In-cylinder pressure smart sampling for efficient data management

Benjamín Pla, Joaquín De la Morena, Pau Bares and Irina Jimenez

Abstract

The present paper proposes an optimized algorithm for minimizing the amount of data processed in order to maintain all critical information from the in-cylinder pressure sensor but with minimum computational cost. The algorithm uses singular value decomposition (SVD) for reducing the number of samples and pivoting QR decomposition to identify the optimal sampling locations.

Experimental data from four engines with different combustion modes, namely Spark ignition (SI), Compression ignition (CI), turbulent jet ignition (TJI), and dual fuel ignition (DFCI), was used to validate the algorithm. The impact of the different sampling methodologies on different metrics for engine performance has been addressed and studied showing negligible information loss when reducing to 25 samples per cycle the acquisition buffer size.

Keywords

Smart sampling, in-cylinder pressure, combustion diagnosis, data processing

Introduction

On-road emissions monitoring of automotive vehicles with portable emission systems (PEMS), as well as the inclusion of new stringent regulations for NOx, particles and CO_2 pushes the automotive industry to better detect aging and combustion dispersion in the fleet in real time by improving the current On-board Diagnostics (OBD) systems with a better analysis of the combustion.[1,2]

In-cylinder pressure sensors offer detailed information of the combustion process. The low frequency content of the in-cylinder pressure signal is critical to analyze the heat release rate (HRR) or the indicated mean effective pressure (IMEP),[3] while the high frequency content is used to estimate the resonance intensity.[4] The HRR is commonly used to estimate the combustion evolution, which is used together with the IMEP to provide feedback from the engine determine the optimal combustion settings for an adequate closed loop combustion control.[5,6] The resonance intensity is frequently estimated through the maximum amplitude pressure oscillation (MAPO) and it can be useful to identify the levels of vibration and even maintain the knock occurrence with stochastic controls in desired levels.[7,8]

The in-cylinder pressure signal can be used as input to feed several models that estimate other control parameters, such as in Khameneian et al.[9] to estimate the temperature and air distribution for a GDI spark ignition engine, in Murić et al.[10] to estimate the emitted NOx from the HRR and the unburned temperature, in d'Ambrosio et al.[11] to estimate the combustion noise, in Bares et al.[12] to predict the knock probability, in Broatch et al.[13] and Youssef[14] to estimate the trapped mass, or in Martos et al.[15] to estimate the soot by semi-empirical models. These models can be used to replace current sensors, to improve the accuracy of the estimation in a sensor data fusion scenario,[16] or to identify the system dynamics in order to design advanced controls.[17]

The potential of the information extracted from the in-cylinder pressure might be critical for the future on-board diagnosis (OBD) systems, which are being adapted to incorporate the connection of the fleet by point-to-point (P2P) or internet of things (IoT) strategies. Current researches are already developing

CMT-Motores Térmicos, Universitat Politècnica de València, Valencia, Spain

Corresponding author:
Pau Bares, CMT-Motores Térmicos, Universitat Politècnica de València, Cami de Vera S/N, Valencia, 46022, Spain.
Email: pabamo@mot.upv.es

algorithms for that future environment: Hamid et al.[18] used five data parameters, engine speed, manifold air pressure, load-fuel, barometric pressure, and engine temperature, to detect anomalies sharing information of a fleet by IoT, Wang et al.[19] used Vehicle to X (V2X) communication to share traffic data information for OBD prediction, Weis et al.[20] used IoT to determine the car range by OBD and the possibilities of refueling and Jeong et al.[21] proposes a self-diagnosis System (ISS) for an Autonomous Vehicle based on an Internet of Things (IoT) Gateway and Deep Learning.

However, processing such amount of information in commercial electronic control units (ECUs), or the interconnection of such units with external computers, might be a challenge for the current technology.[22] On the one hand, in-cylinder pressure sampling is commonly performed crank angle based, as the pressure needs to be properly phased with the volume evolution for a precise HRR computation.[23] On the other hand, sufficient time resolution is required to differentiate resonance (around 5–25 kHz) from the normal combustion.[24] Because of that, an adequate sampling frequency resolution is achieved with five samples per degree (24 kHz at 800 rpm and 60 kHz at 2000 rpm). The world light-duty rest procedure (WLTP) cycle consists in following an engine speed and torque profile over more than 30 min, which results in 18,119 combustion cycles. Assuming an engine with four cylinders and acquiring with 16 bit resolution (3 mbar), the minimum capacity required will be 0.694 GB. Assuming that the WLTP is representative of normal driving, that means a rate of more than 20 MB/minute of memory capacity. Therefore, procedures to compress cylinder pressure signal and extract the sensible information are crucial for processing and sharing combustion features and for real-time combustion control.

The problem addressed in the following paper (smart sampling) consists of finding the number of samples required and the location of such samples to capture the essential information with minimum loss. Because of the complex phenomena involved in combustion, machine learning (ML) seems to be appropriate to capture the main patterns of the pressure trace, while sparse sampling techniques are required to reconstruct the actual signal from a small subset of measurements. Data-driven sparse sampling techniques have been extensively used to compress images and audio files, and more recently, the research of such techniques is being exploited to optimize the sensor placement in dynamic systems.[25] Single value decomposition (SVD) techniques have already being used to determine the fluid dynamics with minimum sampling,[26,27] and pivoted QR decomposition has been proven an efficient technique to reduce the sampling in dynamic systems.[28,29]

The present paper uses single value decomposition for extracting the main features of combustion from the in-cylinder pressure sensor and determines the optimal samples location by using pivoted QR decomposition in order to minimize any loss of information. The method is applied to different engines and combustion concepts, and the trade-off between sampling size and the accuracy in capturing different combustion metrics is analyzed. The content of the article is structured as follows: next section is devoted to present the engines used, section three explains the algorithm developed, then, a discussion of the results is presented in section four, and finally, some conclusions are summarized in the last section.

Experimental set-up

The study presented bellow has been developed with a database obtained in the research institute of CMT-Motores Térmicos. The dataset consists in four parametric studies covering most of the operating range in four engines with different combustion modes, namely SI, CI, DFCI, and TJI.

- Spark Ignited (SI) combustion: A conventional spark ignited gasoline direct injection (GDI) engine was tested at 1000, 2000, and 3000 rpm at various loads by performing variations in the throttle (varying lambda) and the spark advance (varying the combustion phasing). In-cylinder pressure of 54 tests with 200 cycles at each operating condition were recorded and post-processed with a resolution of five samples per crank angle degree (CAD) in the four cylinders.
- Compression Ignited (CI) combustion: A parametric analysis of a CI engine was performed to study the effect of different configurations of exhaust gases recirculation (EGR) by combining the low-pressure (LP) EGR and the high-pressure (HP) EGR. A total of 54,636 cycles of in-cylinder pressure data was obtained from nine operating conditions, at three loads (20%, 40%, and 60%) and at 1750, 2500, and 3000 rpm. At each operating point various configurations of low pressure, high pressure, and variable gas turbine were tested. The pressure was sampled each 0.2 CAD in all four cylinders.
- Turbulent Jet Ignited (TJI) combustion: A classical configuration of SI engine was modified to include a passive pre-chamber in a single-cylinder research configuration to analyze the potential of such combustion mode (TJI).[30] The engine was run at three engine speeds, namely 1350, 2000, and 3000 rpm. At each engine speed the injected fuel, the intake pressure and the spark timing were varied, ranging from misfires to maximum efficiency operation (60,468 cycles). The pressure was sampled each 0.2 CAD in all four cylinders.
- Dual Fuel Compression Ignited (DFCI) combustion: This engine was equipped with gasoline port fuel injection (PFI) and diesel direct injection (DI) to perform different dual fuel strategies of

Table 1. Engines specifications.

Combustion	SI	CI	DFCI	TJI
Displaced volume [cc]	1300	2200	7700	404
Stroke [mm]	81.2	96	135	80.5
Bore [mm]	72	85	110	80
Compression ratio [-]	10.6	16	12.2	13.4
Number of cylinders	4	4	6	1
Fuel injection system	GDI	DI	PFI-DI	PFI

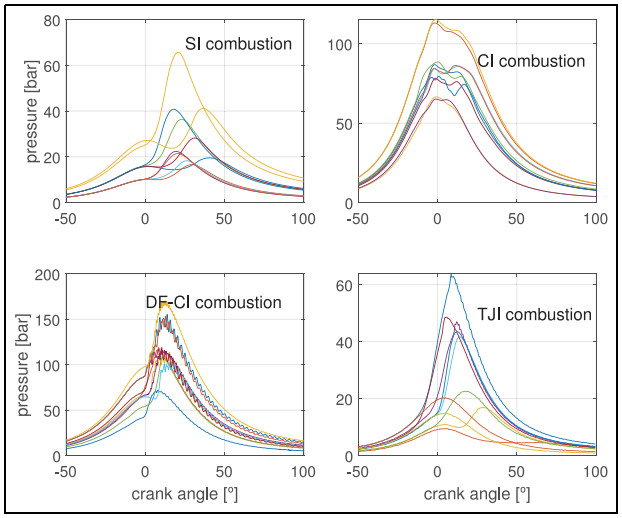

Figure 1. Pressure examples of the four engines.

combustion, ranging from partially-premixed compression ignited combustion (PPCI) to reactivity controlled compression ignition (RCCI).[6] In-cylinder pressure data was sampled each 0.2 CAD at six engine speeds, from 950 to 2200 rpm, by varying the quantity of each fuel and performing different combustion phasing and loads (3591 cycles).

At each operating condition 60% of the cycles were used for training purposes while the remaining cycles were used to compute the final error metrics. The main engine specifications are provided in Table 1 while Figure 1 shows 10 cycles randomly selected from the dataset in order to illustrate each combustion features.

Method description

There are tight in-cylinder pressure sampling requirements to perform combustion analysis and diagnostics. It is generally accepted that a crank angle resolution below 0.5 crank angle degrees is required to accurately compute the heat release rate calculation, and sampling frequencies above 3000 Hz are needed to perform knock evaluation.[31] However, Figure 1 shows well-defined structures suggesting that, any of the in-cylinder pressure sets shown, admits a low rank representation. It is intuitive that the pressure will increase during the compression stroke and decrease during expansion, even more, if during those processes a politropic evolution can be assumed, then pressure signal can be reconstructed from the gas condition in a single point and the politropic coefficient. In this sense, Figure 1 shows that if the pressure traces are represented in the proper coordinate system, the signal can be expressed using a limited number of parameters containing the contributions of the main modes. Admitting the previous statement then, a step further would be to use a limited number of samples and reconstruct the complete signal using previous knowledge from the main modes. The question arising now is where should be those samples placed in the cylinder cycle to reproduce the complete signal with minimum information loss.

According to the previous ideas, the proposed method should address the following challenges:

- Find a low rank representation of the space of the feasible in-cylinder pressures in a given engine.
- Find what are the best sampling points for the chosen basis.
- Reconstruct the pressure evolution in the cycle from a limited set of measurements.

Next sections address both topics by following the method proposed in Manohar et al.[25]

SVD for computing the in-cylinder pressure basis

Consider that a dataset with the in-cylinder pressure signals for n cycles of dimension m (number of samples per cycle) is provided. Hereinafter, we will assume that the set is large enough to cover, to a reasonable extent, all the feasible in-cylinder pressure traces. Then, the following matrix (P) can be constructed:

$$P = [p_1, p_2 \ldots p_n] \tag{1}$$

where the columns of P (that has $m \times n$ elements) are the pressure recordings (with m elements) of any of the n cycles.

The matrix P can be expressed, as the product of two matrices:

$$P = \Psi S = [\psi_1, \psi_2 \ldots \psi_m][s_1, s_2 \ldots s_n] \tag{2}$$

where Ψ has m modes of dimension m, and S is composed from n vectors that determine the intensity of each mode at each cycle (in the case at hand $m < n$). While several matrices Ψ and S may fulfill equation (2) ($\Psi = I$ and $S = P$ is the trivial case), the idea behind such transformation is to find a coordinate system allowing to reduce the number of modes, so the size of Ψ, leading to:

$$P = \Psi_r S_r = [\psi_1, \psi_2 \ldots \psi_r][s_1, s_2 \ldots s_n] \tag{3}$$

where $r < m$, Ψ_r has r modes of dimension m and S_r is composed from n vectors of r values containing the coordinates of the n pressure traces of P in the Ψ_r basis.

Although, in general, P will be a full-rank matrix, the structure shown in Figure 1 should allow the approximation with a lower rank basis. The Eckart–Young theorem[32] shows that the Singular Value Decomposition (SVD) provides the best possible approximation of range r for the matrix P in the sense of minimum least-square error. The result is a representation composed of a orthonormal matrix of m modes (Ψ), a diagonal matrix composed from m singular values (Σ), and n unitary vectors that characterize each cycle (V).[33]

$$P = \Psi \Sigma V \tag{4}$$

Amongst the SVD properties, one may highlight that, as V is unitary, the resulting singular values σ_j offer direct information about each mode importance. Moreover, since the singular values are located in the Σ diagonal in decreasing order, the modes in Ψ are in order of decreasing importance. In this sense, provided a given number of modes (r) it is straightforward to choose the best possible approximation of P with rank r as:

$$P = \Psi_r \Sigma_r V_r \tag{5}$$

where Ψ_r contains the first r columns of Ψ, Σ_r is the upper-left submatrix of order r, and V_r contains the first r rows of V.

Pivoting QR factorization for optimal sample placement

Provided that Ψ_r consists of r linearly independent vectors, r measurements are needed to determine the intensity of the r modes. The problem now is to choose the location of the r samples which allow to reconstruct the complete signal of size m from r measurements with minimum information loss according to the patterns in Ψ_r. To this aim, consider an in-cylinder pressure evolution p of size m. A measurement with r samples can be considered as an extraction of r of the m values in p, or in matrix multiplication form:

$$y = Cp \tag{6}$$

where y is a vector of size r containing the pressure measurements and C is a sparse extraction matrix of size $r \times m$ with a single element equal to one in every row and the rest of elements equal to 0. Note that the ones in C determine the samples taken from p, for example, a one in the entrance ij of C means that the i^{th} measurement is the j^{th} value of p. Assuming that p is in the space of Ψ_r it can be written as linear combination of its rows, leading to:

$$y = C\Psi_r \alpha \tag{7}$$

where α is a vector with r components containing the coefficients of p in the basis Ψ_r (the contribution of any of the Ψ_r modes in p). Provided that the number of measurements has been chosen equal to the number of modes (r), the coefficients in α are perfectly determined from the measurement (y), the basis (Ψ_r) and the measuring (extraction) points (C) as:

$$\alpha = (C\Psi_r)^{-1} y \tag{8}$$

Note that the values of α are dependent on the computation of the inverse of $C\Psi_r$, and therefore on the selection of the measurement locations (C). To make the inverse of $C\Psi_r$ as well-conditioned as possible, C should be chosen in a way that minimizes its condition number, that is, the ratio between the maximum and minimum singular value. Taking into account that the condition number is related to the spectral characteristics of the matrix, Optimal Design of Experiments are aimed to maximize different matrix metrics that tend to minimize the condition number. In this sense D-Optimal design maximizes the determinant of the matrix $C\Psi_r$, while M-Optimal maximizes its trace and A-Optimal maximizes its minimum singular value. Both Optimal Design of Experiments or directly finding the sampling points that minimize the condition number of the matrix $C\Psi_r$ are optimization problems with high computation cost. Pivoting QR decomposition has been proposed as an alternative that tend to maximize the determinant of $C\Psi_r$, so leading to a well conditioned matrix.[25]

QR factorization decomposes a matrix with full column rank into the product of a unitary matrix and an upper triangular matrix, therefore, the determinant of the original matrix is the product of the diagonal of the triangular matrix.[33] While there are several QR decomposition methods, pivoting QR factorization introduces a column permutation matrix to order the elements in the triangular matrix diagonal in decreasing order. Recalling $M = \Psi_r^T$, the pivoting QR factorization leads to:

$$MC^T = QR \tag{9}$$

$$|R_{11}| \leqslant |R_{22}| \leqslant \ldots |R_{rr}| \tag{10}$$

where C is the column permutation matrix, Q is unitary, and R_{jj} are the elements in the diagonal of the upper triangular matrix R. The pivoting method is based on maximizing $|R_{ii}|$ at each step, thus globally tending to maximize the determinant of R.[34,35] In this sense, the matrix C obtained from the pivoting QR factorization of Ψ_r is a good candidate for the extraction matrix, leading to measuring positions, since maximizes the determinant of $C\Psi_r$.

Pressure signal reconstruction

Once the matrix of measurement locations (C) is chosen and the measurement vector y is taken, the intensity of the modes in Ψ_r can be obtained by applying equation (8). Finally, the complete pressure signal p can be reconstructed as:

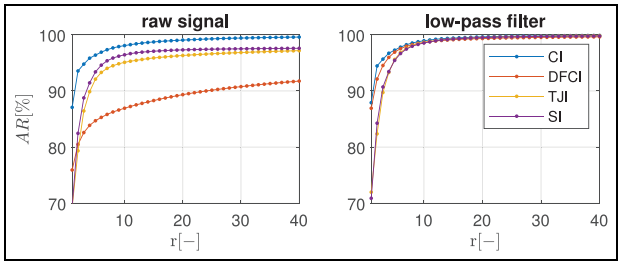

Figure 2. Accumulated relevance of modes: Left plot with no filter, right plot with lowpass filter.

$$p = \psi_r \alpha \tag{11}$$

Results and discussion

For comparison purposes the signal of the combustion stroke (from −180° to 180°) has been processed. Furthermore, the dataset has been re-sampled to 5 samples/° in all the engines by using linear extrapolation in engines B and C. As consequence, all the pressure traces used for training have 1800 samples per cycle and were also used for benchmarking.

The heat release rate (HRR) has been estimated by using the apparent heat release with $\kappa = 1.3$, such as:

$$AHRR = \frac{\kappa}{\kappa - 1} pdV + \frac{1}{\kappa - 1} Vdp \tag{12}$$

where V is the chamber volume, p the pressure evolution, and the derivatives (dp and dV) where obtained as a first order forward differentiation.

The apparent heat release rate (AHHR) was used to estimate the combustion evolution and find the position where the 10%, 50%, and 90% of combustion (CA10, CA50, and CA90 respectively) was reached. Also, the indicated mean effective pressure (IMEP) was computed by integrating the pressure and the volume derivative between 180° before and after the top dead center (TDC).

The combustion features, namely CA10, CA50, CA90, and IMEP, are characterized by the low frequency content of the pressure signal so it is normal to apply a low-pass filter with 3000 kHz cutoff frequency as the content above use to be only associated to resonance or noise.[36] Nevertheless, the intensity of the pressure oscillations might be critical for SI engines to avoid excessive knock,[4] or in CI engines to analyze the emitted noise. The pressure oscillation intensity is commonly measured by MAPO which has been estimated by identifying the maximum of the high-frequency pressure content (above 3 kHz).

According to the previous paragraphs, the present work is aimed to analyze the error in the combustion parameters that are usually computed from the in-cylinder pressure (peak pressure, IMEP, CA10, CA50, CA90, and MAPO) when the method proposed in section 3 is applied to different engine types.

Figure 2 shows the accumulated relative relevance (AR) of r modes, which is computed by adding all the weight of the first r modes and dividing by the total weight of all the modes according to:

$$AR[\%] = \frac{\sum_{j=1}^{r} \sigma_j}{\sum_{j=1}^{m} \sigma_j} 100 \tag{13}$$

Left plot shows the result when applying SVD over the raw signal, while right plot shows the result of the same procedure but using only the content of the frequencies below 3 kHz.

The high frequency content does not perturb the intensity distribution of the modes in traditional combustion modes, such as SI or CI combustion, but it has a complete different effect in the TJI or DFCI combustion that are subject to pressure variations at higher frequencies as shown in the examples of Figure 1.

Figure 2 shows that 25 modes can cover most of the information (above 95%) in the raw pressure signal in SI or CI combustion but loses part of the information in engines with a fast combustion, such as TJI and specially DFCI. This can be explained by the resonance excitation: In the case of SI combustion a flame front covers all the combustion chamber without a sudden heat release (if there is no knock) and in CI engines the combustion is ignited far from the TDC to minimize NO_x, and hence, it has no important resonance excitation. On the contrary, DF-CI performs a nearly homogeneous ignition to benefit from low NO_x and high efficiency and TJI aims to accelerate the flame front evolution with a jet ignition created in a pre-chamber. But such fast combustions heavily excites resonance leading to large oscillations that can be appreciated in Figure 1.

The number of modes retained is a trade-off between the computational cost and size of the database (Ψ_r) versus the accuracy. The singular values are directly linked to the mode, the first four modes are represented in Figure 3. It must be noticed that the first mode is the average of all the cycles while the rest of them configure the final shape of the pressure trace in the various conditions tested.

Finally, pivoting QR decomposition was used to find the adequate location to minimize the error. The optimal locations for determining the modes are shown in Figure 4, where the intensity of the colors determines the preference, ranging from dark blue (more important) to light yellow (less important).

The distribution of the sampling locations depend on the combustion itself:

- In conventional combustions modes (CI and SI) sampling get estranged. On CI ignition, a combination of pilot main and post injections distribute the combustion along the expansion stroke to reduce NOx by maintaining the flame temperature low. On SI ignition the variability of combustion forces

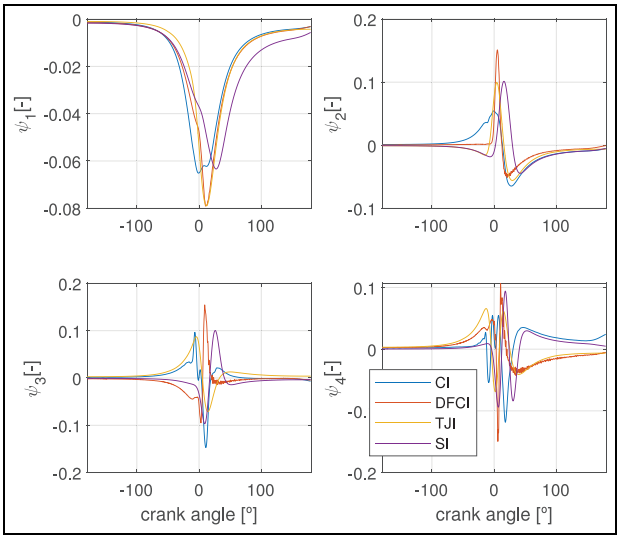

Figure 3. First four modes for each engine by using the raw signal.

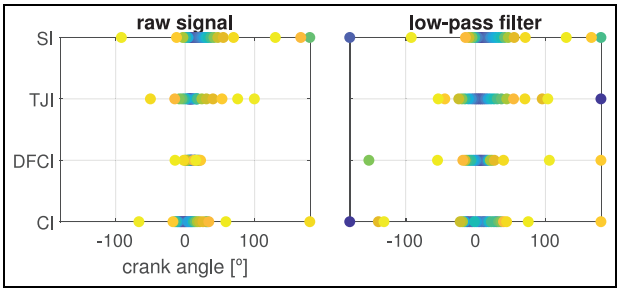

Figure 4. First 25 locations to capture the mean features of each combustion mode.

the algorithm to sample the pressure in several locations leading to a similar distribution that in CI.

- In DFCI or TJI combustion, where heat is intended to be fast released, the sampling is located near the expected combustion phasing, which uses to be some degrees after the TDC.
- Results also indicate that space between samples tend to get shorter when high frequency content needs to be captured, specially in HCCI-like combustion (TJI and DF-CI).

Once the sampling locations and the representative modes of the SVD are defined by a proper training dataset, pressure can be sampled only at such locations to determine the intensity of each mode (ψ_j) and hence, reconstruct the complete pressure signal.

Note, that the method provides the optimal sampling that is able to represent the combustion features though few modes. Henceforth, the information of these few points is also a potential input for applying machine learning control techniques in real time. That means that, once the method have detected the optimal sampling positions, the heat release rate computation can be substituted by other simplified models, such as ANN, and substantially reduce the number of operations.

Figure 5 shows the reconstruction of a randomly selected cycle from the validation dataset in Engine B by using only 25 samples. The Root Mean Square Error (RMSE) is 300 mbar for the raw signal and 26 mbar for the low pressure content. However, it must be said that the error is concentrated on the phase of the oscillation signal, and hence, it does not affect most of the parameters used for combustion (MAPO, IMEP, CAXX).

Figure 6 shows the Root Mean Square Error (RMSE) in % as a function of the number of modes used. The error was normalized by using the time of combustion (CA90−CA10) for the combustion phasing and the average of the measured value in the IMEP, the pmax or the MAPO.

The estimation of the parameters do not substantially differ after 50 samples and 25 samples seems a reasonable value to estimate most of the parameters. The errors obtained in all the engines by sampling the pressure with 25 samples over the validation dataset have been tabbed in Tables 2 to 5. The upper part of the tables contains the RMSE while the low part contains the relative root mean square error in %. In order to benchmark the proposed method, the error obtained with an equispaced sampling is also included in the tables.

It can be observed how the proposed method outperforms the equispaced sampling, since the modes stored in Psi_r provide an sparse basis for the in-cylinder pressure and the pivoting QR algorithm approximates the

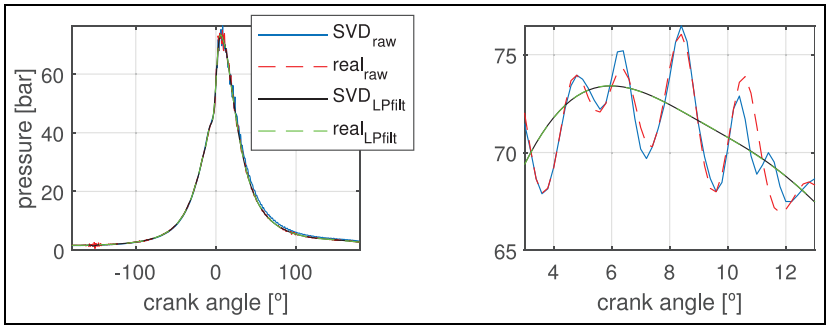

Figure 5. Pressure reconstruction with 25 samples.

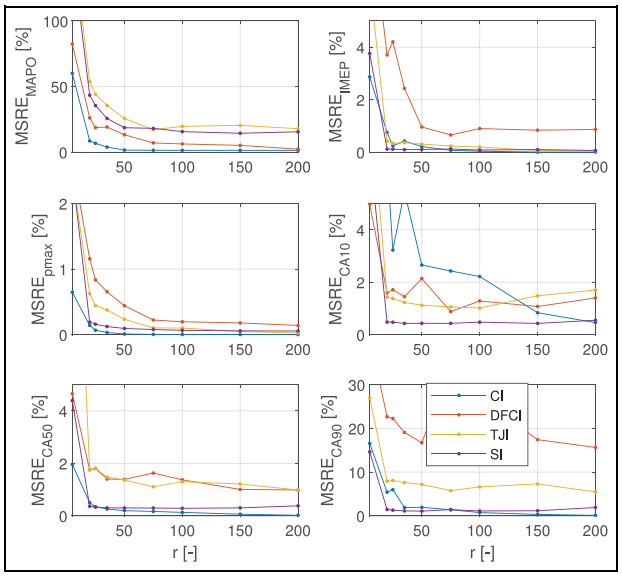

Figure 6. Relative error of the mean parameters as a function of the number of samples and modes used for reconstruction.

Table 2. Errors at the control parameters after reconstruction (CI combustion).

		Equispaced	SVD_{raw}	SVD_{filt}
CA10	[CAD]	5.96	1.88	1.78
CA50	[CAD]	3.02	0.15	0.08
CA90	[CAD]	3.51	2.50	0.71
IMEP	[mbar]	57.02	29.75	19.97
p_{max}	[mbar]	837.54	55.29	10.09
CA10	%	13.57	4.29	4.06
CA50	%	6.88	0.33	0.18
CA90	%	7.99	5.70	1.61
IMEP	%	0.63	0.33	0.22
p_{max}	%	1.02	0.07	0.01

Table 3. Errors at the control parameters after reconstruction (DFCI combustion).

		Equispaced	SVD_{raw}	SVD_{filt}
CA10	[CAD]	1.92	0.29	0.07
CA50	[CAD]	4.30	0.31	0.05
CA90	[CAD]	6.32	3.76	0.25
IMEP	[mbar]	194.64	584.65	19.88
p_{max}	[mbar]	10351.58	98.11	21.64
CA10	%	11.14	1.68	0.38
CA50	%	24.91	1.79	0.29
CA90	%	36.64	21.77	1.44
IMEP	%	1.33	3.98	0.14
p_{max}	%	8.81	0.83	0.02

best sample locations according to the Psi_r basis. Note that the proposed algorithm allows the estimation of the maximum of the in-cylinder pressure with an error below 1% and mostly associated to the oscillations of resonance. Regarding the combustion evolution, the center of the combustion (CA50) is captured with

Table 4. Errors at the control parameters after reconstruction (TJI combustion).

		Equispaced	SVD_{raw}	SVD_{filt}
CA10	[CAD]	2.81	0.36	0.17
CA50	[CAD]	2.33	0.48	0.13
CA90	[CAD]	4.98	2.14	0.67
IMEP	[mbar]	59.76	17.90	5.30
p_{max}	[mbar]	3490.50	149.58	11.67
CA10	%	10.62	1.38	0.65
CA50	%	8.82	1.81	0.48
CA90	%	18.83	8.11	2.53
IMEP	%	1.12	0.34	0.10
p_{max}	%	10.32	0.44	0.03

Table 5. Errors at the control parameters after reconstruction (SI combustion).

		Equispaced	SVD_{raw}	SVD_{filt}
CA10	[CAD]	1.86	0.09	0.07
CA50	[CAD]	1.92	0.06	0.04
CA90	[CAD]	4.48	0.24	0.14
IMEP	[mbar]	103.57	11.56	12.43
p_{max}	[mbar]	3784.10	67.32	30.74
CA10	%	9.92	0.47	0.38
CA50	%	10.22	0.33	0.24
CA90	%	23.85	1.28	0.73
IMEP	%	1.00	0.11	0.12
p_{max}	%	8.94	0.16	0.07

precision in all engines, the start of combustion (CA10) is identified in SI, TJI, and DF-CI combustion with a precision in the range of the crank resolution (0.2 CAD), but in the CI engine an error around 2 CAD might be expected due to the characteristic multiple injections. The end of the combustion (CA90) is accurately estimated in SI combustion but errors between 2 and 4 CAD are expected in the rest because of the slow evolution of the late combustions.

The Maximum Amplitude Pressure Oscillation (MAPO) was estimated by filtering the in-cylinder pressure and computing the maximum of the oscillations with a frequency content above 5 kHz. Henceforth, 25 samples equispaced sampling and SVD method after filtering are not suitable for MAPO evaluation. When comparing the SVD method with 3600 samples equispaced sampling the average error in MAPO is 6.74%, 18.72%, 44.07%, and 35.48% in CI, DFCI, TJI, and CI combustion, respectively. Nevertheless the error is mostly due to the cycle to cycle variability and when a window of 50 samples is used the estimation of MAPO is accurate enough to control the engine. MAPO has a high cycle-to-cycle variability and normally its average value is commonly controlled by stochastic controllers, henceforth, a white noise does not perturb the controller as it is based on stochastic control.

Using a low-pass analog filter would substantially improve the measurements as the sampling location

would be specifically used to determine the parameters associated with combustion, however, information from the oscillation will be lost. Special attention must be paid when combustion heavily excites resonance. In such conditions, a low-pass filter becomes critical to estimate IMEP in order to avoid excessive sampling in the area of combustion where resonance requires from a high sampling rate to identify the high frequency spectrum.

Conclusions

A new methodology for reducing the computational burden of in-cylinder pressure signal processing has been addressed. The Optimal sensor placement algorithm described in Manohar et al.[25] has been adapted for optimal crank-angle sampling of in-cylinder pressure. The algorithm suggests using singular value decomposition for reducing the dimensions of the pressure trace and pivoting QR decomposition to identify the best sampling locations for identifying the modes, which will be finally used for the pressure reconstruction.

The algorithm was used in four engines with different combustions, namely CI, SI, DFCI, and TJI, and the main combustion parameters (CA10, CA50, CA90, IMEP, maximum pressure, and MAPO) were studied to analyze the feasibility of the method for combustion control and diagnosis.

Results demonstrate that it is possible to reduce the number of samples from 3600 samples (0.2 CAD resolution) to 25 with minimum information loss. That means that the computational burden can be reduced from 20 MBytes per minute (estimated shared data along a WLTP cycle) to only 0.14 MB per minute. However, in new combustion modes, such as TJI or DFCI, where fast combustion heavily excite the resonance of the combustion chamber, the modes are driven by the pressure oscillations and the error increases. In such combustion modes, a low-pass filter would reduce the error in one order of magnitude but MAPO information is lost.

References

1. Cha J, Park J, Lee H and Chon MS. A study of prediction based on regression analysis for real-world co2 emissions with light-duty diesel vehicles. *Int J Autom Technol* 2021; 22(3): 569–577.
2. Wei T and Frey HC. Sensitivity of light duty vehicle tailpipe emission rates from simplified portable emission measurement systems to variation in engine volumetric efficiency. *J Air Waste Manag Assoc* 2021; 71: 1127–1147.
3. Tunestål P. Self-tuning gross heat release computation for internal combustion engines. *Control Eng Pract* 2009; 17(4): 518–524.
4. Zhen X, Wang Y, Xu S, et al. The engine knock analysis – an overview. *Appl Energy* 2012; 92: 628–636.
5. Wick M, Bedei J, Gordon D, et al. In-cycle control for stabilization of homogeneous charge compression ignition combustion using direct water injection. *Appl Energy* 2019; 240: 1061–1074.
6. Guardiola C, Pla B, Bares P and Barbier A. Closed-loop control of a dual-fuel engine working with different combustion modes using in-cylinder pressure feedback. *Int J Engine Res* 2020; 21(3): 484–496.
7. Pla BN, Bares P, Jiménez I, Guardiola C, Zhang Y and Shen T. A fuzzy logic map-based knock control for spark ignition engines. *Appl Energy* 2020; 280.
8. Shayestehmanesh S and Peyton Jones JC. Stochastic modeling and prediction of imep for closed loop knock control performance assessment. *Control Eng Pract* 2019; 92.
9. Khameneian A, Wang X, Dice P, et al. Model-based dynamic in-cylinder air charge, residual gas and temperature estimation for a GDI spark ignition engine using cylinder, intake and exhaust pressures. In: *ASME 2020 dynamic systems and control conference, DSCC2020*, 2020, vol. 2. New York: ASME.
10. Murić K, Stenlåås O and Tunestål P. Zero-dimensional modeling of NO_X formation with least squares interpolation. *Int J Engine Res* 2014; 15(8): 944–953.
11. d'Ambrosio S, Ferrari A and Jin Z. Time-frequency analysis application to the evaluation of instantaneous combustion noise. *Fuel* 2022; 312.
12. Bares P, Selmanaj D, Guardiola C and Onder C. Knock probability estimation through an in-cylinder temperature model with exogenous noise. *Mech Syst Signal Process* 2018; 98: 756–769.
13. Broatch A, Guardiola C, Pla B and Bares P. A direct transform for determining the trapped mass on an internal combustion engine based on the in-cylinder pressure resonance phenomenon. *Mech Syst Signal Process* 2015; 62-63: 480–489.
14. Youssef B. Trapped mass estimation in automotive diesel engines based on in-cylinder pressure signal projection. *Int J Adapt Control Signal Process* 2020; 34(12): 1751–1767.
15. Martos FJ, Martín-González G and Herreros JM. Semi-empirical model for indirect measurement of soot size distributions in compression ignition engines. *Measurement* 2018; 124: 32–39.
16. Guardiola C, Pla B, Bares P and Peyton Jones JC. Integration of intermittent measurement from in-cylinder

pressure resonance in a multi-sensor mass flow estimator. *Mech Syst Signal Process* 2019; 131: 152–165.
17. Norouzi A, Shahpouri S, Gordon D, et al. Deep learning based model predictive control for compression ignition engines. *Control Eng Pract* 2022; 127.
18. Hamid AFA, Rahman MTA, Khan SF, et al. Connected car: engines diagnostic via internet of things (IOT). *J Phys Conf Ser* 2017; 908.
19. Wang PW, Yu HB, Xiao L and Wang L. Online traffic condition evaluation method for connected vehicles based on multisource data fusion. *J Sens* 2017; 2017: 1–11.
20. Weis A, Strandskov M, Yelamarthi K, et al. Rapid deployment of IOT enabled system for automobile fuel range and gas price location. In: *IEEE international conference on electro information technology*, 2017, pp.452–455. New York: IEEE.
21. Jeong Y, Son S, Jeong E and Lee B. An integrated self-diagnosis system for an autonomous vehicle based on an IOT gateway and deep learning. *Appl Sci* 2018; 8(7): 1164.
22. Maurya RK. Estimation of optimum number of cycles for combustion analysis using measured in-cylinder pressure signal in conventional ci engine. *Measurement* 2016; 94: 19–25.
23. Maurya RK and Agarwal AK. Investigations on the effect of measurement errors on estimated combustion and performance parameters in HCCI combustion engine. *Measurement* 2013; 46(1): 80–88.
24. Guardiola C, Pla B, Bares P and Barbier A. An analysis of the in-cylinder pressure resonance excitation in internal combustion engines. *Appl Energy* 2018; 228: 1272–1279.
25. Manohar K, Brunton BW, Kutz JN, et al. Data-driven sparse sensor placement for reconstruction: demonstrating the benefits of exploiting known patterns. *IEEE Control Syst* 2018; 38(3): 63–86.
26. Al Mamun S, Lu C and Jayaraman B. Extreme learning machines as encoders for sparse reconstruction. *Fluids* 2018; 3(4): 88.
27. Jayaraman B and Mamun SM. On data-driven sparse sensing and linear estimation of fluid flows. *Sensors* 2020; 20(13): 3752.
28. Barrault M, Maday Y, Nguyen NC and Patera AT. An 'empirical interpolation' method: application to efficient reduced-basis discretization of partial differential equations. *C R Math* 2004; 339(9): 667–672.
29. Sommariva A and Vianello M. Computing approximate Fekete points by QR factorizations of Vandermonde matrices. *Comput Math Appl* 2009; 57(8): 1324–1336.
30. Novella R, Pla B, Bares P, et al. Closed-loop combustion control by extremum seeking with the passive-chamber ignition concept in SI engines. SAE Technical paper 2020, 2020.
31. Luo Y, Maldonado B, Liu S, et al. Portable in-cylinder pressure measurement and signal processing system for real-time combustion analysis and engine control. *SAE Technical paper* 2020, 2020.
32. Eckart C and Young G. The approximation of one matrix by another of lower rank. *Psychometrika* 1936; 1(3): 211–218.
33. Strang G. *Linear algebra and learning from data*. Wellesley, MA: Wellesley-Cambridge Press Cambridge, 2019.
34. Engler H. The behavior of the qr-factorization algorithm with column pivoting. *Appl Math Lett* 1997; 10(6): 7–11.
35. Bujanović Z and Drmač Z. New robust ScaLAPACK routine for computing the QR factorization with column pivoting. *arXiv preprint arXiv*:191005623, 2019.
36. Payri F, Broatch A, Tormos B and Marant V. New methodology for in-cylinder pressure analysis in direct injection diesel engines—application to combustion noise. *Meas Sci Technol* 2005; 16(2): 540–547.