

Data-Driven Virtual Sensing for Probabilistic Condition Monitoring of Solenoid Valves

Victor Vantilborgh¹, Tom Lefebvre¹, Kerem Eryilmaz², and Guillaume Crevecoeur¹, *Member, IEEE*

Abstract—There is an emerging industrial demand for predictive maintenance algorithms that exhibit high levels of predictive accuracy. Such condition monitoring tools must estimate dynamic quantities, such as Remaining Useful Lifetime (RUL) and the State of Health (SOH), based on a, typically, restricted set of measurements that can be obtained in an operational setting. These quantities exhibit inherent stochasticity and can only be approximately determined a posteriori to system failure. This paper proposes a generic prognostic tool for probabilistic condition monitoring of mechatronic systems, with the aim to improve the probabilistic prediction of condition metrics, specifically RUL and SOH. Therefore we propose to identify a Hidden Markov Model (HMM) from a fully instrumented measurement set, that is only available for a restricted set of run-to-failure experiments, typically gathered in an R&D setting. Although being artificial and retrospectively constructed metrics, we interpret RUL and SOH as physical measurements with the purpose to identify accurate degradation dynamics. Once the degradation model is identified, we practice the mathematical flexibility of the HMM framework to estimate several of the no longer available dynamic quantities of interest in real-time, from the limited set of measurements that are available in an operational setting. This modelling paradigm is known as virtual sensing. Predictive performance and computational efficiency are further improved by domain knowledge based pre-processing of the measurements. We apply our methodology to solenoid valves (SV), a widely used and often critical component in many industrial systems, which display a large variation in useful lifetime. Benchmark results show that the predictive capabilities of the presented methodology compares with prognostic techniques that are more computationally and memory demanding.

Note to Practitioners—The motivation for this research is twofold. First there is a pending industrial need for improved diagnostic and prognostic tools. Second there is the observation that lifetime tests usually take place in an R&D setting and that expert labelling of Remaining Useful Lifetime (RUL) or State

of Health (SOH) of a component or system is often based on measurement data that is not available in the industrial setting where the prognostic tools are to be deployed in the end. These two observations suggest that there is large potential in methods that can correlate the expert labelling, in particular RUL & SOH signals, with measurement data that is available in the industrial setting. Our approach has been tested in detail on the case of Solenoid Valves, which are widely used in industry and that are often safety critical. Our experiments demonstrate that the method compares with brute force approaches that overpower ours both in terms of computational as well as memory requirements. The method is furthermore generic and there is no reason to assume it would not work for other applications.

Index Terms—Predictive maintenance, condition monitoring.

NOMENCLATURE

Acronyms

AC	Alternating Current.
BCNN	Bayesian Convolutional Neural Network.
CNN	Convolutional Neural Network.
DL	Deep Learning.
EM	Expectation-Maximization.
EOL	End Of Life.
HGMM	Hidden Gaussian Markov Model.
HMM	Hidden Markov Model.
L2R	Left-to-right.
MAE	Mean absolute error.
MAP	Maximum A Posteriori.
MLE	Maximum Likelihood Estimation.
MSD	Mass-spring-damper.
NRUL	Normalized Remaining Useful Lifetime.
R&D	Research and Development.
RUL	Remaining Useful Lifetime.
SOH	State of Health.
SV	Solenoid valve.
UQ	Uncertainty Quantification.

Symbols

Δ	Time span.
ϵ	Arbitrary threshold value.
\mathcal{D}	Dataset.
\mathcal{T}	Similarity transformation matrix.
\mathcal{Y}	Data sequences of y .
\mathcal{Z}	Data sequences of z .
μ_t	State mean vector at time t .
Σ_t	state covariance matrix at time t .
$\mathcal{Q}(\cdot, \cdot)$	Evidence Lower Bound.

Manuscript received 3 March 2023; revised 12 May 2023; accepted 30 May 2023. This article was recommended for publication by Editor P. Zhou upon evaluation of the reviewers' comments. This work was supported in part by the Flanders Make Projects QUASIMO and HAIEM and in part by the Flemish Government [Artificial Intelligence (AI) Research Program]. (Victor Vantilborgh and Tom Lefebvre contributed equally to this work.) (Corresponding author: Victor Vantilborgh.)

Victor Vantilborgh, Tom Lefebvre, and Guillaume Crevecoeur are with the Dynamic Design Laboratory (D²LAB), Department of Electromechanical, Systems and Metal Engineering, Ghent University, 9052 Ghent, Belgium, and also with the Core Laboratory MIRO, Strategic Research Centre for the Manufacturing Industry, Flanders Make, 3920 Flanders, Belgium (e-mail: victor.vantilborgh@ugent.be; tom.lefebvre@ugent.be; guillaume.crevecoeur@ugent.be).

Kerem Eryilmaz is with the Strategic Research Centre for the Manufacturing Industry, Decisions, Flanders Make, 3920 Flanders, Belgium.

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TASE.2023.3287598>.

Digital Object Identifier 10.1109/TASE.2023.3287598

θ	Model parameters.
A	State transition matrix.
C	Emission matrix.
I	Identity matrix.
$I_{m, fund}$	Impedance spectrum magnitude at the fundamental AC-source frequency.
$I_{d, fund}$	Inductance spectrum magnitude at the fundamental AC-source frequency.
K	Total number of data sequences.
l	Health label.
n_y	Dimension of the vector y .
n_z	Dimension of the vector z .
n_x	Dimension of the vector x .
Q	System covariance matrix.
q_{leak}	Leakage flow during solenoid off-phase.
q_{out}	Outlet flow during solenoid on-phase.
R	Emission covariance matrix.
T	Length of data sequence.
T_d	Failure kick-in time instant.
T_f	Deterioration kick-in time instant.
t_h	Time of first hit during valve opening.
T_{surf}	Plunger surface temperature.
x	Latent state vector.
y	Measurements assumed to be always available.
z	Measurement available for restricted set of experiments.

Operators

$\mathbb{H}(\cdot)$	Entropy.
$\mathcal{N}(\cdot, \cdot)$	Normal distribution.
$\mathcal{Q}(\cdot, \cdot)$	Evidence Lower Bound.
$p(\cdot \cdot)$	Conditional probability density function.
$p(\cdot)$	Probability density function.
$p(\cdot)$	Probability density function.

Subscripts

t	Time index.
-----	-------------

Superscripts

*	Optimal value.
i	i -th iteration.
k	k -th data sequence.

Notations

$\dot{\cdot}$	Given sequence until time t .
---------------	---------------------------------

I. INTRODUCTION

THE operational efficiency of industrial processes can drastically benefit from the integration and optimization of maintenance strategies in the logistic decision making process. The practice of maintenance strategies allows for maximized operation time of equipment, avoidance of unexpected machine failure and reduction of overall operational costs [1]. In today's industry, it is already common practice to rely on domain knowledge based approaches, such as Fault Tree Analysis (FTA) [2] and Failure Modes and Effect Analysis (FMEA) [3], for reliability engineering. Though, such rule-based methods require a priori knowledge about

the degradation dynamics, they cannot cope with temporal dependencies and effects, and rely heavily on the actual expertise of practitioners [4].

Meanwhile, industry is undergoing the *Fourth Industrial Revolution*, which is characterised by the increased integration of digital systems in physical production environments. This evolution, in combination with the steady price drop of sensory equipment, is accompanied by increased instrumentation and data logging. Subsequently, measurement data has become abundantly available for (real-time) statistical analysis and (post-)processing [5]. In turn this has increased the expectations and requirements that industrial practitioners have with regard to monitoring and estimation procedures in the context of maintenance, better known as Predictive Maintenance [6].

A central objective of predictive maintenance algorithms is to predict the Remaining Useful Lifetime (RUL) of a machine (component) [7] or to estimate its State of Health (SOH). The RUL of a machine is the expected usage time remaining before the machine requires repair or replacement. The SOH on the other hand is an artificial metric related to the condition of the system compared to ideal conditions, often the conditions at the time of manufacturing or when it was first commissioned. Both physics-based and data-driven approaches can be pursued to estimate RUL and SOH.¹

Classical physics-based approaches rely on a comprehensive understanding of the system's behavior and potential failure mechanisms to predict the time evolution of critical states, which can be used to indirectly deduce the RUL [8]. However, such physics-based approaches rely on intimate knowledge of the process as well as how systems can degrade or fail. There is furthermore a limit on the extent that long horizon forecasts can be trusted [19]. So instead of relying on exact physical models and probabilistic simulations, rather one emulates the degradation dynamics directly using a discrete HMMs (dHMMs) where the latent state-space occupies a discrete set. For the purpose of RUL and SOH estimation, one imposes a so called left-to-right structure on the dHMM where the latent states represent degrading states of the system with the left most state(s) representing healthier and the right most state(s) faultier states. The RUL is estimated by calculating the remaining time of absorption conditioned onto the observation sequence up to the present time instant t . Whilst the computational procedure resembles that of a physics based approach, no true physical meaning can be attributed to the latent space. Useful references in this line of work are [20], [21], [22], [23], [24], [25], [26], [27], and [28]. In conclusion we may note that lifetime tests are required to fit the imposed model structure making the distinction with data driven approaches more ambiguous.

Data-driven approaches, on the other hand, typically rely on a set of lifetime tests and associated condition monitoring data to estimate the RUL directly. These rely exclusively on a (limited) set of (accelerated or historical) lifetime tests and associated condition monitoring data. This library is then used to detect and compare trends in the real-time measurements,

¹The following overview is non-exhaustive as we limit here to a number of approaches that are closest related, in spirit, or computationally, to ours.

and, process the time series data into a direct estimate of the RUL. Similarity and survival models are two families of methods that rely on historical lifetime tests to predict statistical estimates with associated uncertainty [9], [10]. These methods however do not involve an *offline* pre-processing step and generate an *online* RUL estimate based on a similarity assessment of the present time signals with recorded signals. More recently, the use of Deep Learning (DL) techniques has also been proposed [29]. It is argued that DL approaches are well positioned to automate the pre-processing and extraction of useful features to a high degree [11]. The information from the library is captured in a single purpose model, mapping the input into a deterministic RUL estimate. Extensions have been made that capture uncertainties, mostly by relying on Bayesian Neural Networks (BNNs), and thus are able to express their confidence in the estimate [12]. Convolutional (BCNNs) architectures can be used to treat high-dimensional time series data, capturing both spatial and temporal features, representing the sensor signals as images. These approaches however yield no insight in the complex mechanisms underlying the process; and, once learned, the model cannot be repurposed. Important references are [12], [30], [31], [32], [33], [34], [35], [36], [37], [38], [39], [40], and [41].

To overcome some of the deficiencies of strictly DL approaches whilst exploiting the principles and well understood theory of physics-based modelling, in this work we propose a data-driven virtual sensing² approach tailored to RUL and SOH estimation. Our data-driven virtual sensing approach assumes access to a highly instrumented time series data set from dedicated run-to-failure experiments, which may only be obtained in an R&D setting. In agreement with the virtual sensing paradigm, we aim to estimate several of the measurements that are impractical or impossible in an industrial setting, from a limited set of measurements that can be obtained. We therefore rely on Hidden Markov Model (HMM) theory as an explanatory causal framework, interpreting the data as (artificial) measurements from a dynamic latent state. A HMM is identified using a regularized implementation of the Expectation-Maximization algorithm. The learned HMM is then deployed in a model-based state estimation framework. By reinterpretation of the RUL and SOH as physical measurements –which are only available once the system has failed– this strategy can be repurposed to produce probabilistic estimates of both these quantities of interest. Moreover, we obtain a figure of the uncertainty as a silent feature originating from the HMM framework. To the best of our knowledge this is the first publication to address predictive maintenance by repurposing the virtual sensing paradigm. We consider the SOH and RUL as metrics that are a function of latent variables. This enables closer correspondence with the physical reality of the machine or component degradation mechanism that demands monitoring. This to ultimately reach for higher levels of predictive accuracies.

We validate our approach to predict the RUL and SOH with uncertainty quantification of Solenoid Valves (SVs). SVs are

commonly used components in many industrial processes and often perform safety critical tasks. Accelerated lifetime tests show that there is a large variation in their lifetime. We aim to base the RUL prediction solely on the measurement of the AC current of the SVs, as to come up with a noninvasive prognostics tool whilst only relying on industrially available sensory equipment. Given the arbitrary linear character of the RUL and SOH over time, in this preliminary study we used a linear state-space model to model the deterioration dynamics of the SVs and make prediction about their lifetime and health. Furthermore, expert domain knowledge is leveraged to pre-process the time-series data in the form of useful feature extraction. Our goal with this case study is to show that by leveraging a limited amount of domain knowledge to replace the automatic pre-processing and feature extraction a much simpler linear time-invariant model is able to perform at similar level or above the BCNN approach from [12].

The contributions of this paper are: (1) repurposing Hidden Gaussian Markov Models (HGMM) in the context of condition monitoring by interpreting posterior RUL estimates of a historical dataset as measurement variables and treating the problem as a probabilistic system identification problem, (2) demonstrating the advantage of using multiple measurement variables during training which need not be reconstructed during real-time monitoring, apart from those that are of interest at that time, (3) validating and analyzing our proposed method in terms of predictive capabilities about the remaining useful lifetime of solenoid valves.

II. METHODOLOGY

A. Mathematical Problem Formulation

First let us specify the distinction between the two types of measurements mentioned in the introduction. Measurements in general will be denoted with the symbol $y \in \mathbb{R}^{n_y}$ or $z \in \mathbb{R}^{n_z}$, when a time label is available, a subscript is added, for example z_t . For brevity we introduce the format $\underline{y}_t = \{y_1, \dots, y_t\}$ for leading subsequences of temporal data. We silently assume that all sequences start at discrete time $t = 1$. Finally, we emphasize that the concept of a *measurement* is used in a broad sense, meaning that measurements can represent unprocessed sensory outputs but can also represent processed information or so called feature variables.

The first type of measurements, y , are assumed to be available always, meaning that whatever the circumstances and time of the experiment are, a value for the measurement is available. For the second type of measurements, z , this assumption is no longer valid, meaning that the corresponding measurement values are only available for a restricted set of experiments, for example the type of experiments that can be (and were) conducted in a research and development (R&D) facility but that cannot be replicated in an industrial environment for one reason or the other. The combination of these two measurements determine the complete measurement, $\{y, z\}$. It is further assumed that the complete measurement is available in a restricted set of experiments so that correlations can be drawn between y and z .

As stated before, for many applications of practical interest it would be considered highly purposeful if we could somehow

²Sometimes also referred to as “soft sensing”.

estimate the measurements that are inaccessible in an industrial experimental setting; equivalently, that are only available in the R&D setting, based on the measurements values that are available in an industrial setting. More specifically, if we were able to construct a function that maps, y , into, z .

Provided the stochastic nature of measurements and system behaviour per se (aleatoric) and the uncertain reasoning inherent to modelling from data (epistemic), we adopt a probabilistic framework, and replace the deterministic function described above by a probabilistic model, providing a description of the probability density of z rather than a point estimate; all in all, yielding a more informative description of z . Formally such an estimate could be expressed as the probability of measuring, z , conditioned on the known measurement value, y , that would be $p(z|y)$. Now, when a measurement sequence is available, meaning we have multiple measurements of the first category, e.g. \underline{y}_t , taken at different time instants in the (near) past preceding our current estimate about z_t , we can extend our probabilistic model accordingly. Of course, we could store all of these measurements, notation-wise, in a single measurement, y ; though making the dynamic nature explicit will turn out to be very useful for computational purposes. In summary, the goal of this work is to compute the following probability from data.

$$p(z_t|\underline{y}_t) \quad (1)$$

B. Definition of RUL and SOH Signals

In this work we focus on the case where z represents some abstract monitoring metrics, in particular Remaining Useful Lifetime (RUL) and State of Health (SOH) signals. We aim to estimate RUL or SOH from measurements, y , that are easily accessed on the system when deployed in practice. Such are highly application specific and are not specified further here.

The RUL is a subjective estimate of the remaining time that an item, component, or system is able to function in accordance with its intended purpose before warranting replacement. In the ideal case the RUL can be associated to some physical quantity that exceeds some limit. One could then estimate a value for that quantity and calculate the probability that it exceeds the limit or even calculate the probability of that estimate exceeding the prescribed numerical limit within a certain period of time, the latter being the closest to the objective definition of RUL.

$$\begin{aligned} \text{RUL}(t) &= \arg \max_T T \\ \text{s.t. } P(f(x_{t+T}) < \epsilon_f | \underline{y}_t) &> \epsilon_p \end{aligned} \quad (2)$$

Here x_t represents a physical variable describing the full state of the system, f denotes a function of that state that if exceeding a value ϵ_f implies the system has failed and finally $P(f(x_{t+T}) < \epsilon_f | \underline{y}_t)$ denotes the probability of the predicted value $f(x_{t+T})$ exceeding ϵ_f conditioned on the measurements \underline{y}_t and on its turn must exceed a threshold value ϵ_p .

Unfortunately such intimate knowledge of the system is rare. In practice such an explicit physical definition of the end of life is usually absent. Therefore we may exploit the *benefit of hindsight*. For experiments that we *know* to have

failed, we can calculate an RUL *signal* retrospectively, defining its value simply as the remaining time between its present evaluation and the moment that the system has failed.³

$$\text{RUL}(t) = \max\{T_f - t, 0\}, \quad t > 0, T_f > 0 \quad (3)$$

where the time instant T_f refers to the first time failure is observed. Because this implies large variability in the initial RUL value, $\text{RUL}(0)$, instead we use here the normalized RUL (NRUL). Consequently, any prognostics model should only predict the slope of *degradation* and no longer the initial bias.

$$\text{NRUL}(t) = \frac{\text{RUL}(t)}{T_f} \quad (4)$$

From the NRUL we can still compute the RUL

$$\text{RUL}(t) = \frac{\text{NRUL}(t)}{1 - \text{NRUL}(t)} t, \quad \text{NRUL} \in [0, 1) \quad (5)$$

The SOH is a subjective estimate of the *health index or condition* of the system compared to the *ideal* situation. As opposed to the RUL it is difficult to determine a physical definition for the SOH other than the observation that the system finally broke down. To highlight correspondence with the (N)RUL signals we adopt a similar approach, yet recognize a healthy section in the system's lifetime before degradation and ultimately failure kick in. In retrospect, the NRUL predictor in (4) thus assumes the system instantly starts degrading from the onset of first commissioning.

$$\text{SOH}(t) = \begin{cases} 1, & t < T_d \\ \max\left\{\frac{T_f - t}{T_f - T_d}, 0\right\}, & t \geq T_d \end{cases} \quad (6)$$

Here T_d refers to the time instant where deterioration kicks in. Now assume that we have access to a limited set of lifetime tests consisting of a set of measurements $y(t)$ that can also be accessed in practice and a set of auxiliary measurements $z(t)$. Using the approach described above we can then extend the auxiliary measurement set with signals $\text{RUL}(t)$ and/or $\text{SOH}(t)$. Corresponding to the formulation in section II-A we are interested in identifying a probabilistic model that can estimate the RUL/SOH signal from on-site measurements, \underline{y}_t . This approach circumvents the need for a physical cause of failure or a definition of the health of the system elegantly.

C. Virtual Sensing Using HMMs as an Explanatory Framework

The question that remains unanswered is how to compute this relation efficiently? Probabilistically we are trying to model the correlation between two (a primary and a secondary) measurement sequences, y and z respectively. The computational complexity of such a model is heavily influenced by the conditional independencies that are imposed on the variables that constitute the model. To foster further insight, it helps to visualize the conditional dependencies that govern the model

³Remark that, in the absence of uncertainty, if f , in the example above, is a monotonically increasing function of time - which is a reasonable assumption provided that degradation is a progressive process that cannot reverse its action - the RUL in definition (2) coincides with the definition provided in (3). That is $\text{RUL}(t) = \{T(t) | f(t+T) = \epsilon_f\} = f^{-1}(\epsilon_f) - t = T_f - t$.

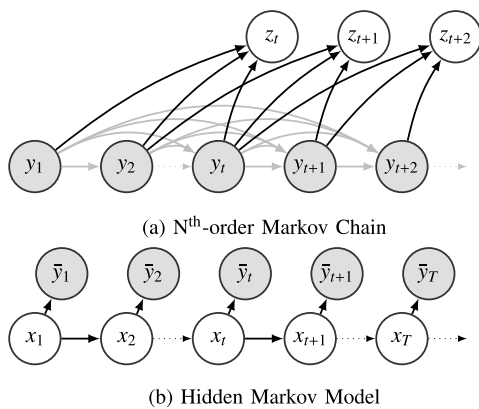


Fig. 1. Graphical representation of probabilistic models for virtual sensing applications. Dotted lines represent an undefined number of repetitions of the same graphical structure.

graphically [13]. In Fig. 1a a graphical representation of the most straightforward probabilistic model implied by (1) is illustrated. Here arrows indicate a conditional dependency between the child and parent variables. Hence, the secondary measurement z_t is thought to be conditionally dependent on the primary measurement sequence of the (exemplary) length t . This model seems to suggest a causal relation between the variables in \underline{y}_t and the variable z_t . Under the assumption that the relation is stationary and that a sequence of length t contains all the required information to predict z_t from the secondary measurement sequence, indeed the most straightforward way is to model the function $p(z_t|\underline{y}_t)$ directly and shift the argument one time instant to compute z_{t+1} and so on [12]. Though, clearly this is not the only way of modelling the correlation between two measurement sequences y and z and certainly not the one with the most computationally rich structure. As an example, assume that we become interested in computing the probability $p(z_{t+\Delta}|\underline{y}_t)$, $\Delta \in \mathbb{N}$ for that matter? Extending the previous modelling approach would then imply learning separate forward models for each of these distributions, not being able to lean on conditional independencies and associated factorization properties to simplify the calculations [13].

Therefore we propose another way of modelling the correlation between the two measurement sequences. We assume that there exists a single underlying process characterised by the latent state, $x_t \in \mathbb{R}^{n_x}$, that spawns both measurement sequences independently. The probabilistic graph corresponding with this point of view is represented in Fig. 1b. This model is known as a Hidden Markov Model (HMM) [14], [15]. As suggested by the arrows, characterisation of the HMM requires specification of the following local probability models.

- initial probability $p(x_1)$
- transition model $p(x_{t+1}|x_t)$
- measurement models $p(y_t|x_t)$ and $p(z_t|x_t)$

Relying on the HMM imposed on the dynamics of the experiments, our estimate, $p(z_t|\underline{y}_t)$, can now be decomposed by means of the latent state, x_t , which circumvents many of the conditional dependencies that made the direct

modelling approach complicated and computationally unattractive. Instead, the HMM offers a practical computational procedure. Specifically, the prediction of $z_{t+\Delta}$, $\Delta \in \mathbb{N}$ by conditioning on \underline{y}_t , then becomes trivial once we have calculated $p(x_{t+\Delta}|\underline{y}_t)$.

$$p(z_{t+\Delta}|\underline{y}_t) = \int p(z_{t+\Delta}|\underline{x}_{t+\Delta})p(x_{t+\Delta}|\underline{y}_t)dx_{t+\Delta} \quad (7)$$

So why bother with estimating the second class of measurements, z , if we can also access a state estimate? The answer is related to the means by which we shall arrive at a model to practice the probabilistic inferences described above. Within this work, we assume that such a model is not existent and cannot be constructed in any meaningful sense by a team of experts. Especially this is the case when the measurements represent *processed* sensory output or features. The causality of the process and hence the temporal causality of the sensory output will be transferred to the features but the underlying physical phenomena will be obscured by the operations wielded to construct those features. Thus, instead we assume only to have access to a set of experimental data, $\{\mathcal{Y}, \mathcal{Z}\}$, existing out of K measurements sequences, \underline{y}_T and \underline{z}_T , i.e. $\mathcal{Y} = \{y_T^k\}_k$ and $\mathcal{Z} = \{z_T^k\}_k$. From those measurements we shall identify a HMM. As we will come to explain later, in this case the state may be uninterpretable and therefore useless for any practical purpose *other than calculation*. In summary, we exploit the structure of the HMM only to have a more insightful model of the correlation between y and z , rather than we hope to extract any useful information from x itself.

D. Filtering & Predicting of Latent Variables

In the previous subsection, we boasted the computational flexibility of the HMM, demonstrating how $p(z_{t+\Delta}|\underline{y}_t)$ could be computed efficiently once we had calculated $p(x_{t+\Delta}|\underline{y}_t)$. Here we explain how to compute the probability $p(x_{t+\Delta}|\underline{y}_t)$.

In HMM theory, the probability $p(x_t|\underline{y}_t)$ is known as the filtering distribution and can be calculated recursively [16]

$$p(x_{t+1}|\underline{y}_{t+1}) \propto p(y_{t+1}|x_{t+1}) \int p(x_{t+1}|x_t)p(x_t|\underline{y}_t)dx_t \quad (8)$$

The probability $p(x_{t+\Delta}|\underline{y}_t)$, $\Delta > 0$ is known as the predictive distribution and can be calculated recursively once the filtering distribution has been calculated.

$$p(x_{t+\Delta+1}|\underline{y}_t) = \int p(x_{t+\Delta+1}|x_{t+\Delta})p(x_{t+\Delta}|\underline{y}_t)dx_{t+\Delta} \quad (9)$$

E. Identification of H(G)MMs

Before we can practice the inferences described above, we require a probabilistic model, say \mathcal{M} , capturing the conditional probabilities that govern the HMM. Such is usually obtained by parametrising the probabilities with a parameter θ . Formally, we write $p(\cdot|\cdot, \theta)$, more in particular, we search quantitative local probabilistic models, $p(x_{t+1}|x_t, \theta)$, $p(y_t|x_t, \theta)$ and $p(z_t|x_t, \theta)$. In this section we further treat the question of how to characterize a HMM given the data, $\mathcal{D} = \{\mathcal{Y}, \mathcal{Z}\}$.

A fully Bayesian approach would imply calculation of the a posteriori probability over the parameters θ conditioned on the data \mathcal{D} , i.e. $p(\theta|\mathcal{D})$. Assuming that we can evaluate the probability $p(\theta|\mathcal{D})$, the model is then determined by expressing the joint probability $p(\theta, \cdot|\mathcal{D}) = p(\cdot|\theta)p(\theta|\mathcal{D})$ and marginalizing out the parametric dependency, rendering the model $p(x_{t+1}|x_t, \mathcal{D})$, $p(y_t|x_t, \mathcal{D})$ and $p(z_t|x_t, \mathcal{D})$. This approach captures both aleatoric uncertainty about the system itself (as governed by $p(\cdot|\theta)$) and epistemic uncertainty about the true parameters. Unfortunately the marginalization and hence the approach in general is intractable but for a handful of toy problems. Alternatively, a Maximum A Posteriori (MAP) estimate, θ^* , can be made by maximizing the probability of the parameters conditioned on the data, i.e. $\theta^* = \arg \max_{\theta} \log p(\theta|\mathcal{D})$. The MAP estimate can then be substituted into the model $p(\cdot|\theta^*)$ so to obtain an estimate of the true Bayesian model. This approach still requires that we evaluate $p(\theta|\mathcal{D})$ which, using Bayes' rule, can be rewritten as $p(\mathcal{D}|\theta)p(\theta)$. Application of the MAP modelling approach, thus requires us to define a prior probability for θ , which, in most cases, cannot be established in any meaningful way so that the uniform is chosen. The MAP modelling approach with uniform prior collapses onto a Maximum Likelihood Estimation (MLE). The MLE parameter is estimated as $\theta^* = \arg \max_{\theta} \log p(\mathcal{D}|\theta)$. This objective expresses the likelihood of the measurements as a function of the parameters.

Applying the MLE approach to the model in Fig. 1b, thus requires to solve the following optimization problem

$$\max_{\theta} \log p(\mathcal{Y}, \mathcal{Z}|\theta) = \max_{\theta} \sum_k \log p(\{y, z\}_T^k | \theta) \quad (10)$$

where

$$p(\{y, z\}_T^k | \theta) = \int p(\underline{x}_T, \{y, z\}_T^k | \theta) d\underline{x}_T \quad (11)$$

It is cumbersome to treat this optimization problem directly. Instead, the Expectation-Maximization (EM) algorithm estimates the parameters iteratively by considering a surrogate objective $Q(\theta, \theta')$, updating the parameters as follows

$$\theta^{i+1} = \arg \max_{\theta} Q(\theta, \theta^i) \quad (12)$$

The objective, $Q(\theta, \theta')$, is often referred to as the Evidence Lower Bound (ELBO). A brief overview of the EM algorithm is given in and Appendix B. For details we refer to [14], [15].

$$Q(\theta, \theta') = \sum_k \int p(\underline{x}_T | \{y, z\}_T^k; \theta') \log p(\underline{x}_T, \{y, z\}_T^k; \theta) d\underline{x}_T \quad (13)$$

It can be shown that the series in (12) converges monotonically to the solution, θ^* , of problem (10). The surrogate ELBO objective and the original MLE objective are then related as

$$\log p(\mathcal{Y}, \mathcal{Z}|\theta^*) = Q(\theta^*, \theta^*) + \sum_k \mathbb{H} \left[p(\underline{x}_T | \{y, z\}_T^k; \theta^*) \right] \quad (14)$$

1) *Hidden Gauss-Markov Models*: Solving (12) for arbitrary probabilistic state-space models remains intractable or requires the introduction of additional approximations. For linear state-space models however, (12) can be solved explicitly. A linear continuous HMM is referred to as a Hidden Gauss-Markov Model (HGMM). Characterisation of a time invariant HGMM requires specification of the following local probabilities.

The parameter θ is thus determined by $\{\mu_1, \Sigma_1, A, Q, C, R\}$.

$$\begin{aligned} p(x_1) &= \mathcal{N}(x_1; \mu_1, \Sigma_1) \\ p(x_{t+1}|x_t) &= \mathcal{N}(x_{t+1}; Ax_t, Q) \\ p(\{y, z\}_t|x_t) &= \mathcal{N}(\{y, z\}_t; Cx_t, R) \end{aligned} \quad (15)$$

Solution of (12) for given θ^i then yields the required parameter updates. The solution of this problem for HGMMs is well studied in for example [16], [17], and [18]. In particular the parameters $\{\mu_1, \Sigma_1, A, Q, C, R\}$ can be expressed as a function of the smoothing distribution parameters which are also Gaussian, $p(x_t|\{y, z\}_T^k) = \mathcal{N}(x_t; \bar{\mu}_t^k, \bar{\Sigma}_t^k)$. Similar to the filtering distributions (see sec. II-D), the smoothing distributions can be computed efficiently. We refer to Appendix C for computational details regarding the parameter updates.

2) *Regularization*: The measurement likelihood, $p(\mathcal{Y}, \mathcal{Z}|\theta)$, of a time invariant HGMM is subject to an important parameter invariance structure that may impact convergence [17]. It is well known that the latent space can be determined up to a similarity transform, $x' = \mathcal{T}x$, where \mathcal{T} is a non-singular square matrix. One verifies that the parameter set θ can be transformed in an equivalent parameter set θ' without affecting the value of (11). Therefore, the EM algorithm may converge to any member of this manifold, depending on the initial values for the parameter estimates. Theoretically one cannot express any preference to any of these members, however, for numerical reasons it is advised to constrain the model family in (12) by imposing restrictions on θ .

In this work we imposed the following constraints.

- The covariance matrices Q and R are constrained to the diagonal set. The update is performed as per usual but only the diagonal elements are maintained in Q' and R'.
- After each standard update, the transformation \mathcal{T} is chosen so that $C' = C\mathcal{T}^{-1} = I$. In case that $n_x > n_y + n_z$, the leading $n_y + n_z$ latent states are selected out. The case where $n_x < n_y + n_z$ was not considered since this would imply that $n_y + n_z - n_x$ measurements are linearly depend of the other n_x . Finally, it is worth mentioning that determining the transformation matrix every update, generally yielded better results than fixing $C = I$.

If these constraints are satisfied, the local probability models, $p(y_t|x_t)$, and, $p(z_t|x_t)$, are in fact uncorrelated, so that

$$\begin{aligned} p(y_t|x_t) &= \mathcal{N}(y_t; x_{t,\{1,n_y\}}, \mathbf{R}_{\{1,n_y\}}) \\ p(z_t|x_t) &= \mathcal{N}(z_t; x_{t,\{n_y+1,n_y+n_z\}}, \mathbf{R}_{\{n_y+1,n_y+n_z\}}) \end{aligned} \quad (16)$$

We list here also our solution initialization procedures.

- We initialize the elements of A and C uniformly in $[0, 1)$.
- The covariance matrix Σ_1 is determined as $\Sigma_1 = \mathbf{P}\mathbf{P}^T$ where P is initialized similar to A and C.

- The covariance matrices Q and R are initialised with diagonal elements uniformly in $[0, 10^{-2}]$.

F. Comparison to State-of-the-Art

Now that we presented our approach to calculate the probability (1), we can revisit some of the other modeling formalism's that address similar purposes, in the particular case where z contains RUL and SOH signals. Therefore we recall from the introduction the work on discrete HMMs [20], [21], [22], [23], [24], [25], [26], [27], [28] and on data driven black-box approaches [12], [30], [31], [32], [33], [34], [35], [36], [37], [38], [39], [40], [41]. Our presented data-driven virtual sensing methodology for probabilistic condition monitoring can be positioned with respect to the above state-of-the-art as follows.

First we may note that the expressiveness of a discrete HMM is limited. In a left-to-right dHMM it is assumed that each state relates to some level of degradation of the associated object. Hence this implies that the associated observations can only depend on the degradation level of the object, i.e. the hidden state, which is by virtue of the HMM modelling assumptions. As a result either the range of measurement associated to a degradation state is limited or the variance of the emission model is large. Secondly, the brute force data-driven approaches aim to recognise patterns between the data and the new experiment excluding the possibility to exploit any physical insight nor allow any physical interpretation.

Our approach is hence situated in between either lines of work discussed above. Strictly speaking, the HMM structure simply offers a computational structure to calculate the probability (1) efficiently. Therewith our approach closely resembles the strictly data driven approaches. On the other hand, the HMM structure can also be compared with a state-space model that governs the physical dynamics of the system. So although it is difficult to attribute a physical interpretation to the identified latent state space, it does cover the dynamics of the latent physical system dynamics. In that sense our approach is closest related with the first line of work. With the exception that in our approach the internal state of the degrading object is represented by n continuous state variables, significantly improving the expressiveness of the model. The observations need not to depend exclusively on the degradation state but can rely on any combination of the latent state variables. In conclusion we emphasize that we know of no other works that can and do consider the use of other measurements signals in the vector z to stimulate the discovery of a physical latent space; which, as we will show with our experiments, turns out to be highly purposeful.

III. SOLENOID VALVE DATASET

We tested the approach from sec. II on a data set containing accelerated lifetime tests of Solenoid Valves (SVs). SVs are commonly used components in many industrial processes, often performing safety critical tasks. These tests show large variation in SV lifetime. Development of predictive maintenance algorithms for SVs could significantly benefit SV reliant applications. In this section we discuss earlier work on diagnostics and prognostics of SVs. Further we provide details regarding the set-up and data that we used in sec. IV.

A. Related Work on SV Diagnostics & Prognostics

Model-based approaches have been proposed by [42], [43], and [44]. These works focus on deterministic fault detection and diagnosis. Data-driven approaches have been proposed for SOH diagnostics of SVs and were successful in anomaly detection and failure clustering. All of these methods, however, fail to give an indication of the End of (useful) Life (EOL). Tod et al. [45] presents a hybrid physics-based data-driven tool that is capable of diagnosing failure modes using a physically interpretable model. Mazeav et al. [11] proposed a Deep Learning method using an ensemble of Convolutional Neural Networks (CNNs) to construct a health index of SVs, which can be extrapolated to a deterministic RUL prediction, based on current measurements in the power electronics circuit feeding the SVs. Later this work was continued and expanded [12] by combining Bayesian CNNs with physical features as described in [45], as to be the first and to our knowledge the only, to obtain accurate RUL predictions with uncertainty quantification (UQ). These approaches obtain accurate results but tend to rely on computationally demanding learning methods and memory demanding models.

B. Set-up

An experimental test set-up, as shown in Fig. 2, was used to perform accelerated lifetime tests of 48 solenoid valves under realistic conditions. The tested valves are direct acting 3/2 way normally closed AC solenoid valves. The SVs are switched on and off at a rate of 1Hz by an AC voltage of 110V at 50Hz. During the open or "ON" state, fluidum can flow through the valve for approximately 0.5 seconds, followed by a transition to the closed or "OFF" state, during which fluidum is blocked for another 0.5 seconds during each cycle. The time-series data of one on/off cycle was captured every hour over the span of 6 weeks (~ 1150 h). All solenoids were supplied by a line pressure of 8bar and the tests were conducted at an ambient temperature of $\sim 25^\circ\text{C}$.

These tests were performed in an R&D setting so that apart from easily accessible measurements, such as SV currents and ambient temperature. The currents are easily extracted from the power electronics circuit. Yet, the data set also contains less convenient measurement data that would not be available during normal operation, such as air supply temperature and pressure, valve surface temperatures and air flows at the outlet ports and ventholes (blow-off holes). The ground truth of the lifetime is measured in number of on/off cycles.

C. Solenoid Valve Degradation & Detection

There is a clear physical difference between a new valve, that is before operation, and a used valve at the end of its useful lifetime. The used valve shows signs of wear at its plunger and deformation of its valve seal, as can be seen on Fig. 3a and Fig. 3b respectively. Although, clearly, SV degradation is physically detectable we aim to develop a method that is practical for SVs during normal deployment. Therefore, our prognostics model will, in the end, solely rely on the current measurements that were mentioned above. Any prognostics

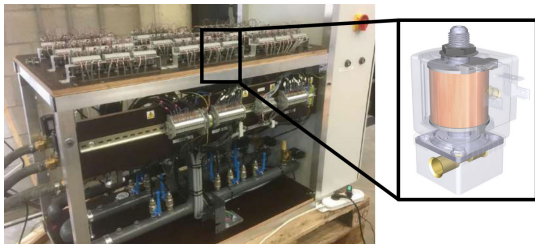


Fig. 2. Test set-up for the accelerated lifetime tests of 48 solenoid valves. The figure on the right depicts the CAD model of a SV.

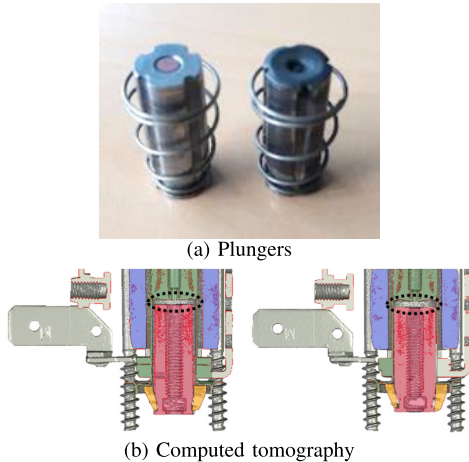


Fig. 3. New (left) versus used (right) solenoid valve. The used valve shows signs of wear at its plunger and valve seals.

can be realised in a noninvasive manner and without the need for inconvenient sensors. Temperature and flow measurement are required to determine the exact moment of EOL of the SVs in our R&D data set.

Three stages in the lifetime of a SV can be identified. The SVs are given a label, $l \in \{\text{healthy, degrading, faulty}\}$, by an expert based on all measurements mentioned above. There is no one clear criterion to determine these labels. Fig. 4 shows the measured signal of the outlet flow and current signal during the switch-on and the outlet flow during the switch-off of an on/off cycle for each of these three stages. Luckily, pronounced differences emerge from the visual comparison of these signals from the three different lifetime stages.

The ‘outlet flow - ON’ plot shows the outlet flow of the fluidum through the valve starting at the beginning of the open state, i.e. from the moment the plunger is being opened. The plunger is opened by applying the AC current signal as depicted by the ‘current - ON’ plot. The ‘outlet flow - OFF’ plot shows the outlet flow from the moment the plunger is being closed. The spring loaded plunger is closed from the moment the AC current signal is no longer applied and typically occurs faster than the opening.

In the healthy operating condition, there is no leakage when the valve is closed and throughflow steadily starts rising from the moment the valve is opened. When the valve is closed, the throughflow steadily decreases back to zero.⁴ As mentioned

⁴The outlet flow is measured using a hot wire anemometer principle. This type of sensor is prone to laggy measurements resulting from the thermal stabilization process of the wire, explaining why the outlet flow does not immediately rises to nominal throughflow when opening the valve and does not drop back to zero after closing the valve.

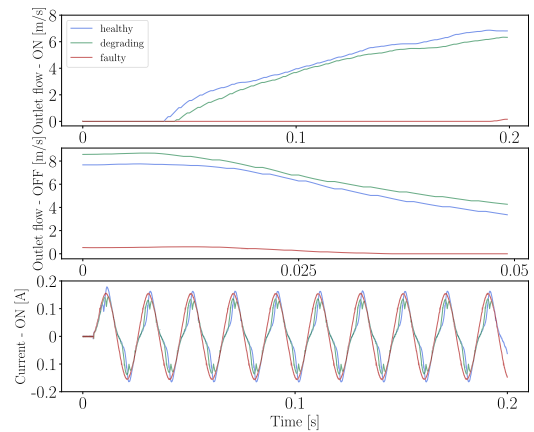


Fig. 4. Evolution of a SV outlet flow during both on- and off-phase and the current signal during on-phase for three cycles over the course of a lifetime. The first cycle (blue) is during healthy operation, the second cycle (green) is during the degrading phase of the solenoid and the third cycle (red) is at the end of useful life of the SV.

before, faulty behaviour can express itself in various forms (failing, other blockages, plunger blocked in end position ...). The valve in its faulty operating condition used for Fig. 4 for example shows no sign of leaking air when it is closed, but on the contrary completely blocks throughflow of air when it should be opened. The opposite behaviour or an intermediate form would be classified as equally faulty behavior, demonstrating the non-trivial character of the problem.

A solid opening of a healthy SV corresponds to a single hit of the plunger, which translates to a single *notch* in the current measurement. A degrading plunger has a more bumpy movement of the plunger during opening, resulting a multitude of these *notches* during opening. Finally, the plunger at its EOL is stuck in this case and therefore displays none of these *notches*. Apart from that the sinusoidal-like waveform of the current signals clearly varies over the three lifetime stages.

Although described measurement sets are highly indicative for the SV degradation and general SOH, these are usually not available in an industrial setting, necessitating the development or identification of alternative features.

D. Domain Knowledge Based Features

These time-series data carry information that can be more conveniently summarized in a number of features. From the discussion above, two crucial features related to the flow measurements can be determined

- the measured flow when the valve is still closed just before opening, hence the first point in the ‘outlet flow - ON’ measurements q_{leak} ,
- the measured flow when the valve is still open just before closing, hence the first value in the ‘outlet flow - OFF’ measurements q_{out} .

The first one indicates whether the supposed closed valve is leaking fluidum, the second one indicates whether the supposed open valve is hindered in supplying fluidum.

From this, three current features were extracted

- The position of the (first) characteristic *notch* in the current signal (corresponding to the first hit),

- The impedance spectrum magnitude at the fundamental AC-source frequency $I_{m,\text{fund}}$,
- The inductance spectrum magnitude at the fundamental AC-source frequency $I_{d,\text{fund}}$.

Finally, The valve surface temperature can be reduced to the average value over one on/off cycle, as it merely varies over this short timespan. A strong increase of the valve surface temperature over a longer time period also indicates deterioration.

IV. RESULTS

A. Mass-Spring-Damper System

In order to verify the proposed methodology in sec. II, we treat a mechanical example first. The system, as schematically depicted in Fig. 5, consist of two connected masses of which the first mass is also connected to a fixed reference. Both connections are established by a parallel spring and damper. The corresponding parameter values used in the simulation are listed in Table II. Its state is characterized by the positions x_1 and x_2 of both masses and their respective velocities v_1 and v_2 . The system is described by a linear time-invariant model which exactly fits the aforementioned HGMM framework and can thus be used to validate our method. The mode and model parameters used for simulation is detailed in Appendix A.

A HGMM was learned from $K = 20$ experiments assuming full yet noisy state observations and according the approach detailed in sec. II-E. To validate the virtual sensing framework we then aim to estimate the position and velocity of both masses whilst only having access to measurements for the first mass. More formally, $y = \{x_1, v_1\}$, $z = \{x_2, v_2\}$.

First, we assess the performance of the model identification procedure. For this simple mechanical example, we know that the state that completely describes the systems' dynamics is 4 dimensional. However, for a practical use case, the dimensionality of the latent state is often unknown. Therefore, it is required to do a search over a range of latent state dimensions, as depicted in Fig. 6. Given that the 4 observed variables are independent, it is senseless to assess lower dimensions of the latent state, as these could never yield all the information contained in the complete observation (recall the discussion in sec. II-E2). We consider a range between 4 and 15. Clearly, a latent state dimension of 4 is most consistent and on average obtains the best maximization of the objective function (13). Second, we validate the filtering framework tailored to virtual sensing. Fig. 7 depicts filtered and predicted measurement sequences for both masses. Clearly, the state measurement of the second mass are successfully reconstructed from the state measurements of the first mass.

B. Solenoid Valves

Here we will deploy the virtual sensing strategy on the industrial problem of RUL and SOH prediction for SVs.

It is recommended to comment on the following issue before we address the estimation problem itself. It is anticipated that the supporting linear theory will produce artefacts that defy the definition of the RUL and SOH (recall sec. II-B). Specifically, the estimated values, and in particular the associated confidence interval, may adopt values in the set $(-\infty, 0) \cup (1, +\infty)$.

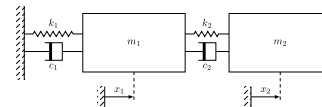


Fig. 5. Schematic representation of the mass-spring-damper system.

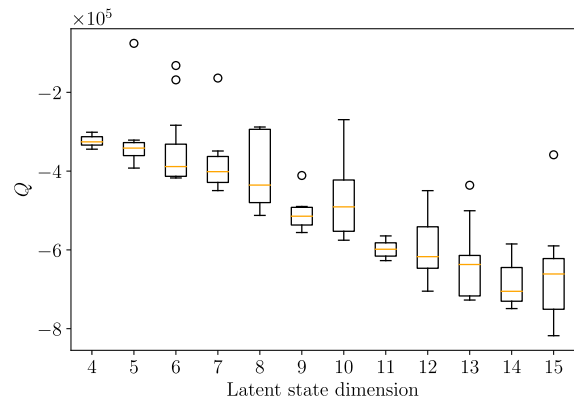


Fig. 6. Comparison of the maximum ELBO value, $Q(\theta^*, \theta^*)$, for different latent state dimensions for the MSD system.

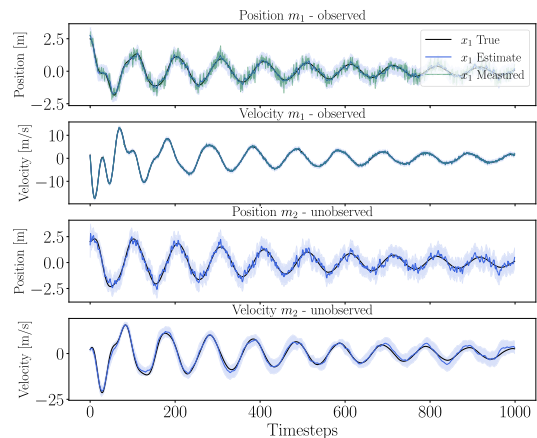


Fig. 7. Comparison of the estimated trajectory to the true underlying trajectory for the mass-spring-damper system based only on the measurement of the position and velocity of first mass.

This is an insurmountable byproduct of the linear theory used to embed the degradation or failure dynamics. Resolving this issue by intrinsically adapting the emission distribution would destroy the linear context. The only workable resolution would be to post-process the output distribution by means of e.g. clipping or truncating the distribution. However then this would be an aesthetic intervention rather than an intrinsic property of the approach. So rather we accept these non-physical artefacts as a compromise to the computational efficiency.

With this final issue clarified, we can present our results. Recall that we aim to obtain a noninvasive RUL/SOH prognostics tool for condition monitoring of these valves. As discussed, therefore we only want to rely on current measurements as inputs for the operational model. However, for the system identification, multiple other features from the R&D data can be used. More formally, $y = \{I_{d,\text{fund}}, I_{m,\text{fund}}, \text{notch}\}$ and z can be any combination of the following features

RUL Remaining Useful Lifetime
SOH State of Health

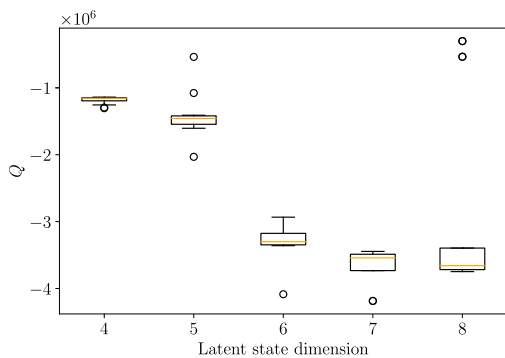


Fig. 8. Comparison of the RUL estimate, $\log p(\text{RUL}|y)$, for different latent state dimensions on the SV data set for the case where $y = \{I_{d,\text{fund}}, I_{m,\text{fund}}, \text{notch}\}$ and $z = \{\text{RUL}\}$.

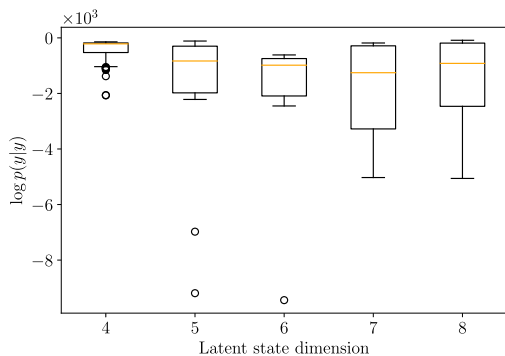


Fig. 9. Comparison of the loglikelihood, $\log p(Y = y|y)$, for different latent state dimensions on the SV data set for the case where $y = \{I_{d,\text{fund}}, I_{m,\text{fund}}, \text{notch}\}$ and $z = \{\text{RUL}\}$.

- q_{out} outlet flow during on-phase
- q_{leak} leakage flow during off-phase
- t_h time of first hit during valve opening
- T_{surf} plunger surface temperature

All results discussed below are obtained with 5-fold cross validation [15].

1) *Latent State Dimensionality Study:* As for the spring-mass-damper system, in order to guarantee that we will eventually obtain the best possible model feasible within our proposed framework, we first have to perform a search for the optimal latent state dimensionality. Fig. 8 shows the complete-data loglikelihood for different latent state dimensionalities in case $z = \{\text{RUL}\}$. All dimensionalities were assessed 50 times. Again, we see that increasing the latent state dimension above the number of observations during training, 4 in this case, results in decreased performance on average. Although some outliers seem to be able to exploit the increased model complexity. However, if we consider Fig. 9 and Fig. 10, which respectively show the loglikelihood of observing $Y = y$ and RUL during operation, given only y , the model only seems to be able to properly fulfill its function as a filter for the measured observations and fails at estimating the RUL. The ability to increase the model complexity could have its benefits in case the presented set of features are a lower dimensional embedding of a higher dimensional physical manifold. The fact that the RUL/SOH features lack physical interpretability could explain why this opportunity is not exploited.

2) *Feature Analysis:* A key idea of our framework is that we can include information during the system identification procedure that is not observable during operation. This has the

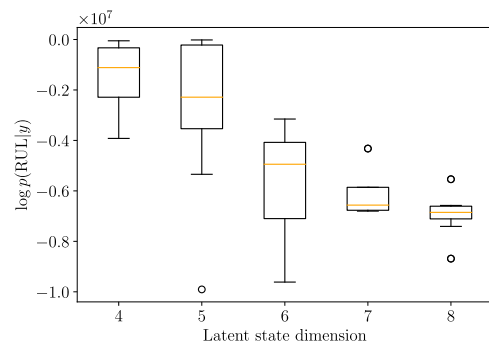


Fig. 10. Comparison of the RUL estimate, $\log p(\text{RUL}|y)$, for different latent state dimensions on the SV data set for the case where $y = \{I_{d,\text{fund}}, I_{m,\text{fund}}, \text{notch}\}$ and $z = \{\text{RUL}\}$.

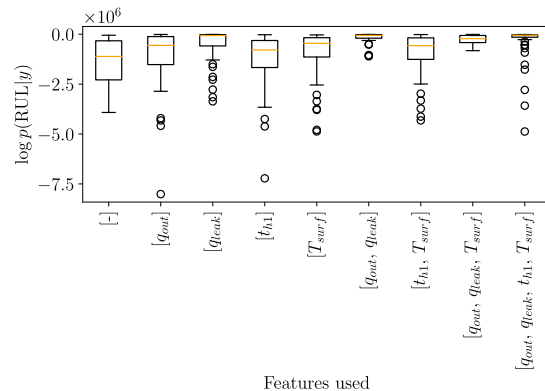


Fig. 11. Comparison of RUL models for including additional features in z apart from RUL and SOH. The optimal latent state dimension was determined using the procedure described in sec. IV-B1.

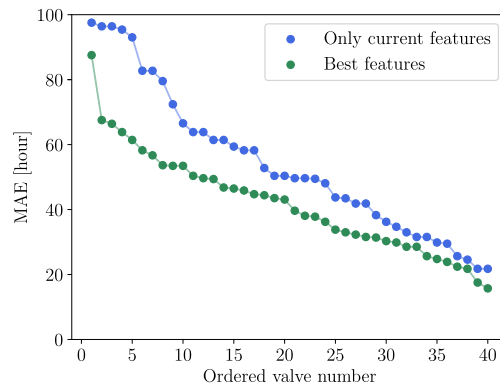


Fig. 12. Comparison of MAE for RUL predictions averaged over the entire lifetime per solenoid valve. The optimal latent state dimension was determined using the procedure described in sec. IV-B1.

advantage that a more accurate dynamics model can be learned and deployed with fewer inputs during the filtering phase. Fig. 11 shows the comparison of models deployed on test SVs, that were trained with different features combinations. The base case is always $y = \{I_{d,\text{fund}}, I_{m,\text{fund}}, \text{notch}\}$, z contains at least RUL/SOH. For each of these combinations of model inputs we first performed a latent state dimensionality study as discussed in the previous section. However, a general tendency could be observed to choose the dimension equal to the input dimensionality. Each combination was assessed 50 times.

Clearly, including either one of the other four available features from the R&D dataset has an impact on the model performance. The optimal features combination turns out to be $\{q_{\text{out}}, q_{\text{leak}}, T_{\text{surf}}\}$. Fig. 12 shows that including additional

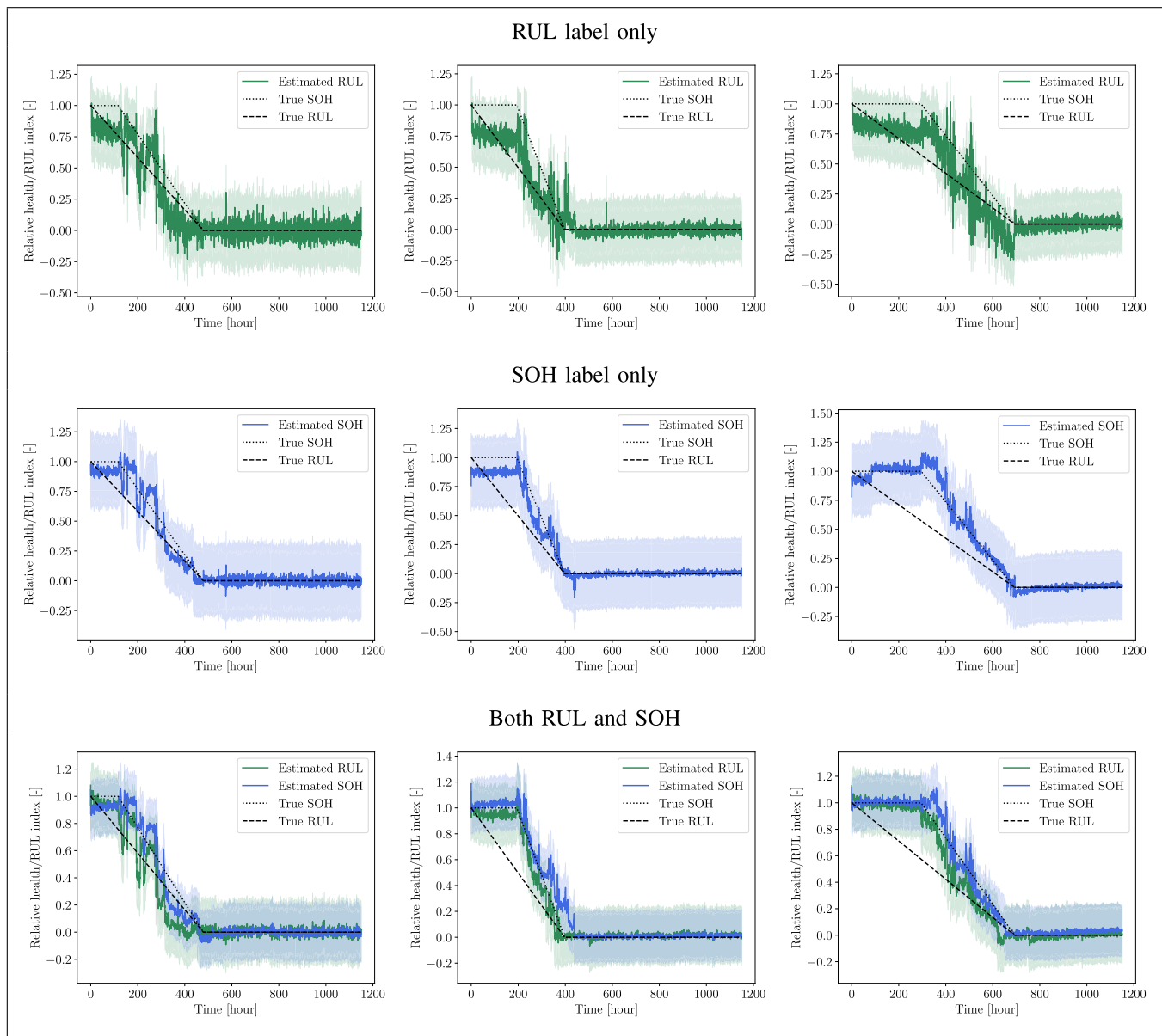


Fig. 13. Comparison of SOH and/or RUL predictions for three different SVs. The upper figures show the RUL predictions of a model trained without observability of the SOH labels, the middle figures show SOH predictions in the opposite case and the lower figures show both RUL and SOH predictions in case both labels are available during the model learning phase.

features during model learning outperforms the base case. We hypothesize that a superior underlying dynamic model can be identified, including the temporal dependencies of the various features included, whilst, on the other hand, the 3 current features remain sufficient to estimate the underlying state and thus the quantities of interest such as the RUL and SOH.

3) *RUL Prediction*: Some examples of both RUL and SOH predictions and the corresponding the 95% confidence intervals are shown in Fig. 13, for three different SVs. These predictions are obtained with the best possible model, resulting from the analysis in the previous section. The RUL predictions are relative values, with the ground truth scaled between 0 and 1. On top of the figure, RUL predictions, given by a model that was trained without the SOH labels, are shown. In the

middle, SOH predictions, given by a model trained with SOH and without the RUL labels, are shown and the bottom row shows the results for a model where both labels were available during the model identification phase.

Even in case the model has never seen SOH labels, the behavior of the RUL estimate typically consists out of 3 phases: a rather constant RUL estimate around the value of 1 at the beginning of the valves' lifetime, followed by a phase of accelerated deterioration and finally a correct estimation of no remaining useful lifetime when the SV has failed. Naturally, the RUL predictions clearly tends towards the shape of the SOH characteristic. The behavior of these RUL predictions could be summarised as an underestimation of the SOH characteristic up to the point of EOL, forced by the learned monotonically decreasing RUL character. This learned

TABLE I
MAE AVERAGED PER VALVE, AT 70% AND AT 90% DEGRADATION

Method	average MAE	MAE at 70%	MAE at 90%
Similarity model, current features	93.3 ± 56.2	63.4 ± 49.4	43.2 ± 35.5
FCNN, raw current signals	69.5 ± 44.5	53.8 ± 39.5	-
BCNN, raw current signals	63.2 ± 34.8	39.0 ± 28.8	-
HGMM, current features	55.3 ± 28.5	42.8 ± 27.2	19.8 ± 12.1
HGMM, optimal feature set	43.6 ± 25.4	34.6 ± 15.4	16.4 ± 8.9

monotonically decreasing character also result in predictions below 0 for the last solenoid, to only correct to zero once the model is confident that the system has failed.

For the second case, the model is clearly much more capable of correctly identifying the SOH dynamics, conforming our hypothesis. Although the variance on the prediction is not significantly smaller, the model seems much more robust as the prediction is less noisy compared to the previous case.

Finally, in case both labels are available during training, the RUL and SOH estimates are clearly highly correlated. Again, the RUL predictions tend to better approximate the SOH than the actual RUL characteristic. However, towards the end of the useful lifetime, the RUL predictions are accurate and more stable than in the first case.

One could argue that our model fails to capture the behavior of the RUL characteristic. However, as stated before, these empirical RUL signals are artificial constructs and are supported by little physical insights other than the observation that the system finally broke down. Reconstructing the corresponding RUL value as a signal that decays linearly with time is an arbitrary choice and any other decaying carrier signal would have been just as valid. Despite the forced bias of having to put forth a monotonically decreasing estimate of the RUL over the lifetime of a SV, the model persists in predicting a relatively constant value of the RUL, and thus deviating from the *true* RUL characteristic, in the early stage of the lifetime until it observes actual signs of deterioration. It can thus be stated that the model successfully identifies the true underlying degradation dynamics, to a certain extent, even given a poor indication of those dynamics.

Table I summarizes and compares our methodology to the results of closest related *state-of-the-art* methods, in terms of the mean absolute error (MAE) between the predicted RUL and the true RUL of the solenoid valves until failure. We compare the mean and standard deviation of the average MAE, the MAE at 70% and at 90% degradation per solenoid valve. This degradation corresponds to the respective 70% and 90% of the useful lifetime of the solenoid valve, which has been retrospectively determined for each solenoid valve. The *raw current signals* refers to the high frequency current measurements before extraction of any features as depicted in Fig. 4, *current features* refers to the extracted current features based on expert knowledge as described in III-C, the

optimal feature set refers to the optimal features combination is described in IV-B2. The implementation of the similarity model is based on [9] and [47]. The results from the frequentist CNN (FCNN) and Bayesian CNN (BCNN) are by [12] Mazaev et al. It is important to mention that the FCNN and BCNN automatically extract features from the raw current measurement, while the similarity model and the HGMM model use preprocessed current features. Furthermore, it is important to mention that additional improved results are obtained by [12] as they include two physical forces as input to the model, identified by Tod et al. [45] by leveraging a detailed physical model of the solenoid valves.

We can observe that our HGMM approach in combination with the current features outperforms the similarity model and performs in terms of predictive capabilities at a similar level as the end-to-end CNN models that deduce RUL estimates directly from the raw current signals. The possibility of including additional features during the identification of the degradation model clearly enables the HGMM to outperform all other models during deployment.

V. CONCLUSION

In this article we proposed a data-driven framework for probabilistic condition monitoring. Our aim is to model the correlations between two measurement sequences based on time series data from dedicated run-to-failure experiments, that are obtained in a R&D setting. This means the system is fully instrumented and measurements are logged that are impractical or even unachievable in a practical and/or industrial setting. Based on this model it is then possible to estimate the unfeasible measurements from the limited set of available measurements, a concept known as virtual sensing. This allows for real-time estimation of the unfeasible measurements in an industrial setting. Specifically, our approach identifies a Hidden Gauss-Markov Model from the complete measurement set and uses Bayes filtering techniques to reconstruct the unfeasible measurements in practice.

We showed that, although the RUL/SOH signals are not objectively physical quantities, our framework is capable of estimating these artificial signals with high precision. By default a measure of uncertainty is obtained as an inherited property of the Hidden Markov Model theory. The uncertainty on the RUL/SOH measurements can be interpreted as epistemic uncertainty on the prediction, which represents, according to the Bayesian interpretation, our belief about the RUL/SOH.

Two practical tendencies were observed. First it is useful to include additional features during training other than the available measurements and the features that we want to predict. This is unsurprising given that the more information that we grant to the training process about the underlying dynamics, the better the final identified model. Though it was also observed that certain features destabilize the training procedure, as measured in terms of prediction accuracy, possibly on account of the assumed linearity of the model. Second, we observed that for most but a few feature combinations, the optimal latent state dimension equals the number of features. This implies that the optimal underlying dynamics should

TABLE II
THE SPRING-MASS-DAMPER'S PARAMETERS

Parameter	Value	Unit
m_1	0.01	[kg]
m_2	0.01	[kg]
k_1	1	[N/m]
k_2	1	[N/m]
c_1	0.01	[Ns/m]
c_2	0.01	[Ns/m]
σ_w	0.002	[-]
σ_v	0.002	[-]

be as complex as the measurements and adding additional complexity does not result in improved predictive capabilities.

Future work may include extension to forecasting rather than filtering.

APPENDIX A MASS-SPRING-DAMPER MODEL

The equations describing the mass-spring-damper system dynamics, with the variables defined as in Fig. 5, are given by the following continuous-time state space model

$$\begin{aligned}\dot{x}(t) &= Ax(t) + w(t) \\ y(t) &= Cx(t) + v(t)\end{aligned}\quad (17)$$

where

$$\begin{aligned}A &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\frac{k_1+k_2}{m_1} & \frac{k_1}{m_1} & -\frac{c_1+c_2}{m_1} & \frac{c_1}{m_1} \\ \frac{k_2}{m_2} & -\frac{k_2}{m_2} & \frac{c_2}{m_2} & -\frac{c_2}{m_2} \end{bmatrix} \\ C &= I_4 \\ w(t) &\sim \mathcal{N}(0, \sigma_w I_4) \\ v(t) &\sim \mathcal{N}(0, \sigma_v I_4)\end{aligned}\quad (18)$$

To be able to calculate the filtering and smoothing densities, this model is discretized as described in [46]. The simulations with the discretized linear state-space model were performed with a sample time $T_s = 0.01s$.

APPENDIX B EXPECTATION-MAXIMIZATION

Consider a parametric probabilistic model existing of the two variable sets, Z , and, X . The model is described by the local probabilities, $p(X|\theta)$, and, $p(Z|X, \theta)$. Then the marginal distribution, $p(Z|\theta)$, can be decomposed as

$$p(Z|\theta) = \int p(Z|X, \theta)p(X|\theta)dX \quad (19)$$

Ad hoc, we can express the logarithm of $p(Z|\theta)$ by introducing an auxiliary *inference* distribution, $q(X)$.

$$\begin{aligned}\log p(Z|\theta) &= \int q(X) \log \frac{p(X, Z|\theta)}{q(X)} dX \\ &+ \int q(X) \log \frac{q(X)}{p(X|Z, \theta)} dX\end{aligned}\quad (20)$$

The second term equals the non-negative relative entropy between the inference distribution, $q(X)$, and, the a posteriori distribution, $p(X|Z, \theta)$. As a result $\log p(Z|\theta)$ is always

greater than the first term. This observation implies that if we were to maximize the first term with respect to θ rather than $\log p(Z|\theta)$, the associated parameter update is also an ascent direction for the true objective. Consequently, the first term, often referred to as the Evidence Lower Bound (ELBO), can be used as a surrogate objective, and, the inference distribution can be chosen so to minimize the approximation error, equivalently, to minimize the relative entropy. The relative entropy is zero only between equivalent distributions, so that

$$q^*(X) = p(X|Z, \theta) \quad (21)$$

Substituting the optimal inference distribution in the ELBO, omitting the denominator, yields the surrogate objective from sec. II-E. Note that the inference distribution does not depend on the parameter. Thus, the former parameter estimate is used to determine the inference distribution, the present parameter estimate follows from optimizing the ELBO; hence, suggesting an iterative procedure which is known as the EM algorithm.

It is well known that the EM algorithm converges to the solution of the reciprocal MLE problem. Once the parameter has reached its final value, therefore we have that

$$\begin{aligned}\log p(Z|\theta^*) &= \int p(X|Z, \theta^*) \log \frac{p(X, Z|\theta^*)}{p(X|Z, \theta^*)} dX \\ &= Q(\theta^*, \theta^*) - \int p(X|Z, \theta^*) \log p(X|Z, \theta^*) dX \\ &= Q(\theta^*, \theta^*) + \mathbb{H}[\log p(X|Z, \theta^*)]\end{aligned}\quad (22)$$

APPENDIX C PARAMETER UPDATES FOR HGMMs

From [16, Theorem 12.4] and [17], we adopt the following expression for $Q(\theta, \theta')$, which applies to time invariant Hidden Gauss-Markov Models

$$Q = Q_1 + Q_X + Q_{Y,Z} \quad (23)$$

where

$$\begin{aligned}Q_1 &= -\frac{K}{2} \log |2\pi \Sigma_1| \\ &- \frac{1}{2} \sum_{k=1}^K \text{tr} \left(\Sigma_1^{-1} \left(\bar{\Sigma}_1^k + (\mu_1 - \bar{\mu}_1^k)(\mu_1 - \bar{\mu}_1^k)^\top \right) \right) \\ Q_X &= -\frac{TK}{2} \log |2\pi Q| \\ &- \frac{1}{2} \text{tr} \left(Q^{-1} (AV_X^\top + V_X A^\top - AU_X A^\top - U_X) \right) \\ U_X &= \sum_{k=1}^K \sum_{t=1}^T \bar{\Sigma}_t^k + \bar{\mu}_t^k \bar{\mu}_t^{k,\top} \\ V_X &= \sum_{k=1}^K \sum_{t=2}^T \bar{\Sigma}_{t,t-1}^k + \bar{\mu}_t^k \bar{\mu}_{t-1}^{k,\top} \\ Q_Z &= -\frac{TK}{2} \log |2\pi R| \\ &- \frac{1}{2} \text{tr} \left(R^{-1} (CV_Z^\top + V_Z C^\top - CU_Z C^\top - U_Z) \right) \\ U_Z &= \sum_{k=1}^K \sum_{t=1}^T \{y, z\}_t^k \{y, z\}_t^{k,\top} \\ V_Z &= \sum_{k=1}^K \sum_{t=1}^T \{y, z\}_t^k \bar{\mu}_t^{k,\top}\end{aligned}\quad (24)$$

where $\bar{\mu}_t^k$ and $\bar{\Sigma}_t^k$ denote the smoothing distribution parameters. The smoothing distributions are defined as [16]

$$p(x_t | \{y, z\}_T^k) = \mathcal{N}(x_t; \bar{\mu}_t^k, \bar{\Sigma}_t^k) \quad (25)$$

The parameters are governed by efficient recursive procedures. We refer to [16] for details.

From the expressions above one easily verifies that the surrogate Q is maximized for the following parameter updates

$$\begin{aligned} \mu_1 &= \frac{1}{K} \sum_{k=1}^K \mu_1^k \\ \Sigma_1 &= \frac{1}{K} \sum_{k=1}^K \Sigma_1^k + (\mu_1 - \underline{\mu}_1^k)(\mu_1 - \underline{\mu}_1^k)^\top \\ A &= V_X U_X^{-1} \\ Q &= \frac{1}{K} (U_X - V_X U_X^{-1} V_X^\top) \\ C &= V_Z U_Z^{-1} \\ R &= \frac{1}{K} (U_Z - V_Z U_X^{-1} V_Z^\top) \end{aligned} \quad (26)$$

REFERENCES

- [1] J. Endrenyi et al., "The present status of maintenance strategies and the impact of maintenance on reliability," in *IEEE Trans. Power Syst.*, vol. 16, no. 4, pp. 64–638, Nov. 2001.
- [2] E. Ruijters and M. Stoelinga, "Fault tree analysis: A survey of the state-of-the-art in modeling, analysis and tools," *Comput. Sci. Rev.*, vols. 15–16, pp. 29–62, Feb. 2015.
- [3] N. Xiao, H.-Z. Huang, Y. Li, L. He, and T. Jin, "Multiple failure modes analysis and weighted risk priority number evaluation in FMEA," *Eng. Failure Anal.*, vol. 18, no. 4, pp. 1162–1170, Jun. 2011.
- [4] F. Mohrle, M. Zeller, K. Hofig, M. Rothfelder, and P. Liggesmeyer, "Automated compositional safety analysis using component fault trees," in *Proc. IEEE Int. Symp. Softw. Rel. Eng. Workshops (ISSREW)*, Nov. 2015, pp. 152–159.
- [5] M. Xu, J. M. David, and S. H. Kim, "The fourth industrial revolution: Opportunities and challenges," *Int. J. Financial Res.*, vol. 9, no. 2, pp. 90–95, 2018.
- [6] T. P. Carvalho, F. A. A. M. N. Soares, R. Vita, R. D. P. Francisco, J. P. Basto, and S. G. S. Alcalá, "A systematic literature review of machine learning methods applied to predictive maintenance," *Comput. Ind. Eng.*, vol. 137, Nov. 2019, Art. no. 106024.
- [7] J. Lee, F. Wu, W. Zhao, M. Ghaffari, L. Liao, and D. Siegel, "Prognostics and health management design for rotary machinery systems—Reviews, methodology and applications," *Mech. Syst. Signal Process.*, vol. 42, nos. 1–2, pp. 314–334, Jan. 2014.
- [8] Y. Lei, N. Li, L. Guo, N. Li, T. Yan, and J. Lin, "Machinery health prognostics: A systematic review from data acquisition to RUL prediction," *Mech. Syst. Signal Process.*, vol. 104, pp. 799–834, May 2018.
- [9] P. Wang and D. W. Coit, "Reliability prediction based on degradation modeling for systems with multiple degradation measures," in *Proc. Annu. Symp. Rel. Maintainability (RAMS)*, 2004, pp. 302–307.
- [10] A. K. S. Jardine, D. Lin, and D. Banjevic, "A review on machinery diagnostics and prognostics implementing condition-based maintenance," *Mech. Syst. Signal Process.*, vol. 20, no. 7, pp. 1483–1510, Oct. 2006.
- [11] G. Mazaev, A. P. Ompusunggu, G. Tod, G. Crevecoeur, and S. Van Hoecke, "Data-driven prognostics of alternating current solenoid valves," in *Proc. Prognostics Health Manage. Conf.*, May 2020, pp. 109–115.
- [12] G. Mazaev, G. Crevecoeur, and S. V. Hoecke, "Bayesian convolutional neural networks for remaining useful life prognostics of solenoid valves with uncertainty estimations," *IEEE Trans. Ind. Informat.*, vol. 17, no. 12, pp. 8418–8428, Dec. 2021.
- [13] F. Pernkopf, R. Peharz, and S. Tschitschek, "Introduction to probabilistic graphical models," in *Academic Press Library in Signal Processing*, vol. 1. Amsterdam, The Netherlands: Elsevier, 2014, pp. 989–1064.
- [14] C. M. Bishop and N. M. Nasrabadi, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006.
- [15] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*. Cambridge, MA, USA: MIT Press, 2012.
- [16] S. Särkkä, *Bayesian Filtering and Smoothing*. Cambridge, U.K.: Cambridge Univ. Press, 2013.
- [17] P. L. Ainsleigh, N. Kehtarnavaz, and R. L. Streit, "Hidden Gauss-Markov models for signal classification," *IEEE Trans. Signal Process.*, vol. 50, no. 6, pp. 1355–1367, Jun. 2002.
- [18] S. Chen et al., "An M-estimator for reduced-rank system identification," *Pattern Recognit. Lett.*, vol. 86, pp. 76–81, Jan. 2017.
- [19] E. Zio, "Reliability engineering: Old problems and new challenges," *Rel. Eng. Syst. Saf.*, vol. 94, no. 2, pp. 125–141, Feb. 2009.
- [20] D. A. Tobon-Mejia, K. Medjaher, N. Zerhouni, and G. Tripot, "A mixture of Gaussians hidden Markov model for failure diagnostic and prognostic," in *Proc. IEEE Int. Conf. Autom. Sci. Eng.*, Aug. 2010, pp. 338–343.
- [21] D. A. Tobon-Mejia, K. Medjaher, N. Zerhouni, and G. Tripot, "Hidden Markov models for failure diagnostic and prognostic," in *Proc. Prognostics Syst. Health Management Conf.*, May 2011, pp. 1–8.
- [22] A. Giantomassi, F. Ferracuti, A. Benini, G. Ippoliti, S. Longhi, and A. Petrucci, "Hidden Markov model for health estimation and prognosis of turbofan engines," in *Proc. ASME/IEEE Int. Conf. Mech. Embedded Syst. Appl., A B*, vol. 54808, Jun. 2012, pp. 681–689.
- [23] K. Medjaher, D. A. Tobon-Mejia, and N. Zerhouni, "Remaining useful life estimation of critical components with application to bearings," *IEEE Trans. Rel.*, vol. 61, no. 2, pp. 292–302, Jun. 2012.
- [24] C. Piao, Z. Li, S. Lu, Z. Jin, and C. Cho, "Analysis of real-time estimation method based on hidden Markov models for battery system states of health," *J. Power Electron.*, vol. 16, no. 1, pp. 217–226, Jan. 2016.
- [25] Z. Chen, Y. Li, T. Xia, and E. Pan, "Hidden Markov model with auto-correlated observations for remaining useful life prediction and optimal maintenance policy," *Rel. Eng. Syst. Saf.*, vol. 184, pp. 123–136, Apr. 2019.
- [26] W. Li and T. Liu, "Time varying and condition adaptive hidden Markov model for tool wear state estimation and remaining useful life prediction in micro-milling," *Mech. Syst. Signal Process.*, vol. 131, pp. 689–702, Sep. 2019.
- [27] G. Zhiyong, L. Jiwu, and W. Rongxi, "Prognostics uncertainty reduction by right-time prediction of remaining useful life based on hidden Markov model and proportional hazard model," *Eksploracjia i Niezawodność Maintenance Rel.*, vol. 23, no. 1, pp. 154–164, Mar. 2021.
- [28] T. Aggab, P. Vrignat, M. Avila, and F. Kratz, "Remaining useful life estimation based on the joint use of an observer and a hidden Markov model," *Proc. Inst. Mech. Eng., O, J. Risk Rel.*, vol. 236, no. 5, pp. 676–695, Oct. 2022.
- [29] O. Fink, Q. Wang, M. Svensén, P. Dersin, W.-J. Lee, and M. Ducoffe, "Potential, challenges and future directions for deep learning in prognostics and health management applications," *Eng. Appl. Artif. Intell.*, vol. 92, Jun. 2020, Art. no. 103678.
- [30] S. Zheng, K. Ristovski, A. Farahat, and C. Gupta, "Long short-term memory network for remaining useful life estimation," in *Proc. IEEE Int. Conf. Prognostics Health Manage. (ICPHM)*, Jun. 2017, pp. 88–95.
- [31] T. Jing, X. Tian, H. Hu, and L. Ma, "Deep learning-based cloud-edge collaboration framework for remaining useful life prediction of machinery," *IEEE Trans. Ind. Informat.*, vol. 18, no. 10, pp. 7208–7218, Dec. 2021.
- [32] G. Zhang, W. Liang, B. She, and F. Tian, "Rotating machinery remaining useful life prediction scheme using deep-learning-based health indicator and a new RVM," *Shock Vibrat.*, vol. 2021, pp. 1–14, May 2021.
- [33] Z. Tian, L. Wong, and N. Safaei, "A neural network approach for remaining useful life prediction utilizing both failure and suspension histories," *Mech. Syst. Signal Process.*, vol. 24, no. 5, pp. 1542–1555, Jul. 2010.
- [34] T. Jing, P. Zheng, L. Xia, and T. Liu, "Transformer-based hierarchical latent space VAE for interpretable remaining useful life prediction," *Adv. Eng. Informat.*, vol. 54, Oct. 2022, Art. no. 101781.
- [35] Y. Qin, N. Cai, C. Gao, Y. Zhang, Y. Cheng, and X. Chen, "Remaining useful life prediction using temporal deep degradation network for complex machinery with attention-based feature extraction," 2022, *arXiv:2202.10916*.
- [36] M. Kraus and S. Feuerriegel, "Forecasting remaining useful life: Interpretable deep learning approach via variational Bayesian inferences," *Decis. Support Syst.*, vol. 125, Oct. 2019, Art. no. 113100.
- [37] H. Li, W. Zhao, Y. Zhang, and E. Zio, "Remaining useful life prediction using multi-scale deep convolutional neural network," *Appl. Soft Comput.*, vol. 89, Apr. 2020, Art. no. 106113.

- [38] R. Liu, B. Yang, and A. G. Hauptmann, "Simultaneous bearing fault recognition and remaining useful life prediction using joint-loss convolutional neural network," *IEEE Trans. Ind. Informat.*, vol. 16, no. 1, pp. 87–96, Jan. 2020.
- [39] B. Wang, Y. Lei, T. Yan, N. Li, and L. Guo, "Recurrent convolutional neural network: A new framework for remaining useful life prediction of machinery," *Neurocomputing*, vol. 379, pp. 117–129, Feb. 2020.
- [40] T. S. Kim and S. Y. Sohn, "Multitask learning for health condition identification and remaining useful life prediction: Deep convolutional neural network approach," *J. Intell. Manuf.*, vol. 32, no. 8, pp. 2169–2179, Dec. 2021.
- [41] M. Ma and Z. Mao, "Deep-convolution-based LSTM network for remaining useful life prediction," *IEEE Trans. Ind. Informat.*, vol. 17, no. 3, pp. 1658–1667, Mar. 2021.
- [42] S. Jo, B. Seo, H. Oh, B. D. Youn, and D. Lee, "Model-based fault detection method for coil burnout in solenoid valves subjected to dynamic thermal loading," *IEEE Access*, vol. 8, pp. 70387–70400, 2020.
- [43] J. Liniger, S. Stubkier, M. Soltani, and H. C. Pedersen, "Early detection of coil failure in solenoid valves," *IEEE/ASME Trans. Mechatronics*, vol. 25, no. 2, pp. 683–693, Apr. 2020.
- [44] A. Sarwar and X. Lu, "Fault isolation and diagnosis of high pressure fuel pump solenoid valves using current feedback," in *Proc. IEEE Int. Conf. Prognostics Health Manage. (ICPHM)*, Jun. 2018, pp. 1–8.
- [45] G. Tod, G. Mazaev, K. Eryilmaz, A. P. Ompusunggu, E. Hostens, and S. V. Hoecke, "A convolutional neural network aided physical model improvement for AC solenoid valves diagnosis," in *Proc. Prognostics Syst. Health Manage. Conf. (PHM-Paris)*, May 2019, pp. 223–227.
- [46] J. Hartikainen, A. Solin, and S. Särkkä, "Optimal filtering with Kalman filters and smoothers—A manual for MATLAB toolbox EKF/UKF," Dept. Biomed. Eng. Comput. Sci., Helsinki Univ. Technol., Tech. Rep. FIN-02015 TKK, Espoo, Finland, 2011.
- [47] Q. Sun, P. Chen, D. Zhang, and F. Xi, "Pattern recognition for automatic machinery fault diagnosis," *J. Vibrat. Acoust.*, vol. 126, no. 2, pp. 307–316, May 2004.



Victor Vantilborgh received the M.Sc. degree in electromechanical and control engineering from Ghent University, Belgium, in 2021. He is currently pursuing the Ph.D. degree in electromechanical engineering with the Department of Electromechanical, Systems and Metal Engineering. In September 2021, he joined the Department of Electromechanical, Systems and Metal Engineering. His current research interests include modeling, control, and uncertainty quantification of dynamic systems by combining physical and data-driven approaches. He is also an affiliate member of Flanders Make.



Tom Lefebvre received the M.Sc. degree in control engineering and automation from Ghent University, Belgium, in 2015, and the Ph.D. degree in electromechanical engineering in 2019. Since 2019, he has been a post-doctoral research assistant. He worked on stochastic optimal control, gradient and stochastic trajectory optimization, uncertainty quantification, and black-box optimization. His current research interests include probabilistic modeling, identification, and control. He is currently an affiliate member of Flanders Make.



Kerem Eryilmaz received the B.Sc. degree in computer engineering from Bilkent University in 2008 and the M.Sc. degree in cognitive science from Middle East Technical University in 2011. He is currently pursuing the Ph.D. degree with the Artificial Intelligence (AI) Laboratory, Vrije Universiteit Brussel, Brussels, Belgium, on the Markovian modeling of combinatorial artificial languages. Since 2017, he has been with Flanders Make as a Research Engineer in the monitoring domain, specializing in predictive maintenance, vibration monitoring, and machine learning.



Guillaume Crevecoeur (Member, IEEE) received the M.Sc. and Ph.D. degrees in engineering physics from Ghent University in 2004 and 2009, respectively. In 2009, he became a Post-Doctoral Fellow with Research Foundation Flanders. In 2014, he was appointed as an Associate Professor with the Faculty of Engineering and Architecture, Ghent University. His research interests are the modeling, optimization, and control of mechatronic and industrial robotic applications. He is an active member of Flanders Make and the Strategic Research Center for the Manufacturing Industry.