain frequency (cutoff frequency), therefore a unit transfer function is used for all harmonics below this frequency. From this frequency, the harmonics of the transfer function are gradually attenuated until it becomes equal to zero from another frequency (stop frequency) [15–17,19]. The spacing between these two frequencies determines the filter order and affects the final results [16].

In the thermodynamic cycle that takes place in an *ICE*, the characteristic harmonics of the process vary depending on the crank angle (cycle instant), since the phenomenons are different during the compression, combustion and expansion processes. During compression and expansion, the frequencies of the characteristic harmonics are lower than during combustion. In addition, in all cases the frequencies depend on the engine speed, and, in the case of combustion, the frequencies also depend on the properties of the mixture that reacts. These properties greatly influence the speed at which combustion occurs.

Another technique used for signal filtering is the weighted-average based filters [20]. One of their advantages is the high application speed. They consist of assigning to a given point the weighted average of a signal points series around the point where the filter is applied. The filtering characteristics depend on the weighting values (filter coefficients) and the number of used data (window length).

The Savitzky-Golay filters (*SGF*) belong to this type of filter. These filters are based on using minimum square interpolation polynomials of a certain order and a certain number of points around the point where the filter is applied [21]. The particularity of this filter is that the value of the polynomial at the central point of the interpolated data is obtained using weighted-average of the interpolated data. Therefore, they are not based on frequency decomposition. *SGF* have already been used to process pressure signals in *ICE*. In [22] four *SGF* with order 3 and with fixed numbers of points (21, 51, 101 and 151) are studied, afterwards the *RoHR* of raw and filtered data are analyzed with the aim of choosing an optimal filter configuration.

Other works, not related to *ICE*, propose the possibility of modifying the *SGF* configuration by signal zones (adaptive filtering). In [23,24] procedures are proposed to adaptively select the order of the polynomial at each instant maintaining a fixed window length.

Results obtained are compared with those obtained with *SGF* with fixed numbers of points and fixed orders. Algorithms for window size selection of *SGF* with fixed polynomial orders are proposed in [25,26]. This work proposes the modification of the order and number of points based on the frequency analysis by successive intervals.

In [27], *SGF* transfer function is analyzed. The polynomial order, n, and the number of points to perform the interpolation, $2m+1$, determine the filtering transfer function. The higher the polynomial order and the lower the number of points taken, the higher the transfer function cutoff frequency. In this work a methodology for calculating the transfer function has been performed.

To analyze the filtering performance, it is common to take a noise-free signal (usually simulated), add noise and apply different filters [23–26]. Since the original signal is available, it is possible to quantify the removed noise, and the information lost.

1.2. Motivation of this work

The motivation of this work is the need of a continuous and derivable function from the filtering of discrete data records. The function will be used as a boundary condition in complex diagnostic models based on integration of differential equations.

In order to obtain a continuous and derivable function in the interval between two data, a polynomial fit with order n can be performed by means of minimum square interpolation using m data on each side of the interval [13]. This filtering method has similar properties to an *SGF* with the same order and number of points, but with a higher computational cost, and on the other hand, a continuous function in the whole interval is obtained.

The filtering method developed in this work, from now on adaptive polynomial filter (*APF*), consists of using interpolation polynomials between two consecutive data point. The interpolation polynomials are chosen based on the cut-off frequency calculated by a previous analysis of the signal. After filtering, the polynomials are modified to obtain a continuous and derivable function in the entire data point domain. For this purpose, the following novel works have been developed:

- Methodology to determine the cut-off frequency at which the data must be filtered in each interval between two measurement points. The criterion for choosing the cutoff frequency is based on the convolution integral value of the signal and a reference function. This transformation quantifies the importance of each frequency for each interval between two data points.
- Methodology for choosing the order and number of points of the interpolation polynomials to be used for filtering the signal.
- Calculation of the transfer function of each interpolation polynomial using the DFT of the *SGF* coefficients.
- Once the interpolation polynomials have been calculated for each interval between two data, new polynomials are calculated. The new polynomials ensure continuity and derivability at the interval limits, and also the same value of the original polynomial and its derivative at the midpoint of the interval.
- Development of a method for evaluating the quality of filtering for cyclic signals. This method is based on the calculation, for each cycle point, of the cycle averaged value of the difference between the original and filtered data.

The novelty of this work lies in the combination of both frequency analysis for signal filtering and polynomial interpolation to obtain a continuous and derivable filtered function in the whole domain. This is a novel contribution in the case of in-cylinder pressure records in *ICE*. It is interesting that the methodology only uses in-cylinder pressure records as input having the same configuration parameters for extreme test conditions with a moderate computational cost. This methodology can be especially useful when studying cycle-to-cycle variations.

2. Methodology: *APF* description

Polynomial minimum square interpolation of data is a tool to remove part of the signal noise. The filtering level applied is controlled by the polynomial order (n) and the number of points used to calculate the polynomial.

Once n is selected, there is a lower limit to the number of points to make the fitting. If the number of points is $n+1$, the polynomial

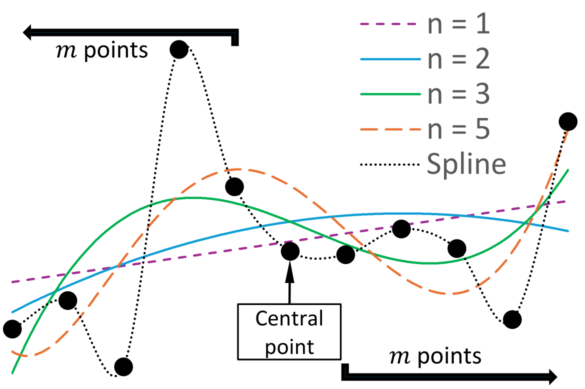

Fig. 1. Representation of the result of different filtering of a signal with different filtering orders.

will pass through all points becoming a spline, with a minimum filtering level. The higher the number of points used and the lower the order of the polynomial, the lower is the frequency from which the harmonics are attenuated [27].

Savitzky-Golay filters [23] are used to calculate the value of the interpolation polynomial at the central point of the $2m+1$ interpolated data. They allow this value to be calculated as the weighted average of the data with coefficients tabulated based on the order of the polynomial n and the odd number of points used $2m+1$, as shown in Fig. 1.

The filtered value at the central point h_i corresponds to Eq. (1).

$$h_i = \sum_{l=-m}^{l=m} a_l\, f_{i+l} \tag{1}$$

Where f_{i+l} are experimental values at the points around the position i to be calculated, and a_l are the coefficients of the filter used to calculate the weighted average, the sum of all a_l is one and their values depend on the number of points used and the polynomial order. The coefficients values a_l can be found in [21]. The coefficients for two symmetrical points respect the center point are equal. Odd-order polynomials have the same coefficients as the previous even-order polynomials [21,27]. From now on only results for odd-order polynomials will be presented, since the results are exactly the same as those of lower even order.

2.1. Savitzky-Golay polynomials transfer function

The filter transfer function (*TF*), or gain, allows to analyze what range of frequencies are removed or attenuated when applying the filter to data. It is calculated as the ratio between the frequency decomposition $H(\omega)$ of the filtered function $h(x)$ and the frequency decomposition $F(\omega)$ of the original function $f(x)$. This section shows how to calculate the transfer function of a weighted-average based filter from the filter coefficients.

If a transformation based on the calculation of the weighted average is applied to a function $f(x)$, the new function $h(x)$ would be expressed as Eq. (2), where Δx is the distance between two consecutive points used for the transformation.

$$h(x) = \sum_{l=-m}^{l=m} a_l\, f(x + l\,\Delta x) \tag{2}$$

The transfer function of the applied filter is the Fourier transform of the function $h(x)$ divided by the Fourier transform of the function $f(x)$, Eq. (3).

$$\begin{gathered} H(\omega) = \mathscr{F}[h(x)] = \sum\nolimits_{l=-m}^{l=m} a_l\, \mathscr{F}[f(x + l\,\Delta x)] = \sum\nolimits_{l=-m}^{l=m} a_l\, \mathscr{F}[f(x)] e^{j\omega\, l\,\Delta x} \\ H(\omega) = \sum\nolimits_{l=-m}^{l=m} a_l\, F(\omega)\, e^{j\omega\, l\,\Delta x} \end{gathered} \tag{3}$$

In Eq. (3) time-translation property of the Fourier transform [28] has been applied. The Fourier transform of $f(x)$ can be extracted out of the summation being the transfer function as Eq. (4).

$$TF(\omega) = \frac{H(\omega)}{F(\omega)} = \sum_{l=-m}^{l=m} a_l\, e^{j\omega\, l\,\Delta x} \tag{4}$$

Turning to discrete variable, the filter is applied to a N data series sampled every Δx. Assuming that these data correspond to a periodic function of period $N\Delta x$, if the frequencies for which the transfer function is wanted to be known are multiples of the fundamental frequency of the periodic function, $\omega_F = 2\pi/N\Delta x$, the values of the transfer function for any multiple frequency $\omega_k = k\omega_F$, hereafter harmonic k, are given by Eq. (5).

$$TF(\omega_k) = \sum_{l=-m}^{l=m} a_l\, e^{j\frac{2\pi k}{N\Delta x}\, l\,\Delta x} = \sum_{l=-m}^{l=m} a_l\, e^{j\frac{2\pi k}{N}\, l} \tag{5}$$

The obtained *TF* corresponds to the discrete Fourier transform (*DFT*) of a series of N points, of which, $2m+1$ coincide with the filter coefficients and the rest are zero.

Values of k higher than $N/2$ do not make sense according to the Shannon theorem [30]. In order to make the transfer function independent of the sampling frequency, the dimensionless frequency, $V = \omega_k/\omega_N$ is defined, being $\omega_N = N\omega_F$ the angular data sampling frequency. Thus $V = k/N$, and *TF* being Eq. (6).

$$TF(V) = \sum_{l=-m}^{l=m} a_l\, e^{j\,2\pi V\, l} \tag{6}$$

Where the acceptable V values range from 0 to 0.5.

In case the filter coefficients are symmetric with respect to the central point, as is the case of *SGF* coefficients, Eq.(1), the imaginary part of the coefficients of the summation of the corresponding values with positive l are equal and opposite in sign to those with negative l. Therefore, the transfer function has no imaginary part and can be expressed as Eq. (7).

$$TF(V) = a_0 + \sum_{l=1}^{l=m} 2\, a_l \cos(2\pi V l) \tag{7}$$

Eq. (7) applied to a *SGF* for an order 5 polynomial with different values of m is shown in Fig. 2. As can be seen, there is a value of V below which the gain is close to 1 and above which the gain decreases. In this work, the cutoff frequency, V_c, is the value of V at which the gain takes a value below 0.95. The stop frequency, V_s, is the lowest value of V at which the transfer function takes value 0.

One characteristic of the filter that is important in the obtained results is the slope (dTF/dV) at which frequencies are attenuated from the cutoff frequency. In this work, the value $V_s - V_c$ is used to characterize the filter attenuation slope. This parameter is inversely related with the commonly called filter order, i.e. how fast the transfer function value decreases from the value 0.95 to 0.

The *SGF* has a transfer function similar to a low-pass filter with the particularity that above V_s the signal is not completely attenuated, and thereafter the gain changes sign as it attenuates.

2.2. Selection of n and m from V_c and V_s

As V_c and V_s have been defined, each *SGF* (n, m) has unique determined values of V_c and V_s, which can be calculated using Eq.*(7)*. For any value (n, m) there are values of V_c and V_s, however, due to (n, m) are integer values, not all values of V_c and V_s, have the corresponding values (n, m).

As will be seen later, the methodology proposed in this work requires choosing an *SGF*, a pair of values (n, m), from selected values of V_c and $V_s - V_c$. In this section we discuss how this choice is made.

The first step is to choose a range of polynomial orders. In this work the lower limit is taken with $n = 3$ and the upper limit is taken with $n = 11$. Fig. 9 shows the results of the methodology applied to data . It can be seen that values $n > 7$ are not necessary.

Fig. 3a shows the value of V_c as a function of the number of points m used on each side of the point where the filtered signal is calculated and using the order of the polynomial n as a parameter.

Fig. 3b shows the value of the cutoff frequency V_c as a function of $V_s - V_c$. The figure has been obtained using as parameter n and varying m. It can be observed a linear trend between $V_s - V_c$ and V_c when m is modified for a fixed n value. The slope increases as n increases.

If a range of $V_s - V_c$ is selected, Fig. 3b shows the definition of V_{c2} as the cut-off frequency above which the $V_s - V_c$ of the higher order polynomial is greater than the selected range. Similarly, V_{c1} is the frequency below which the $V_s - V_c$ of the lower order polynomial is less than the proposed range.

In Fig. 3b it can be observed that, if a value of V_c is selected, there are only polynomials whose $V_s - V_c$ value is within the selected $V_s - V_c$ range if $V_{c2} > V_c > V_{c1}$.

Fig. 3b illustrates the criterion for choosing the pair (n, m) from V_c and the delimited $V_s - V_c$ range. Three situations, points A, B and C in Fig. 3b, can occur depending on the value of V_c:

- Point A: When $V_c < V_{c1}$ and all polynomials have a value of $V_s - V_c$ below the lower limit of the interval. In this case the polynomial of lower order is chosen because it is the one with the highest value of $V_s - V_c$ below V_{c1}. The number of points will be the one that has the cutoff frequency nearest to the selected V_c.

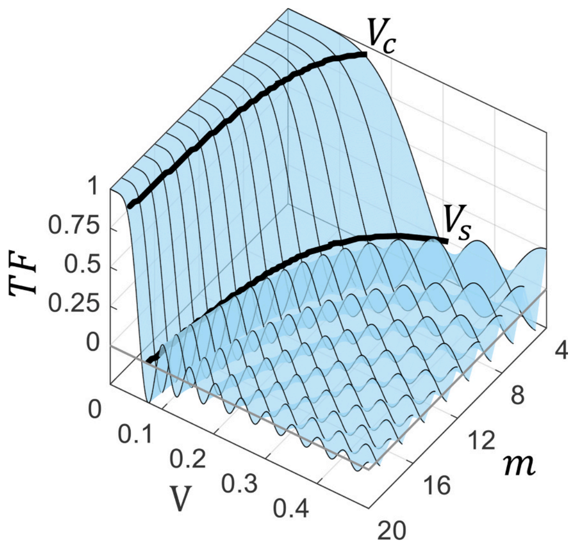

Fig. 2. *TF* plot of *SGF* with order $n = 5$ and different number of points m. It shows V_c calculated when $TF = 0.95$ and V_s is calculated at the first pass through 0 of *TF*.

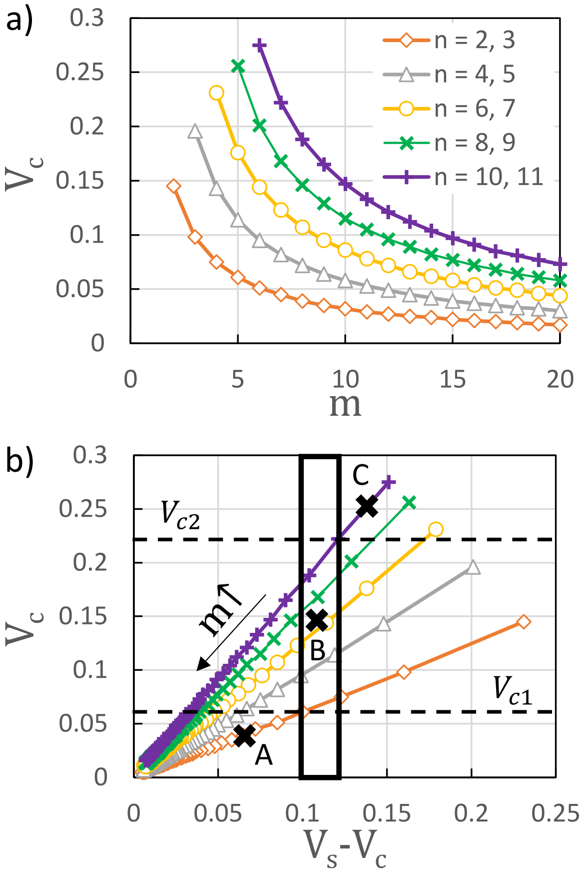

Fig. 3. V_c of *SGF* as a function of different parameters. (a) V_c, as a function of the number of points m used on each side of the filtered signal calculation point. (b) Value of the cutoff frequency V_c as a function of the filter order $V_s - V_c$. (c) Selection approach used in this work to determine the polynomial pair (n, m) from V_c and the range $V_s - V_c$ selected.

- Point B: When $V_{c1} < V_c < V_{c2}$, there are polynomials with a value of $V_s - V_c$ within the desired limits. In this case, the order and number of points chosen are closest to the selected V_c.
- Point C: When $V_c > V_{c2}$ and all polynomials have a value of $V_s - V_c$ above the higher limit of the interval. In this case the polynomial of higher order is chosen because it is the one with the lowest value of $V_s - V_c$ above V_{c2}. The number of points will be the one that has the cutoff frequency nearest to the selected V_c.

In short, for a given V_c and a given $V_s - V_c$ the (n, m) pair of the *SGF* to apply is obtained. Closing the problem requires defining a criterion for V_c and $V_s - V_c$ selection at each point.

2.3. V_c calculation from the convolution of the signal.

This section shows how to determine the value of V_c at each point of a discrete signal.

Any filtering system based on low-pass filters aims to remove frequencies above a certain value and thus eliminate high-frequency noise. To perform this type of filtering it is usual to consider that the signal is periodic, to apply the *DFT* to obtain the signal harmonic decomposition and to multiply this transformation by a *TF* with low or zero gain from V_c value onwards [16]. A selection criterion for selecting V_c is taking the one in which the amplitude of the harmonics of the signal *DFT* (normalized to the unit of the original data) is negligible compared to the original signal levels [17].

In processes in which their characteristic harmonics vary significantly over time, it is convenient to filter with different V_c adapted to the characteristic harmonics of the process in each zone. This eliminates as much noise as possible in the zones where the phenomenon does not have characteristic harmonics of high frequency, without eliminating these characteristic harmonics in the zones where they are representative of the phenomenon. In these last zones, the noise in those characteristic harmonics is not eliminated. The signal to noise ratio is expected to be high in this area.

The problem of this filtering methodology is to identify the characteristic harmonics in each zone. To do this, it is necessary to analyze the signal by intervals in the frequency domain. In each interval, V_c is selected looking for the signal harmonic which amplitude is lower than a certain value.

Convolution allows to quantify the similarity between a function $f(x)$ and the pattern function $g(x)$, obtaining a value representative of the similarity in all the $f(x)$ domain, Eq. (8). Generally, far of the origin, $g(x) = 0$.

$$c(x_0) = \int_{-\infty}^{\infty} f(x)\, g(x - x_0)\, dx \tag{8}$$

The non-integrated variable x_0, shifts the origin of the function $g(x)$ to the position x_0. The value of $c(x_0)$ gives a comparison value to $f(x)$ and $g(x-x_0)$. Only the values of $f(x)$ in the region where $g(x-x_0)$ is not zero, i.e. around x_0, are important for this comparison.

The comparison is done in terms of zero crossing, so that if the two functions have the same sign in the same areas around x_0, the value $c(x_0)$ is high. On the other hand, if the functions take values of opposite signs in the same areas, $c(x_0)$ takes very negative values.

If the integral of $g(x)$ in all integration domain is zero (its mean value is zero), then the value of $c(x_0)$ is independent of the mean value of the function $f(x)$ in the region where $g(x-x_0)$ is non-zero. In that case, the comparison is made not on the basis of the zero crossing of $f(x)$ but on the basis of $f(x)$ crossing its mean value in the region where $g(x-x_0)$ is non-zero.

The units of convolution are not the same as those of $f(x)$.

Functions $g(x)$ with several zero crossing are often used, for example the Morlet wavelet $g_M(t)$.

$$g_M(t) = C_\sigma \pi^{-\frac{1}{4}} e^{-\frac{1}{2}t^2}\left(e^{j\sigma t} - K_\sigma\right) \qquad K_\sigma = e^{-\frac{1}{2}\sigma^2} \qquad C_\sigma = \left(1 + e^{-\sigma^2} - 2e^{-\frac{3}{4}\sigma^2}\right)^{-\frac{1}{2}} \tag{9}$$

This function takes significant values in the interval $t \in [-\pi, \pi]$, outside of it the values are very close to zero. In this interval the function has 2σ zero crossing. K_σ is a coefficient so that the mean value of the real part of $g_M(t)$ equals zero in the interval. It is usually taken $\sigma > 5$, in that case the term K_σ becomes negligible and $C_\sigma \cong 1$. In repetitive functions such as pressure records in ICE, the variable t is replaced by kx/X. Where X corresponds to the duration of the engine cycle, two revolutions in four strokes ICE. k is an integer and determines the frequency multiple of the frequency of the first characteristic harmonic, to be analyzed with $g_M(kx/X)$.

In the case of pressure recordings in ICE, the resolution of the angular encoders is usually greater than 0.1°. The combustion process occurs in a few degrees and therefore there are important pressure variations in small angular intervals, this being more significant when the engine speed is low (in the same angular interval the time interval is greater when the speed is lower) and the combustion process is fast. All this makes the frequencies of the characteristic harmonics of the process approach half the sampling frequency.

In order to obtain information at the highest possible frequencies compared to the sampling frequency, the comparison function $g(x)$ must have a low number of zero crossings in the interval considered. In this work the comparison function $g(x)$ has been chosen with only 2 zero crossings, i.e. what for the Morlet wavelet function would mean $\sigma = 1$. Under these conditions $\sigma < 5$, the amplitude of the Morlet wavelet functions, $C_\sigma \pi^{-1/4}$, varies substantially with σ.

The unit circle function, Eq. *(9)*, has a single zero crossing and its modulus always equals 1 regardless of the k value, so the value of the convolution will have the same units as the function $f(x)$ if the result of the integral is divided by the integration interval. For this reason, the unit circle function has been chosen instead of the Morlet wavelet function. Now k directly denotes the multiple of the frequency of the first characteristic harmonic being analyzed.

$$g(x,k) = \begin{cases} e^{jkx/X} = cos(kx/X) + jsin(kx/X) & -\pi \le kx/X \le \pi \\ 0 & another\ case \end{cases} \tag{10}$$

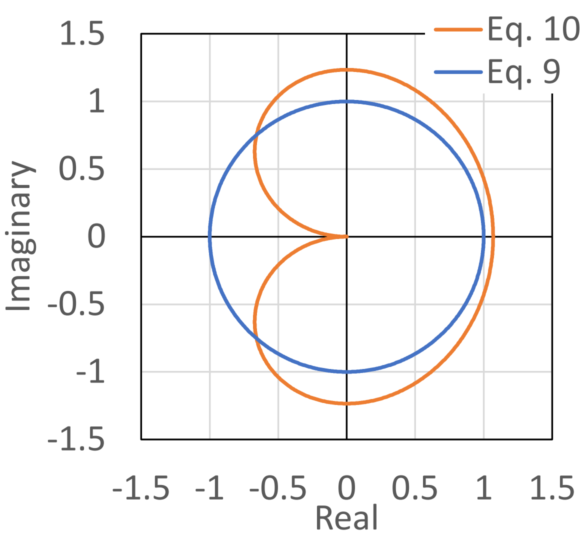

Fig. 4. Complex plane representation of $g(x,\omega)$ Eq. (9) and Eq. (10).

The real part of the comparison function is not continuous at $\frac{kx}{X} = -\pi$ and $\frac{kx}{X} = \pi$. In this work, it is proposed to use Eq. (11) in the interval. In short, it is proposed to use a $\cos(kx/2X)$ window to fix the continuity problem, and to correct the cosine function with an offset of 1/3 (similar to the K_σ parameter of the Morlet function) to make zero the mean value of the real part. Additionally, it is necessary to multiply by $\sqrt{18/7}$ (similar to the C_σ parameter of the Morlet function) to equal the area of the new comparison function to that of the complex unit circle. With this function, the area of the function in independent of k. Both functions have been plotted in the complex plane in Fig. 4.

$$g(x,k) = \sqrt{\frac{18}{7}}\cos(kx/2X)\left[\cos(kx/X) - \frac{1}{3} + j\,sin(kx/X)\right] \qquad -\pi \le kx/X \le \pi \tag{11}$$

In this work pressure records in the combustion chamber of an *ICE* sampled at constant crank angle intervals $\Delta\alpha = 0.6^\circ$ are used. Each work cycle, which corresponds to two engine revolutions, can be considered a periodic function. In each cycle the number of sampled data is $N = 720^\circ/\Delta\alpha$. Eq. (12) is used to apply the convolution in a discrete form to a pressure data series $p(i)$, where i denotes the pressure data index.

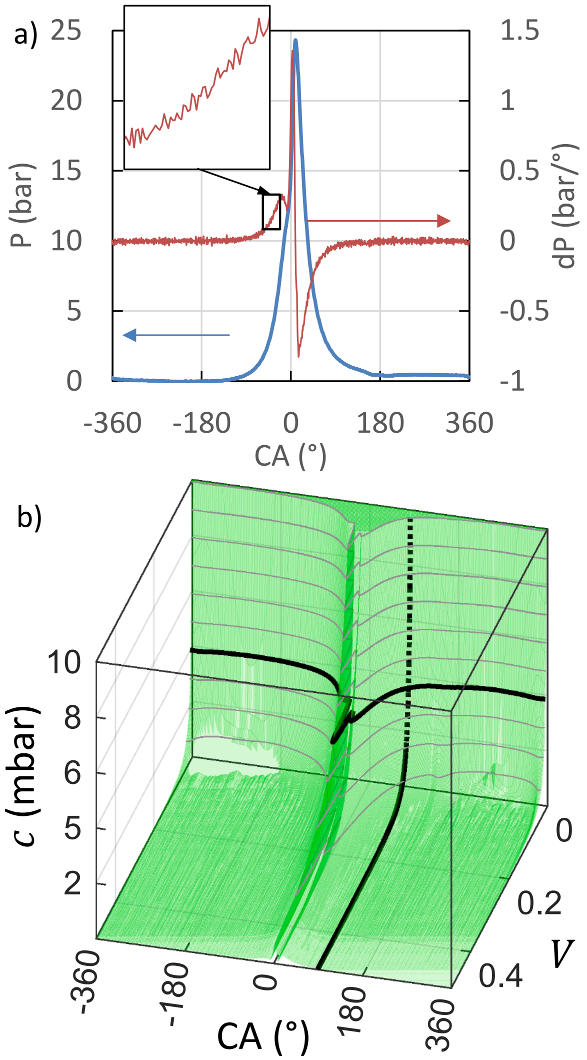

Fig. 5. (a) Experimental pressure signal and its derivative calculated using finite differences. (b) Convolution of an experimental pressure signal.

$$c(i,k) = \frac{k}{N}\sqrt{\frac{18}{7}}\sum_{l\geq -\frac{N}{2k}}^{l\leq \frac{N}{2k}} p(i+l)\ [Re(k,l) + j\,\mathrm{Im}(k,l)]$$
$$Re(k,l) = \cos\left(\frac{\pi k}{N}l\right)\left[\cos\left(\frac{2\pi k}{N}l\right) - \frac{1}{3}\right] \tag{12}$$
$$Im(k,l) = \cos\left(\frac{\pi k}{N}l\right)\left[\sin\left(\frac{2\pi k}{N}l\right)\right]$$

k takes values from 1 to $N/2$. So that $c(i,k)$ has units of pressure, the summation must be divided by the integration interval N/k. A high value of $|c(i,k)|$ indicates that k harmonic is present at $x = i\Delta\alpha$ position, either in the form of sine or cosine. Thus, for a given crank angle (a fixed i value), if $|c(i,k)|$ takes values below a certain level, the value of $k/2N$ will be used to calculate the value of V_c in that point.

When k values are close to $N/2$ the number of summands of Eq. (12) decreases to 3 and it is not possible to guarantee that the mean value of the comparison function is zero. In this case, the continuous component of the function $f(x)$ would have influence on the value of $c(i,k)$.

The cosine values in the real part of $g(x,k)$, for the same absolute values of l, have the same value and sign, therefore, the mean value of the function does not have to be zero. In the case of the imaginary part, the values of the sine, symmetric with respect to the origin, are inverse so the mean value of the function is always zero.

To ensure that in the real part the mean of all coefficients is zero, all real values of the comparison function are subtracted by their mean value $\overline{Re}$ (Eq. (13)).

$$\overline{Re}(k) = \frac{1}{N/k+1}\sum_{l\geq -N/2k}^{l\leq N/2k} Re(k,l) \tag{13}$$

In Fig. 5 (a) an experimental pressure signal and its derivative have been plotted. Fig. 5 (b) shows the result of the convolution $c(x, V)$ applied to one cycle of the experimental data. The variable $x = i\Delta\alpha$ is the crank angle CA, and $V = k/N$.

If the convolution is to be calculated at an intermediate point between two experimental data, which is how it has been used in this work, the expression to be applied in each interval z between two experimental points,$p(z^-)$ and $p(z^+)$, corresponds to that indicated in Eq. (13).

$$c(z,k) = \frac{k}{N}\sqrt{\frac{18}{7}}\sum_{l=0}^{l\leq \frac{N}{2k}+\frac{1}{2}}[p(z^+ + l) + p(z^- - l)]Re^{'}(k,l) + j[p(z^+ + l) - p(z^- - l)]\,\mathrm{Im}'\left(k,l\right)$$
$$Re^{'}(k,l) = \cos\left[\frac{\pi k}{N}\left(l+\frac{1}{2}\right)\right]\left\{\cos\left[\frac{2\pi k}{N}\left(l+\frac{1}{2}\right)\right] - \frac{1}{3}\right\} - \overline{Re}(k) \tag{14}$$
$$Im'\left(k,l\right) = \cos\left[\frac{\pi k}{N}\left(l+\frac{1}{2}\right)\right]\sin\left[\frac{2\pi k}{N}\left(l+\frac{1}{2}\right)\right]$$

The summation has been reduced to half terms considering the symmetry of the sine and cosine functions.

Considering that when the convolution takes values below a selected level (filter level) there is no information of the pressure signal and only noise is present. This transformation can be used as a tool to determine the cutoff frequency to be applied at each crank angle using a *SGF*.

Under the assumption that the convolution decreases with frequency, which is only false when the convolution values are near zero, once the filtering level has been chosen, it is not necessary to calculate all the values of $c(i,k)$ at each crank angle to determine the corresponding cutoff frequency V_c. It is enough to find the two values of the module of $c(i,k)$ above and below the level and interpolate

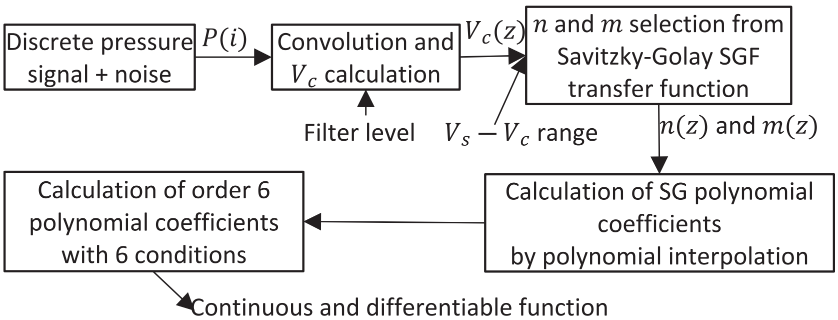

Fig. 6. Scheme of the procedure followed in this work to obtain a continuous and derivable function in the whole domain from discrete data.

to calculate V_c. At the next angular position, the V_c search is started at the value of the previous data point. This procedure decreases the number of $c(i,k)$ calculations to be made.

With this V_c calculation strategy, the application of this adaptive method for the choice of the *SGF* at each point would involve the realization of a number of weighted averages, N_{wa}, before applying the weighted average of the *SGF*. Therefore, with respect to a fixed *SGF*, the computational cost is multiplied by $N_{wa} + 1$.

2.4. Using APF and obtention of a continuous and derivable signal in the whole domain

The final result of the application of the methodology described in this article is an interpolation polynomial between every two experimental data obtained from the filtering of the experimental data. The set of all the polynomials (each one used in its interval) is a continuous and derivable function in the entire domain. Fig. 6 schematizes this methodology, which is described below.

For this purpose, in each interval z between two experimental points, z^- and z^+, as described in section 2.3, the cut-off frequency V_c is calculated as the frequency from which the value of the convolution $c(z,k)$, in this case $c(CA,V)$, (in the same units of the original signal) is less than a certain chosen level.

The criterion used to select this level with the data used in this work is presented in section 3.1, and it is illustrated in Fig. 12, where also the selection of $V_s - V_c$ range value is illustrated. From the V_c and $V_s - V_c$ selected values for each interval, according to the methodology explained in section 2.2, the order and number of points are chosen.

This choice of n and m will implies the same filtering level of filtering of the signal in the interval, as if an *SGF* were used. Note that a SG filtering with an average weighted by the SG coefficients is not applied, but rather it is necessary to calculate the coefficients of the interpolation polynomial.

As a different polynomial is chosen for each interval, on the intervals ends, i.e. at the experimental data points, the polynomial (n, m) changes so continuity and derivability are not ensured.

With each polynomial it is possible to calculate six data, three values and three derivatives, at the two interval ends and at the center of the interval. At each experimental point two polynomials take values, to avoid a possible discontinuity, the mean value of these two values is assigned. In each interval, these 6 conditions suppose 6 linear equations that allow to determine the 6 coefficients of a polynomial of order 5. These polynomials defined for each interval have the property that the polynomials of two consecutive intervals have the same value and the same derivative at the point of union of the two intervals, which guarantees the continuity and derivability in the whole domain. A graphical representation of what this new polynomial supposes is presented in Fig. 7.

3. Results and discussion

3.1. APFanalysis using simulation-generated data

Using a predictive model developed in AVL BOOST 2021, two *SIE* simulations have been carried out. One of them corresponds with low combustion speed and high engine speed, this means that the characteristic harmonics of the phenomenon are at low frequencies. The other one with high combustion speed and low engine speed, consequently with higher frequency characteristic harmonics. This results in two pressure diagrams with different gradients in the combustion zone. In the slow frequency combustion process (slow cycle) natural gas is used as fuel with 42.9 % excess air at 2500 rpm, this corresponds to a combustion duration of 110° (NG_0.7_2500). In the fast frequency combustion process (fast cycle), the combustion of hydrogen (a fuel with a much higher combustion speed) has been simulated with 100 % excess air at 1000 rpm, resulting in a combustion duration of 55° (H2_0.5_1000). The simulation results are obtained with the same angular resolution as the experimental pressure records ($\Delta\alpha = 0.6^\circ$). These two configurations are common in natural gas (NG) and hydrogen (H2) engines [29] in order to limit NOx emissions and to have stable operation.

A random noise of ± 25 mbar has been added to the simulated pressure signals. This noise is of the order of the noise found in the experimental tests of [2,30]. To analyze *APF* efficiency, different filters have been applied to the noisy pressure signals and then processed with an inverse predictive model to calculate the rates of heat released (*RoHR*) during combustion Eq. (15). The *RoHR* obtained after filtering have been compared with the original *RoHR* without noise that was used to obtain the pressure diagrams.

Fig. 8 shows the original pressure plots of the two cycles with and without noise and their derivatives.

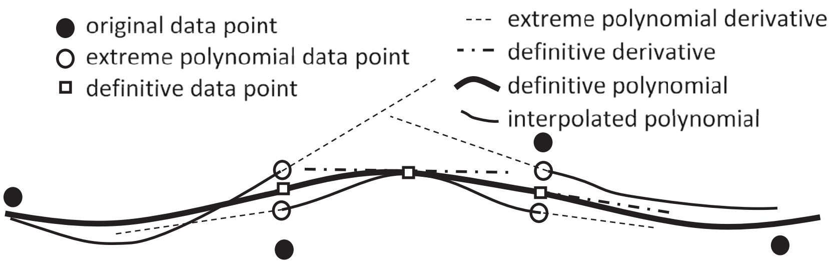

Fig. 7. Graphical representation of the continuous and derivative polynomial construction. The extreme definitive data points are the mean of extreme polynomial data point. The same for derivative values.

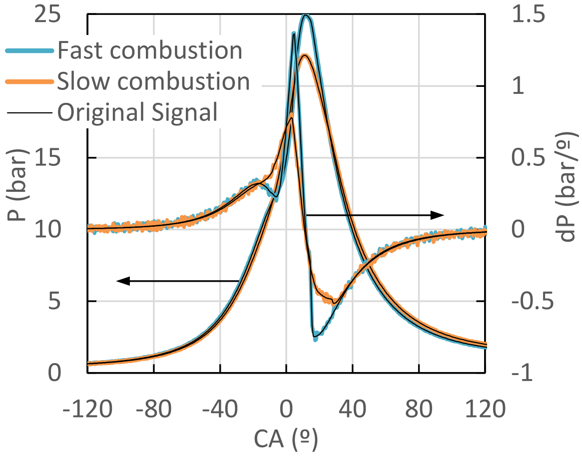

Fig. 8. Pressure signals and their derivatives of the simulated and randomly noisy, slow and fast cycles.

These two simulations with added noise have been filtered with the methodology of this work. Fig. 9a show the dimensionless harmonic V_c from which the function $c(CA, V)$ takes values lower than 5 mbar (this value selection will be explained onwards with Fig. 12). Also shows the order n and the number of points m selected to have a transfer function with a cutoff frequency value as close as possible to V_c for a range of $V_s - V_c = [0.06, 0.08]$, for a fast combustion H2_0.5_1000, and for a slow combustion NG_0.7_2500.

It can be observed that the maximum V_c is obtained around 0°, where the highest-pressure gradients exist. This zone is where combustion occurs. Away from this zone the V_c values are much lower, indicating that there are no process characteristic harmonics of high frequency. In the results corresponding to fast combustion, the V_c values are higher during the combustion process, which results in higher n and lower m values.

Fig. 9b shows the maximum of V_c calculated for each of the 500 cycles. It can be seen how each cycle has a different maximum V_c due to the cycle by cycle variation. CCV makes each cycle under the same engine operating conditions to have a different combustion speed.

Eq. (15) presents the relationship between the two variables in the differential equation used in the two models, each variable being cleared in its case. It can be seen how the *RoHR* depends linearly on both the pressure and its derivative.

$$RoHR = \frac{\gamma p \frac{dV}{dt} + V \frac{dp}{dt}}{\gamma - 1} - \dot{Q}_w \tag{15}$$

In order to evaluate the efficiency of the filtering methodology, this section follows the process schematized in Fig. 10. Firstly, pressure signals are simulated from *RoHR*. Latter noise is added to the simulated signals. The methodology is applied to the noisy signals obtaining a filtered pressure signal. Finally, the combustion diagnostics is performed to obtain *RoHR* that can be compared with the original simulated *RoHR*.

Fig. 11 shows the *RoHR* obtained using the original pressure signal, the original pressure signal with noise, the filtered signal with a *SGF* of $n = 5$ and $m = 10$ (*SGF*-5–10), the filtered signal with a *FF* with normalized cutoff frequency $V_c = 0.1$ and normalized filter order $V_s - V_c = 0.11$ (*FF*-0.1–0.1). The signal filtered with *APF* using a filter level of 5 mbar and a $V_s - V_c = [0.1, 0.12]$ is also presented. Fig. 11a shows the comparison of the *RoHR* of slow cycle and Fig. 11b of fast cycle.

In Fig. 11, three zones can be distinguished. In the first one there is no combustion and it is divided in two parts: before combustion during the compression process and after combustion during the expansion process. The second zone is the beginning and end of combustion, where the *RoHR* takes low values and consequently the pressure gradients are not very high. Finally, the third zone is located where the highest-pressure gradients exist and therefore the highest *RoHR* values. In the third zone about 90 % of the fuel is burned.

In order to compare the results of the different filtering techniques, it must be kept in mind that the only adaptive filter is the APF and the other two have a fixed transfer function for all points of the cycle. An FF filter with parameters $V_c = 0.061$ and $V_s - V_c = 0.07$ has a transfer function practically equal to that of the SG-5–10 filter. These values can be calculated from n and m in Fig. 3.

Fig. 11 must be understood as the comparison of an APF filter and two FF filters, one FF with a higher V_c, $V_c = 0.1$, and $V_c = 0.1$ (legend *FF* −0.1 −0.1), and another with a lower V_c (legend *SGF*-5–10). In zone 1, where pressure gradients are low, the *APF* removes lower frequencies than the other filters, so the result is closer to the original *RoHR* signal.

Something similar happens in zone 2, however, as the high frequency harmonics of the signal increase their value, the *APF* increases V_c, thus the three filters behave in a similar way. At the end of the combustion, as the pressure gradients decrease, the *APF* performance improves with respect to the other two filters.

In the third zone, in the case of fast combustion speed (Fig. 11b), the *FF* and the *APF* have a similar behavior. However, in the slow combustion (Fig. 11a), possibly the cutoff frequency of the *FF* would be too high so it has a worse performance than the *APF*.

In order to compare the *APF* and the *FF* in a quantitative way, the root mean square error (*RMSE*) between the *RoHR* of the filtered

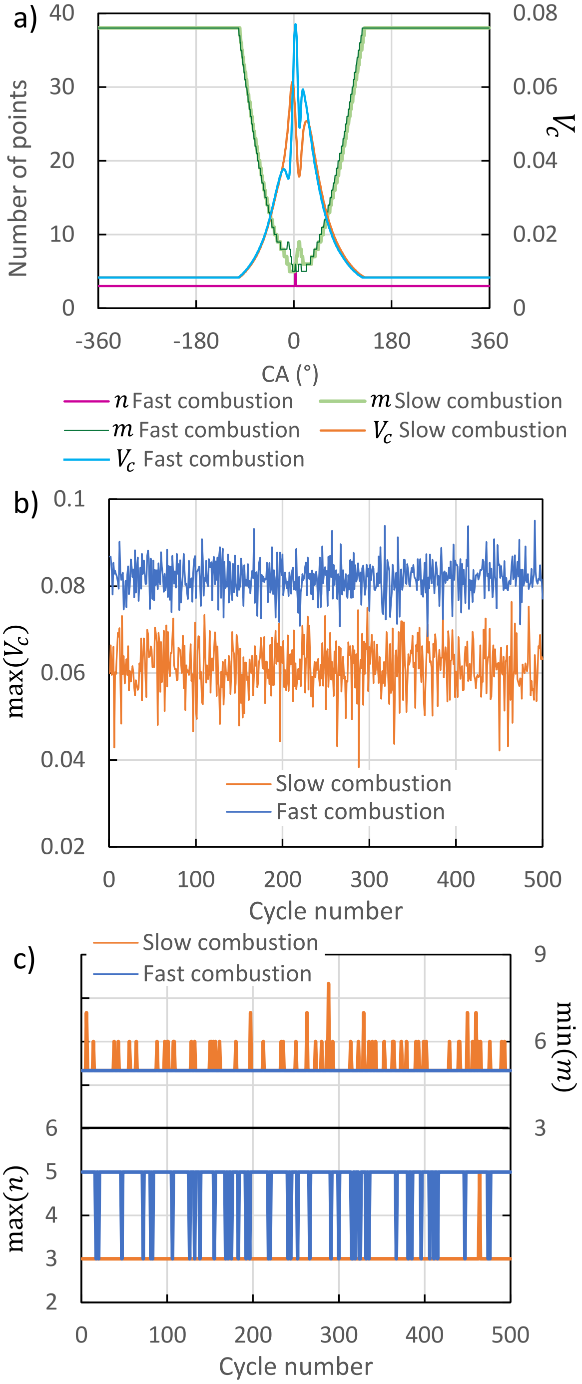

Fig. 9. a) V_c, m and n obtained with the proposed procedure for a value of $c(z,k) = 5$ *mbar* as a function of the crank angle, applied to an experimental signal of slow combustion, NG_0.7_2500, and fast combustion, H2_0.5_1000, (section 3.2). b) Maximum V_c for each cycle for slow and fast combustion. c) Minimum number of points m and maximum polynomial order n, for each cycle, for slow and fast combustion.

signal and the original signal without noise was calculated. The lowest *RMSE* value means the best filtering. Fig. 12 shows the results obtained, on the one hand, the *RMSE* values of the *APF* versus its filtering level (upper horizontal axis), and on the other hand, the *RMSE* of the *FF* versus its cutoff frequency V_c (lower horizontal axis). When the filtering level in the *APF* decreases, the value of V_c increases, so the direction of growth of the horizontal axes is different. With both filters, the more to the right, the less the signal is filtered.

Different values of $V_s - V_c$ have been used for both filters. Fig. 12a shows the results obtained in the slow cycle and Fig. 12b the

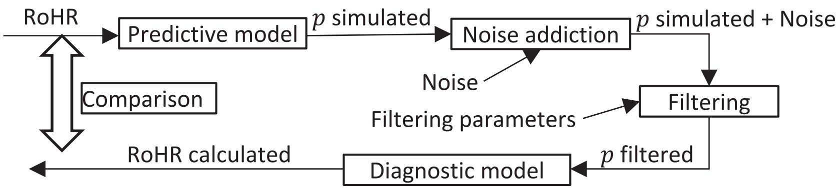

Fig. 10. Schematic of the process followed to quantify and evaluate the different filters on the simulated pressure signals with noise.

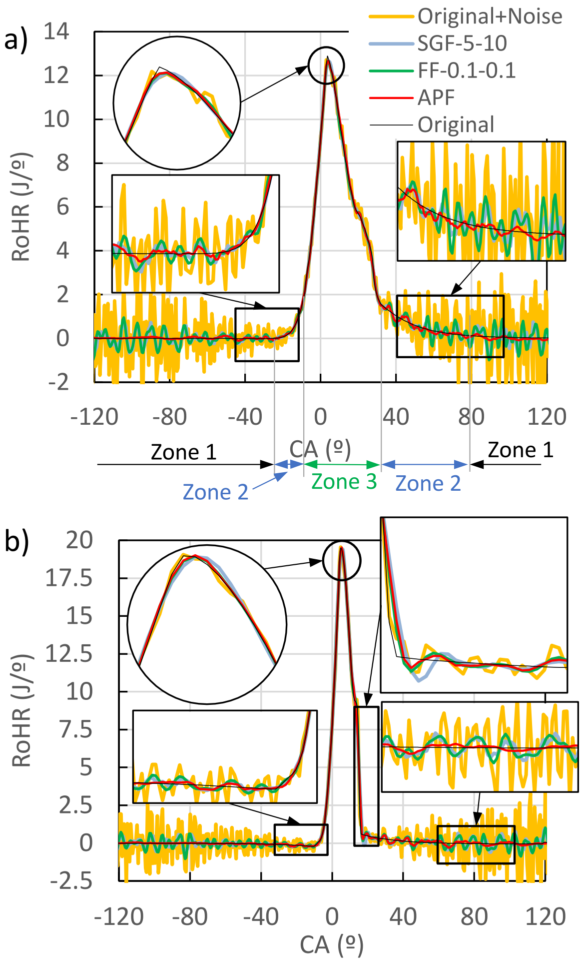

Fig. 11. *RoHR* as a function of crank angle, original, with noise and obtained with the filters: *SGF*-5–10, *FF*-0.1–0.1 and *APF* with a filter level of 5 mbar and $V_s - V_c = [0.1, 0.12]$. (a) Slow combustion process. (b) Fast combustion process.

results obtained in the fast cycle. It has also been indicated in the two figures where the SG-5–10 filter would be with the corresponding values of $V_c = 0.06$ and $V_s - V_c = 0.07$.

Fig. 12 shows that the errors are lower in the slow cycle than in the fast cycle. The filter order $V_s - V_c$ affects the *FF* results, even changing its trend. However, it has much less influence on the *APF*.

The optimal cutoff frequency in the case of the *FF* depends on how the signal is. It is different for the fast cycle (Fig. 12b) than for the slow cycle (Fig. 12a). However, in *APF*, the optimal filtering level is independent whether the process is fast or slow and is situated for the data used at around 5 mbar. This is a great advantage and novelty of the filter used in this work, since it is possible to have a single criterion for all signals in contrast to [31,32] that propose cutoff and stop frequencies dependent on the engine speed and load.

For *APF* applied to signals with high pressure gradients (Fig. 12b), low $V_s - V_c$ values lead to inconsistent results when varying the filtering level. It is not advisable to use low $V_s - V_c$ values.

In accordance with the results obtained, a value of 5 mbar as the filter level and a $V_s - V_c = [0.1, 0.12]$ have been chosen.

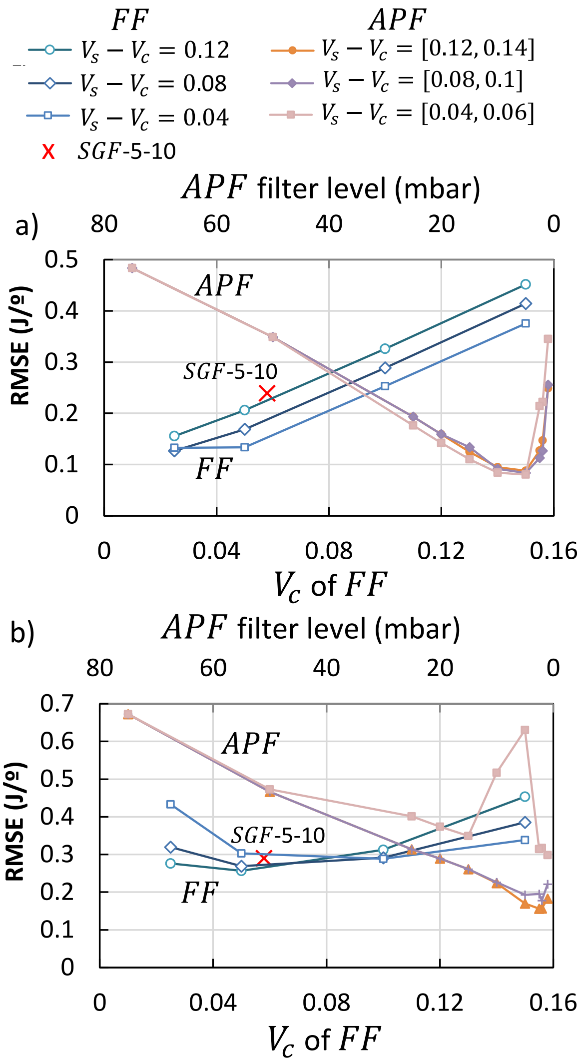

Fig. 12. Comparisons between the *RMSE* of the *APF* versus its filtering level and of the *FF* versus its V_c, for different values of $V_s - V_c$. (a) Slow combustion process. (b) Fast combustion process. Filled markers correspond to the series of *APF* results. The empty markers correspond to the *FF* results series.

3.2. APFanalysis using experimental data.

In this section, the results of different filters applied to experimental data are compared. When experimental data are filtered, the noise-free signal is not known, so evaluating the quality of the filtering is very difficult. In the first section, a novel methodology, only valid for cyclic processes, is presented, based on the use of two parameters to compare the filters. In the second section, a qualitative comparison is made of the information obtained once the experimental data has been processed for the particular case of this work, pressure records used to calculate the *RoHR*.

3.2.1. Quantitative comparison of filters

The difference between the experimental and filtered signals is the sum of the noise (*noise*) plus the error in the filtering process (*error*). This difference is defined as $L(\alpha)$:

$$L(\alpha) = p_{exp}(\alpha) - p_{filt}(\alpha) = p_{exp}(\alpha) - p_{real}(\alpha) + p_{real}(\alpha) - p_{filt}(\alpha) = noise + error$$

The attenuation is defined as the mean value of the absolute value of this difference, Eq. (16).

$$Atten = \frac{1}{N_{fin} - N_{ini}} \sum_{\alpha_{ini}}^{\alpha_{fin}} |L(\alpha)| \tag{16}$$

Where $N_{fin} - N_{ini}$ corresponds to the points number used to calculate the attenuation. The attenuation quantifies the severity of the applied filter, but not the quality because it includes filtering error and noise.

The higher the attenuation, the more noise is removed, but it is possible that information about the process being measured is also removed. The problem is to determine the lost process information da (*error*). If the mean value of $L(\alpha)$ is calculated over the entire domain, it will be equal to the mean value of the error, because the mean value of the noise is zero $\overline{L(\alpha)} = \overline{noise} + \overline{error} = 0 + \overline{error}$.

Probably $\overline{error} = 0$, however, if a cyclic signal is being excessively filtered and removing relevant information from the process, it is very likely that at one point in the cycle a non-zero error is being committed in mean value, that is, at that point, the real signal will be in most cycles, either above or below the poorly filtered signal.

If in that case, an average is made by differentiating each instant of the cycle, the mean cyclic difference $MCD(\alpha)$, Eq. (17), the mean error will not be zero at that point.

$$MCD(\alpha) = \frac{1}{N_{cycles}} \sum_{i=1}^{i=N_{cycles}} L(\alpha, i) \tag{17}$$

$MCD(\alpha)$ quantifies the average error that is being systematically committed at a point in the cycle. The lower the $MCD(\alpha)$, the better the quality of the filtering, while the attenuation quantifies the amount that has been filtered, the higher the attenuation, the more severe the filtering. Note that if nothing is filtered, $MCD(\alpha) = 0$ and $Atten = 0$.

Table 1 shows the attenuation value of three *SGF* for 100 cycles in the crank angle range from -120° to 120°. It is found that the more severe the filter, the greater the attenuation.

In order to determine N_{cycles} to be used in Eq. (17), Fig. 13 presents the maximum absolute value of $MCD(\alpha)$ within crank angle interval from -120° to 120° as a function of the number of cycles analyzed for NG_0.7_2500 and H2_0.5_1000. This zone is the one of greatest interest since it involves the compression, combustion and expansion processes, see Fig. 15. It can be observed how after 100 cycles the value remains practically constant, in the order of 24 mbar in H2_0.5_1000 and 12 mbar in NG_0.7_2500. Therefore, a value of $N_{cycles} = 100$ has been chosen for the calculation of $MCD(\alpha)$.

To compare two filters on the same signal, one criterion for selecting filter settings is that they have the same attenuation and then compare the $MCD(\alpha)$ value.

Fig. 14a shows the attenuation value of *APF* as a function of the filtering level with $V_s - V_c = [0.1, 0.12]$ and the attenuation value of *FF* as a function of V_c for $V_s - V_c = 0.11$ in the crank angle range from -120° to 120°. Fig. 14b shows for the same configuration on both filters the maximum absolute value of $MCD(\alpha)$.

The attenuation of *FF* with $V_c = 0.05$ is similar to that of *APF* with a filtering level of 5 mbar (Fig. 14). On the other hand, a *SGF*-5–10 (Table 1) presents attenuation values similar to those of *APF* with a filtering level of 5 mbar. $MCD(\alpha)$ of these three filters are comparable each other since they attenuate the signal at similar levels.

Fig. 15 shows the value of $MCD(\alpha)$ for the filters *APF*, *FF* and *SGF* with a similar attenuation. In NG_0.7_2500 there are no remarkable differences between the filters, all of them have absolute values of $MCD(\alpha)$ below 10 mbar. In this case no filter has systematic errors. In H2_0.5_1000, the same can be said in the compression and expansion zones. However, in the combustion zone, all filters have the same tendencies in the systematic errors that exist, but the amplitude of the *APF* oscillations are smaller.

3.2.2. RoHRcomparison using different filters.

Fig. 16 shows a comparison of *RoHR* results obtained after applying the diagnostic model, Eq. (15), from filtered pressure signal. Results are presented for the experimental signal without filtering and with different filters: *SGF*-5–10 that would correspond with a $V_c = 0.058$ and $V_s - V_c = 0.061$ (Fig. 3); *APF* with a filtering level of 5 mbar and $V_s - V_c = [0.1, 0.12]$; and *FF* with $V_c = 0.05$ and $V_s - V_c = 0.11$. The three filters have similar attenuation levels.

Results of a *FF* obtained by applying the procedure proposed by Payri et al. [17] to the selection of V_c are also presented. In this case, the cutoff frequency results in $V_c = 0.16$ in H2_0.5_1000 and $V_c = 0.04$ in NG_0.7_2500. The same $V_s - V_c = 0.11$ has been used.

In the compression and expansion zone, zone 1 in Fig. 11, the real value of *RoHR* is known that is zero because there is no combustion. *APF* has the lowest fluctuations around the real value in this zone. This is due that *APF* selects lower cutoff frequencies for these zones than any other filter. The behavior of the rest of the filters is dependent on the cutoff frequency they use.

In the combustion zone, zone 3 in Fig. 11, the real value of *RoHR* is not known and it is not possible to decide which filter is better. The analysis made in the previous section indicates that the *APF*, in the case of fast combustion, has a better performance than the rest of the cycles in this zone.

In the start and end of combustion zone, zone 2 in Fig. 11, it is known that combustion processes are slow and therefore the *RoHR*

Table 1
Attenuation values (mbar) for 100 cycles from -120° to 120° obtained with different *SGF*: *SGF* with $n = 6$ and $m = 4$ (*SGF*-6–4), *SGF* with $n = 5$ and $m = 10$ (*SGF*-5–10) and *SGF* with $n = 3$ and $m = 20$ (*SGF*-3–20). For tests NG_0.7_2500 and H2_0.5_1000.

Filter	NG_0.7_2500	H2_0.5_1000
SGF-6–4	5.57	4.9
SGF-5–10	8.41	9.08
SGF-3–20	22.67	70.22

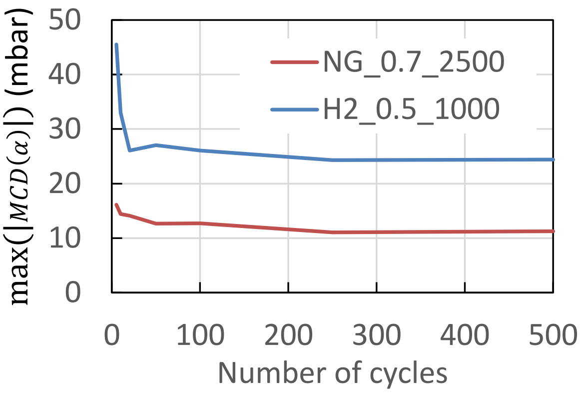

Fig. 13. Maximum value of $|MCD(\alpha)|$ in the crank angle interval between -120° and 120° versus the number of cycles used for the calculation of $MCD(\alpha)$. Data filtered with the *APF* with a filtering level of 5 mbar and $V_s - V_c = [0.1, 0.12]$ have been used in GN_0.7_2500 and H2_0.5_1000 tests.

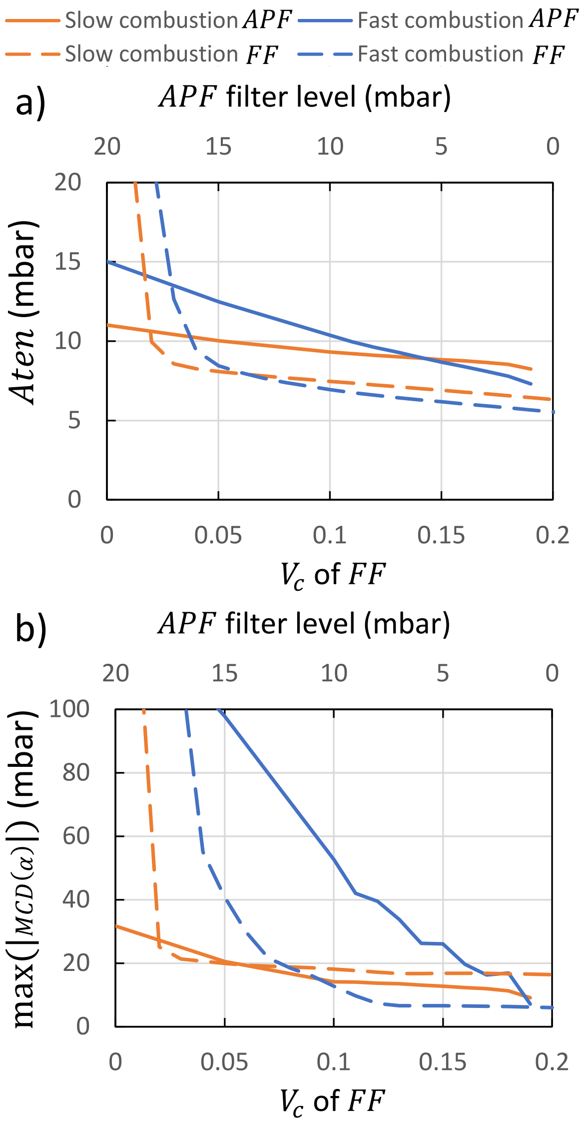

Fig. 14. Filter attenuation (a) and maximum value of $|MCD(\alpha)|$ (b) in the crank angle interval -120° to 120° for NG_0.7_2500 and H2_0.5_1000. *APF* vs filtering level with $V_s - V_c = [0.1, 0.12]$. *FF* vs V_c with $V_s - V_c = 0.11$.

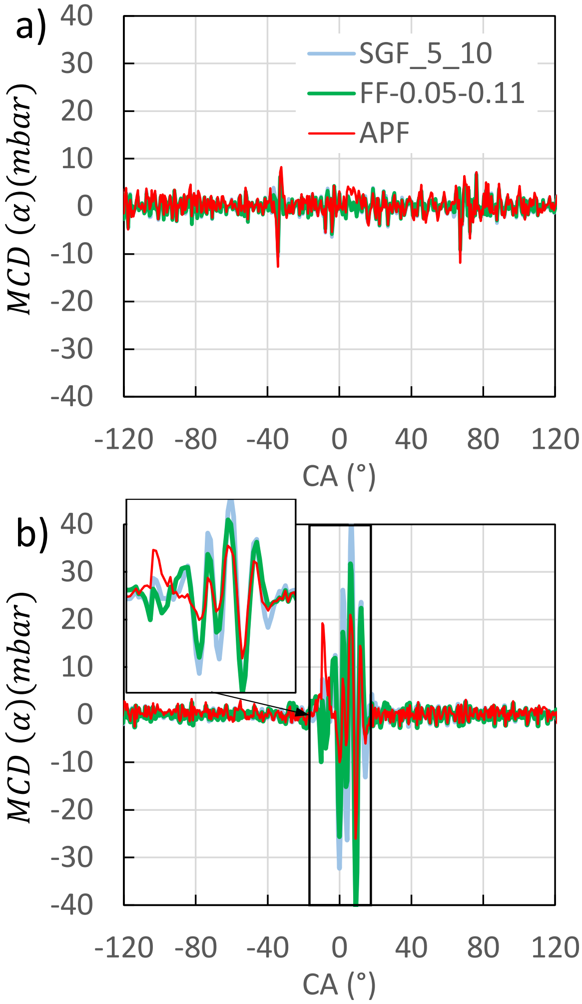

Fig. 15. $MCD(\alpha)$as a function of crank angle for the filters: *APF* with a filtering level of 5 mbar and $V_s - V_c = [0.1, 0.12]$, *FF* with $V_c = 0.05$ and $V_s - V_c = 0.11$ and *SGF* with $n = 5$ y $m = 10$. (a) NG_0.7_2500. (b) H2_0.5_1000.

does not have fluctuations. The filter that behaves best in this area is also the *APF* since it is the one in which the signal has the least fluctuations.

4. Conclusions

The main objective of this work is to obtain a continuous and derivable function that can be used to perform data diagnosis. A novel methodology has been developed to not only obtain a continuous and derivable function but also to filter the noise present in the signal. For this, the work carried out in this paper can be summarized in the following points:

- Firstly, the filtering properties of polynomial fits have been studied in the frequency domain. The information obtained is used in the methodology to select interpolation polynomials by controlling the cutoff frequency.
- To select the cutoff frequency applied at each time instant a novel convolution transformation is proposed.
- Finally, in order to obtain a continuous and derivable function in the whole domain, a transformation of the polynomials is performed. The last transformation consist in fitting in each interval between two experimental data points a 5-order polynomial fulfilling 6 conditions (no degrees of freedom) guarantying continuity and derivability in the interval extremes and in the central point.

This methodology is intended to be used to filter pressure signals recorded in the combustion chamber of an internal combustion engine *ICE*. Two types of results have been obtained, results from simulated data and results from experimental data.

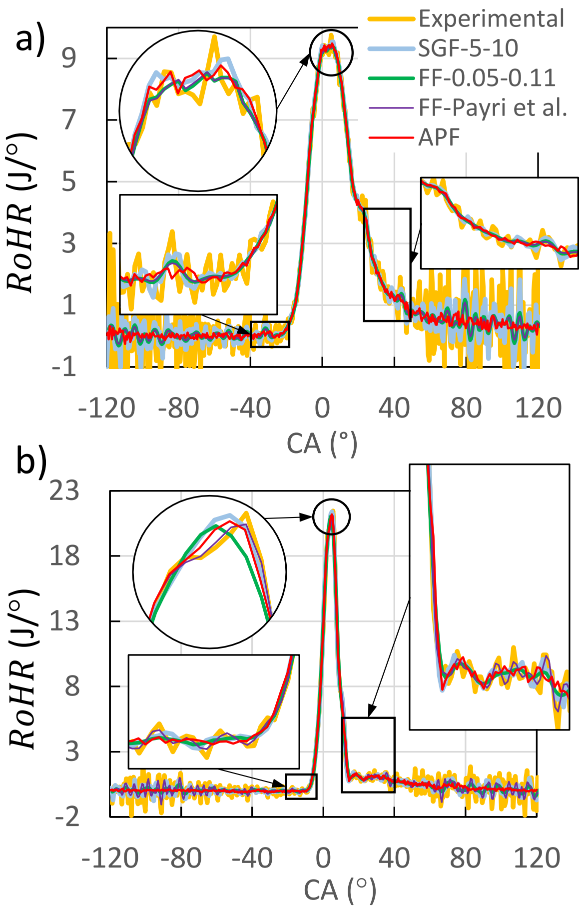

Fig. 16. *RoHR* versus crank angle, experimental and obtained with filters: *SGF*-5–10, *FF*-0.05–0.11, *FF*-Payri et al. and *APF* with a filtering level of 5 mbar and $V_s - V_c = [0.1, 0.12]$. (a) NG_0.7_2500. (b) H2_0.5_1000.

- Two signals have been generated through process simulation in order to validate the method developed in this work. One signal corresponds to a slow combustion process and the other corresponds to a fast combustion process. Noise of a level similar to the experimental observed was artificially added to the simulated signals. The noisy signal has been filtered with: the filter proposed in this work *APF*, filters based on frequency decomposition *FF* and Savitzky-Golay filters *SGF*. The results obtained after the filtered signals diagnosis have been evaluated and compared with the original simulated signal results (without noise).
- The filter proposed in this work has also been used on the experimental pressure signals of [2,30]. In order to evaluate the result fit of the different filters to the experimental data, the use of two quantifiers has been proposed. The first quantifier proposed, mean cyclic dispersion ($MCD(\alpha)$), serves to analyze the cycle averaging of the noise removed from the original signal with each filter. The second quantifier proposed, attenuation *Aten*, serves to analyze the information removed. The *MCD* of different filters (*APF*, *FF* and *SGF*) with the same attenuation, has been analyzed. The best results are obtained using *APF*.

The conclusions obtained by the analysis of the results are:

- The relation between the pairs [cutoff frequency, filter order] and [polynomial order, number of interpolation points] is determined by performing the discrete Fourier transform (*DFT*) of the *SGF* coefficients.
- The use of a preferred filter order for the selection of polynomial order and number of interpolation points, prevents abrupt changes in the order and number of points of the polynomial for consecutive data. In case of only attending to the polynomial with the closest cutoff frequency, abrupt changes could occur (Fig. 3).
- The convolution applied to experimental data allows to perform a frequency decomposition at each angular interval between two data. The convolution result, obtained at each interval, is used to select the cutoff frequency of the filter to apply. The convolution units are the same as the original data if the comparison pattern function $g(x)$ is properly selected. Thus, the convolution value used to choose the cutoff frequency has a quantitative physical meaning.
- The computational cost of performing the convolution for all frequencies is high. But it is not necessary to calculate it at all frequencies for each point. The filter computation time can be drastically reduced by starting the search for the cutoff frequency for

each interval from the cutoff frequency of the previous interval. The *APF*, compared to a fixed *SGF*, would multiply the computational cost by the number of times the convolution was performed at each point plus 1, but would have the great advantage of being adaptive. In addition, if a continuous and derivable function is desired, this implies an additional computational cost of a multilinear regression and the resolution of a system of six linear equations.
- Applying *APF* on simulated signals gives better results than using *FF* or *SGF*. In addition, the optimal *APF* cutoff level is practically independent of the chosen filter order $V_s - V_c$ and the signal characteristics (fast or slow combustion process). This is not true for *FF* since the optimal cutoff frequency depends on the signal characteristics to be filtered and the filter order.
- $MCD(\alpha)$ is only applicable to cyclic signals sampled at the same instant of the cycle, as is the case for chamber pressure data from an *ICE*.
- When filtering experimental data, the interesting point is $MCD(\alpha)$ taking values close to zero while the attenuation takes high values. $MCD(\alpha)$ decreases as fewer cycle characteristic frequencies are removed, i.e. the higher is the cutoff frequency. On the other hand, the attenuation increases as more noise is removed, i.e. the lower the cutoff frequency. The improvement of both quantizers implies opposite trends in the cutoff frequency, so a compromise solution concerning the filtering level must be found.
- All filters based on polynomials let some of high frequencies pass through. *APF* is the filter that filters out the high frequencies most uniformly. *FF* eliminates them completely.
- Unlike the other analyzed filters, *APF* treats each cycle zone differently, therefore, it is able to remove more noise in the zones of the cycle where the signal-to-noise ratio is lower, while it does not remove characteristic harmonics of the process when the signal-to-noise ratio is higher.

To sum up, the result of the methodology proposed in this work is a continuous and derivable function that allows detailed analysis of specific areas. For instance, when using the data as a boundary condition for differential equations systems integration with variable step numerical methods. Moreover, the methodology only uses data from the pressure signal without considering other operating conditions. In addition, the same methodology configuration is optimum for extreme test conditions. The proposed method offers a moderate computational cost solution for filtering periodic signals with cycle-to-cycle variations. For all these reasons, it can be a useful tool for pressure data treatment in *ICE*.

Future works could deal with the use of this filtering methodology under different engine operating conditions and operating with different fuels. Since the methodology uses the signal recorded, if the pressure variations of new tests are in the range of the pressure variations considered in this work this methodology should work without needing any change. If the pressure variations of new tests are outside the range of the pressure variations considered in this work it would be necessary to study the results of this methodology. Also, the simulation of the pressure data acquisition system is interesting in order to study the effect of pressure signal distortions in the methodology proposed.

CRediT authorship contribution statement

Pedro Gabana: Writing – original draft, Visualization, Software, Resources, Project administration, Investigation, Conceptualization. **Blanca Giménez:** Writing – original draft, Visualization, Resources, Investigation, Formal analysis, Conceptualization. **Andrés Melgar:** Software, Methodology, Investigation, Formal analysis. **Alfonso Horrillo:** Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Much of the research work carried out has been financed with funds from the Spanish government through the project “Analysis and characterization of dual fuel combustion for the reduction of CO2 emissions in the transport sector” (PID2019-106957RB-C22).

P. Gabana has been funded by the 2021 call for predoctoral contracts from the University of Valladolid, co-financed by Banco Santander.

The authors would like to thank AVL Advanced Simulation Technologies for allowing the use of the AVL BOOST program in carrying out this work.

Data availability

Data will be made available on request.

References

[1] D. Descieux, M. Feidt, One zone thermodynamic model simulation of an ignition compression engine, Appl. Therm. Eng. 27 (2007) 1457–1466, https://doi.org/10.1016/j.applthermaleng.2006.10.002.

[2] B. Giménez, A. Melgar, A. Horrillo, F.V. Tinaut, A correlation for turbulent combustion speed accounting for instabilities and expansion speed in a hydrogen-natural gas spark ignition engine, Combust. Flame 223 (2021) 15–27, https://doi.org/10.1016/j.combustflame.2020.09.026.
[3] X. Duan, B. Deng, Y. Liu, Y. Li, J. Liu, Experimental study the impacts of the key operating and design parameters on the cycle-to-cycle variations of the natural gas SI engine, Fuel 290 (2021) 119976, https://doi.org/10.1016/j.fuel.2020.119976.
[4] R. Davis, J. Duncan, S. Gopujkar, J. Worm, Evaluation of Cylinder Pressure Transducer Performance Including the Influence of Mounting Location and Thermal Protection, In (2022), https://doi.org/10.4271/2022-01-5014.
[5] X. Jiang, K. Kim, S. Zhang, J. Johnson, G. Salazar, High-Temperature Piezoelectric Sensing, Sensors 14 (2013) 144–169, https://doi.org/10.3390/s140100144.
[6] M. Lapuerta, O. Armas, J.J. Hernández, Diagnosis of DI Diesel combustion from in-cylinder pressure signal by estimation of mean thermodynamic properties of the gas, Appl. Therm. Eng. 19 (1999) 513–529, https://doi.org/10.1016/S1359-4311(98)00075-1.
[7] J. Gomesantunes, R. Mikalsen, A. Roskilly, An investigation of hydrogen-fuelled HCCI engine performance and operation, Int. J. Hydrogen Energy 33 (2008) 5823–5828, https://doi.org/10.1016/j.ijhydene.2008.07.121.
[8] . Hamedović, F. Raichle, J. Breuninger, W. Fischer, W. Dieterle, M. Klenk, J.F. Böhme, IMEP-Estimation and In-Cylinder Pressure Reconstruction for Multicylinder SI-Engine by Combined Processing of Engine Speed and One Cylinder Pressure, in: 2005. doi: 10.4271/2005-01-0053.
[9] F. Liu, G.A.J. Amaratunga, N. Collings, A. Soliman, An Experimental Study on Engine Dynamics Model Based In-Cylinder Pressure Estimation, in: 2012. doi: 10.4271/2012-01-0896.
[10] Y. Shimasaki, M. Kobayashi, H. Sakamoto, M. Ueno, M. Hasegawa, S. Yamaguchi, T. Suzuki, Study on Engine Management System Using In-cylinder Pressure Sensor Integrated with Spark Plug, In (2004), https://doi.org/10.4271/2004-01-0519.
[11] M.F.J. Brunt, C.R. Pond, J. Biundo, Gasoline Engine Knock Analysis using Cylinder Pressure Data, in: 1998. doi: 10.4271/980896.
[12] H. Xu, X. Ni, X. Su, C. Weng, C. Yao, The effect of ignition intensity and in-cylinder pressure on the knock intensity and detonation formation in internal combustion engines, Appl. Therm. Eng. 200 (2022) 117690, https://doi.org/10.1016/j.applthermaleng.2021.117690.
[13] E. Pipitone, A. Beccari, Determination of TDC in internal combustion engines by a newly developed thermodynamic approach, Appl. Therm. Eng. 30 (2010) 1914–1926, https://doi.org/10.1016/j.applthermaleng.2010.04.012.
[14] J.A. Gatowski, E.N. Balles, K.M. Chun, F.E. Nelson, J.A. Ekchian, J.B. Heywood, Heat Release Analysis of Engine Pressure Data, in: 1984. doi: 10.4271/841359.
[15] R. Kumar Maurya, D.D. Pal, A. Kumar Agarwal, Digital signal processing of cylinder pressure data for combustion diagnostics of HCCI engine, Mech. Syst. Sig. Process. 36 (2013) 95–109, https://doi.org/10.1016/j.ymssp.2011.07.014.
[16] F. Payri, J.M. Luján, J. Martín, A. Abbad, Digital signal processing of in-cylinder pressure for combustion diagnosis of internal combustion engines, Mech. Syst. Sig. Process. 24 (2010) 1767–1784, https://doi.org/10.1016/j.ymssp.2009.12.011.
[17] F. Payri, P. Olmeda, C. Guardiola, J. Martín, Adaptive determination of cut-off frequencies for filtering the in-cylinder pressure in diesel engines combustion analysis, Appl. Therm. Eng. 31 (2011) 2869–2876, https://doi.org/10.1016/j.applthermaleng.2011.05.012.
[18] J. Cuisano, F. Flores, L. Chirinos, A. Vaudrey, In-cylinder pressure statistical analysis and digital signal processing methods for studying the combustion of a natural gas/diesel heavy-duty engine at low load conditions, Energy Convers Manag 269 (2022) 116089, https://doi.org/10.1016/j.enconman.2022.116089.
[19] E.L. Kosarev, E. Pantos, Optimal smoothing of "noisy" data by fast Fourier transform, J. Phys. E 16 (1983) 537–543, https://doi.org/10.1088/0022-3735/16/6/020.
[20] C. McGurk, H. Ahmed, M. Foo, A. Pike, Q. Lu, D.S. Laila, A comparative analysis of moving average filter and Kalman filter for large diesel engine test cell back-pressure control, Int. J. Engine Res. (2023) 146808742211457, https://doi.org/10.1177/14680874221145749.
[21] Abraham. Savitzky, M.J.E. Golay, Smoothing and Differentiation of Data by Simplified Least Squares Procedures, Anal. Chem. 36 (1964) 1627–1639, https://doi.org/10.1021/ac60214a047.
[22] I.M. Gogolev, J.S. Wallace, Study of Assisted Compression Ignition in a Direct Injected Natural Gas Engine, J Eng Gas Turbine Power 139 (2017), https://doi.org/10.1115/1.4036968.
[23] Phillip. Barak, Smoothing and Differentiation by an Adaptive-Degree Polynomial Filter, Anal. Chem. 67 (1995) 2758–2762, https://doi.org/10.1021/ac00113a006.
[24] H. Huang, S. Hu, Y. Sun, A Discrete Curvature Estimation Based Low-Distortion Adaptive Savitzky–Golay Filter for ECG Denoising, Sensors 19 (2019) 1617, https://doi.org/10.3390/s19071617.
[25] S.R. Krishnan, C.S. Seelamantula, On the Selection of Optimum Savitzky-Golay Filters, IEEE Trans. Signal Process. 61 (2013) 380–391, https://doi.org/10.1109/TSP.2012.2225055.
[26] M. Sadeghi, F. Behnia, R. Amiri, Window Selection of the Savitzky–Golay Filters for Signal Recovery From Noisy Measurements, IEEE Trans. Instrum. Meas. 69 (2020) 5418–5427, https://doi.org/10.1109/TIM.2020.2966310.
[27] R.W. Schafer, On the frequency-domain properties of Savitzky-Golay filters, in: 2011 Digital Signal Processing and Signal Processing Education Meeting (DSP/SPE), IEEE, 2011: pp. 54–59. doi: 10.1109/DSP-SPE.2011.5739186.
[28] A. V. Oppenheim, R. W. Schafer, Discrete: Time Signal Processing, 3rd ed., Pearson, Upper Saddle River, NJ, 2010.
[29] Dr. P. Kapus, Dr. R. Heindl, M. Weißbäck, Dr. P. Grabner, High Efficiency Hydrogen Internal Combustion Engine – Carbon Free Powertrain for Passenger Car Hybrids and Commercial Vehicles, in: 43rd International Vienna Motor Symposium, Vienna Motor Symposium, Vienna, 2022.
[30] B. Giménez, A. Melgar, A. Horrillo, P. Gabana, Prediction of the flame kernel growth rate in spark ignition engine fueled with natural gas, hydrogen and mixtures, Fuel 339 (2023) 126908, https://doi.org/10.1016/j.fuel.2022.126908.
[31] J. Martín, Aportación al diagnóstico de la combustión en motores Diesel de inyección directa, Universidad Politécnica de Valencia, 2007.
[32] O. Armas, Diagnostico experimental del proceso de combustión en motores Diesel de inyección directa, Universidad Politécnica de Valencia, 1998.