A Composite Failure Precursor for Condition Monitoring and Remaining Useful Life Prediction of Discrete Power Devices

Zhao, Shuai; Chen, Shaowei; Yang, Fei; Ugur, Enes; Akin, Bilal; Wang, Huai

Published in:
IEEE Transactions on Industrial Informatics

DOI (link to publication from Publisher):
10.1109/TII.2020.2991454

Publication date:
2021

Document Version
Accepted author manuscript, peer reviewed version

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2020.2991454, IEEE Transactions on Industrial Informatics

A Composite Failure Precursor for Condition Monitoring and Remaining Useful Life Prediction of Discrete Power Devices

Shuai Zhao, *Member, IEEE*, Shaowei Chen, *Member, IEEE*, Fei Yang, *Student Member, IEEE*, Enes Ugur, *Student Member, IEEE*, Bilal Akin, *Senior Member, IEEE*, and Huai Wang, *Senior Member, IEEE*

***Abstract*—In order to prevent catastrophic failures in power electronic systems, multiple failure precursors have been identified to characterize the degradation of power devices. However, there are some practical challenges in determining the suitable failure precursor which supports the high-accuracy prediction of remaining useful life (RUL). This paper proposes a method to formulate a composite failure precursor (CFP) by taking full advantage of potential failure precursors, where CFP is directly optimized in terms of the degradation model to improve the prediction performance. The RUL estimations of the degradation model are explicitly derived to facilitate the precursor quality calculation. For CFP formulation, a genetic programming method is applied to integrate the potential failure precursors in a nonlinear way. As a result, a framework that can formulate a superior failure precursor for the given RUL prediction model is elaborated. The proposed method is validated with the power cycling testing results of SiC MOSFETs.**

***Index Terms*—Condition monitoring, power devices, genetic programming, composite failure precursor, SiC MOSFETs, remaining useful life.**

I. Introduction

POWER semiconductors have been extensively used in transportation, renewable systems, and industrial automation, where unexpected failures may lead to safety issues or operational losses. In order to minimize potential failures and avoid costly shutdowns, reliability requirements have been drastically increasing in power devices and systems [1]. As an enabling approach, prognostics and health management (PHM) has been proposed to guarantee the safety and reliability of the power electronics so as to mitigate potential risks and prevent unexpected failures.

To implement PHM, various condition monitoring (CM) techniques have been developed to track the varying characteristics of the power devices including power metal-oxide-semiconductor field-effect transistors (MOSFETs) [2], insulated-gate bipolar transistors (IGBTs) [3], [4], [5], [6], and silicon carbide (SiC) MOSFETs [7], [8]. Effective CM technique is essential to the identification of failure precursors, which are the basis for reliability analysis, lifetime estimation, and remaining useful life (RUL) prediction [9], [10], [11]. For power MOSFETs, with an accelerated aging test (AAT) performed on multiple power MOSFETs [11], it is observed that the drain-to-source on-state resistance $R_{\mathrm{DS(on)}}$ is one of the remarkable failure precursors. With this precursor, a model based on Kalman filter is implemented for RUL prediction. Considering the implementation challenges and cost associated with the $R_{\mathrm{DS(on)}}$ monitoring during power converter operation, a software frequency response analysis method is proposed where the variation of $R_{\mathrm{DS(on)}}$ is indirectly revealed by the plant model response at the double pole frequency [2]. Considering the difficulties in precise measurements, a list of viable aging precursors including $R_{\mathrm{DS(on)}}$, body diode voltage drop V_{SD}, parasitic capacitance (input capacitance C_{iss}, output capacitance C_{oss}, and reverse capacitance C_{rss}), and gate threshold voltage V_{th}, is reported in [12]. And the evolution of V_{th} is modeled to estimate the RUL of switches. In [13], the miller plateau voltage is identified as an aging precursor which can support the online CM for power MOSFETs subjected to gate oxide degradation. Using the failure precursor $R_{\mathrm{DS(on)}}$, a random sample consensus algorithm [14] is proposed to implement the RUL prediction of power MOSFETs. For IGBTs, the Mahalanobis distance calculated by the on-state collector-emitter voltage $V_{\mathrm{CE,on}}$ and the current $I_{\mathrm{CE,on}}$ is applied to the anomaly detection of IGBTs in [3], and $V_{\mathrm{CE,on}}$ is applied to the RUL prediction. The characteristic variations of discrete IGBTs due to package degradation triggered by thermal stress are comprehensively investigated in [15]. Specifically, it analyzes the potential failure precursors (PFPs) including saturation voltage, gate threshold voltage, transfer capacitance, and gate charge. Using the failure precursor $V_{\mathrm{CE,on}}$, a Gaussian process regression method combining a Bayesian inference method [4] is proposed for the RUL estimation. With the identified failure precursor $V_{\mathrm{CE,on}}$, a computationally efficient algorithm based on the Monte-Carlo simulation is developed in [9] to support the real-time prediction of the RUL for IGBTs subjected to wire bond failure. In addition, it also presents an RUL prediction approach based on the time delay neural network [10] to improve the prediction accuracy. For SiC MOSFETs, the variations of the SiC MOSFETs electrical parameters are investigated in [8] through the thermal cycling experiments to assess their correlation with the device degradation. In [16], a degradation model based on the non-homogeneous gamma process is proposed to characterize the evolution of threshold

Shuai Zhao and Huai Wang are with the Department of Energy Technology, Aalborg University, Aalborg 9220, Denmark. (Email: szh@et.aau.dk; hwa@et.aau.dk) (*Corresponding author: Shuai Zhao*)

Shaowei Chen is with the School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710072, China. (Email: cgong@nwpu.edu.cn)

Fei Yang, Enes Ugur, and Bilal Akin are with the Department of Electrical and Computer Science, University of Texas at Dallas, Richardson, TX 75080 USA (Email: fei.yang1@utdallas.edu; enesugur@gmail.com; bilal.akin@utdallas.edu).

voltage V_{th} of SiC MOSFETs and estimate its RUL. In addition to the conventional parametric measurements such as on-state resistance, threshold voltage, leakage currents, etc., a real-time CM technique is developed in [7] by investigating the switching characteristics of aged SiC devices.

It should be noted that mostly only a single failure precursor is applied to the RUL prediction of power devices (e.g., the $V_{\text{CE,on}}$ of IGBTs in [4], the V_{th} of SiC MOSFETs in [16], and the $R_{\text{DS(on)}}$ of power MOSFETs in [11]). On the one hand, the selection of the suitable failure precursor is mainly determined as the one which shows an obvious variation during the device aging, and the selection is challenging in the presence of a group of potential failure precursors. Although the empirically selected failure precursor can show a clear degradation trend during the aging, the information contribution of this single failure precursor will be quite limited for the RUL prediction. Moreover, although these potential failure precursors can indicate the degradation status of power MOSFETs and IGBTs, the change of the precursor value can also be due to other external factors including the environmental and operational conditions (e.g., temperature, voltage, current, frequency, etc). The precursor accuracy is therefore affected by other noise factors, such as measurement errors, circuit parasitic elements, etc., which will make the following RUL prediction less robust. Thus, it is desirable to investigate the degradation behavior by using the information of multiple potential failure precursors to provide a comprehensive and robust reliability evaluation. On the other hand, for practical applications, the RUL prediction accuracy is further determined by the mutual adaptability of the failure precursor and the selected RUL prediction model. In most studies, the empirically selected failure precursor is directly applied for a certain RUL prediction model while the adaptability between the selected failure precursor and the RUL prediction model has been seldom investigated. Thus, it is necessary to develop a framework to leverage multiple potential failure precursors for a composite failure precursor (CFP) in terms of the selected RUL prediction model to facilitate a highly accurate RUL estimation.

Motivated by the above discussions, this paper aims to synthesize a group of potential failure precursors of power devices to derive a composite failure precursor to improve the RUL prediction performance. The degradation behavior of power devices is driven by a stochastic degradation model with a Bayesian updating mechanism. For the stochastic degradation model, the reliability characteristics are explicitly derived to facilitate the evaluation of the RUL prediction performance. By using the information fusion technique, the genetic programming method is applied to formulate a composite failure precursor where all the useful precursor information is exploited to improve the prediction performance. The proposed method is validated by a case study on SiC MOSFETs. The contributions of this study are twofold: Firstly, the data fusion method is applied to the condition monitoring of power devices to improve the health evaluation performance. Secondly, in most of the existing frameworks of data fusion-based RUL prediction (e.g., [17], [18]), the data fusion is applied to optimize the features of failure precursors, such as monotonicity, trendability, etc. In these cases, the performance improvement of the final prediction is heuristic. This issue is firstly proposed in [19] and the direct optimization is suggested in terms of the final prediction performance by combining the fusion and the RUL prediction, and the respective advantages are illustrated. Thus, from the methodology perspective, with the explicitly derived RUL estimations, in this study the nonlinear data fusion method and the degradation model are closely interacted as an efficient prediction framework, where the genetic programming is directly performed in terms of the final prediction performance of the stochastic degradation model instead of maximizing the favorable features of precursors. As a result, the proposed method is more efficient.

The rest of this paper is organized as follows. Section II describes the accelerated aging test experiment of SiC MOSFETs and a group of potential failure precursors. Section III presents an RUL prediction model with the reliability characteristic calculation. In Section IV, a data fusion method for the composite failure precursor formulation is presented. In Section V, the performance of the developed composite failure precursor is comprehensively illustrated and compared with the potential failure precursors. Finally, conclusion is given in Section VI.

II. Accelerated Aging Test for Power Devices

Due to the low leakage current superiority at the high temperature, SiC power MOSFETs are very promising for applications such as automotive powertrain, oil down-hole tractors, etc. To obtain the device aging characteristics in a reasonable period of time, power cycling test is a widely used method. In the power cycling test, the device under test (DUT) without the heat sink is actively heated up by injecting half of the rated current and cooled down by forced air-cooling using fans. Fig. 1 shows the customized setup for the accelerated aging test experiment. This experiment setup supports a total of seven devices subjected to accelerated aging simultaneously where each device can be independently controlled. Note that the 7th device is behind the DAQ system. Thermocouples are directly attached to the metal tab of the device to measure the device case temperature so as to monitor the variation of the junction temperature, which is applied for a closed-loop system for temperature regulation during the heating and cooling stages. More details of the accelerated aging test experiment can refer to [8].

The devices used in the test are generation-II SiC MOSFETs (1.2kV/10A). For each power cycling in the AAT, devices are injected with 4.5A constant current until its junction temperature goes up to its upper limit 200°C and then the fan is triggered to cool down the device until the device temperature falls to 30°C. Note that the solution of obtaining the junction temperature from the case temperature measured by the thermocouples can be found in Section II in [8], where the drain current during the junction temperature measurement period keeps small and IR camera is periodically applied to calibrate the temperature measurement accuracy. In the experiment, each power cycling takes around 3 minutes, where the heating lasts for 1 minute and the cooling lasts for 2 minutes. After every 250 cycles, the device is removed and

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2020.2991454, IEEE Transactions on Industrial Informatics

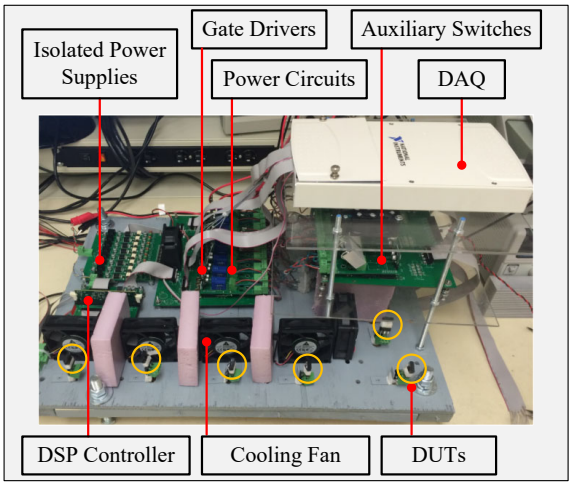

Figure 1. Setup of the accelerating aging testing for multiple SiC MOSFETs.

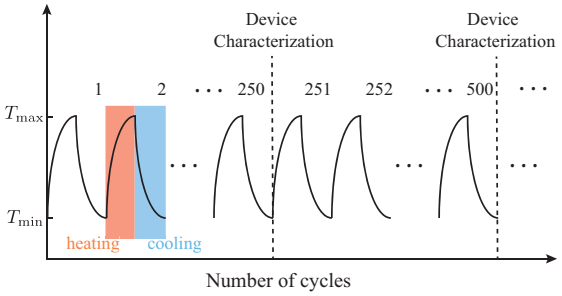

Figure 2. Procedures of power cycling and device characterization in AAT.

characterized by Keysight B1506A curve tracer at the ambient temperature. The procedure details of the power cycling and the device characterization are shown in Fig. 2. The CM data of a total of seven devices are recorded, which last for 8 weeks. As a summary, a total of six electrical parameters are monitored during the AAT including the input capacitance C_{iss}, the output capacitance C_{oss}, the reverse transfer capacitance C_{rss}, the drain-source on-state resistance $R_{\text{DS(on)}}$, the gate threshold voltage $V_{\text{GS(th)}}$, and the diode forward voltage V_{SD}. The electrical parameters are measured after every 250 power testing cycles until the end of the experiment, i.e., 10,000 cycles. The details of the failure mechanism analysis, the original CM data, and the device decapsulation for the post-experiment analysis can be found in [8]. In order to eradicate the static deviations among the units and measurements caused by manufacturing imperfections, sensor and analog-to-digital converter tolerances, for each electrical parameter, the initial value is subtracted and then the data are standardized [17].

For the RUL prediction, the time-to-failure of each device is first determined. Generally, the device is considered as properly operating during the first several CM periods, and the performance degradation in such a period is almost negligible. The CM information of these periods can be determined as a health cluster. In this way, the cluster distance measure of Mahalanobis distance [3] between the health cluster and the following CM information will increase smoothly unless a significant performance fluctuation occurs, which can be considered as device failure. In this case, the device failure is

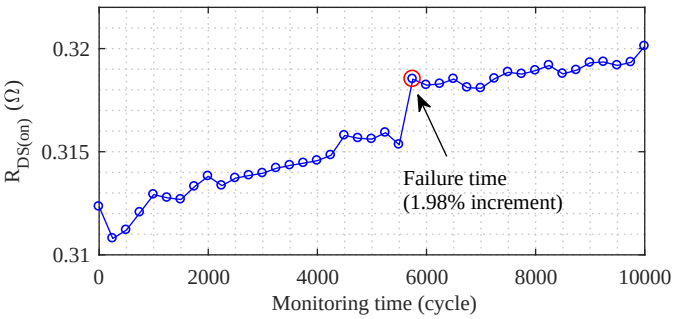

Figure 3. Evolution of $R_{\text{DS(on)}}$ and failure time determination for device #6.

determined when there is an abrupt change of the Mahalanobis distance between the current CM information and the healthy cluster, which is comprised of the CM information of the first five CM times of each device. As a result, the failure times of these seven devices are determined as 5250, 5000, 6250, 6500, 6750, 5750, 6250 cycles, respectively. As a result, the dataset details are summarized in Table I. It can be seen that the dataset consists of a total of 1044 data points.

Table I
DETAILS OF THE DATASET FROM THE POWER CYCLING EXPERIMENT OF SIC MOSFETS

Device number	1	2	3	4	5	6	7
Number of potential failure precursors	6	6	6	6	6	6	6
Number of observations	22	21	26	27	28	24	26
Number of data points of each device	132	126	156	162	168	144	156
Total number of data points of the dataset	1044						

Generally, the $R_{\text{DS(on)}}$ is an effective precursor indicating the device health status. When its degradation rate changes from linear to exponential, i.e., an abrupt evolution stage, the device performance will be quickly deteriorating and then will run to failure. Therefore, in this case, the evolution of $R_{\text{DS(on)}}$ of each device at the failure time is further checked to justify the method for time-to-failure determination. The percentage increases at the failure time are 2.19%, 1.45%, 1.48%, 1.82%, 2.31%, 1.98%, 2.09%, respectively. As an illustration, the evolution of $R_{\text{DS(on)}}$ and the failure time determination for device #6 is shown in Fig. 3. Such an agreement validates the effectiveness of the time-to-failure determination method.

Fig. 4 shows the variations of six potential failure precursors for a total of seven devices until the failure. It is worth mentioning that these PFPs are dimensionless and possess no physical unit due to the applied data preprocessing technique. As can be seen, all the six PFPs show variations indicating the device degradation with the power cycling progressing. For the reliability evaluation, however, it is challenging to determine the most suitable failure precursor from these PFPs and the respective degradation model which can better facilitate the RUL prediction. Moreover, the RUL prediction based on a single failure precursor will be less accurate and less robust to the external factors (e.g., the high-level heterogeneity of

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2020.2991454, IEEE Transactions on Industrial Informatics

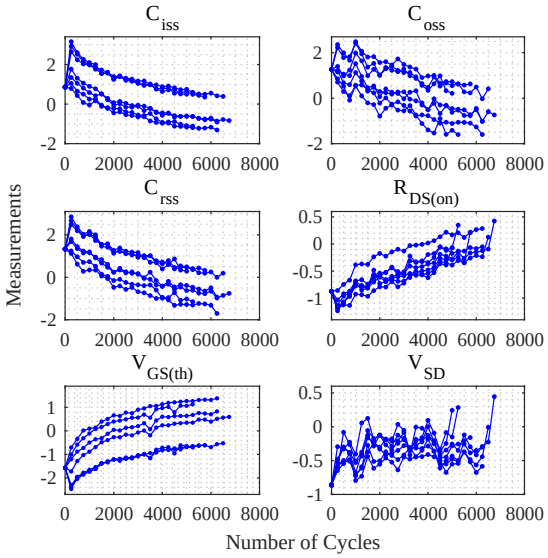

Figure 4. Multiple potential failure precursors throughout the accelerated aging test.

C_{oss}, the high-level noise of V_{SD}). It is therefore beneficial to propose a method which can not only mitigate the selection issue of the failure precursor but also provide a superior performance of RUL prediction by fully exploiting the multiple potential failure precursors.

III. Methodology

In this section, a method is proposed to fully exploit the condition monitoring information of power devices so as to improve the performance of RUL prediction and the method practicability. The framework of the proposed method is given in Fig. 5, which is comprised of two parts, i.e., the part of CFP development with the RUL prediction model, and the part of the RUL prediction for a new power device subjected to the condition monitoring. In the first part, a training dataset from a batch of power devices is used to develop the CFP by optimizing the RUL prediction error in terms of the RUL prediction model. In the second part, with a Bayesian update mechanism and the determined fusion formula of the CFP, the degradation model is updated to its posterior given the latest data of a new testing device. Using the posterior model, the reliability characteristics including the RUL and the 95% confidence interval (CI) can be obtained. Subsequently, a stochastic degradation model is developed to provide a unified basis for the calculation and the comparison of the RUL prediction.

A. Stochastic Degradation Model Formulation

In order to evaluate the prediction performance of PFPs and the CFP simultaneously, a stochastic degradation model facilitating a unified RUL prediction in terms of multiple failure precursors is applied here. It is considered as the RUL prediction model in the framework in Fig. 5. For the degradation model formulation, considering the CM data characteristics of SiC MOSFETs in Fig. 4 and the generality of the degradation model, we follow [17], [19] and the references (e.g., Gebraeel,

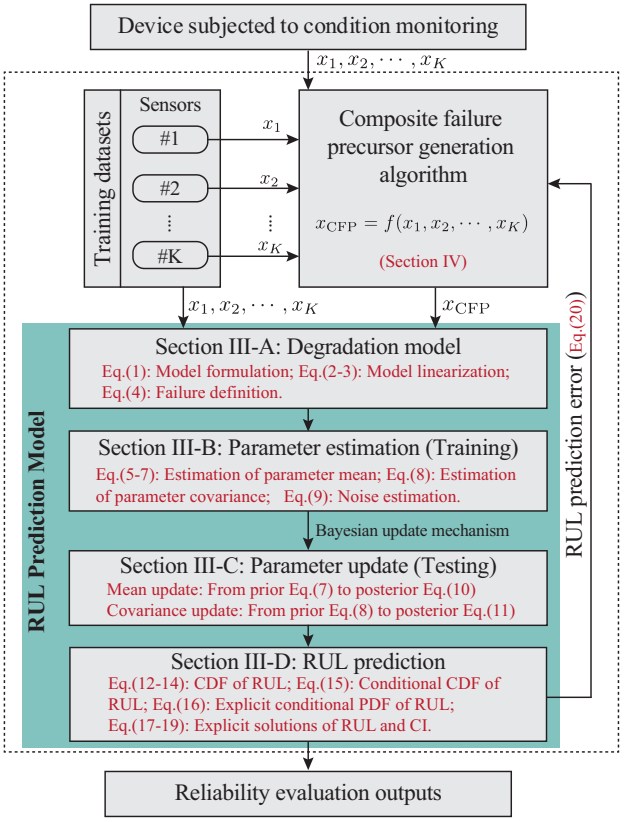

Figure 5. Framework of the proposed method.

2006; Liu et al., 2013) therein, and the degradation model of the CM data $x_{i,s,j}$ for the s^{th} precursor of device i at monitoring time t_j is formulated as

$$x_{i,s,j} = \kappa_s + \beta_{i,s} \exp(\alpha_{i,s} t_j + \epsilon_{i,s}(t_j) - \frac{\sigma_s^2}{2}), \qquad (1)$$

where $i = 1, \cdots, M$, $s = 1, \cdots K$, and $j = 1, \cdots, n_i$. M denotes the number of devices, K denotes the number of potential failure precursors, and n_i denotes the number of CM time points for device i. κ_s is the deterministic constant for s^{th} PFP for describing the sensor characteristic. $\beta_{i,s}$ and $\alpha_{i,s}$ are the random coefficients of the degradation model to cover the device heterogeneity. $\epsilon_{i,s}(t_j)$ is the measurement noise variable which is normally distributed with mean 0 and variance σ_s^2. For mathematical tractability and derivation convenience of the reliability characteristics, the *log* transformation is applied to make (1) linear and the transformed CM data $D_{i,s,j}$ is obtained as

$$D_{i,s,j} = \ln(x_{i,s,j} - \kappa_s) = \theta_{i,s} + \alpha_{i,s} t_j + \epsilon_{i,s}(t_j), \qquad (2)$$

where $\theta_{i,s}$ is defined as $\ln \beta_{i,s} - \frac{\sigma_s^2}{2}$. Without loss of generality, the random parameters $\theta_{i,s}$ and $\alpha_{i,s}$ are assumed to be jointly distributed to cover the heterogeneity in the batch of devices and follow a multivariate normal distribution as

$$\mathbf{\Gamma}_{i,s} = \begin{bmatrix} \theta_{i,s} & \alpha_{i,s} \end{bmatrix}^T \sim N_2\left(\boldsymbol{\mu}_0^s, \mathbf{\Sigma}_0^s\right), \qquad (3)$$

where $\boldsymbol{\mu}_0^s$ is the mean vector and $\mathbf{\Sigma}_0^s$ is the covariance matrix. Given the degradation model, the failure criterion is

determined subsequently. For the RUL prediction using s^{th} PFP, the remaining useful life l at the time t_j is defined as the first passage time when the degradation process $D_s(t)$ exceeds the failure threshold w_s, i.e.,

$$l = \inf\left\{l : D_s(t_j + l) \geq w_s \mid \mathbf{D}_{s,1:j}\right\}, \tag{4}$$

where the CM sequence is $\mathbf{D}_{1:j} = [D(t_1), \cdots, D(t_j)]^T$. w_s is the failure threshold for s^{th} precursor which is assumed to be normally distributed as $w_s \sim N(\eta_s, \vartheta_s^2)$, where η_s is the mean and ϑ_s^2 is the corresponding variance.

B. Parameter Estimation with Training Dataset

The training dataset consists of the CM data of a batch of power devices. With the CM data of the training dataset, the parameter estimation of the degradation model (2) is formulated subsequently. In this case, the parameters of $\mathbf{\Gamma}_{i,s}$ can be obtained as the prior to characterize the population features of the devices in the training dataset. According to regression theory, it can be estimated by fitting the degradation path of each device with (2) as

$$\tilde{\mathbf{\Gamma}}_{i,s} = \left(\mathbf{\Psi}_i^T \mathbf{\Psi}_i\right)^{-1} \mathbf{\Psi}_i^T \mathbf{D}_{i,s}, \tag{5}$$

where $\mathbf{\Psi}_i$ is defined as

$$\mathbf{\Psi}_i = \begin{bmatrix} 1 & \cdots & 1 & \cdots & 1 \\ t_1 & \cdots & t_j & \cdots & t_{n_i} \end{bmatrix}^T \in \mathbb{R}^{n_i \times 2}. \tag{6}$$

Considering the properties of the multivariate normal distribution, with algebraic manipulations, the statistics of $\mathbf{\Gamma}_{i,s}$ including the mean $\boldsymbol{\mu}_0^s$ and the covariance matrix $\mathbf{\Sigma}_0^s$ can be thus obtained [19] as

$$\boldsymbol{\mu}_0^s = \frac{1}{M} \sum_{i=1}^{M} \tilde{\mathbf{\Gamma}}_{i,s}, \tag{7}$$

and

$$\mathbf{\Sigma}_0^s = \frac{1}{M-1} \sum_{i=1}^{M} \left(\tilde{\mathbf{\Gamma}}_{i,s} - \boldsymbol{\mu}_0^s\right) \left(\tilde{\mathbf{\Gamma}}_{i,s} - \boldsymbol{\mu}_0^s\right)^T - \frac{1}{M} \sum_{i=1}^{M} \frac{\left(\mathbf{D}_{i,s} - \mathbf{\Psi}_i \tilde{\mathbf{\Gamma}}_{i,s}\right)^T \left(\mathbf{D}_{i,s} - \mathbf{\Psi}_i \tilde{\mathbf{\Gamma}}_{i,s}\right)}{n_i - 2} \left(\mathbf{\Psi}_i^T \mathbf{\Psi}_i\right)^{-1}. \tag{8}$$

Subsequently, the variance of the noise term $\epsilon_{i,s}$ can be estimated with (5) as

$$\sigma_s^2 = \frac{1}{M} \sum_{i=1}^{M} \frac{\left(\mathbf{D}_{i,s} - \mathbf{\Psi}_i \tilde{\mathbf{\Gamma}}_{i,s}\right)^T \left(\mathbf{D}_{i,s} - \mathbf{\Psi}_i \tilde{\mathbf{\Gamma}}_{i,s}\right)}{n_i - 2}. \tag{9}$$

As a result, the parameters of the degradation model can be explicitly estimated in the presence of the training dataset, and the shared degradation features of a batch of devices can be well characterized by the model with the prior parameters.

C. Parameter Update with Latest Data of Test Device

Once the prior parameters of $\mathbf{\Gamma}_{i,s}$ in the degradation model are estimated, they are updated to the posterior with the latest CM data obtained from the testing device. In this way, the degradation model can be elaborately tailored and individualized to the testing device subjected to the CM, which will significantly improve the RUL prediction accuracy. The parameter update is involved with Bayesian update mechanism and it is beneficial to online applications. For the test device i, denote $\mathbf{D}_{i,s,1:j}$ as its CM history for precursor s up to the CM time t_j. With the properties of the multivariate normal distribution, the posterior distribution of $\mathbf{\Gamma}_{i,s}$ given the information $\mathbf{D}_{i,s,1:j}$ is a multivariate normal distribution as $\mathbf{\Gamma}_{i,s} \mid \mathbf{D}_{i,s,1:j} \sim N_2\left(\boldsymbol{\mu}_p^{i,s}, \mathbf{\Sigma}_p^{i,s}\right)$ [17]. The mean and the covariance matrix can be derived [19] as

$$\boldsymbol{\mu}_p^{i,s} = \left(\frac{\mathbf{\Psi}_i^T \mathbf{\Psi}_i}{\sigma_k^2} + \left(\mathbf{\Sigma}_0^s\right)^{-1}\right)^{-1} \left(\frac{\mathbf{\Psi}_i^T \mathbf{D}_{i,s,1:j}}{\sigma_k^2} + \left(\mathbf{\Sigma}_0^s\right)^{-1} \boldsymbol{\mu}_0^s\right), \tag{10}$$

and

$$\mathbf{\Sigma}_p^{i,s} = \left(\frac{\mathbf{\Psi}_i^T \mathbf{\Psi}_i}{\sigma_k^2} + \left(\mathbf{\Sigma}_0^s\right)^{-1}\right)^{-1}. \tag{11}$$

In this way, the prior distribution with mean (7) and covariance (8) of $\mathbf{\Gamma}_{i,s}$ is updated to the posterior one with mean (10) and covariance (11). As a result, the degradation model learned from the training dataset will be more specific to the testing device.

D. Remaining Useful Life Prediction

With the degradation model (2), the formulations of the RUL and the confidence interval are presented. It can be seen that the CM data at the monitoring time t_{j+l} given the information $\mathbf{D}_{i,s,1:j}$ is normally distributed with mean $\mu_{i,s,j+l}$ and variance $\sigma_{i,s,j+l}^2$, which can be estimated as

$$\tilde{\mu}_{i,s,j+l} = \boldsymbol{\psi}(t_{j+l}) \boldsymbol{\mu}_p^{i,s}, \tag{12}$$

$$\tilde{\sigma}_{i,s,j+l}^2 = \boldsymbol{\psi}(t_{j+l}) \mathbf{\Sigma}_p^{i,s} \boldsymbol{\psi}(t_{j+l})^T + \sigma_s^2, \tag{13}$$

where $\boldsymbol{\psi}(t)$ is denoted as $\begin{bmatrix} 1 & t \end{bmatrix}$. Subsequently, with the normal properties of the CM data and the failure threshold, the cumulative distribution function (CDF) of RUL for the device i using s^{th} PFP given the CM data $\mathbf{D}_{i,s,1:j}$ can be developed [19] as

$$\begin{aligned} &F_{L_{i,s}|\mathbf{D}_{i,s,1:j}}(l) = P(L_{i,s} \leq l \mid \mathbf{D}_{i,s,1:j}) \\ =&P(D_{i,s,j+l} \geq w_s \mid \mathbf{D}_{i,s,1:j}) \\ =&1 - \Phi\left(\frac{\eta_s - \tilde{\mu}_{i,s,j+l}}{\sqrt{\tilde{\sigma}_{i,s,j+l}^2 + \vartheta_s^2}}\right) = 1 - \Phi\left(z(l)\right), \end{aligned} \tag{14}$$

where $\Phi(\cdot)$ is the CDF of the standard normal distribution, and $z(l)$ is defined as $z(l) = (\eta_s - \tilde{\mu}_{i,s,j+l}) / \sqrt{\tilde{\sigma}_{i,s,j+l}^2 + \vartheta_s^2}$. For online applications, the CM data of the testing device are

monitored continuously. Incorporating the latest information timely will increase the accuracy of RUL prediction. Thus, it is necessary to extend (14) to the conditional CDF of RUL given the latest CM data and the failure information. The CDF of the RUL for device i using s^{th} PFP conditioning on the fact that the device does not fail at the monitoring time t_j can be derived [17], [19] as

$$\begin{aligned} &P(L_{i,s} \le l \mid \mathbf{D}_{i,s,1:j}, L_{i,s} \ge 0) \\ =&\frac{P(L_{i,s} \le l \mid \mathbf{D}_{i,s,1:j}) - P(L_{i,s} < 0 \mid \mathbf{D}_{i,s,1:j})}{P(L_{i,s} \ge 0 \mid \mathbf{D}_{i,s,1:j})} \\ =&\frac{\Phi\left(z(0)\right) - \Phi\left(z(l)\right)}{\Phi\left(z(0)\right)}. \end{aligned} \tag{15}$$

More details of the degradation model and its update mechanism can be found in [19] and references (e.g., Gebraeel, 2006; Liu et al., 2013) therein. Another useful measure of RUL is its probability density function (PDF). Subsequently, taking the first derivative of (15), the PDF of the RUL can be explicitly derived as

$$\begin{aligned} &f_{L_{i,s}|\mathbf{D}_{i,s,1:j}}(l \mid \mathbf{D}_{i,s,1:j}) = -\frac{1}{\Phi\left(z(0)\right)}\frac{d(\Phi\left(z(l)\right))}{dl} \\ =&\frac{\varphi(z(l))}{\Phi\left(z(0)\right)\left(\tilde{\sigma}^2_{i,s,j+l} + \vartheta^2_s\right)}\Bigg(\boldsymbol{\psi}^{'}(t_j+l)\boldsymbol{\mu}_p^{i,s}\sqrt{\tilde{\sigma}^2_{i,s,j+l} + \vartheta^2_s} \\ &+ \frac{(\eta_s - \tilde{\mu}_{i,s,j+l})\begin{pmatrix} \boldsymbol{\psi}^{'}(t_j+l)\boldsymbol{\Sigma}_p^{i,s}\boldsymbol{\psi}^{T}(t_j+l) \\ +\boldsymbol{\psi}^{T}(t_j+l)\boldsymbol{\Sigma}_p^{i,s}\left(\boldsymbol{\psi}^{'}(t_j+l)\right)^{T} \end{pmatrix}}{2\sqrt{\tilde{\sigma}^2_{i,s,j+l} + \vartheta^2_s}}\Bigg), \end{aligned} \tag{16}$$

where $\varphi(\cdot)$ denotes the PDF of standard normal distribution. Due to the fact that the CDF of RUL is skewed, the estimate of the RUL $\tilde{l}_{i,s,j}$ can be reasonably determined by making the CDF of RUL (15) be equal to 0.5. Additionally, to evaluate the prediction uncertainty, the lower limit $\tilde{c}_l$ and upper limit $\tilde{c}_u$ of 95% confidence interval are considered and they can be obtained by making the CDF of RUL (15) be equal to 0.025 and 0.975, respectively.

Note that the estimates of the RUL and the confidence interval involve the inverse function of (15), which is computationally extensive and is challenging for the further CFP development. Thus, it is necessary to develop an explicit form of the RUL estimates so as to make the framework more efficient. Here, an approximation of the CDF of the standard normal distribution $\Phi\left(\cdot\right)$ is applied according to [20] as

$$\Phi\left(x\right) \approx \frac{1}{1+\exp(-1.702\cdot x)}, x \in \mathbb{R}. \tag{17}$$

Note that the absolute error of this approximation is less than 9.49×10^{-3}, which is sufficient and safe for our application. With the algebra manipulations, the l_p which can facilitate $P(L_{i,s} \le l_p \mid \mathbf{D}_{i,s,1:j}, L_{i,s} \ge 0) = p$ can be derived as

$$l_p \approx \frac{\left(\begin{array}{c} (2ef - 2\eta_s f - b - c) + \\ g \cdot \sqrt{\begin{array}{c}(2ef - 2\eta_s f - b - c)^2 - 4(f^2 - d)\cdot \\ (\eta_s^2 - 2\eta_s e + e^2 - A^2\sigma_s^2 - A^2\vartheta_s^2 - a)\end{array}}\end{array}\right)}{2(f^2 - d)} \tag{18}$$

where $A = \ln\left(\left(\exp(-1.702\cdot z(0)) + p\right)/(1-p)\right)/1.702$, $\boldsymbol{\mu}_p^{i,s} = [\ e \quad f\ ]^T$, g equals 1 if $A \ge 0$ otherwise equals -1, and a, b, c, d are variables such that

$$A^2\boldsymbol{\Sigma}_p^{i,s} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}. \tag{19}$$

In this way, the estimated RUL, the lower limit, and the upper limit of the 95% CI can be obtained efficiently by making p be equal to 0.5, 0.025, and 0.975, respectively.

In this section, the RUL prediction model consisting of the degradation model and the respective parameter estimation is presented. The model is able to update adaptively to a specific testing device by incorporating the latest CM data with the Bayesian update mechanism. Note that the reliability characteristics can be calculated with multiple failure precursors at the same time. As a result, the RUL estimates can be efficiently calculated to facilitate the following CFP development.

IV. Composite Failure Precursor Development

From the AAT in Section II, it can be seen that there are multiple failure precursors indicating the degradation behavior of SiC MOSFETs. Each failure precursor can only provide a partial detail about the device degradation. As one of the relevant applications of artificial intelligence in power electronics [21], data fusion method is an effective tool to improve the quality of condition monitoring of power devices by providing richer insights. By aggregating a group of potential failure precursors, a data fusion technique is applied in this section to develop a composite failure precursor. In this way, the developed CFP can be more suitable for the given RUL prediction model compared to any of the individual PFP.

A. Fusion Objective

To improve the RUL prediction performance, various optimization objectives have been proposed to improve the precursor features including monotonicity [17], trendability [18], prognosability [18], etc. Note that these optimization objectives are intuitively formulated and heuristic. Even though the prediction performance can be improved, the improvement of final prediction performance is unpredictable and the contribution of each desirable feature cannot be quantitatively determined. Thus, the optimization and the fusion are indirect and inefficient. This fact is first put forward in [19] and the indirect supervised learning is proposed for efficient optimization, i.e., optimizations in terms of the final RUL prediction performance, which makes the fusion more direct and predictable.

Considering the proposed explicit form of RUL prediction of the degradation model in (18), in this paper, the fusion

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2020.2991454, IEEE Transactions on Industrial Informatics

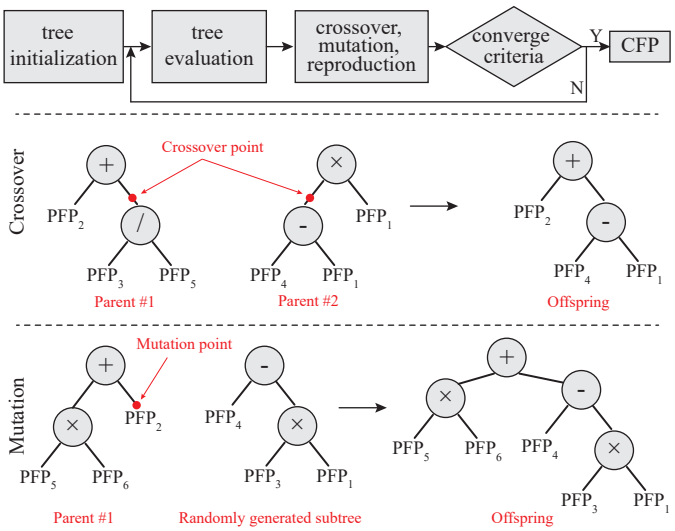

Figure 6. Flowchart of GP algorithm and evolution procedures of the tree structure.

objective is directly formulated in terms of the final prediction performance, i.e., the mean absolute error (MAE) of RUL. To assess the RUL prediction performance quantitatively, the evaluations at every ten percent of the whole life cycle are selected. Particularly, when using s^{th} failure precursor, the MAE of RUL is formulated with (18) as

$$MAE_s = \frac{1}{9M}\sum_{i=1}^{M}\sum_{\gamma=0.1}^{0.9}\left|\tilde{l}_{i,s,\lfloor\gamma\cdot n_i\rfloor} - l_{i,s,\lfloor\gamma\cdot n_i\rfloor}\right|, \qquad (20)$$

where $\lfloor x\rfloor$ is the floor function which outputs the greatest integer less than or equal to x, and $\gamma = 0.1, 0.2, \cdots, 0.9$. $l_{i,s,\lfloor\gamma\cdot n_i\rfloor}$ is the real RUL and $\tilde{l}_{i,s,\lfloor\gamma\cdot n_i\rfloor}$ is the predicted one for the device i. Note that (20) can be applied to the performance evaluation for both the developed CFP and the PFPs. In this way, the performance of the developed CFP and the PFPs can be quantitatively compared.

B. Genetic Programming for Data Fusion

By leveraging each useful PFP information, the data fusion is an effective technique for the development of CFP which can better characterize the degradation progression [18], [22]. Regarding the data fusion technique, we are interested in 1) determining the precursor candidates which should be incorporated into the fusion process, and 2) the nonlinear and flexible aggregation method between the selected precursor candidates. Genetic Programming (GP) can explore the implicit data structure by maximizing the predefined objective. It is one of the metaheuristic methods [23] in the applications of artificial intelligence. Compared to other fusion methods, GP is able to automatically select the relevant precursors from the group of potential PFPs and determine their combination form in a nonlinear way, with which the above two requirements can be well served. Therefore, in our case GP is applied to develop the CFP.

Generally, GP is a nonlinear optimization technique driven by the Darwinian evolution idea. In GP, the individuals are organized as syntax trees consisting of the allowed arithmetic operators, i.e., *addition*, *minus*, *times*, *divide*, *sine*, *cosine*, *logarithm*, *exponential*, *square root* in our case, and the terminal set, i.e., a group of PFPs. Each individual tree structure is composed of multiple branches and a root node that acts as the glue to connect the branches. The flowchart of GP algorithm is given in Fig. 6, and the algorithm is performed as follows. First, the initial population of trees is randomly generated. When it is executed against the terminals, i.e., the PFPs, it will produce numerical values, i.e., a possible CFP. An example of the individual tree is also given in Fig. 6, where the tree structure of the offspring for the mutation can be interpreted as $\text{PFP}_5 \times \text{PFP}_6 + \text{PFP}_4 - \text{PFP}_3 \times \text{PFP}_1$. Second, the quality of the individual tree is evaluated with the fusion objective (20) in our case. Note that the better individual tree with a higher fusion objective will have a higher chance of producing the children than the inferior individual trees. To guarantee the population diversity, the tournament selection scheme is applied in this case. Third, the evolution of the generation will be implemented through crossover, mutation, and reproduction with the predefined probabilities in order to find the individual with a superior fitness, and the details are also shown in Fig. 6. For the crossover, the procedure randomly selects the crossover points in both of the parent trees, and then it will produce an offspring tree by combining the copies of the subtree rooted at the crossover point of the parent trees. For the mutation, an offspring tree will be generated by connecting a copy of the subtree rooted at the mutation point of a parent tree with a randomly generated subtree. For the reproduction, based on the evaluated fitness, the selected parent tree will be simply inserted in the next generation. Note that the sum of the probabilities in terms of the crossover, the mutation, and the reproduction is 1. Fourth, the algorithm will be terminated when the value of the fusion objective function (20) is not increasing and its maximum value is obtained by at least one individual.

More details of the GP algorithm can be found in [24]. As a result, a CFP which can minimize the MAE of the RUL prediction can be obtained by applying GP to the selected candidate failure precursor from a group of PFPs and then fusing them by performing a series of mathematical operations.

V. Performance Illustration and Comparison

A. Developed Composite Failure Precursor

In this section, the proposed method is applied to the SiC MOSFETs as an example for the method validation. The PFP which shows a decreasing trend will be reversed to adapt to the stochastic degradation model. Considering the size of the PFPs, the population size in GP should be sufficiently large for the population diversity, and it is determined as 300 in our case. For the fusion objective (20) calculation, the cross validation technique [9] is applied by changing the test device one by one after the comparison scores are calculated for each individual test device. Specifically, the first six SiC MOSFETs devices are chosen as the training dataset and the #7 device as the tested one. For the next iteration, the first training device is shifted to the second training device and the second training

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2020.2991454, IEEE Transactions on Industrial Informatics

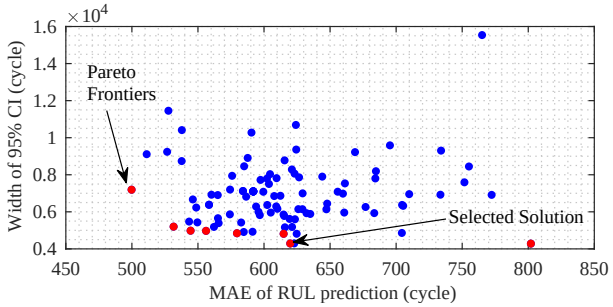

Figure 7. Pareto frontier analysis and the final solution determination.

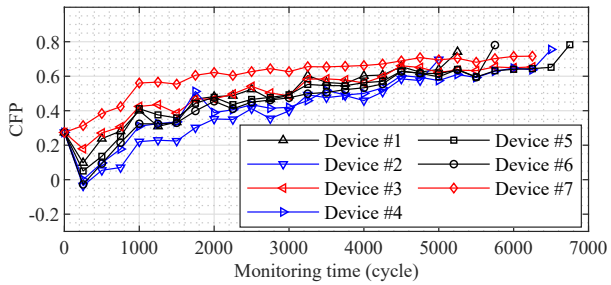

Figure 8. Degradation trends of the developed composite failure precursors.

device is shifted to the third and so on, until the first one reaches the last training device. The test device will then be rotated one by one for all seven SiC MOSFETs. Note that the GP algorithm initializes the population randomly. Due to the limited population size and the flexible nonlinear fusion, each run is converged with a different optimal solution. Considering the prediction uncertainty is another critical measure for performance evaluation, similar to (20), the prediction uncertainty level is calculated with the mean width of the 95% CI formulated as

$$W_s = \frac{1}{9M}\sum_{i=1}^{M}\sum_{\gamma=0.1}^{0.9}\left(\tilde{c}^{u}_{i,s,\lfloor\gamma\cdot n_i\rfloor} - \tilde{c}^{l}_{i,s,\lfloor\gamma\cdot n_i\rfloor}\right). \tag{21}$$

As a result, the solutions of a total of 100 runs are obtained and shown in Fig. 7. Note that the Pareto frontier is applied to balance the prediction accuracy and the prediction uncertainty. Meanwhile, considering the solution complexity, i.e., the size of the tree structure, the final optimal CFP is determined as $x_{\text{CFP}} = \sin(R_{\text{DS(on)}}) \cdot \cos(\ln(V_{\text{GS(th)}}))$. The respective MAE is 620 cycles and the mean width of 95% CI is 4291 cycles.

It can be seen that only the drain-source on-state resistance $R_{\text{DS(on)}}$ and the gate threshold voltage $V_{\text{GS(th)}}$ are selected from a total of six PFPs for the mathematical calculation of CFP. Note that only the basic operations including *times*, *sine*, *cosine,* and *log* are involved. These simple operations are computationally efficient and there is no specific requirement for the computation platform. It suggests that the developed CFP can be implemented on the regular embedded platform and is applicable to the *in-situ* health assessment. Note that the developed CFP is a dimensionless precursor possessing no specific physical meaning. As a result, Fig. 8 shows the developed CFP for a total of seven devices.

Note that the proposed stochastic degradation model supports the RUL prediction in terms of multiple failure precursors simultaneously. Combined with the PFPs, the developed CFP is considered as one of the multiple failure precursors and applied to the model for performance evaluation. Table II shows the MAE of RUL and the mean width of the 95% CI for the PFPs and the developed CFP. As can be seen, the MAE of RUL using the developed CFP is improved by 35.3% ((958-620)/958) compared to the best performance provided by PFPs, i.e., 958 cycles by using $R_{\text{DS(on)}}$. Additionally, compared to the best uncertainty performance provided by using the $R_{\text{DS(on)}}$, the mean width of the 95% confidence interval is reduced by 16.3% ((5125-4291)/5125) to 4291 cycles.

Table II
MAE OF RUL AND MEAN WIDTH OF THE 95% CI (UNIT IS CYCLE)

	C_{iss}	C_{oss}	C_{rss}	$R_{\text{DS(on)}}$	$V_{\text{GS(th)}}$	V_{SD}	CFP
RUL	3314	5477	2664	958	1426	5006	620
95% CI	13554	14732	12651	5125	7932	58108	4291

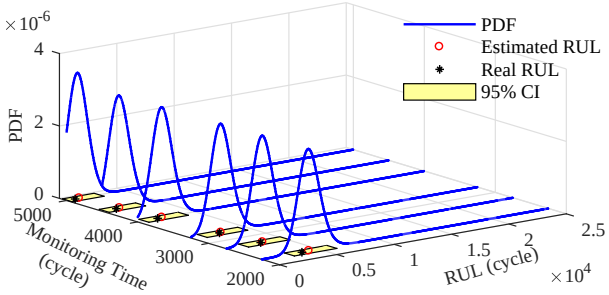

Figure 9. Evolution of RUL and 95% CI using CFP for device #6.

Next, *in-situ* RUL estimation for a specific test device subjected to the CM is performed in order to further show the superiority of the proposed method. As an illustration, device #6 is randomly selected as the test device and the rest devices are deployed as the training dataset. Fig. 9 shows the evolution of the estimated RUL and the 95% CI for device #6 using the CFP as the cycles progress. It can be seen that the predicted RUL is quite close to the real one and the width of the 95% CI is narrow.

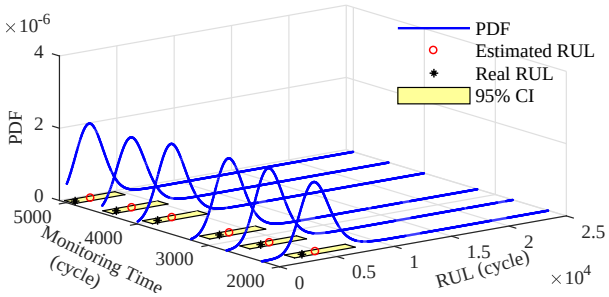

Figure 10. Evolution of RUL and 95% CI using $R_{\text{DS(on)}}$ for device #6.

Considering the $R_{\text{DS(on)}}$ is widely selected as the failure precursor in various studies (e.g., [2], [12], [14]) and provides the best performance among the PFPs, in this case the $R_{\text{DS(on)}}$ is used as a benchmark to show the superiority of the proposed CFP. Fig. 10 shows the evolution of RUL and 95% CI for device #6 using $R_{\text{DS(on)}}$, and the performance comparison of

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2020.2991454, IEEE Transactions on Industrial Informatics

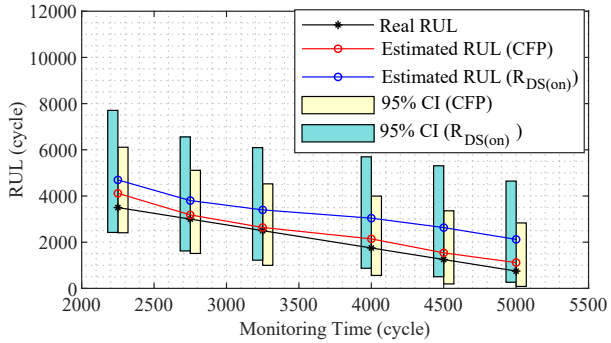

Figure 11. Performance comparison of RUL and 95% CI using the $R_{\text{DS(on)}}$ and the CFP for device #6.

the RUL and the 95% CI by using $R_{\text{DS(on)}}$ and the developed CFP is given in Fig. 11. It can be seen that the estimated RUL by using the $R_{\text{DS(on)}}$ is not quite close to the real RUL, even near the end-of-life of the device. In addition, the width of the 95% confidence interval is quite large, which suggests that the prediction uncertainty is high. While the prediction accuracy and the uncertainty level by using the developed CFP is much better. Note that the scales of z-axis of Fig. 9 and Fig. 10 are the same. It can be seen that the PDF provided by the CFP is more sharp than that of the $R_{\text{DS(on)}}$, which indicates the better performance of the prediction uncertainty. Specifically, Table III shows the details of the RUL prediction at multiple monitoring times using the $R_{\text{DS(on)}}$ and the developed CFP. For example, at CM time 4500 cycles, the estimated RUL is 2633 cycles with the 95% CI [505, 5309] using the $R_{\text{DS(on)}}$. Even the 95% CI covers the real RUL value, the width of the 95% CI, i.e., 4804 cycles, and the absolute error, i.e., 1383 cycles, are too large. While with the developed CFP, the estimated RUL is calculated as 1539 cycles, which is quite close to the real RUL 1250 cycles considering the CM period is 250 cycles. Additionally, compared to the $R_{\text{DS(on)}}$, the 95% CI covers the real RUL with a short width, i.e., 3166 cycles, which indicates that a more accurate prediction result has been obtained.

Table III
MAE OF RUL AND 95% CI AT MULTIPLE MONITORING TIMES FOR DEVICE #6 USING $R_{\text{DS(ON)}}$ AND DEVELOPED CFP (UNIT IS CYCLE)

CM	Real	Estimates with CFP		Estimates with $R_{\text{DS(on)}}$	
Time	RUL	RUL	95% CI	RUL	95% CI
2250	3500	4117	[2409, 6112]	4697	[2422, 7707]
2750	3000	3182	[1513, 5113]	3802	[1621, 6558]
3250	2500	2638	[998, 4526]	3399	[1222, 6094]
4000	1750	2144	[562, 3996]	3040	[871, 5697]
4500	1250	1539	[193, 3359]	2633	[505, 5309]
5000	750	1117	[85, 2834]	2121	[266, 4646]

B. Discussions

To further justify the proposed method, several suggestions and considerations of the proposed method in practical applications are presented. First, note that the dataset obtained from the power cycling test is aimed to the validation of the proposed method. The temperature stress of the device under the power cycling test is higher than the rated specifications and multiple failure mechanisms are triggered during the test. As an underlying assumption, the PFP variations due to the different failure mechanisms are not separately investigated. For example, the increase of the $R_{\text{DS(on)}}$ can be attributed to both gate interface degradation and package related degradation. As one of the future works, the considerations of the failure mechanisms and the accessibility of the PFPs will be incorporated into the CFP development to improve the engineering practicability. Second, it is worth mentioning that although an effective CFP is developed for the SiC MOSFETs with dataset from the power cycling test, such a CFP cannot be directly applied to SiC MOSFETs with a different model type due to the inconsistent degradation characteristics. For a different device model, the accelerated aging test experiment needs to be performed accordingly to generate the dataset for the degradation characteristics learning with the proposed method. Moreover, the proposed method is not limited to discrete power devices and can be extended to power modules or converters, where the PFPs can be obtained from the system-level electrical measurements. Third, it is worth mentioning that this paper aims to provide an integrated framework under which multiple potential failure precursors of the power device can be efficiently fused and exploited. The stochastic degradation model applied in this paper is capable of modeling the multiple precursors simultaneously. In this way, the RUL prediction performance in terms of each precursor can be fairly compared to highlight the superiority of the CFP. For practical applications, considering the degradation characteristics of power device, the degradation model in the framework can be flexibly replaced with other advanced models, e.g., Gamma or Wiener models.

VI. CONCLUSIONS

This paper presents a composite failure precursor formulation method for the condition monitoring and the RUL prediction of SiC MOSFETs. By interacting with the RUL prediction model directly, the genetic programming method is applied to the nonlinear fusion of the potential failure precursors to improve the RUL prediction performance. The findings are experimentally validated using the aging data of SiC MOSFETs exposed to power cycling tests. It is shown that the developed composite failure precursor is able to improve the RUL prediction accuracy by 35.3% and reduce the prediction uncertainty by 16.3%, compared to the directly measured physical failure precursors. With this approach, the strenuous works of selecting the suitable failure precursor of power devices and the respective degradation model can be mitigated, and the available condition monitoring data can be fully exploited to improve the RUL prediction performance. Also, the proposed method provides an effective solution to determine an efficient and pertinent failure precursor for other PHM applications, e.g., anomaly detection and fault diagnosis. This approach can be readily extended to other power devices, power modules, and converters.

REFERENCES

[1] H. Wang, M. Liserre, and F. Blaabjerg, "Toward reliable power electronics: Challenges, design tools, and opportunities," *IEEE Ind. Electron. Mag.*, vol. 7, no. 2, pp. 17–26, Jun. 2013.

[2] S. Dusmez, M. Bhardwaj, L. Sun, and B. Akin, "In situ condition monitoring of high-voltage discrete power MOSFET in boost converter through software frequency response analysis," *IEEE Trans. Ind. Electron.*, vol. 63, no. 12, pp. 7693–7702, Jul. 2016.

[3] N. Patil, D. Das, and M. Pecht, "A prognostic approach for non-punch through and field stop IGBTs," *Microelectronics Reli.*, vol. 52, no. 3, pp. 482–488, Mar. 2012.

[4] S. H. Ali, M. Heydarzadeh, S. Dusmez, X. Li, A. S. Kamath, and B. Akin, "Lifetime estimation of discrete IGBT devices based on gaussian process," *IEEE Trans. Ind. Appl.*, vol. 54, no. 1, pp. 395–403, Feb. 2018.

[5] A. Singh, A. Anurag, and S. Anand, "Evaluation of v_{ce} at inflection point for monitoring bond wire degradation in discrete packaged IGBTs," *IEEE Trans. Power Electron.*, vol. 32, no. 4, pp. 2481–2484, Oct. 2017.

[6] D. Astigarraga, F. M. Ibanez, A. Galarza, J. M. Echeverria, I. Unanue, P. Baraldi, and E. Zio, "Analysis of the results of accelerated aging tests in insulated gate bipolar transistors," *IEEE Trans. Power Electron.*, vol. 31, no. 11, pp. 7953–7962, Dec. 2016.

[7] S. Pu, E. Ugur, and B. Akin, "Real-time degradation monitoring of SiC-MOSFETs through readily available system microcontroller," in *IEEE 5th Workshop Wide Bandgap Power Devices Appl. (WiPDA),*, Dec. 2017, pp. 378–382.

[8] E. Ugur, F. Yang, S. Pu, S. Zhao, and B. Akin, "Degradation assessment and precursor identification for SiC MOSFETs under high temp cycling," *IEEE Trans. Ind. Appl.*, vol. 55, no. 3, pp. 2858–2867, Jan. 2019.

[9] A. Alghassi, S. Perinpanayagam, M. Samie, and T. Sreenuch, "Computationally efficient, real-time, and embeddable prognostic techniques for power electronics," *IEEE Trans. Power Electron.*, vol. 30, no. 5, pp. 2623–2634, Oct. 2015.

[10] A. Alghassi, S. Perinpanayagam, and M. Samie, "Stochastic RUL calculation enhanced with TDNN-based IGBT failure modeling," *IEEE Trans. Rel.*, vol. 65, no. 2, pp. 558–573, Dec. 2016.

[11] S. Dusmez, H. Duran, and B. Akin, "Remaining useful lifetime estimation for thermally stressed power MOSFETs based on on-state resistance variation," *IEEE Trans. Ind. Appl.*, vol. 52, no. 3, pp. 2554–2563, Jan. 2016.

[12] S. Dusmez, S. H. Ali, M. Heydarzadeh, A. S. Kamath, H. Duran, and B. Akin, "Aging precursor identification and lifetime estimation for thermally aged discrete package silicon power switches," *IEEE Trans. Ind. Appl.*, vol. 53, no. 1, pp. 251–260, Aug. 2017.

[13] X. Ye, C. Chen, Y. Wang, G. Zhai, and G. J. Vachtsevanos, "Online condition monitoring of power MOSFET gate oxide degradation based on miller platform voltage," *IEEE Trans. Power Electron.*, vol. 32, no. 6, pp. 4776–4784, Aug. 2017.

[14] S. Dusmez, M. Heydarzadeh, M. Nourani, and B. Akin, "Remaining useful lifetime estimation for power MOSFETs under thermal stress with RANSAC outlier removal," *IEEE Trans. Ind. Informatics*, vol. 13, no. 3, pp. 1271–1279, Feb. 2017.

[15] S. H. Ali, S. Dusmez, and B. Akin, "A comprehensive study on variations of discrete IGBT characteristics due to package degradation triggered by thermal stress," in *Proc. IEEE Energy Convers. Congr. Expo.*, Sep. 2016, pp. 1–6.

[16] T. Santini, S. Morand, M. Fouladirad, L. V. Phung, F. Miller, B. Foucher, A. Grall, and B. Allard, "Accelerated degradation data of SiC MOSFETs for lifetime and remaining useful life assessment," *Microelectronics Reli.*, vol. 54, no. 9-10, pp. 1718–1723, Oct. 2014.

[17] K. Liu, N. Z. Gebraeel, and J. Shi, "A data-level fusion model for developing composite health indices for degradation modeling and prognostic analysis," *IEEE Trans. Autom. Sci. Eng.*, vol. 10, no. 3, pp. 652–664, Apr. 2013.

[18] P. Baraldi, G. Bonfanti, and E. Zio, "Differential evolution-based multi-objective optimization for the definition of a health indicator for fault diagnostics and prognostics," *Mech. Syst. Signal Process.*, vol. 102, pp. 382–400, Mar. 2018.

[19] C. Song and K. Liu, "Statistical degradation modeling and prognostics of multiple sensor signals via data fusion: A composite health index approach," *IISE Trans.*, vol. 50, no. 10, pp. 853–867, May 2018.

[20] S. R. Bowling, M. T. Khasawneh, S. Kaewkuekool, and B. R. Cho, "A logistic approximation to the cumulative normal distribution," *J. Ind. Eng. Manage.*, vol. 2, no. 1, 2009.

[21] B. K. Bose, "Artificial intelligence techniques in smart grid and renewable energy systems-some example applications," *Proc. IEEE*, vol. 105, no. 11, pp. 2262–2273, 2017.

[22] L. Liao, "Discovering prognostic features using genetic programming in remaining useful life prediction," *IEEE Trans. Ind. Electron.*, vol. 61, no. 5, pp. 2464–2472, May 2014.

[23] S. E. De León-Aldaco, H. Calleja, and J. A. Alquicira, "Metaheuristic optimization methods applied to power converters: A review," *IEEE Trans. Power Electron.*, vol. 30, no. 12, pp. 6791–6803, 2015.

[24] S. Silva and J. Almeida, "GPLAB-a genetic programming toolbox for matlab," in *Proc. Nordic MATLAB conf.*, 2003, pp. 273–278.

Shuai Zhao (S'14-M'18) received the B.S. degree in telcommunication engineering, the M.S. degree in telcommunication and information system, and the Ph.D. degree in information and telcommunication engineering from Northwestern Polytechnical University, Xi'an, China, in 2011, 2014, and 2018, respectively.

He is currently a postdoctoral researcher with the Center of Reliable Power Electronics (CORPE), Department of Energy Technology, Aalborg University, Denmark. From Sep. 2014 to Sep. 2016, he was a visiting Ph.D. Student with the Department of Mechanical and Industrial Engineering at the University of Toronto, Toronto, ON, Canada, with the scholarship from China Scholarship Council (CSC). In August 2018, he was a visiting scholar with the Power Electronics and Drives Laboratory, Department of Electrical and Computer Science at the University of Texas at Dallas, Richardson, TX, USA. His research interests include condition monitoring, data analytics, residual life prediction, and health assessment of power electronic systems.

Shaowei Chen (M'15) is currently an associate professor in the School of Electronics and Information, Northwestern Polytechnical University, Xi'an, China, where he is also the dean of the Department of telcommunication engineering and the Director of Signal Processing and Control Technology Laboratory. He is the principal investigator of several projects supported by the Aeronautical Science Foundation of China, Beijing, China. His research expertise is in the area of the fault diagnosis, sensors, condition monitoring, and prognosis of electronic systems. Prof. Chen is selected to receive several Provincial and Ministerial Science and Technology Awards. He is a senior member of the Chinese Institute of Electronics.

Fei Yang (S'13) received the B.S. degree from Northwestern Polytechnical University, Xi'an, China, in 2011, the M.S. degree from the Center for Ultra-Wide-Area Resilient Electric Energy Transmission Networks (CURENT), the University of Tennessee, Knoxville, TN, USA, in 2017, and the Ph. D. degree from the University of Texas at Dallas, Richardson, TX, USA, in 2020, all in electrical engineering.

He is currently a System Engineer at Texas Instruments, Dallas, TX, USA. From 2012 to 2014, he was working as a researcher at Kettering University, Flint, MI, USA. His research interests include wide bandgap semiconductor device's reliability and application, power module packaging and integration, and motor drive system design.

Dr. Yang received the 2019 Mary and Richard Templeton Fellowship from the University of Texas at Dallas

Enes Ugur (S'11) received the B.Sc. degree in electrical engineering from Istanbul Technical University, Istanbul, Turkey, in 2008 and the M.Sc. degree from Yildiz Technical University, Istanbul, in 2011. He received the Ph.D. degree in Electrical Engineering from The University of Texas at Dallas, Richardson, TX, USA, in 2019. He is currently working as a power electronics systems engineer at Wolfspeed, NC, USA. His research interests include power converters and inverters, wide-bandgap devices and applications.

Bilal Akin (M'08-SM'13) received the Ph.D. degree in electrical engineering from the Texas A&M University, College Station, TX, USA, in 2007.

He was an R&D Engineer with Toshiba Industrial Division, Houston, TX, USA, from 2005 to 2008. From 2008 to 2012, he worked as an R&D Engineer at C2000 DSP Systems, Texas Instruments Incorporated. Since 2012, he has been with University of Texas at Dallas as faculty. Dr. Akin is recipient of NSF CAREER'15 award, IEEE IAS Transactions 1st Place Prize Paper Award and Top Editors Recognition Award from IEEE TVT Society, Jonsson School Faculty Research Award and Jonsson School Faculty Teaching Award. He is an Associate Editor of IEEE Transactions on Industry Applications and IEEE Transactions on Vehicular Technology. His research interests include design, control and diagnosis of electric motors & drives, digital power control and management, fault diagnosis & condition monitoring of power electronics components and ac motors.

Huai Wang (M'12-SM'17) received the B.E. degree in electrical engineering, from Huazhong University of Science and Technology, Wuhan, China, in 2007 and the Ph.D. degree in power electronics, from the City University of Hong Kong, Hong Kong, in 2012. He is currently Professor with the Center of Reliable Power Electronics (CORPE), Department of Energy Technology at Aalborg University, Denmark. He was a Visiting Scientist with the ETH Zurich, Switzerland, from Aug. to Sep. 2014, and with the Massachusetts Institute of Technology (MIT), USA, from Sep. to Nov. 2013. He was with the ABB Corporate Research Center, Switzerland, in 2009. His research addresses the fundamental challenges in modelling and validation of power electronic component failure mechanisms, and application issues in system-level predictability, condition monitoring, circuit architecture, and robustness design.

Dr. Wang received the Richard M. Bass Outstanding Young Power Electronics Engineer Award from the IEEE Power Electronics Society in 2016, and the Green Talents Award from the German Federal Ministry of Education and Research in 2014. He is currently the Chair of IEEE PELS/IAS/IES Chapter in Denmark. He serves as an Associate Editor of IET Electronics Letters, IEEE JOURNAL OF EMERGING AND SELECTED TOPICS IN POWER ELECTRONICS, and IEEE TRANSACTIONS ON POWER ELECTRONICS.