



HAL
open science

Bayesian Estimation and Control of Engine Knocking Level for Transient Operation

Maxime Jean, Thomas Leroy, Fabien Vidal-Naquet

► **To cite this version:**

Maxime Jean, Thomas Leroy, Fabien Vidal-Naquet. Bayesian Estimation and Control of Engine Knocking Level for Transient Operation. IEEE Transactions on Control Systems Technology, 2022, 30 (3), pp.944-955. 10.1109/TCST.2021.3081378 . hal-03694585

HAL Id: hal-03694585

<https://ifp.hal.science/hal-03694585>

Submitted on 13 Jun 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Bayesian Estimation and Control of Engine Knocking Level for Transient Operation

Maxime Jean¹, Thomas Leroy, and Fabien Vidal-Naquet

Abstract—The contribution of this article is to propose a knock controller based on the estimation of the distribution quantile of the knock intensity measurements. Despite the quantile estimation randomness, the corrections undertaken by the controller are moderated by the level of confidence in the quantile estimation. This strategy offers a fast control ensuring a better compromise between the engine efficiency and the prevention of knock phenomenon. The design of the controller reduces the engine torque fluctuations and is suited for engine transient operations.

Index Terms—Bayesian inference, knock control, quantile estimation, spark-ignition engine, transient response.

I. INTRODUCTION

THE spark-ignition timing of four-stroke engines is a critical degree of freedom to be carefully controlled by the electronic control unit. An optimally advanced ignition timing with respect to the top dead center helps to reach higher pressure in the combustion chamber and enables to retrieve more mechanical power from the gas oxidization [1]. For this reason, maximizing engine efficiency encourages to set the spark advance as close as possible to this optimum value. However, excessive temperatures generate the knocking phenomena. It is commonly accepted that the knocking mechanisms rely on the uncontrolled autoignition of the air/gas mixture. Knocking generates undesirable shock waves within the combustion chamber and damages the materials constituting its inner surface. The spark advance has to be set to its optimum value to maximize fuel efficiency. However, it should also occur late enough in order to prevent knocking regimes. The knocking limit of the spark advance depends on engine operating conditions and could result in efficiency losses; this is an engineering compromise. An online controller answering this tradeoff has to be designed to ensure the optimal run of the engine. For given operating parameters, the random occurrence of the knocking phenomena [2] leads to adopt a probabilistic description. Knock intensities would be considered as stochastic quantities. Accordingly, a knocking engine state (operated at steady state) would exhibit both knocking and nonknocking cycles. An operating point is deemed knocking based on the fraction of knocking cycles.

Considering the stochastic nature of the knocking phenomena, feedback loop controllers built on a probabilistic approach have already been developed [3]–[12]. Instead of acting immediately to a punctual high knock intensity measurement, actions can be taken only if their relevance is supported by accumulated measurements stored in a buffer memory. If computed over few samples required to follow the engine dynamics, the inaccuracy and noise of the estimation of the knocking probability lead to irrelevant corrections of the spark-ignition timing. The consequences of these fluctuations are a suboptimal spark advance and engine torque fluctuations. A solution is to moderate the corrective actions undertaken by the controller with the level of confidence in the estimation of the knocking probability.

This article adopts a Bayesian framework. A decision criterion is assessed based on past measurements. If conditions are met with high probability, the controller undertakes a corrective action. Two estimators are derived from the standard Bernoulli and log-normal models. The Bayesian inference based on a log-normal model has been used in [13]. However, the proposed approach derives the posterior probability that the engine is in a knocking state without approximations, offering a simple design. If the knock intensities conform to a log-normal model, it is also shown that using this model improves the precision of the estimation. The controller developed in [14], maintaining the probabilities of different theories, is insensitive to departures from parametric models, but potentially misses valuable information in return. An implementation based on filters is proposed. Promising properties of the proposed controller were presented in [15], without mathematical derivations. The developed method may give sufficient precision for feedback control and may be suited for the engine transient regimes. An improvement of the algorithm is also proposed to undermine older measurements and strengthen newer measurements. The simplicity of the computations allows its online implementation.

This article is organized as follows. In Section II, physical modeling of the knock intensities is described. Section III describes the two Bayesian estimators relying on different sets of assumptions and selects the most appropriate. Section IV describes the weighting of the observations. Section V presents a simple controller based on the proposed Bayesian estimator. Section VI describes its performance.

II. KNOCKING PHENOMENA

A. Knock Intensity Measurement

The shock waves generated by the autoignition excite the vibrational modes of the system. The measurements

from an accelerometer attached to the cylinder block contain the signatures of the corresponding natural frequencies. The accelerometer signal, sent through an adapted time and frequency window function filter, evaluates the magnitude of the excitation. Different methods exist to get a single scalar per cycle. The maximum amplitude pressure oscillation (MAPO) is adopted here. The knock intensity measurement in itself is not in the scope of this article and it is admitted that the MAPO constitutes a sufficient measurement of the knock intensity [16]. In this article, the MAPO experimental data refer to the maximum peak value of the filtered signal of the sensor. The MAPO provides an intensity quantification of the vibrations caused by two combined sources: knock and engine noise. Therefore, no static knock intensity threshold above which a cycle would be deemed as a knocking cycle can be defined over the operating range of the engine. A standard solution is to map this threshold on the engine speed and torque. In this study, such a map is assumed to be known.

In practice, rather than reacting to every single knocking occurrence, a small fraction of knocking cycles is tolerated. It is generally set at 5% [17], but other percentages could also be selected. An operating point exhibiting less than 5% of knocking cycles is seen as energetically suboptimal if its efficiency can be increased at the cost of having more frequent knocking occurrences. More than 5% of knocking cycles is unacceptable.

Other promising knock intensity measurements based on different physical properties exist (e.g., ion current detectors [18], [19], cylinder pressure sensors [16], and microphones [20]). If they provide a better quantification of the intensity of the knocking phenomena, the proposed algorithm may be adapted to use the feedback measurements of these sensors.

B. Engine Knock: A Stochastic Process

Deterministic prediction models of knock occurrences are unreachable due to the stochastic nature of knocking phenomena [21]. Moreover, considering the complexity of the underlying physicochemical processes, it is arguable that knock has an intrinsic probabilistic nature [22]. A statistical description has emerged as relevant to describe experimental observations. This suggests to adopt a probabilistic theoretical framework and to consider the knock intensity measurements as realizations of random variables [3]–[5], [14], [23], [24]. A knock intensity measurement campaign has been performed on engine test bench. The engine under consideration is a 1.2-L three-cylinder turbocharged spark-ignition engine using a direct injection technology. Operating points are defined by the engine speed and load, measured as indicated mean effective pressure (IMEP). For several operating points, the engine actuators are adjusted to achieve optimal efficiency. For each of these operating points, a spark advance sweep is conducted while maintaining the other actuators to their optimal set point. In particular, it affects the IMEP. On each step of spark advance, the knock intensity values of 300 consecutive cycles have been acquired. Measurements of the operating

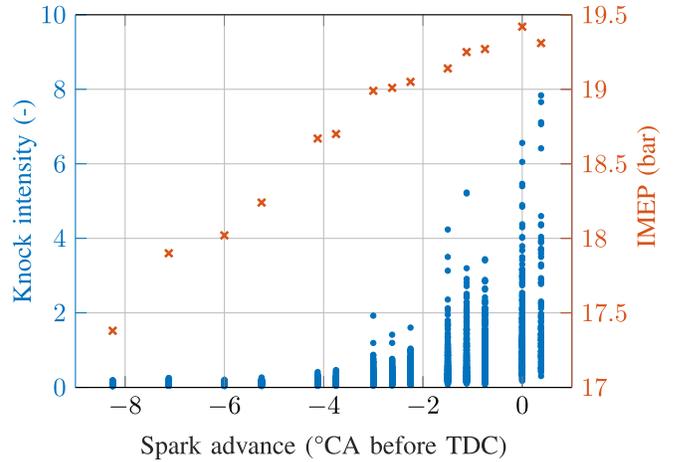


Fig. 1. Knock intensity and mean IMEP for different spark advance timing.

point 1750 rpm and 18 bar of IMEP are shown in Fig. 1. It depicts the increase of standard deviation of the knock intensities with the spark advance. It also shows the compromise between fuel efficiency and knocking occurrences. The spark advance is expressed in degrees of Crankshaft angle ($^{\circ}\text{CA}$) before the top dead center (TDC).

Rephrasing the condition described in Section II-A into a probabilistic perspective is to ensure that the knock intensity of the next cycle, seen as a random variable, has a probability of 5% to exceed the mapped threshold.

1) *Independently and Identically Distributed*: The distribution of these random variables depends on the engine operating point (speed and load) and state (spark advance, temperature, and other unmeasured exogenous parameters). When the engine is working on steady states over relatively short time scales, the abovementioned parameters stay constant and the knock intensities are identically distributed. Furthermore, autocorrelation functions analyzed in [2], [13], [20], [21], [25], [26] clearly show the cycle-to-cycle independence of the knock intensities. Those two considerations fully support the design of a stochastic controller. Even though this hypothesis is legitimate, it does not hold anymore for transient regimes.

2) *Log-Normally Distributed*: In addition, the knock intensities, given a constant engine state, are sometimes assumed to be log-normally distributed. A random variable is log-normally distributed if its logarithm is normally distributed. This hypothesis is presented and challenged in [27]. This assumption has been again highlighted by the measurement campaign previously mentioned. For instance, the empirical probability density functions, obtained with a kernel estimation [28], and corresponding to Fig. 1, are shown in Fig. 2. Their cumulative distribution functions and their theoretical log-normal fits are shown in Fig. 3.

The log-normal distribution is well suited to describe the distribution of the value of interest. However, this is just a convenient model, not relying on any physical explanation. It might not describe correctly the knock intensities for all operating points. Other models are available in the

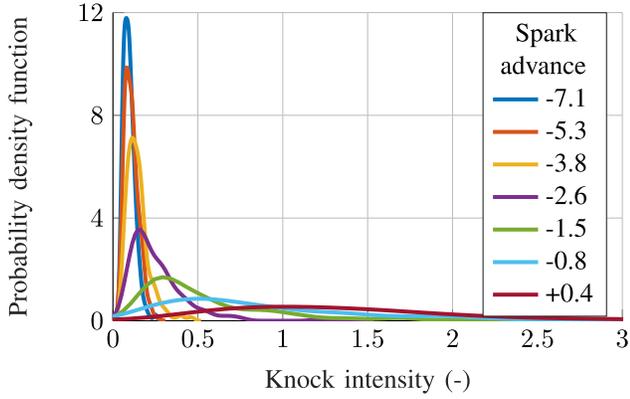


Fig. 2. Experimental probability density functions of the knock intensity measurements for different spark advances.

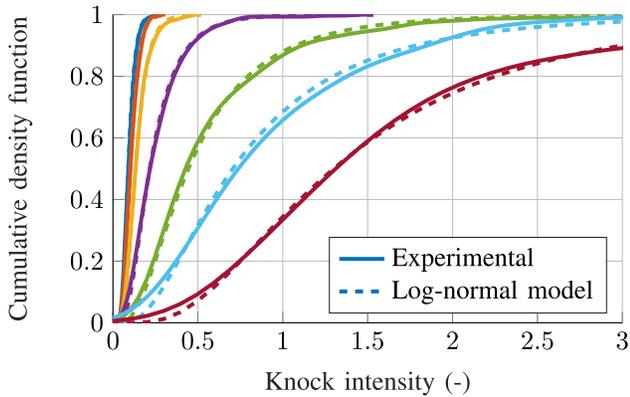


Fig. 3. Cumulative distribution functions (experimental and theoretical) of the knock intensity measurements for different spark advances.

literature [27], but the log-normal model has been chosen for its simplicity and because its good description of the knock intensities is relatively consensual.

C. Spark Advance: A Knock Intensity Control Parameter

The spark advance SA is the single degree of freedom of the proposed controller. The knock intensities decrease when SA decreases and moves further from its optimal value [29]. In general, the spark advance timing is adjusted based on a comparison between the knock intensities and the threshold. If necessary, the spark advance is decreased to reduce the knock intensity; however, there is an associated efficiency loss.

Nonmonotonous spark adjustments at every step should be avoided. They needlessly increase the coefficient of variance of the IMEP (common combustion stability indicator), impacting the driving comfort. Also, they make probabilistic estimations less precise in comparison with constant spark advances.

III. BAYESIAN ESTIMATION

A. Mathematical Modeling

For each engine cycle $i \in \mathbb{N}$, the notation KI_i stands for the random variable of the knock intensity measurement, only depending on the engine state. Its realization, once it has been observed, is written ki_i . The mapped knocking threshold

(generally on the engine speed and load) is θ_i . In order to limit the number of notations, the target value of the fraction of knocking cycles (5%) is directly used in this article. The algorithm would not be changed if another value was selected. It leads to the definition of the decision criterion H_i at cycle i

$$H_i : \Pr(KI_i > \theta_i) > 0.05. \quad (1)$$

Also, its contrary is

$$\overline{H}_i : \Pr(KI_i > \theta_i) \leq 0.05. \quad (2)$$

B. Decision Criteria

The state of $H_i \in \{\text{true}, \text{false}\}$ determines which correction the controller has to apply. However, the intrinsic random nature of KI_i makes the estimation of H_i necessarily uncertain. Thus, instead of estimating only the single Boolean H_i , the proposed algorithm estimates the probability, given the past measurements, that H_i is true. If this probability exceeds a specific threshold b (99% in the following sections), then the spark advance should be decreased. The other way around is that if $\Pr(\overline{H}_i) = 1 - \Pr(H_i)$ exceeds b , then an opposite corrective action has to be taken.

According to Section II-B1, if the engine is running in a steady state, $(KI_i)_{i \in \mathbb{N}}$ is a sequence of random variables, independent and identically distributed.

In particular, if the engine state is constant during $[j, i]$, $j < i$, then ki_i and ki_j can be seen as two independent realizations of the same random variable.

Otherwise, the identical property might not hold any more. However, for n much smaller than the characteristic time of variation of the engine state, one can assume that the n random variables KI_{i-k} , $k \in \llbracket 1, n \rrbracket$ are equal in law (or identically distributed). Then, the measurements ki_{i-k} are independent realizations of random variables equal in law to KI_i . It allows to derive Bayesian estimators inferring H_i from the n last knock intensity measurements, stored in a buffer memory. $n \in \mathbb{N}$ needs to be experimentally adjusted. This assumption obviously does not hold during rapid transients. n should then be kept as small as possible. Bayesian, over frequentist, approaches provide better estimation since limit theorems cannot be used.

C. Bernoulli Distribution Model

1) *Model*: In this method, no further assumptions are made on the distribution of KI_i . It is a common framework established in the literature in order to derive the probability of H_i or at least its likelihood. The method relies on the statistics $Y_i := \{1 \text{ if } KI_i > \theta_i, 0 \text{ else}\}$. Its realization is written as y_i . In this formalism, there exists a probability $p_i \in [0, 1]$ such that Y_i follows a Bernoulli distribution law written as

$$Y_i \mid p_i \sim \text{Be}(p_i). \quad (3)$$

Note in particular that

$$\begin{aligned} H_i \text{ is true} &\Leftrightarrow \Pr(KI_i > \theta_i) > 0.05 \\ &\Leftrightarrow \Pr(Y_i = 1) > 0.05 \\ &\Leftrightarrow p_i > 0.05. \end{aligned} \quad (4)$$

2) *Bayesian Inference*: For all i , the probability that H_i is true should be inferred from the n last measurements of knock intensities.

Property 1: The posterior distribution of p_i derived with a noninformative reference prior, given the n last measurements, is

$$\pi(p_i | Y_{i-1} = y_{i-1}, \dots, Y_{i-n} = y_{i-n}) \propto p_i^{z_i - (1/2)} (1 - p_i)^{n - z_i - (1/2)} \quad (5)$$

with $\pi(\cdot)$ the probability density function. Or equivalently

$$p_i | y_{i-1}, \dots, y_{i-n} \sim \text{Beta}\left(z_i + \frac{1}{2}, n - z_i + \frac{1}{2}\right) \quad (6)$$

with $z_i = \sum_{k=1}^n y_{i-k}$.

It derives from the Bayes formula applied to the independent and identically distributed random variables Y_{i-k} , $k \in \llbracket 1, n \rrbracket$. It is a common result and can be found in [30]. The selection of the noninformative prior distribution of p_i demands, however, some justification. It is chosen such that it has a minimal impact on the posterior distribution. It would then ensure that the Bayesian inference derived from the n last observations is not biased by potentially irrelevant prior information. The reference prior introduced by José-Miguel Bernardo [31], [32] has been selected. It maximizes in some particular way the information entropy of the posterior distribution. A rigorous proof of its derivation is given in [33]. It provides the following improper form of the prior distribution:

$$\pi(p_i) \propto p_i^{-(1/2)} (1 - p_i)^{-(1/2)}. \quad (7)$$

Property 2: The posterior confidence interval $[p_i^{\min}, p_i^{\max}]$ of p_i , such that $\Pr(p_i > p_i^{\min}) = \Pr(p_i < p_i^{\max}) = b$, is simply written as

$$\left[\underbrace{F_{p_i}^{-1}(1 - b)}_{p_i^{\min}}, \underbrace{F_{p_i}^{-1}(b)}_{p_i^{\max}} \right] \quad (8)$$

with F_{p_i} the cumulative distribution function of p_i .

3) *Decision Criterion*: The confidence interval is shown in Fig. 4 for $n = 10$ and $z_i \in \llbracket 1, n \rrbracket$. This leads to the expression of the following decision criterion:

$$\begin{aligned} \Pr(H_i \text{ is true}) > b &\Leftrightarrow \Pr(p_i > 0.05) > b \\ &\Leftrightarrow 1 - F_{p_i}(0.05) > 1 - F_{p_i}(p_i^{\min}) \\ &\Leftrightarrow p_i^{\min} > 0.05. \end{aligned} \quad (9)$$

In a similar way, $\Pr(\overline{H}_i) > b \Leftrightarrow p_i^{\max} < 0.05$. In fact, $[p_i^{\min}, p_i^{\max}]$ is the centered 98% confidence interval of p_i . If the target 0.05 lies outside of the confidence interval, one can be 99% sure that p_i is lower or higher (depending on its relative position to the confidence interval) than the target.

For example, Fig. 4 shows that if $z_i \geq 3$ (out of ten samples, three or more exceed the threshold), then one can be 99% sure that H_i is true and that a corrective action should be taken. The other way around is that even if $z_i = 0$ (none of the measurements exceed the threshold), one cannot be sure at 99% that \overline{H}_i is true. In fact, with only ten samples, p_i^{\max} always exceeds 0.05.

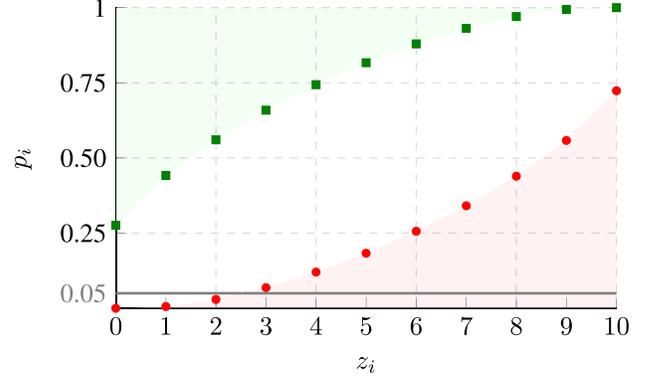


Fig. 4. p_i^{\min} (dots) and p_i^{\max} (squares) as a function of z_i for $n = 10$. Probability of 98% for p_i to be in the white area.

4) *Robustness and Performance*: This method does not rely on specific assumptions about the knock intensities, besides the cycle-to-cycle independence and identical distribution. It offers a robust model. However, because Y_i provides a partial description of KI_i , the method has a slow convergence. If the engine was always successfully controlled, a prior stating that p_i lies in the vicinity of 0.05 could have been used. However, in sharp transients, it is unreachable and would undesirably bias the estimation. Such prior was then not selected. In Section III-D, a second method is proposed.

D. Log-Normal Distribution Model

This section introduces the derivation of the posterior probability that the engine is in a knocking state.

1) *Model*: The condition H_i can be rephrased in terms of distribution quantile. Introducing the 95% quantile q_i of the distribution of KI_i , one has by definition

$$\Pr(KI_i > q_i) = 0.05. \quad (10)$$

Then

$$\begin{aligned} H_i \text{ is true} &\Leftrightarrow \Pr(KI_i > \theta_i) > 0.05 \\ &\Leftrightarrow \Pr(KI_i > \theta_i) > \Pr(KI_i > q_i) \\ &\Leftrightarrow \theta_i < q_i. \end{aligned} \quad (11)$$

The assumption of log normality of KI_i states that

$$\forall i \in \mathbb{N}, \exists (\mu_i, \sigma_i) \in \mathbb{R} \times \mathbb{R}^+, \log(KI_i) \sim \mathcal{N}(\mu_i, \sigma_i^2). \quad (12)$$

For example, the three log-normal probability density functions shown in Fig. 5 share the same quantile 95%. This quantity has to be estimated. Once again, KI_i being independent and identically distributed during stationary regimes, μ_i and σ_i are constant. In the following sections, the notations $X_i = \log(KI_i)$ and $x_i = \log(k_i)$, its realization, are used. The observation of the buffer memory is the event $B_i = (X_{i-1} = x_{i-1}, \dots, X_{i-n} = x_{i-n})$. Its sample mean is $\bar{x}_i = (1/n) \sum_{k=1}^n x_{i-k}$ and its sum of squared deviations is $S_i = \sum_{k=1}^n (x_{i-k} - \bar{x}_i)^2$.

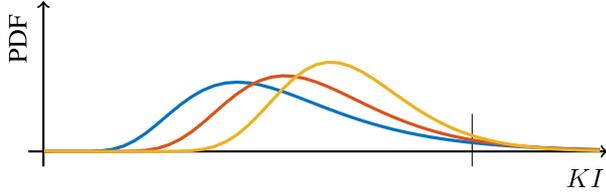


Fig. 5. Log-normal distributions with the same quantile 95%.

2) *Bayesian Inference*: For a normal distribution, the 95% quantile is directly given by the knowledge of μ_i and σ_i through the relation

$$q_i = \mu_i + r\sigma_i \quad (13)$$

with $r = F_{\mathcal{N}}^{-1}(0.95)$ and $F_{\mathcal{N}}$ the normal centered reduced cumulative distribution function.

Property 3: The posterior distribution of (μ_i, σ_i^2) derived with a noninformative reference prior, given the n last measurements, is

$$\begin{aligned} \pi((\mu_i, \sigma_i^2) | B_i) \\ \propto \frac{1}{\sigma_i} \exp\left(-\frac{\kappa_i(\mu_i - m_i)^2}{2\sigma_i^2}\right) \left(\frac{1}{\sigma_i^2}\right)^{\alpha_i+1} \exp\left(-\frac{\beta_i}{\sigma_i^2}\right) \end{aligned} \quad (14)$$

with $m_i = \bar{x}_i$, $\kappa_i = n$, $\alpha_i = (n - 1/2)$, $\beta_i = (1/2)S_i$. Or equivalently

$$(\mu_i, \sigma_i^2) | B_i \sim \mathcal{NIG}(m_i, \kappa_i, \alpha_i, \beta_i). \quad (15)$$

\mathcal{NIG} standing for normal-inverse-gamma and defined by

$$\begin{cases} \mu_i | \sigma_i^2 \sim \mathcal{N}\left(m_i, \frac{\sigma_i^2}{\kappa_i}\right) \\ \sigma_i^2 \sim \mathcal{G}^{-1}(\alpha_i, \beta_i). \end{cases} \quad (16)$$

Proof: The Bayes formula states

$$\begin{aligned} \overbrace{\pi((\mu_i, \sigma_i) | X_{i-1} = x_{i-1}, \dots, X_{i-n} = x_{i-n})}^{\text{Posterior distribution}} \\ = \frac{\overbrace{\pi(X_{i-1} = x_{i-1}, \dots, X_{i-n} = x_{i-n} | (\mu_i, \sigma_i))}^{\text{Likelihood}} \overbrace{\pi(\mu_i, \sigma_i)}^{\text{Prior distribution}}}{\underbrace{\pi(X_{i-1} = x_{i-1}, \dots, X_{i-n} = x_{i-n})}_{\text{Normalization factor}}}. \end{aligned} \quad (17)$$

- 1) *Normalization Factor*: It does not depend on (μ_i, σ_i^2) and can be seen as a multiplicative constant.
- 2) *Likelihood*: Because of the assumption of conditional independence given the parameters (μ_i, σ_i^2) , the likelihood factor can be written as

$$\begin{aligned} \pi(X_{i-1} = x_{i-1}, \dots, X_{i-n} = x_{i-n} | (\mu_i, \sigma_i^2)) \\ = \prod_{k=1}^n \pi(X_i = x_{i-k}) \\ = \left(\frac{1}{\sigma_i \sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2\sigma_i^2} \sum_{k=1}^n (x_{i-k} - \mu_i)^2\right) \\ \propto \left(\frac{1}{\sigma_i^2}\right)^{\frac{n}{2}} \exp\left(-\frac{S_i}{2\sigma_i^2} - \frac{n(\mu_i - \bar{x}_i)^2}{2\sigma_i^2}\right). \end{aligned} \quad (18)$$

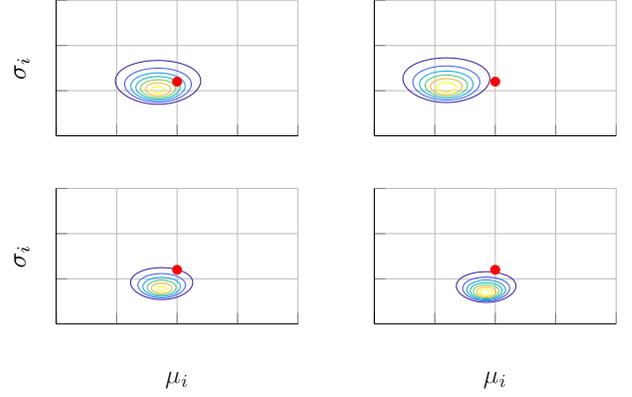


Fig. 6. Joint probability density function of $(\mu_i, \sigma_i) | B_i$ for the four sets of observations.

- 3) *Prior Distribution*: The prior distribution of (μ_i, σ_i^2) is again selected to ensure a distribution as much noninformative as it can be. It is [32]

$$\pi(\mu_i, \sigma_i^2) \propto \frac{1}{\sigma_i^2}. \quad (19)$$

- 4) *Posterior Distribution of (μ_i, σ_i^2)* : Combining (18) and (19) gives (14).

The joint probability density function of (μ_i, σ_i) is shown in Fig. 6 for four different sets of ten observations. The center dot represents the true pair (μ_i, σ_i) from which the n random realizations are drawn.

The computation of the posterior distribution of q_i is analytically nontractable, but its confidence interval can be derived.

Property 4: The posterior confidence interval $[q_i^{\min}, q_i^{\max}]$ of q_i , such that $\Pr(q_i > q_i^{\min}) = \Pr(q_i < q_i^{\max}) = b$, is

$$(m_i + ru\sqrt{\beta_i}) + [t^{\min}, t^{\max}] \cdot \sqrt{\beta_i(u^2/\kappa + r^2v^2)} \quad (20)$$

with $(t^{\min}, t^{\max}) \in \mathbb{R}^2$ two constants, Γ the Euler Gamma function, and $u = (\Gamma(\alpha_i - (1/2))/\Gamma(\alpha_i))$ and $v = ((1/\alpha_i - 1) - u^2)^{1/2}$.

The omission of some indices i , explained in the following proof, only concerns variables depending on the number of observations n and not on the values taken by those observations.

Proof: In order to find some invariant characteristics of the joint distribution, new variables are introduced

$$\begin{cases} M_i = \frac{\mu_i - \mathbb{E}(\mu_i)}{\mathbb{E}(\sigma_i)/\sqrt{\kappa_i}} \\ \Sigma_i = \frac{\sigma_i - \mathbb{E}(\sigma_i)}{\text{std}(\sigma_i)}. \end{cases} \quad (21)$$

Computations detailed in Appendix VII lead to

$$\begin{cases} \mathbb{E}(\mu_i) = m_i \\ \mathbb{E}(\sigma_i) = u_i \sqrt{\beta_i} \\ \mathbb{E}(\sigma_i^2) = \frac{\beta_i}{\alpha_i - 1} \\ \text{std}(\sigma_i) = v_i \sqrt{\beta_i}. \end{cases} \quad (22)$$

Note that $E(\sigma_i)/(\kappa)_i^{1/2}$ is the conditional standard deviation of μ_i knowing that $\sigma_i = E(\sigma_i)$. The joint distribution of $(M_i, \Sigma_i | B_i)$ is as follows:

$$\pi(M_i, \Sigma_i | B_i) = \pi(\mu_i, \sigma_i^2 | B_i) |(\det J)^{-1}(\mu_i, \sigma_i^2)| \quad (23)$$

with $J(\mu_i, \sigma_i^2)$ being the Jacobian matrix of the transformation $(\mu_i, \sigma_i^2) \mapsto (M_i, \Sigma_i)$ and \det standing for the determinant of a square matrix. The posterior probability density function of (M_i, Σ_i) can be derived as

$$\pi(M_i, \Sigma_i | B_i) = \frac{2u_i v_i \exp\left(-\frac{2+u_i^2 M_i^2}{2(u_i+v_i \Sigma_i)^2}\right)}{\sqrt{2\pi} \Gamma(\alpha_i) (u_i + v_i \Sigma_i)^{2\alpha_i+2}} \quad (24)$$

which does not depend on m_i, κ_i , and β_i , but only on α_i . α_i is not a function of the observations, but only of the number n of samples used in the inference. The joint distribution of (M_i, Σ_i) conditional to the observations is then invariant to changes in the observations B_i and one has

$$\pi(M_i, \Sigma_i | B_i) = \pi(M_i, \Sigma_i | n). \quad (25)$$

It is defined for $\Sigma \geq -(u_i/v_i)$. For the sake of simplicity, the conditional dependence $(\cdot | n)$ will be omitted in the next sections, along with the index i for the variables depending only on the constant n , i.e., κ, α, u, v, M , and Σ .

In order to compute the bounds of the confidence interval of q_i , its cumulative distribution function F_i is introduced

$$F_i : \mathbb{R} \rightarrow [0, 1] \quad s \mapsto \Pr(q_i \leq s | B_i). \quad (26)$$

The bounds of the confidence interval of q_i are defined such that $F_i(q_i^{\min}) = 1 - b$ and $F_i(q_i^{\max}) = b$. The relation (13) leads to the expression in the integral form

$$F_i(s) = \int_{\mathcal{D}_i(s)} d\mu_i d\sigma_i \pi(\mu_i, \sigma_i | B_i) \quad (27)$$

with $\mathcal{D}_i(s) = \{(\mu_i, \sigma_i) \in \mathbb{R} \times \mathbb{R}^+ | \mu_i + r\sigma_i \leq s\}$. Or equivalently

$$F_i(s) = \int_{\mathcal{E}_i(s)} dM d\Sigma \pi(M, \Sigma) \quad (28)$$

with

$$\mathcal{E}_i(s) = \left\{ (M, \Sigma) \in \mathbb{R} \times \left[-\frac{u}{v}, \infty\right) \mid \right. \\ \left. \times \frac{u\sqrt{\beta_i}}{\sqrt{\kappa}} M + m_i + r\sqrt{\beta_i}(v\Sigma + u) \leq s \right\}. \quad (29)$$

Using $\phi = \text{atan2}(u/(\kappa)^{1/2}, rv)$, atan2 being the four-quadrant inverse tangent, the above inequality can also be written as

$$\begin{pmatrix} \cos(\phi) \\ \sin(\phi) \end{pmatrix}^T \begin{pmatrix} M \\ \Sigma \end{pmatrix} \leq \frac{1}{\sqrt{u^2/\kappa + r^2v^2}} \left(\frac{s - m_i}{\sqrt{\beta_i}} - ru \right).$$

Defining the variable change t as the RHS of the above inequality, one gets

$$F_i(s_i(t)) = \int_{\mathcal{F}(t)} dM d\Sigma \pi(M, \Sigma) \quad (30)$$

with

$$\mathcal{F}(t) = \left\{ (M, \Sigma) \in \mathbb{R} \times \left[-\frac{u}{v}, \infty\right) \mid \left| \begin{pmatrix} \cos(\phi) \\ \sin(\phi) \end{pmatrix}^T \begin{pmatrix} M \\ \Sigma \end{pmatrix} \right| \leq t \right\}$$

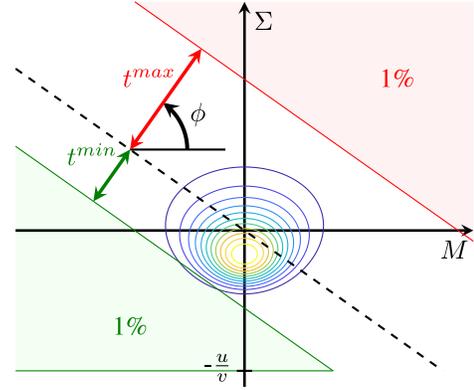


Fig. 7. Joint probability density function of (M, Σ) . Areas of probability 1%, 98%, and 1% are delimited by the distances t^{\min} and t^{\max} .

and

$$s_i(t) = t\sqrt{\beta_i(u^2/\kappa + r^2v^2)} + \left(m_i + ru\sqrt{\beta_i}\right). \quad (31)$$

Notice that the index i is omitted in the definition of \mathcal{F} because it does not depend on the realizations. The same goes for the entire RHS of (30), which will be named $F(t)$. The values t^{\min} and t^{\max} such that $F(t^{\min}) = 1 - b$ and $F(t^{\max}) = b$ can be numerically found by a dichotomy search. The situation is shown in Fig. 7 and F is represented in Fig. 8. Finally, one has

$$F_i(q_i^{\max}) = b = F(t^{\max}) = F_i(s_i(t^{\max})). \quad (32)$$

Since F_i is an injective function

$$q_i^{\max} = s_i(t^{\max}). \quad (33)$$

Following the same logic, $q_i^{\min} = s_i(t^{\min})$. The values of t^{\min} and t^{\max} are constant, while the number of observations n is kept unchanged. This leads to the easily implementable algorithm

$$[q_i^{\min}, q_i^{\max}] = [s_i(t^{\min}), s_i(t^{\max})] \quad (34)$$

with s_i defined in (31).

Note in particular that

$$m_i + ru\sqrt{\beta_i} = E(\mu_i) + rE(\sigma_i) = E(q_i). \quad (35)$$

3) *Decision Criteria:* The confidence interval of q_i indicates whether the probabilities of H_i or $\overline{H_i}$ exceed b . In fact

$$\Pr(H_i \text{ is true}) > b \Leftrightarrow \Pr(q_i > \theta_i) > \Pr(q_i > q_i^{\min}) \\ \Leftrightarrow \theta_i < q_i^{\min} \quad (36)$$

and its symmetric inequality $\Pr(\overline{H_i}) > b \Leftrightarrow \theta_i > q_i^{\max}$. The computation of the boundaries of the confidence interval is intensive and does not suit the requirements for an online use. In order to make its computation far less complex, a change of variables was introduced in Section III-D2.

Once the values of t^{\min} and t^{\max} are stored in the embedded control unit, the simple computation $[q_i^{\min}, q_i^{\max}]$ can be performed from m_i and β_i . If $q_i^{\min} > \theta_i$ (the bottom configuration in Fig. 9), then the probability H_i is true, which exceeds b .

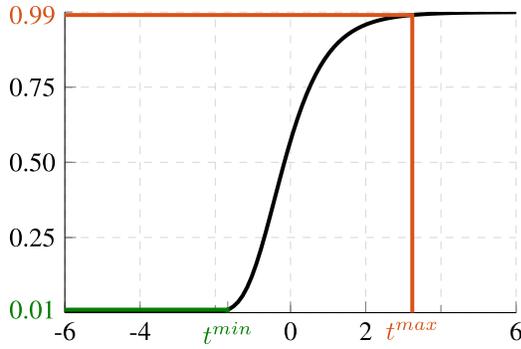


Fig. 8. $t \mapsto F(t)$.

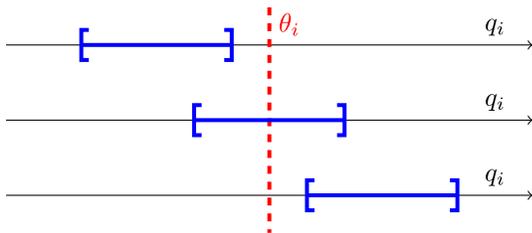


Fig. 9. Three different configurations of the target θ and the confidence interval of q_i .

The other way around is that if $q_i^{\max} < \theta_i$ (top configuration in Fig. 9), then the probability \overline{H}_i is true, which exceeds b . It offers a decision criterion to take a corrective action. If $q_i^{\min} < \theta_i < q_i^{\max}$ (middle configuration in Fig. 9), one cannot conclude with certainty (more than 99% of probability in the meaning of Bayesian statistics), and then, no action is performed.

4) *Robustness and Performance*: Experimental measurements tend to confirm the precision of the log-normal distribution model, but deviations could still occur. In such cases, overestimation of the 95% quantile has generally been observed. If confirmed, it would induce a slightly more protective behavior of the controller fed by this estimator. Moreover, if the estimator is used to map the knocking threshold, this bias would also be incorporated into the definition of the threshold and would become transparent.

During sharp transients, the identical distribution model of the buffered measurements can be violated. It is dealt with by using only the most recent observations (no prior from past knowledge and shortest possible buffer).

E. Out Performance of the Log-Normal Distribution Model

In order to facilitate the understanding of the properties of the two proposed estimators, a third, simple, estimator is introduced. It simply makes the comparison between the threshold θ and the quantile estimator given by (sample mean + $r \times$ sample standard deviation). It does not use the confidence interval.

To benchmark the three estimators, Monte Carlo simulations have been performed and are shown in Fig. 10. It represents the proportion of positive tests, over 10^6 samples of buffers

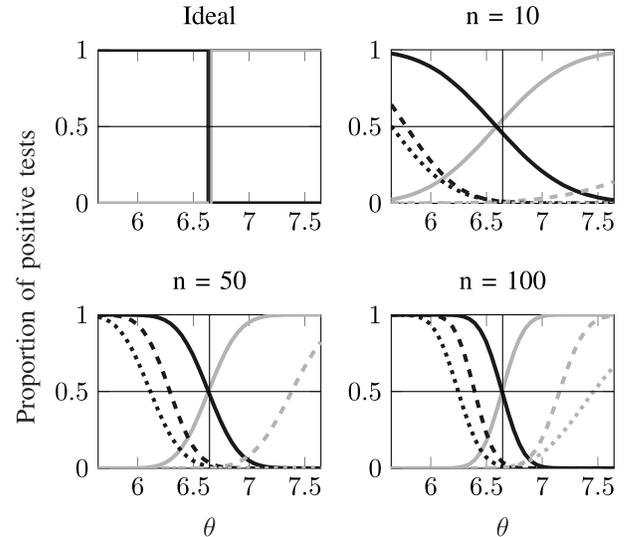


Fig. 10. Monte Carlo simulations. Light color: proportion of estimations of too many knocking cycles. Dark color: proportion of estimations of not enough knocking cycles. Solid: log-normal model without confidence interval estimator. Dashed: log-normal model with confidence interval estimator. Dotted: Bernoulli model with confidence interval estimator.

of n measurements (n taking the values 10, 50, and 100), and for a range of target threshold θ . The measurements are drawn from a normal random generator of mean 5 and variance 1, this setting having no influence on the results; 5% of the generated numbers are above 6.7. Dashed lines are used for the test too many knocking cycles and solid lines are used for the test not enough knocking cycles. The top-left plot represents the performance of an ideal estimator, in dotted lines those of the estimator using the Bernoulli model, in dashed lines those of the estimator using the log-normal model and the confidence interval, and finally in solid lines those of the estimator using the log-normal model without the confidence interval. Notice that for the third model, each sample is either in the category too many knocking cycles or in not enough knocking cycles and the sum of the proportion of the two tests is 1. Due to the construction of the Bayesian estimators, if the target θ is set to 6.7, then exactly 98% of the samples induce the estimation that H_i is neither true or false with at least 99% of probability. On the contrary, the estimator without confidence interval estimates erroneously $\sim 50\%$ of the time that H_i is true and $\sim 50\%$ of the time H_i is false. An integral controller based on this estimation would systematically impose irrelevant changes of spark advance at each engine cycle and generate an unnecessary increase in the dispersion of the spark advance. This problem is alleviated by the use of confidence intervals.

When θ is below 6.7, an ideal estimator (top-left plot, dark line) would always estimate that H_i is true. Obviously, this is not reachable, but the highest proportion of detection is desired. The same goes with the estimation of \overline{H}_i if θ is above 6.7. The estimator using the log-normal model (dashed lines) outperforms the one using the Bernoulli model (dotted lines), as shown in Fig. 10. For example, if $\theta = 6$, the first

one has a 28% detection probability that the true quantile (6.7) exceeds θ , while the second one only has a 21% probability.

Consequently, Sections IV–VI focus only on the log-normal distribution model and the Bernoulli model is left out for the proposed controller.

IV. WEIGHTING OF THE OBSERVATIONS

The proposed algorithm is designed to be implemented in engine control units (ECUs). They have stringent requirements in terms of memory and computational power. The proposed algorithm needs a buffer memory of size n . As exposed later in Section VI, five samples may agree with the needs of the engine. The parametric model and the Bayesian estimation allow to estimate the 95% quantile with much less than 20 samples. However, n has been calibrated on simulations. Discrepancies with actual engines might impose longer buffers and exceed the available resources. A turnaround is proposed in this section.

In Section III, it was assumed that each measurement carries the same amount of information about the distribution parameters. However, the Bayesian framework also offers the possibility to weight the different observations. The same posterior distribution function should be derived if one observation is replaced by two equal measurements with a halved reliability. It is achieved by weighting the corresponding terms of the log likelihood by $1/2$. Similarly, it can be extended to all sets of positive weights. The derivation of q_i performed in Section III-D holds with little changes. Computations, not detailed in this section, indicate that the variables have to be replaced as follows:

$$\begin{aligned} m_i &\leftarrow \frac{1}{W_i} \sum_{k=1}^n w_{i,k} x_{i-k} \\ \kappa_i &\leftarrow W_i \\ \alpha_i &\leftarrow \frac{W_i - 1}{2} \\ \beta_i &\leftarrow \frac{1}{2} \sum_{k=1}^n w_{i,k} (x_{i-k} - m_i)^2 \\ t^{\min} &\leftarrow t^{\min}(W_i) \\ t^{\max} &\leftarrow t^{\max}(W_i) \end{aligned} \quad (37)$$

with $(w_{ij})_{(i,j) \in \mathbb{N} \times \llbracket 1, n \rrbracket}$, the weight, at time i , of the measurement performed at time $i - j$, and $W_i = \sum_{j=1}^n w_{i-j}$.

The two maps $W_i \mapsto t^{\min}(W_i)$ and $W_i \mapsto t^{\max}(W_i)$ have to be numerically built, following the construction method presented in Section III-D and shown in Fig. 11. The benefit from this approach is that the buffer memory is no longer required since m_i , κ_i , α_i , and β_i can be implemented with finite impulse response filters. The number of stored values is given by the order of the selected filter (potentially first order filter). Another benefit is the possibility to use a forgetting factor, which is a function of the engine transient speed. In particular, the Taylor expansion of (24) leads to the limit

$$\pi(M_i, \Sigma_i | B_i) \xrightarrow{W_i \rightarrow \infty} \frac{1}{2\pi} \exp\left(-\frac{M_i^2 + \Sigma_i^2}{2}\right). \quad (38)$$

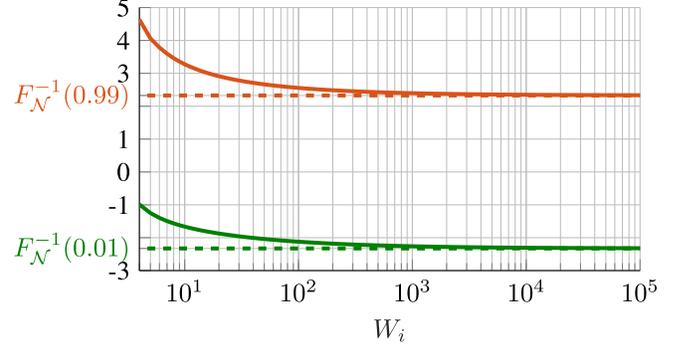


Fig. 11. Functions $W_i \mapsto t^{\min}(W_i)$ (lower curve) and $W_i \mapsto t^{\max}(W_i)$ (upper curve).

$M_i, \Sigma_i | B_i$ converges in law toward two independent standardized normally distributed random variables. Then, the limits of the boundaries of the confidence interval converge as follows:

$$[t^{\min}, t^{\max}] \xrightarrow{W_i \rightarrow \infty} [F_{\mathcal{N}}^{-1}(1-b), F_{\mathcal{N}}^{-1}(b)]. \quad (39)$$

Note that $F_{\mathcal{N}}^{-1}(1-b) = -F_{\mathcal{N}}^{-1}(b)$. One also has

$$u^2 \xrightarrow{W_i \rightarrow \infty} \frac{1}{\alpha_i - 1}$$

and the upper bound of the confidence interval of (34) converges, as $W_i \rightarrow \infty$, toward

$$\underbrace{(m_i + ru\sqrt{\beta_i})}_{E(q_i)} + F_{\mathcal{N}}^{-1}(b) \underbrace{\sqrt{\beta_i} \sqrt{\frac{1}{\kappa_i(\alpha_i - 1)} + r^2 v^2}}_{\text{std}(q_i)}. \quad (40)$$

Details of the derivations of $E(q_i)$ and $\text{std}(q_i)$ are given in Appendix VII. The same goes with the other lower bound of the confidence interval and one gets, when $W_i \rightarrow \infty$, an interval of the form $E(q_i) \pm F_{\mathcal{N}}^{-1}(b) \text{std}(q_i)$, as one would expect from the asymptotic normal behavior of Bayesian estimators stated by the Bernstein–von Mises theorem [34].

V. CONTROL STRATEGY

At each cycle, a correction ΔSA proportional to the error between the expected value of the quantile and its target is added to the spark-ignition timing of the preceding cycle, forming a discrete-time integral controller. However, the correction is inactive if, according to the confidence interval, one cannot be sure at 99% that a correction should be applied

$$\Delta SA = \begin{cases} 0, & \text{if } \theta \in [q^{\min}, q^{\max}] \\ K(\bar{q} - \theta), & \text{otherwise} \end{cases} \quad (41)$$

where $\bar{q} = E(q)$ (35) and K is a gain to be experimentally adjusted. Indeed, if $\theta < q^{\min}$, one can be sure at 99% that the spark ignition occurs too soon. Then, its set point needs a correction to delay the spark-ignition timing. The other way around is that if $\theta > q^{\max}$, one can be sure at 99% that the spark ignition occurs too late. Then, it needs a correction to advance the spark-ignition timing. The structure of the controller is summarized in the block diagram in Fig. 12.

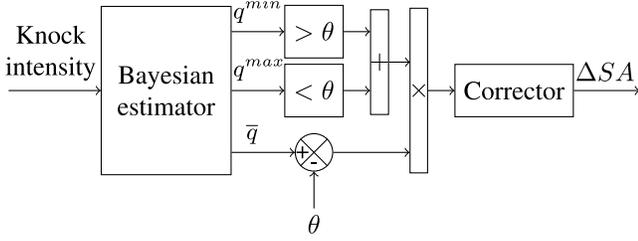


Fig. 12. Controller block diagram.

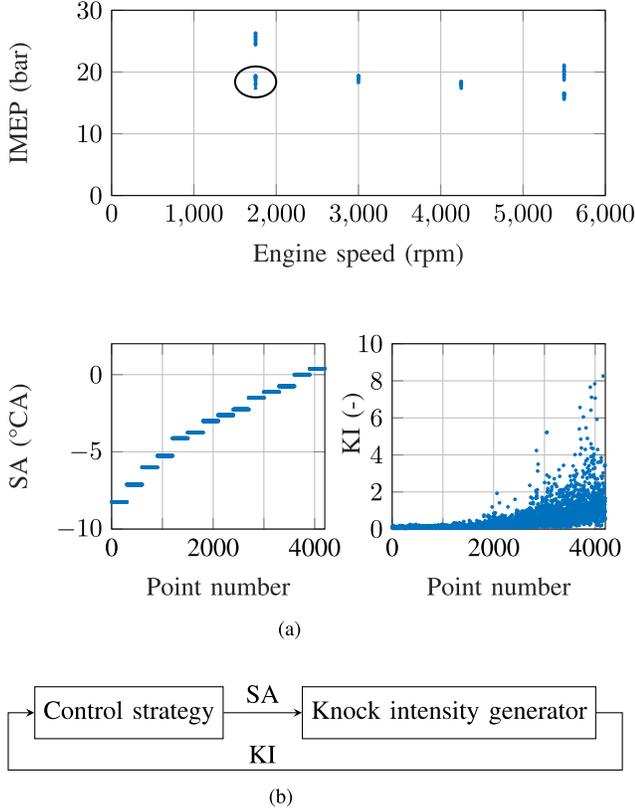


Fig. 13. Methodology. (a) Experimental database. The spark advance sweep presented in the two bottom plots corresponds to the circled operating point in the top plot. (b) Stochastic control.

VI. PERFORMANCE

A. Setup

1) *Methodology*: Even under identical conditions, each test is unique. To allow for a more extensive validation and a better performance assessment of the new stochastic controller, based on reproducible results, the behavior of the engine regarding knock is simulated.

In order to obtain knock intensities reflecting real distributions, experimental measurements have been acquired on different operating points, as shown in Fig. 13(a) (top). The sweep of spark advance, conducted for each operating point, is shown in Fig. 13(a) (bottom left) for the operating point highlighted by a circle (1750 rpm and 18 bar) and the corresponding knock intensity measurements are shown in Fig. 13(a) (bottom right). The measurements are performed

on different spark advance ranges, depending on the operating points. Empirical cumulative density functions have been estimated for each operating point and each spark advance timing. Knock intensities are then simulated with a random number generator whose probability density function changes depending on engine speed, torque, and spark advance timing. They are simulated with the inversion method [35] as follows:

$$KI_i = F_i^{-1}(U_i) \quad (42)$$

with $(U_i)_{i \in \mathbb{N}}$ a generated series of random numbers independent and identically distributed, following a uniform distribution between 0 and 1, and F_i^{-1} the inverse of the cumulative distribution function of the knock intensities at cycle i , computed with a kernel density estimator. Log-normal hypothesis is then not considered here. The performance of the controller can then be assessed within a feedback loop design, as shown in Fig. 13(b). Note that the proposed controller would be suitable for any engine technologies (EGR, VVT, and so on), as long as the knock intensities are approximately independent and log-normally distributed.

2) *Control Algorithm*: To improve the spark advance control during engine speed and load changes, conventional controllers always adopt both feedforward and feedback designs. The spark advance should be mapped beforehand as a function of these two parameters and then corrected online by the knock feedback algorithm presented in this article, potentially combined with an online learning model. However, in order to assess its adaptation abilities, no prepositioning maps were used in the simulations of the following sections.

The controller must be able to cope with two extreme situations: stationary states and the transient operation presented in Section VI-B2, deemed as representing the class of realistic, fastest, and dangerous transients. It would indicate that the controller is also suitable for all intermediate transients. A suitable calibration has a fast transient response in the protective direction (less than five cycles) and should not be much slower than conventional controllers (described in Section VI-A3) in the opposite direction. The added value of the controller, on top of being at rest during 98% of the cycles during stationary states, will be seen in reduction of the dispersion of the spark advance. The selected calibration $n = 5$ and $K = 0.1$ has the shortest n such that the controller fulfills the previous criteria on simulation. The map of the knocking threshold was also built from experimental measurements.

3) *Reference Strategy*: This new stochastic knock controller is compared with a common reference strategy: a sawtooth-shaped spark advance controller. The spark advance is increased at each time step by a small δSA and suddenly decreased by a large ΔSA if the knock intensity measurement exceeds the threshold θ . This strategy ensures the shortest reaction time to knock events while maintaining, on average, an implicit knocking frequency of $(\delta SA / (\delta SA + \Delta SA))$. With the calibration $\delta SA = 0.1^\circ$ and $\Delta SA = 1.9^\circ$, the knocking frequency is set at 5%.

B. Simulation Results

1) *Steady-State Operation*: The proper functioning of both algorithms is first confirmed on stationary operating points.

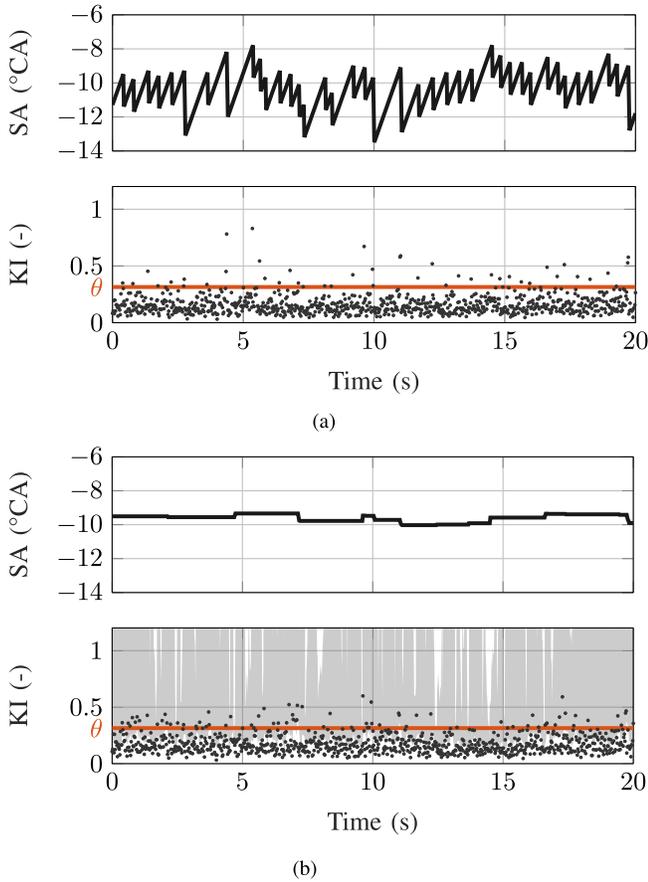


Fig. 14. Stationary regime. Bottom plot: Solid line—threshold. Bottom plot: dots—knock intensity measurements. Bottom plot: filled area—quantile confidence interval. (a) Sawtooth control. (b) Stochastic control ($n = 5$ and $K = 0.1$).

The evolution of the confidence interval of q is shown in Fig. 14(b). If the quantile target θ steps out of this interval, since one is 99% sure that a change needs to be applied on the spark advance, then a corrective action is taken (exactly at $t = 7.1$ s and $t = 14.5$ s). The behavior of the sawtooth knock controller is shown in Fig. 14(a). The spark advance timing is continuously increased until a knock intensity measurement exceeds the threshold θ , in which case it is suddenly decreased.

Simulations during 10^5 consecutive cycles have been performed on a Simulink platform to reliably estimate the dispersion of the controlled spark advance, as shown in Fig. 15. The proposed controller has a smaller spark advance dispersion than the sawtooth controller, reducing the engine IMEP undesirable variations. The average values of the spark advance of the proposed controller (-9.9°) and the sawtooth controller (-10.3°) indicate an improved engine efficiency. While both controllers maintain 5% of knocking cycles, those knocking cycles are less intense for the proposed controller than for the sawtooth controller, ensuring a better engine protection. The three properties (smaller IMEP variations, improved efficiency, and limited knock intensities) come from the ability of the proposed controller to maintain the spark advance closer to its ideal value, without constant exploration of its vicinity. Sweeps of the threshold θ have been performed in [15] and show an efficiency gain of 1% at constant knocking frequency.

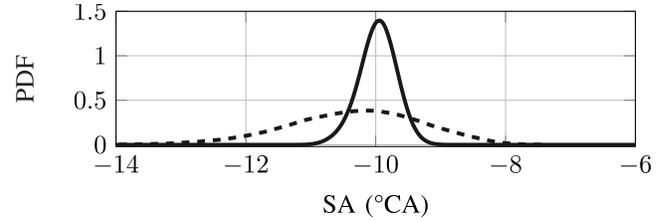


Fig. 15. Spark advance probability density function during stationary operations. Dashed: sawtooth controller. Solid: proposed controller.

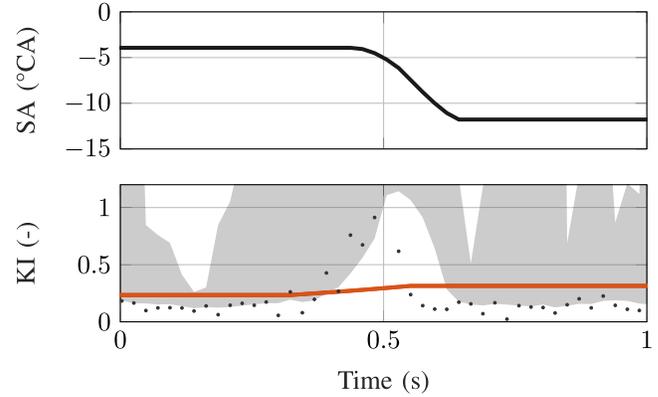


Fig. 16. Transient operation from 19 bar to 25 bar of IMEP. Bottom plot: solid line—threshold. Bottom plot: dots—knock intensity measurements. Bottom plot: filled area—quantile confidence interval.

2) *Transient Regime*: The behavior of the controller was simulated on a transient state shown in Fig. 16, i.e., a step of charge from 19 bar to 25 bar of IMEP at the constant speed of 1750 rpm. Since the knock intensities are randomly generated, different results might be observed if the simulation was repeated. It should be seen as an indicative test of the controller response. No spark advance prepositioning is used, artificially generating a 6° error in spark advance during the transient. The controller immediately starts to decrease the spark advance and reaches its final value within 0.2 s. Response time is consistent with the number of cycles used for the knock estimation, i.e., $n = 5$. Note that the controller is satisfyingly reactive in the protection direction.

The response to transients in both directions (19 bar to 25 bar and then 25 bar to 19 bar of IMEP) of the reference controller and the proposed controller, with different calibrations, are shown in Fig. 17. The variations of the threshold θ indicate the transient timings. The response time depends on the calibrations of K and n . It can be seen that the controller is more sensitive in the protective direction than in the opposite direction, potentially leading to efficiency losses, but far less problematic than failures to protect the engine. It is slower in the opposite direction than the reference controller calibrated in Section VI-A3. It can be seen that a shorter buffer makes the controller to jump from time to time and that a longer buffer tends to smooth its response. Like integral controllers, excessive gain combined with a long buffer makes the controller to oscillate (for example $n = 10$ and $K = 0.45$ at time 6 s).

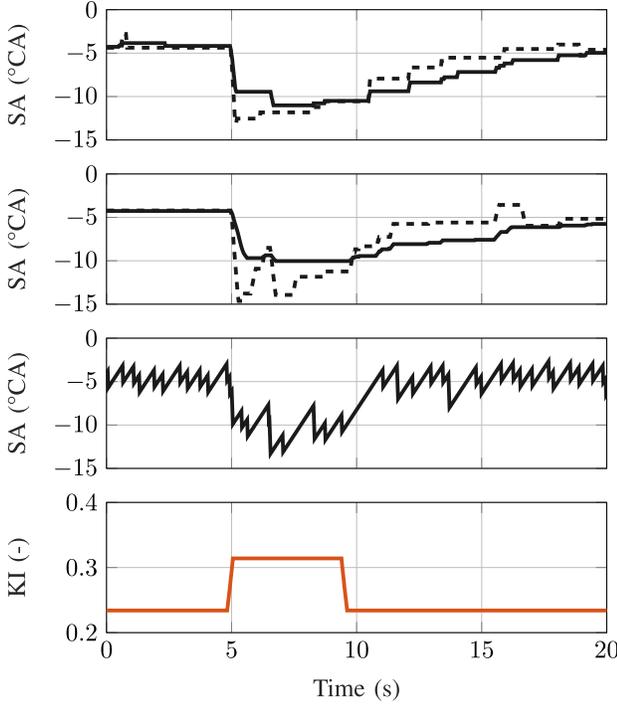


Fig. 17. Transient operations between 25 bar and 19 bar of IMEP proposed controller spark advance with different calibrations. First plot: solid line—($n = 5$ and $K = 0.1$) and dashed line—($n = 5$ and $K = 0.7$). Second plot: solid line—($n = 10$ and $K = 0.1$) and dashed line—($n = 10$ and $K = 0.45$). Third plot: sawtooth controller spark advance. Fourth plot: knock intensity threshold θ indicating transient timings.

VII. CONCLUSION

The stochastic controller proposed in this article estimates the confidence interval of the 95% quantile of the knock intensity distribution. Based on the position of the mapped 95% quantile target and the quantile confidence interval, an adapted spark advance correction is undertaken. This controller is compared with a standard sawtooth-shaped spark advance controller in simulations. The engine knock intensities are simulated based on a random number generator following experimental distributions. The stochastic controller reduces the dispersion of the spark advance timing during stationary operations, leading to smaller engine torque fluctuations. It allows to maintain, on average, the spark advance timing closer to the knock limit while preventing knocking cycles. It is expected to improve the engine efficiency. In addition, the rapidity of the controller and the requirement of only few calibration parameters makes it suited for engine transient operations. Experimental validations, calibration, tuning of the weighting method, and assessment of the impact of departures from probabilistic assumptions on a real engine test bench over the full operating range are planned in the future works.

APPENDIX

EVALUATION OF EXPECTED VALUES AND VARIANCES

A. Evaluation of the Expected Value and Variance of μ_i

Since $\mu_i | \sigma_i \sim \mathcal{N}(m_i, (\sigma_i/\sqrt{\kappa_i}))$

$$E(\mu_i) = E(E(\mu_i | \sigma_i)) = E(m_i) = m_i.$$

Also, $\text{std}(\mu_i | \sigma_i = E(\sigma_i)) = (E(\sigma_i)/\sqrt{\kappa_i})$.

B. Evaluation of the Expected Value and Variance of σ_i

Since $\sigma^2 \sim \mathcal{G}^{-1}(\alpha_i, \beta_i)$, $E(\sigma_i^2) = (\beta_i/\alpha_i - 1)$

$$E(\sigma_i) = \int_{\sigma_i \in \mathbb{R}^+} \sigma_i \pi(\sigma_i) d\sigma_i$$

with the variable change $\sigma_i \mapsto \sigma_i^2 = u$

$$\begin{aligned} E(\sigma_i) &= \int_{u \in \mathbb{R}^+} \sqrt{u} \pi(u) du \\ &= \int_{u \in \mathbb{R}^+} \sqrt{u} \frac{\beta_i^{\alpha_i}}{\Gamma(\alpha_i)} u^{-\alpha_i-1} \exp\left(\frac{-\beta_i}{u}\right) du \\ &= \frac{\beta_i^{\alpha_i}}{\Gamma(\alpha_i)} \int_{u \in \mathbb{R}^+} u^{-\alpha_i-\frac{1}{2}} \exp\left(\frac{-\beta_i}{u}\right) du \end{aligned}$$

with the variable change $\alpha'_i = \alpha_i - (1/2)$

$$E(\sigma_i) = \frac{\beta_i^{\alpha_i}}{\Gamma(\alpha_i)} \frac{\Gamma(\alpha'_i)}{\beta_i^{\alpha'_i}} \int_{u \in \mathbb{R}^+} \frac{\beta_i^{\alpha'_i}}{\Gamma(\alpha'_i)} u^{-\alpha'_i-1} \exp\left(\frac{-\beta_i}{u}\right) du.$$

The integral is the integral of the inverse gamma probability density function with parameters (α'_i, β_i) . It is equal to 1

$$E(\sigma_i) = \frac{\beta_i^{\alpha_i}}{\Gamma(\alpha_i)} \frac{\Gamma(\alpha'_i)}{\beta_i^{\alpha'_i}} = \frac{\beta_i^{\alpha_i}}{\Gamma(\alpha_i)} \frac{\Gamma(\alpha_i - \frac{1}{2})}{\beta_i^{\alpha_i - \frac{1}{2}}} = \frac{\Gamma(\alpha_i - \frac{1}{2})}{\Gamma(\alpha_i)} \sqrt{\beta_i}.$$

Finally

$$\text{Var}(\sigma_i) = E(\sigma_i^2) - (E(\sigma_i))^2$$

and

$$\text{std}(\sigma_i) = \sqrt{\beta_i} \underbrace{\sqrt{\frac{1}{\alpha_i - 1} - \left(\frac{\Gamma(\alpha_i - \frac{1}{2})}{\Gamma(\alpha_i)}\right)^2}}_{v_i}.$$

C. Evaluation of the Variance of q_i

$$\begin{aligned} \text{Var}(q_i) &= \text{Var}_{\sigma_i}(E_{\mu_i|\sigma_i}(q_i)) + E_{\sigma_i}(\text{Var}_{\mu_i|\sigma_i}(q_i)) \\ &= \text{Var}_{\sigma_i}(\mu_i + r\sigma_i) + E_{\sigma_i}\left(\frac{\sigma_i^2}{\kappa_i}\right) \\ &= r^2 v_i^2 \beta_i + \frac{1}{\kappa_i} \frac{\beta_i}{\alpha_i - 1}. \end{aligned}$$

REFERENCES

- [1] J. B. Heywood, *Internal Combustion Engine Fundamentals* (Automotive technology). New York, NY, USA: McGraw-Hill, 1988.
- [2] J. M. Spelina, J. C. Peyton Jones, and J. Frey, "Stochastic simulation and analysis of a classical knock controller," *Int. J. Engine Res.*, vol. 16, no. 3, pp. 461–473, Apr. 2015.
- [3] J. C. P. Jones, J. Frey, K. R. Muske, and D. J. Scholl, "A cumulative-summation-based stochastic knock controller," *Proc. Inst. Mech. Eng., D, J. Automobile Eng.*, vol. 224, no. 7, pp. 969–983, Jul. 2010.
- [4] J. C. Peyton Jones, J. M. Spelina, and J. Frey, "Likelihood-based control of engine knock," *IEEE Trans. Control Syst. Technol.*, vol. 21, no. 6, pp. 2169–2180, Nov. 2013.
- [5] J. M. Spelina, J. C. Peyton Jones, and J. Frey, "Recent advances in knock analysis, simulation, and control," *SAE Int. J. Eng.*, vol. 7, no. 2, pp. 947–955, Apr. 2014.

- [6] J. C. P. Jones, J. M. Spelina, and J. Frey, "An optimal cumsum-based knock controller," *IFAC Proc. Volumes*, vol. 46, no. 21, pp. 372–377, 2013.
- [7] J. C. P. Jones and J. Frey, "Closed loop knock intensity characteristics of a classical knock control system," *IFAC-PapersOnLine*, vol. 48, no. 15, pp. 167–173, 2015.
- [8] J. C. Peyton Jones, S. Shayestehmanesh, and J. Frey, "Closed loop statistical performance analysis of N-K knock controllers," *Mech. Syst. Signal Process.*, vol. 94, pp. 253–266, Sep. 2017.
- [9] A. Thomasson, H. Shi, T. Lindell, L. Eriksson, T. Shen, and J. C. P. Jones, "Experimental validation of a likelihood-based stochastic knock controller," *IEEE Trans. Control Syst. Technol.*, vol. 24, no. 4, pp. 1407–1418, Jul. 2016.
- [10] A. Stotsky, "Statistical engine knock control," *IFAC Proc. Volumes*, vol. 41, no. 2, pp. 7064–7065, 2008.
- [11] J. C. P. Jones, J. B. Roberts, and K. J. Landsborough, "Stochastic prediction of cycle-by-cycle cylinder pressure fluctuations in a spark ignition engine," *Proc. Inst. Mech. Eng., D, J. Automobile Eng.*, vol. 214, no. 4, pp. 435–451, Apr. 2000.
- [12] J. C. Peyton Jones, J. Frey, and S. Shayestehmanesh, "Stochastic simulation and performance analysis of classical knock control algorithms," *IEEE Trans. Control Syst. Technol.*, vol. 25, no. 4, pp. 1307–1317, Jul. 2017.
- [13] K. Zhao and T. Shen, "Normal-gamma distribution-based stochastic knock probability control scheme for spark-ignition engines," *Proc. Inst. Mech. Eng., D, J. Automobile Eng.*, vol. 234, no. 7, pp. 1986–2000, Jun. 2020.
- [14] S. Shayestehmanesh, J. Jones, and J. Frey, "A Bayesian knock event controller," *IEEE Trans. Control Syst. Technol.*, vol. 28 no. 6, pp. 1627–1637, Sep. 2020.
- [15] M. Jean, T. Leroy, and F. Vidal-Naquet, "Trustworthy estimation and control of engine knocking level for transient operation," in *Proc. Int. Conf. Knocking Gasoline Eng.* Berlin, Germany: Springer, 2017, pp. 267–278.
- [16] J. Naber, J. R. Blough, D. Frankowski, M. Goble, and J. E. Szpytman, "Analysis of combustion knock metrics in spark-ignition engines," SAE, Warrendale, PA, USA, Tech. Rep. 2006-01-0400, 2006.
- [17] A. Stahr, P. Langfritz, M. Guenther, and M. Kratzsch, "The delta knocking control—The necessary paradigm shift for engines with high power density," in *Proc. 4th Int. Conf. Knocking Gasoline Eng.*, 2013, pp. 1–6.
- [18] G. G. Zhu, I. Haskara, and J. Winkelman, "Stochastic limit control and its application to knock limit control using ionization feedback," SAE, Warrendale, PA, USA, Tech. Rep. 2005-01-0018, 2005.
- [19] N. Cavina, G. Po, and L. Poggio, "Ion current based spark advance management for maximum torque production and knock control," in *Eng. Syst. Des. Anal.*, vol. 42517, pp. 537–545, Mar. 2006.
- [20] M. Otaka, T. Kasahara, and K. Komaba, "Development of a device that detects knocking by analyzing engine radiation noise," SAE, Warrendale, PA, USA, Tech. Rep. 2016-01-1069, 2016.
- [21] J. M. Spelina, J. C. Peyton Jones, and J. Frey, "Characterization of knock intensity distributions: Part 1: Statistical independence and scalar measures," *Proc. Inst. Mech. Eng., D, J. Automobile Eng.*, vol. 228, no. 2, pp. 117–128, Feb. 2014.
- [22] A. Karvountzis-Kontakiotis, A. Dimaratos, L. Ntziachristos, and Z. Samaras, "Exploring the stochastic and deterministic aspects of cyclic emission variability on a high speed spark-ignition engine," *Energy*, vol. 118, pp. 68–76, Jan. 2017.
- [23] J. C. P. Jones, J. Frey, and K. R. Muske, "A statistical likelihood based knock controller," *IFAC Proc. Volumes*, vol. 43, no. 7, pp. 809–814, Jul. 2010.
- [24] S. Shayestehmanesh and J. C. Peyton Jones, "Stochastic modeling and prediction of IMEP for closed loop knock control performance assessment," *Control Eng. Pract.*, vol. 92, Nov. 2019, Art. no. 104130.
- [25] F. Bozza, D. Siano, and E. Torella, "Cycle-by-cycle analysis, knock modeling and spark-advance setting of a 'downsized' spark-ignition turbocharged engine," *SAE Int. J. Eng.*, vol. 2, no. 2, pp. 381–389, 2010.
- [26] N. Ozdor, M. Dulger, and E. Sher, "Cyclic variability in spark ignition engines a literature survey," *SAE Trans.*, vol. 4, pp. 1514–1552, May 1994.
- [27] J. M. Spelina, J. C. Peyton Jones, and J. Frey, "Characterization of knock intensity distributions: Part 2: Parametric models," *Proc. Inst. Mech. Eng., D, J. Automobile Eng.*, vol. 227, no. 12, pp. 1650–1660, Dec. 2013, doi: 10.1177/0954407013496234.
- [28] P. D. Hill, "Kernel estimation of a distribution function," *Commun. Statist.-Theory Methods*, vol. 14, no. 3, pp. 605–620, 1985.
- [29] F. Millo and C. Ferraro, "Knock in si engines: A comparison between different techniques for detection and control," *SAE Trans.*, vol. 5, pp. 1091–1112, May 1998.
- [30] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and A. B. Rubin, *Bayesian data Analysis*. Boca Raton, FL, USA: CRC Press, 2013.
- [31] J. M. Bernardo, "Reference analysis," *Handbook Stat.*, vol. 25, pp. 17–90, Oct. 2005.
- [32] J. M. Bernardo and A. F. M. Smith, *Bayesian Theory*. Hoboken, NJ, USA: Wiley, 2007.
- [33] J. O. Berger, J. M. Bernardo, and D. Sun, "The formal definition of reference priors," *Ann. Stat.*, vol. 37, no. 2, pp. 905–938, Apr. 2009.
- [34] A. W. V. D. Vaart, *Asymptotic Statistics*, vol. 3. Cambridge, U.K.: Cambridge Univ. Press, 2000.
- [35] W. Hörmann, J. Leydold, and G. Derflinger, *General Principles in Random Variate Generation*. Berlin, Germany: Springer, 2004.



Maxime Jean received the M.Sc. degree from École des Mines ParisTech, Paris, France, in 2013, and from the Université Pierre et Marie Curie, Paris, in 2014.

He is currently a Research Engineer at IFP Energies nouvelles, Rueil-Malmaison, France. His research activities focus on modeling, simulation, and control of systems, such as thermal engines, energy storage processes, and new mobilities. During his studies, he worked as a Mechatronics Engineer at Shell and developed the design of experiment algorithms at Hitachi.



Thomas Leroy graduated from ESSTIN, Vandœuvre-lès-Nancy, France, and received the M.Sc. degree in control theory from Henri Poincaré University, Nancy, France, in 2006, and the Ph.D. degree in control theory and mathematics from the École des Mines ParisTech, Paris, France, in 2010.

He is currently a Research Engineer at IFP Energies nouvelles, Rueil-Malmaison, France. His research activities focus on modeling, control, and simulation of internal combustion engines and hybrid electric powertrains.



Fabien Vidal-Naquet graduated from the École Centrale des Arts et Manufactures, Paris, France, in 2008, specializing in energy and propulsion.

Since April 2010, he has been a Research Engineer in the field of automotive control at IFP Energies nouvelles, Rueil-Malmaison, France, with a focus on thermal engine control, hybrid powertrains energy management, and electric motor control.