IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 22 November 2023, accepted 15 December 2023, date of publication 25 December 2023, date of current version 9 January 2024.

Digital Object Identifier 10.1109/ACCESS.2023.3347035

RESEARCH ARTICLE

Reliability Demonstration Test Planning for Systems Using Prior Knowledge

ALEXANDER GRUNDLER[1] **AND MARTIN DAZER**[2]
[1]Robert Bosch GmbH, 71701 Schwieberdingen, Germany
[2]Institute of Machine Components, University of Stuttgart, 70569 Stuttgart, Germany
Corresponding author: Martin Dazer (dazer@ima.uni-stuttgart.de)

ABSTRACT Empirical life tests are used for reliability demonstration and determination of the actual reliability of the product. Therefore, engineers are faced with the challenge of selecting the most suitable test strategy out of the possible many and also the optimal parameter setting, e.g. sample size, in order to realize reliability demonstration with limited costs, time and with their available testing resources. It becomes even more challenging due to the stochastic nature of failure times and necessary cost and time being dependent on those. The considerations and guidelines in this paper are intended to simplify this process. Even simple products can fail due to several causes and mechanisms and usually have several components and subsystems. Therefore, this paper provides test planning options for single critical failure mechanisms as well as for systems with multiple failure mechanisms. For this purpose, the Probability of Test Success (Statistical Power of a life test) is used as a central, objective assessment metric. It is capable of indicating the probability of a successful reliability demonstration of a test and thus allows, for example, to answer the question of the required sample size for failure-based tests. The main planning resource is prior knowledge, which is mandatory due to the stochastic lifetime, in order to provide estimates for the Probability of Test Success at all. Therefore, it is also shown how to deal with uncertain prior knowledge and how the underlying information can additionally be used to increase the Probability of Test Success using Bayes' theorem. The guidelines show how the most efficient test can be identified in the individual case and for individual boundary conditions.

INDEX TERMS Bayes' theorem, prior knowledge, reliability demonstration, system reliability, test planning, uncertainty.

I. INTRODUCTION

Today's products are usually characterized by multiple functions. Particularly in the digital age, complex products are often necessary in order to survive successfully on the market. However, such complex products are reflected in an equally large variety of possible causes of failure, which intensify the challenges in product validation. But even conventional and simpler products often have more than one cause of failure that needs to be dealt with. In combination with the constantly increasing market pressure and rising customer requirements, these challenges must be conquered with scientific methods. The reliability is determined with appropriate tests. In the selection and design of such tests, the conflicting goals of accuracy, cost and time must be addressed, which complicates the test planning. An objective assessment of the various tests with respect to the demonstration of the system reliability is necessary to avoid expensive development loops and tests with little chance of success. Current test planning methods are not able to consider complex systems with multiple failure modes.

Physical testing of products prior to market entry is essential for determining actual reliability and comparison with requirements. In order to gain the most accurate information possible, testing must be carried out with a sample as large as possible. However, this statistical requirement contrasts with the possibilities of a company. Resources are always limited and testing should be completed in a short time.

The associate editor coordinating the review of this manuscript and approving it for publication was Haidong Shao.

In addition, there is a large number of different tests available that can be used for testing and reliability demonstration. The concept of Probability of Test Success was first introduced by Dazer et al. [1], [2], [3] and allows to plan the necessary tests in a way that they have a maximum probability of success and at the same time can be implemented with the available resources. Accordingly, it is an objective evaluation of the tests, which significantly supports the planning process. However, Dazer's previous work only provides the basic concept for test planning. It is not possible to evaluate tests of complex systems with more than one failure mode. Furthermore, no prior knowledge can be additionally included in the demonstration test and the uncertainty, inevitably present in the prior knowledge, remains unconsidered.

The objectives of this work can be derived from the problem described. The main focus is on extending the use of the Probability of Test Success for the objective assessment of tests for the demonstration of reliability in realistic, complex scenarios. For this purpose, solutions for the two main challenges in realistic use cases must be developed:

Consideration of…

- several competing failure modes (Chapter V)
- uncertainty in prior knowledge (Chapter VI)
- prior knowledge using Bayes' theorem (Chapter IV)

Due to the fact that prior knowledge is needed for test planning anyway, it is obvious to use the prior information also to improve the result with Bayes' theorem. Therefore, the third main challenge is the consideration of prior knowledge information for P_{ts} calculation using Bayes' theorem.

In order to make the planning of real systems with several failure modes possible, the concept has to be extended to consider the uncertainty of the prior knowledge. Ultimately, the combination of the extensions in a holistic procedure for the evaluation and identification of the most efficient test for the demonstration of the reliability of systems is necessary.

This paper is organized as follows:

Chapter II outlines the state of research and illustrates the necessity of the developed method. The central metric of the method is the Probability of Test Success. In Chapter III, it is explained how it can be calculated for products with a single failure mode, as well as the statistical context of hypothesis testing. In order to additionally consider prior knowledge using Bayes' theorem, Chapter IV is concerned with the calculation of the Probability of Test Success while using Bayes' theorem. Since real products have several failure modes, Chapter V describes the developed methods which are required to use the metric on systems with multiple failure modes. Due to the inevitably present uncertainty in the used prior knowledge, Chapter VI deals with the necessary methods to consider this uncertainty in the planning of reliability tests. In Chapter VII, the influences of the consideration of Bayes' theorem, uncertainty in prior knowledge as well as multiple failure modes on the Probability of Test Success are studied and general findings for the failure based and failure free tests are derived. The interactions of those influences are additionally analyzed in Chapter VIII and conclusions for reliability test planning are drawn. Finally, Chapter IX summarizes the findings.

The aspects of the proposed procedure are shown in Fig. 1.

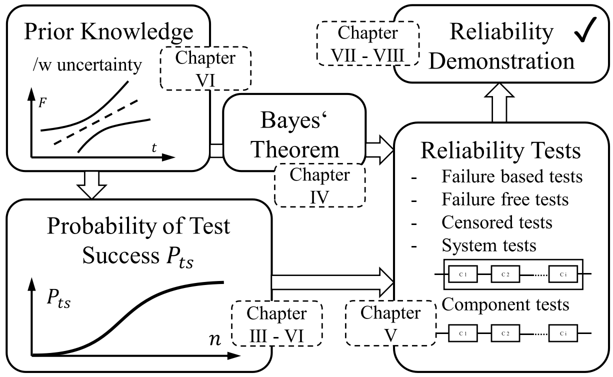

FIGURE 1. Considered aspects of the proposed method.

II. STATE OF RESEARCH

The Probability of Test Success, first introduced by Dazer et al. [1], is defined as the relative frequency of a successful reliability demonstration. Early calculation procedures made use of the law of large numbers and by simulating the tests while incorporating prior knowledge about the failure distribution, the probability could be calculated. In order to establish a broader statistical context, Grundler et al. [4] defined the Probability of Test Success as the statistical power of a reliability demonstration test, since all reliability demonstration tests can be approached as hypothesis tests. By making use of this statistical context, new calculation procedures could be developed [8], [21], [22], [23], [24] e. g. using the asymptotic variance of the maximum likelihood estimation in [4]. Although several studies have been conducted in order to enable the application of the Probability of Test Success for systems with multiple failure modes [8], [21], [24], a proper procedure facilitating a holistic view is still necessary for an efficient planning procedure of reliability demonstration tests. In addition to the studies regarding the consideration of uncertainty [22] as well as the combined approaches for using Bayes' theorem [19], [20], [23], [25] and the concept of the Probability of Test Success [4], the combination of all three aspects in a single holistic procedure has not been tackled yet. Other approaches for reliability demonstration test planning solely consider the statistical error of type I (confidence) in order to derive required sample sizes of EoL tests [28]. Since the confidence bounds vary from test to test, this approach is lacking the consideration of the error of type II (statistical power), which is incorporated in the Probability of Test Success itself. Hamada et al. [26], Tsai et al. [30] as well as Wilson and Farrow [31] are using a similar concepts. However, they are establishing the equations and credibility intervals solely based on a Bayesian view, which is therefore decoupled from the frequentist reliability requirement using a confidence. Systems with multiple failure modes are not considered.

Maisch [27] and Yadav et al. [32] are investigating reliability demonstration tests on different system levels and the additional consideration of prior knowledge using Bayes' theorem. However, they're only considering failure free tests and no statistical error of type II. In addition, the interaction of the confidence intervals of the different tests of the components are not considered and the requirements are assigned to the components separately.

III. CALCULATION OF THE PROBABILITY OF TEST SUCCESS

In principle, life tests can be divided into failure-free and failure-based test strategies. For both representative's own calculation rules for the planning of reliability demonstration tests can be derived. In order to be able to explain the extensions to cover the realistic use cases a basic understanding of the previous calculation of the planning procedure is necessary. For more details, refer to [4].

A. END OF LIFE TESTS

Product development - with all the calculations and service life estimates as well as the design of the product itself - should ensure not only the main and secondary functions but also the specified reliability requirement for the corresponding conditions of use [5], [6]. However, without having actually observed the fulfillment of these requirements, they cannot be assumed to be met. For this reason, a hypothesis about the reliability target can be formulated. This hypothesis must be either rejected or confirmed by appropriately performed reliability tests. Accordingly, a reliability demonstration test can be considered as a hypothesis test. The lifetime quantile t_R obtained by the test at required reliability R_{req} must be greater than or equal to the required lifetime t_{req}. Since a test can only provide information that rejects the hypothesis about the absence of a phenomenon under investigation [7], the null hypothesis H_0 represents the non-fulfillment of the given reliability target. The goal of the test is to gather information so that the null hypothesis can be rejected. Since the statistical power of a test corresponds to the discovery of the alternative, the alternative hypothesis H_1 represents the fulfillment of the reliability target. Accordingly, the hypotheses to be used in a reliability demonstration test are the following [4], [8]:

$$H_0 : t_R < t_{\text{req}} \tag{1}$$

$$H_1 : t_R \geq t_{\text{req}} \tag{2}$$

The significance level at which H_0 should be rejected is given by the maximum accepted error α (type I error). This corresponds to the probability of rejecting the null hypothesis although it is actually true. The required confidence C_{req} of a reliability demonstration test is the complement of the required significance level, since it describes the probability that the null hypothesis is correctly accepted as true, i.e. that the product does actually not achieve the reliability target. The type II error, on the other hand, describes the probability β that the null hypothesis was wrongly accepted. It depends on the type of test, sample size, actual value of the lifetime quantile, failure distribution, sample variance and required confidence level. The complement of the type II error is the statistical power and describes the probability of correctly rejecting the null hypothesis for a certain value of effect size [9]. Thus, in the context of reliability demonstration tests, the statistical power describes the probability of the test to demonstrate the reliability requirement. Since the hypotheses defined according to Eq. 1 and 2 are always identical for a reliability demonstration test, the Probability of Test Success P_{ts} can be understood as the statistical power. A sole consideration of the required confidence for the planning of reliability tests is therefore insufficient, since the investigation regarding the suitability of the test to the current scenario is neglected [1], [2], [3], Since the statistical power - and thus the P_{ts} - can only ever be specified for a certain effect size [7], i.e. a certain value of the actual lifetime quantile of the product, and since the detection probability of the alternative hypothesis changes with the effect size, a suitable quantity must be developed which can indicate the effect size in reliability demonstration tests. For this purpose, the product design safety distance is used, which was introduced by Dazer et al. [1]. It describes the proportional distance between the required t_{req} and the actual service life of the product t_R corresponding to the prior knowledge t_{p}. Both lifetime quantiles have to be obtained for the required reliability. It is defined as:

$$s = 1 - \frac{t_{\text{req}}}{t_R} \quad \text{with} \quad t_R := t_{\text{p}} \tag{3}$$

The product safety distance is equal to zero if prior knowledge states that the required service life is equal to the actual service life. If this quantity is used the effect size τ can be formulated as follows:

$$\tau \equiv t_{\text{p}} - t_{\text{req}} = s \cdot t_{\text{p}} \tag{4}$$

Fig. 2 shows all the necessary relationships.

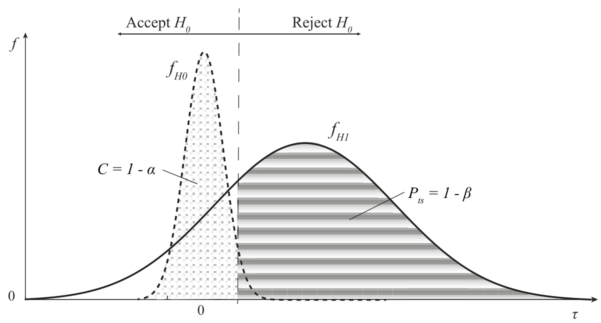

FIGURE 2. Relationships of f_{H_0} & f_{H_1} and the resulting integrals of C and P_{ts} [4].

The hypotheses used for reliability demonstration can then be reformulated using the test statistic:

$$H_0 : \tau < 0 \tag{5}$$

$$H_1 : \tau \geq 0 \tag{6}$$

To calculate the P_{ts}, the distributions of the null hypothesis f_{H_0} and those of the alternative hypothesis f_{H_1} must be calculated. The P_{ts} can then be determined using the integral:

$$P_{\text{ts}} = 1 - \beta = \int_{\tau_{\text{crit}}}^{\infty} f_{H_1}(\tau)\, d\tau \tag{7}$$

Whereas the value τ_{crit} corresponds to the confidence via the following integral:

$$C = 1 - \alpha = \int_{0}^{\tau_{\text{crit}}} f_{H_0}(\tau)\, d\tau \tag{8}$$

The calculation of these distributions can be done in a numerical way or in an analytical-approximative way. In this paper only the analytical approach will be presented. For the numerical solution we refer to [4].

To be able to determine the distributions of τ_{H_0} and τ_{H_1} without bootstrap or MCS, the central limit theorem is used. Accordingly, for finite sample sizes, the normal distribution can be used as an approximation to the distribution of the test statistic. According to the central limit theorem, the distribution of the sample quantile of a known distribution $F(t)$, i.e. the empirically formed quantile, is normally distributed with the parameters [18], DasGupta 2008:

$$\mu = F^{-1}(q) \tag{9}$$

$$\sigma = \sqrt{\frac{q \cdot (1-q)}{n \cdot f\left(F^{-1}(q)\right)^2}} \tag{10}$$

where the q-quantile of the distribution $F(t)$ is formed from its inverse function $F^{-1}(q)$. $f(t)$ is the density function of $F(t)$ and n is the sample size. Using the same test statistic 0f Eq. 4 again and the asymptotic behavior from Eq. 9 & 10, the approximate distributions of τ_{H_0} and τ_{H_1} can be determined as normal distributions as follows:

$$\tau_{H_0} \sim \mathcal{N}\left(0, \sqrt{\frac{R_{\text{req}} \cdot (1-R_{\text{req}})}{n \cdot f_{H_0}(t_{\text{req}})^2}}\right) \tag{11}$$

$$\tau_{H_1} \sim \mathcal{N}\left(t_{\text{p}} - t_{\text{req}}, \sqrt{\frac{R_{\text{req}} \cdot (1-R_{\text{req}})}{n \cdot f_{H_1}(t_{\text{req}})^2}}\right) \tag{12}$$

In this notation $\mathcal{N}(\mu, \sigma)$ is the normal distribution with mean μ and standard deviation σ.

f_{H_0} is the transformed distribution from prior knowledge f_{H_1} according to H_0 [4]. If prior knowledge is formulated using a Weibull distribution with characteristic life T_{H_1} one can calculate the transformed characteristic life T_{H_0} as follows:

$$T_{H_0} = (1-s) \cdot T_{H_1} = \frac{t_{\text{req}}}{t_{\text{p}}} \cdot T_{H_1} \tag{13}$$

The P_{ts} can then be determined very easily via the approximated normal distribution:

$$P_{\text{ts}} = 1 - \Phi\left(\Phi^{-1}\left(C_{\text{req}}; 0, \sqrt{\frac{R_{\text{req}} \cdot (1-R_{\text{req}})}{n \cdot f_{H_0}(t_{\text{req}})^2}}\right); t_{\text{p}} - t_s, \sqrt{\frac{R_{\text{req}} \cdot (1-R_{\text{req}})}{n \cdot f_{H_1}(t_{\text{req}})^2}}\right) \tag{14}$$

In this notation Φ is the cdf of the normal distribution.

Here, the lifetime quantile is determined as an empirical sample quantile.

Due to the multiplicative relation between the failure times of H_1 and H_0, this relationship is also valid for the corresponding life quantiles and therefore also for the relationship between the alternative distribution and null distribution. For this reason, according to Eq. 3, there is the following relationship of the standard deviations of the alternative and null distribution:

$$\sigma_{H_0} = (1-s) \cdot \sigma_{H_1} = \frac{t_{\text{req}}}{T_{H_1}} \cdot \left(-\ln\left(R_{\text{req}}\right)\right)^{-1/b} \cdot \sigma_{H_1} \tag{15}$$

Here, it is assumed that no estimation of the failure distribution takes place, for example, via an MLE. This also means that censoring can only be considered to a very limited extent. To overcome this drawback, the scale and variance of the lifetime quantile can also be calculated using the asymptotic properties of the MLE and the variance-covariance matrix. For a detailed derivation of the equations involved, see [4].

The scale and variance of the distribution of the lifetime quantile under validity of the alternative hypothesis is obtained using the Weibull distribution as:

$$t_{R,H_1} \sim \mathcal{N}\left(T_{H_1}\left(-\ln\left(R_{\text{req}}\right)\right)^{1/b}, \sigma_{H_1}\right) \tag{16}$$

with:

$$\begin{aligned}\sigma_{H_1} = \Bigg(&\left(-\ln\left(R_{\text{req}}\right)\right)^{1/b} \cdot \text{Var}\left(T_{H_1}\right) \\ &+ \frac{T_{H_1}^2}{b^4} \cdot \ln\left(-\ln\left(R_{\text{req}}\right)\right)^2 \cdot \left(-\ln\left(R_{\text{req}}\right)\right)^{2/b} \cdot \text{Var}(b) \\ &- \frac{2T_{H_1}^2}{b^2} \cdot \ln\left(-\ln\left(R_{\text{req}}\right)\right) \cdot \left(-\ln\left(R_{\text{req}}\right)\right)^{2/b} \\ &\cdot \text{Cov}\left(T_{H_1}, b\right)\Bigg)^{1/2}\end{aligned} \tag{17}$$

T_{H_1} and b are the Weibull scale and shape parameter stemming from prior knowledge considering the assumed failure behavior of the H_1 hypothesis. Since the failure mode stays the same, the shape parameter b is valid for both hypotheses.

The transformation using s can also be used here to calculate the statistics under validity of the null hypothesis. The P_{ts} then results in:

$$P_{\text{ts}} = 1 - \Phi\left(\Phi^{-1}\left(C_{\text{req}};\ 0,\ \frac{T_{H_1}}{t_{\text{req}}}\left(-\ln\left(R_{\text{req}}\right)\right)^{1/b}\sigma_{H_1}\right);\ T_{H_1}\left(-\ln\left(R_{\text{req}}\right)\right)^{1/b} - t_{\text{req}},\ \sigma_{H_1}\right) \tag{18}$$

B. SUCCESS RUN TEST

The Success Run (SR) Test is based on a binary classification in which all specimens are tested up to a predefined lifetime. Each specimen is simply assigned as "failed" or "success" for the result. Due to the simple binary classification, the binomial distribution can be used as a planning approach:

$$C = 1 - \sum_{i=0}^{k} \binom{n}{i} \cdot \left(R_{\text{req}}\right)^{n-i} \cdot \left(1 - R_{\text{req}}\right)^{i} \tag{19}$$

The binomial approach assumes that the parameter of the success probability of the binomial distribution p is equal to the complement of the required reliability R_{req}. Therefore:

$$p = 1 - R_{\text{req}} \tag{20}$$

This corresponds to the null hypothesis for the limiting case of $s = 0$. In analogy to Eq. 12, the P_{ts} of a Success Run Test can be calculated analytically and exactly using the following binomial distribution:

$$P_{\text{ts}} = \sum_{i=0}^{k} \binom{n}{i} \cdot \left(R_{\text{p}}\right)^{n-i} \cdot \left(1 - R_{p}\right)^{i} \tag{21}$$

Instead of the required reliability, the reliability at the required lifetime t_{req} corresponding to the prior knowledge $R_{\text{p}}\left(t_{\text{req}}\right) = 1 - F(t_{\text{req}})$ is used here as the complement of the success probability parameter p of the binomial distribution [4]. If the reliability according to prior knowledge is greater than or equal to the required $R_{\text{p}}\left(t_{\text{req}}\right) \geq R_{\text{req}}\left(t_{\text{req}}\right)$, this corresponds to the alternative hypothesis. Due to the relationship between the binomial and beta distributions [16], Eq. 19 & 21 can also be written as beta distributions:

$$C = \int_{R_{\text{req}}}^{1} \frac{R^{n-k-1} \cdot (1-R)^{k}}{\beta\,(n-k, k+1)} dR \tag{22}$$

$$P_{\text{ts}} = \int_{0}^{R_{\text{p}}} \frac{R^{n-k-1} \cdot (1-R)^{k}}{\beta\,(n-k, k+1)} dR \tag{23}$$

It is the same beta distribution $\mathcal{B}\,(n-k, k+1)$ in both cases, because the resulting reliability distribution is defined solely by the number of survivors and the failed specimens.

Since Confidence and P_{ts} are calculated from same distribution they only differ in the integral limits. It is obvious that the P_{ts} becomes the complement of the required Confidence i.e., $P_{\text{ts}} \rightarrow C_{\text{req}}$ when the required reliability approaches the actual reliability, i.e. for $R_{\text{p}} \rightarrow R_{\text{req}}$ and $s \rightarrow 0$, respectively. The relationship of P_{ts} and Confidence and the relevant parameters are shown in Fig. 3.

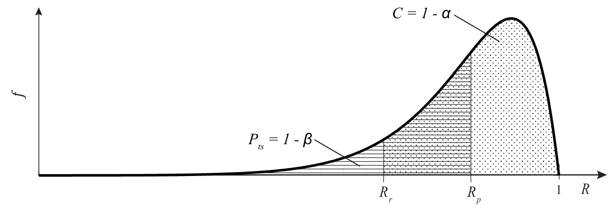

FIGURE 3. Beta distribution of the SR test and the corresponding integrals for C and P_{ts} [4].

TABLE 1. Two types of prior knowledge and their relationships with the test types.

		Type of planned test	
		SR Test	EoL Test
Type of prior knowledge	SR Test (Distribution of Reliability)	OK	Not OK
	EoL Test (Failure Distribution)	OK	OK

IV. CONSIDERATION OF PRIOR KNOWLEDGE USING BAYES' THEOREM AND P_{ts}

Prior knowledge is essential for efficient reliability demonstration. For example, the widely used SR test cannot be evaluated without valid prior knowledge if test times deviate from the required service life. The planning of EoL tests also necessarily requires prior knowledge of the predicted failure behavior of the product. Since prior knowledge must be available anyway in order to calculate a P_{ts} at all on the basis of the stochastic service life, it is obvious to include the prior knowledge also as additional information about reliability in the demonstration test. Bayes' theorem is used for this purpose. The essential feature is that in this case information from prior knowledge is treated as equivalent to information from the test, whereas when using prior knowledge by means of the P_{ts}, prior knowledge is only used to assess the tests and a subsequent evaluation of the performed test is not influenced by the prior knowledge. However, in order to use prior knowledge in these two cases, it must first be available in a suitable form, see Table 1.

A. CALCULATION OF P_{ts} FOR EOL TESTS

In an EoL test, a distinction must be made between the two types of prior knowledge. Prior knowledge in the form of an SR test only contains information about the reliability of the product for a certain service life. The EoL test, on the other hand, provides information about the overall failure behavior. For the calculation of the P_{ts}, it must be estimated where the failures may probably occur. For this reason, prior knowledge of the failure behavior is mandatory for the calculation of the P_{ts} for an EoL test. Prior knowledge from an SR test is not sufficient.

If prior knowledge of the failure distribution is available, a distinction must still be made between two cases. The first case is the knowledge of the entire sample with failure times and in the other case the information about the failure distribution like for example a Weibull distribution with the information about the original sample size n_0 is available. If the sample is known, its likelihood function L_0 can be used as the distribution of prior knowledge regarding the parameters of the failure distribution together with the likelihood function from the current EoL test L_1 in a MAP estimation [10], resulting in the following expression for the resulting likelihood function, which represents the combined information:

$$L_{\text{post}} \propto L_1 \cdot L_0 = \prod_{i=1}^{n_1-z_1} f(t_i) \prod_{j=1}^{z_1} R(t_j) \cdot \prod_{l=1}^{n_0-z_0} f(t_l) \prod_{m=1}^{z_0} R(t_m) \tag{24}$$

In this MAP n_0 and z_0 are the sample size and the number of right-censored run times of prior knowledge and analogously n_1 and z_1 those of the current EoL test (test to be planned).

The MAP represents the application of Bayes' theorem by updating the likelihood function by prior knowledge as a priori distribution over the parameters of the failure distribution with the likelihood function from the current EoL test. Accordingly, the combination of the information takes place via a multiplication of the two likelihood functions. The normalization constant, as it occurs in the theorem of Bayes is not needed here, because the unknown parameters are determined by the maximum of L_{post} and this maximum does not change by a multiplicative factor. This is the reason why Eq. 24 is formulated as a proportional expression. By simply multiplying the two likelihood functions, the application of MAP here is identical to an MLE estimate of the parameters if the sample from prior knowledge were evaluated together with the sample from the current EoL test. This simplifies the application since no further implementation effort is required.

However, if the prior knowledge is available as a failure distribution with underlying sample size n_0, then this combination of samples cannot be easily made. Nevertheless, to obtain a sample for this combination, an approximate approach can serve as an estimate of the expected failure times. The values of the failure times t_i^* of the synthetic sample generated here, are calculated according to the distribution of the order statistics. The distribution of the failure probabilities of the order statistics is used, as in the method for computing the beta-binomial confidence intervals. For the known sample size n_0 and the known failure distribution $F_0(t)$ of prior knowledge, the synthetic failure times t_i^* can thus be calculated as:

$$t_i^* = F_0^{-1}\left(\mathcal{B}\left(0,5; i, n_0 - 1 + 1\right)\right) \approx F_0^{-1}\left(\frac{i - 0.3}{n_0 + 0.4}\right) \tag{25}$$

Here, the approximation to the median of the beta distribution according to Benard was used [11].

However, the sample combined here includes the proportion from prior knowledge, which is fixed and does not vary. For this reason, it does not generate any sample scatter. Only the still uncertain failure times of the not yet performed test provide scatter in the combined sample. This fact is considered by calculating the variance of the quantile of the combined sample Var_{comb} from two estimates. Firstly, there is the variance of the quantile of a sample of size $n + n_0$ of the combined sample Var_{n+n0}, which is the variance of the quantile from the test. To reduce this by the correct proportion of the non-scattering fixed sample of prior knowledge, we first calculate the variance of the quantile of a sample of size n_0 of prior knowledge Var_{n0}, which represents the second variance. To combine these two variances and thus effectively reduce the variance Var_{n+n0}, the two distributions of the quantiles describing these variances are combined by Bayes' theorem. This represents a multiplication of the two distributions (corresponding to $n + n_0$ and n_0), which can also be understood as the weighting of the two distributions, see equation. The variance of the resulting quantile distributions Var_{comb} is then calculated via.

$$\text{Var}_{\text{comb}} = \frac{\text{Var}_{n+n_0} \cdot \text{Var}_{n_0}}{\text{Var}_{n+n_0} + \text{Var}_{n_0}} \tag{26}$$

The equation is obtained by applying Bayes' theorem to two normal distributions where the resulting distribution is again a normal distribution, whose variance has the above relationship to the initial distributions [12].

Eq. 26 can also be adapted to calculate the standard deviation which is necessary to calculate the P_{ts} :

$$\sigma_{\text{comb}} = \sqrt{\frac{\left(\sigma_{n+n_0}\right)^2 \cdot \left(\sigma_{n_0}\right)^2}{\left(\sigma_{n+n_0}\right)^2 + \left(\sigma_{n_0}\right)^2}} \tag{27}$$

The calculation of the standard deviation is still the same using the corresponding sample sizes of $n + n_0$ and n_0. In analogy to Eq. 18 the P_{ts} is then calculated as:

$$P_{\text{ts}} = 1 - \Phi\left(\Phi^{-1}\left(C_{\text{req}}; 0, \frac{T_{H_1}}{t_{\text{req}}}\left(-\ln\left(R_{\text{req}}\right)\right)^{1/b} \cdot \sigma_{\text{comb}}\right); T_{H_1}\left(-\ln\left(R_{\text{req}}\right)\right)^{1/b} - t_{\text{req}}, \sigma_{\text{comb}}\right) \tag{28}$$

B. CALCULATION OF P_{ts} FOR SUCCESS RUN TESTS

For an SR test, the calculation of the P_{ts} does not change because the calculation reflects the probability of survival of a certain sample size given a known probability of survival of the test items. However, taking prior knowledge into account only leads to a reduction of the necessary sample size while the reliability target remains unchanged, see also Chapter III. This means that the calculation of the P_{ts} can be performed unchanged according to Eq. 21 and 23, respectively. Only the sample size changes. A reduction of the required sample size is only possible if the prior knowledge attests a minimum reliability equal to the required, or a greater one.

V. CONSIDERATION OF THE SYSTEM STRUCTURE

Even relatively simple products usually have several components that could potentially fail. A consideration of the system structure and thus a holistic view of the system already in the planning phase of the reliability tests is therefore indispensable for a successful demonstration and verification of the system reliability.

To demonstrate the reliability of a system, it must be clearly defined which elements belong to the system. System elements are subsystems, components or parts. The physical components of a system usually each have at least one failure mechanisms, which is why the system must be analyzed in this respect. Knowledge of all potential failures as well as their combined effect on the system behavior can be obtained, for example, by means of a failure mode and effects analysis (FMEA) or fault tree analysis (see [5]). The failure mechanisms relevant to the system must be documented in a reliability block diagram so that the joint effect on the system is clear and the logical relationship between the failure mechanisms is defined. In this work, only systems that have independent failure modes are considered.

Unlike products or components with a single failure mode, systems with multiple subsystems, assemblies, components, and parts may use more than one test to demonstrate system reliability. Since the P_{ts} is intended to represent the demonstration of system reliability, it must be determined jointly for all tests. For the calculation, a distinction must be made between the two test types EoL test and SR test as well as their combination. In addition, it is important that prior knowledge of the system structure and all critical failure modes is available. For sole SR tests, prior knowledge in the form of the SR test is sufficient. For failure-based tests, knowledge of the failure distribution is necessary.

A. CALCULATION OF P_{ts} FOR EOL TESTS CONSIDERING THE SYSTEM STRUCTURE

The approach presented in Chapter III for the analytical calculation of P_{ts} can be transferred to tests for demonstrating system reliability. In this context, a system comprises several failure modes. However, some adjustments are necessary for this purpose. According to the system structure, the failure distribution of the system is composed of the respective failure distributions of the failure modes. For this reason, the variance-covariance matrix is of higher dimension, it corresponds to the sum of the parameters of all failure distributions of the failure modes. If the K failure distributions of the failure modes are all described by a two-parameter Weibull distribution each, then the variance-covariance matrix $\boldsymbol{V}$ has dimension $2K$, i.e. $\boldsymbol{V} \in \mathbb{R}^{2K}$. If the failure modes are independent, the covariances of the parameters of two different failure modes are zero. As a consequence, the entries of the variance-covariance matrix can be computed analogously as for one failure mode. This also corresponds to the test evaluation procedure: the failure times are assigned to the failure modes and then the parameters of the failure distributions of these failure modes are determined separately. Then, the distributions of the failure modes are multiplicatively combined to form the failure distribution of the system according to the system structure. The variance-covariance matrix $\boldsymbol{V}$ is then composed of the respective variance-covariance matrices $\boldsymbol{V_1}$ of the K failure modes as follows:

$$\boldsymbol{V} = \begin{bmatrix} \boldsymbol{V_1} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ & \boldsymbol{V_2} & \boldsymbol{0} & \boldsymbol{0} \\ & & \ddots & \boldsymbol{0} \\ \text{sym} & & & \boldsymbol{V_K} \end{bmatrix} \tag{29}$$

This matrix is valid for operations with the parameter vector P, which contains all parameters of the failure distributions in the corresponding order. If we are dealing exclusively with two-parameter Weibull distributions, this vector is, for example:

$$\boldsymbol{P} = [T_1, b_1, \ldots T_l, b_l, \ldots T_K, b_K] \tag{30}$$

The general form for calculating the variance of the quantile function t_q analogous to Eq. 17 is:

$$\mathrm{Var}\left(t_q\right) = \left[\frac{\partial t_q}{\partial P}\right]' \boldsymbol{V} \left[\frac{\partial t_q}{\partial P}\right] \tag{31}$$

If this equation is solved considering the two-parameter Weibull distributions the following expression is obtained:

$$\mathrm{Var}\left(t_q\right) = \sum\nolimits_{l=1}^{K} \left(\left(\frac{\partial t_q}{\partial T_l}\right)^2 \cdot \mathrm{Var}\left(T_l\right) + 2\frac{\partial t_q}{\partial T_l} \cdot \frac{\partial t_q}{\partial b_l} \mathrm{Cov}\left(T_l, b_l\right) + \left(\frac{\partial t_q}{\partial b_l}\right)^2 \mathrm{Var}\left(b_l\right) \right) \tag{32}$$

The terms of the variances $\mathrm{Var}(T_l)$, $\mathrm{Var}(b_l)$ and covariances $\mathrm{Cov}(T_l, b_l)$ can be calculated with the corresponding values of the parameters of the failure distribution of the failure modes, which are given by prior knowledge. This is equally true for the terms of the partial derivatives. Since the quantile function cannot be expressed as an explicit equation, the derivatives in $\boldsymbol{P}$ must be calculated by implicit differentiation. For this purpose, it is recommended to use the system equation of reliability in logarithmic form. For a strictly serial system with Weibull distributions the following equation can be used:

$$\ln\left(1-q\right) = -\sum_{l=1}^{K} \left(\frac{t_q}{T_l}\right)^{b_l} \tag{33}$$

If all failure distributions are described by two-parameter Weibull distributions, then the partial derivatives for the series system are as follows

$$\frac{\partial t_q}{\partial T_l} = \frac{b_l t_q \cdot \left(\frac{t_q}{T_l}\right)^{b_l}}{T_l \cdot \sum_{p=1}^{K} b_p \cdot \left(\frac{t_q}{T_p}\right)^{b_p}} \tag{34}$$

$$\frac{\partial t_q}{\partial b_l} = \frac{t_q \cdot \ln\left(\frac{t_q}{T_l}\right)\left(\frac{t_q}{T_l}\right)^{b_l}}{\sum_{p=1}^{K} b_{\mathrm{p}} \cdot \left(\frac{t_q}{T_{\mathrm{p}}}\right)^{b_{\mathrm{p}}}} \tag{35}$$

Together with Eq. 32, we obtain the following expression for a series system

$$\mathrm{Var}\left(t_q\right)=\sum_{l=1}^{K}\left(\frac{t_q\cdot\left(\frac{t_q}{T_l}\right)^{b_l}}{\sum_{p=1}^{K}b_\mathrm{p}\cdot\left(\frac{t_q}{T_\mathrm{p}}\right)^{b_\mathrm{p}}}\right)^2\cdot\left(\left(\frac{b_l}{T_l}\right)^2\mathrm{Var}\left(T_l\right)\right.$$
$$\left.-2\frac{b_l}{T_l}\ln\left(\frac{t_q}{T_l}\right)\mathrm{Cov}\left(T_l,b_l\right)+\left(\ln\left(\frac{t_q}{T_l}\right)\right)^2\mathrm{Var}\left(b_l\right)\right) \tag{36}$$

The asymptotic normal distribution of the lifetime quantile of the system t_q can be determined in the same way as in the case of a single dominant failure mode and is as follows:

$$t_q\sim\mathcal{N}\left(t_q;\sqrt{\mathrm{Var}\left(t_q\right)}\right) \tag{37}$$

However, t_q must be determined from the system structure. Since this cannot be solved explicitly for t_q, it can be solved, for example, by numerically approximation methods. However, a simple Newton-Raphson algorithm [29] is sufficient to solve the following expression if a strictly serial system is present:

$$1-q-e^{\sum_{l=1}^{K}\left(\frac{t_q}{T_l}\right)^{b_l}}=0 \tag{38}$$

The distribution of the lifetime quantile with the validity of the alternative hypothesis is thus

$$t_{R,H_1}\sim\mathcal{N}\left(t_\mathrm{p};\sigma_{H_1}\right) \tag{39}$$

with

$$\sigma_{H_1}=\left(\sum_{l=1}^{K}\left(\frac{t_\mathrm{p}\cdot\left(\frac{t_\mathrm{p}}{T_l}\right)^{b_l}}{\sum_{p=1}^{K}b_\mathrm{p}\cdot\left(\frac{t_\mathrm{p}}{T_\mathrm{p}}\right)^{b_\mathrm{p}}}\right)^2\left(\left(\frac{b_l}{T_l}\right)^2\mathrm{Var}\left(T_l\right)\right.\right.$$
$$\left.\left.-2\frac{b_l}{T_l}\ln\left(\frac{t_\mathrm{p}}{T_l}\right)\mathrm{Cov}\left(T_l,b_l\right)+\left(\ln\left(\frac{t_\mathrm{p}}{T_l}\right)\right)^2\mathrm{Var}\left(b_l\right)\right)\right)^{1/2} \tag{40}$$

t_p, for example, is determined by Eq. 33 and $q=1$ - R_req and the necessary failure times of the respective failure modes are determined by Eq. 25. Thus, t_p is the lifetime quantile from the prior knowledge which corresponds to the required reliability. If the entire system is tested, these failure times must be censored for the respective failure modes according to the system structure [21], [24]. Since the failure modes are multiplicatively related to each other to form the system and the null and alternative distributions are also multiplicatively related, the asymptotic distribution of the lifetime quantiles can be described as follows for validity of the null hypothesis:

$$t_{R,H_0}\sim\mathcal{N}\left(t_\mathrm{req};(1-s)\sigma_{H_1}\right)=\mathcal{N}\left(t_\mathrm{req};\frac{t_\mathrm{req}}{t_\mathrm{p}}\sigma_{H_1}\right) \tag{41}$$

By analogy with Eq. 18, this allows to calculate the P_ts for system reliability demonstration as follows:

$$P_\mathrm{ts}=1-\Phi\left(\Phi^{-1}\left(C_\mathrm{req};0,\frac{t_\mathrm{req}}{t_\mathrm{p}}\sigma_{H_1}\right);t_\mathrm{p}-t_\mathrm{req},\sigma_{H_1}\right) \tag{42}$$

If the system is a strictly parallel structure (redundant structure) the calculation procedure is very similar. Due to the changes in the system structure the calculation for system reliability has to be changed as well. The logarithmic quantile of the strictly parallel structure can be calculated as follows:

$$\ln\left(q\right)=-\sum_{l=1}^{K}\ln\left(1-e^{-\left(\frac{t_q}{T_l}\right)^{b_l}}\right) \tag{43}$$

From this, the partial derivatives of the quantile function t_q can be calculated:

$$\frac{\partial t_q}{\partial T_l}=\frac{b_l\cdot\left(\frac{t_q}{T_l}\right)^{b_l}}{T_l\cdot\left(e^{\left(\frac{t_q}{T_l}\right)^{b_l}}-1\right)\sum_{l=1}^{K}\frac{b_l\cdot\left(\frac{t_q}{T_l}\right)^{b_l}}{t_q\left(e^{\left(\frac{t_q}{T_l}\right)^{b_l}}-1\right)}} \tag{44}$$

$$\frac{\partial t_q}{\partial b_l}=\frac{ln\left(\frac{t_q}{T_l}\right)\left(\frac{t_q}{T_l}\right)^{b_l}}{\left(1-e^{\left(\frac{t_q}{T_l}\right)^{b_l}}\right)\sum_{l=1}^{K}\frac{b_l\cdot\left(\frac{t_q}{T_l}\right)^{b_l}}{t_q\left(e^{\left(\frac{t_q}{T_l}\right)^{b_l}}-1\right)}} \tag{45}$$

Using those two equations $\mathrm{Var}(t_q)$ and σ_{H_1} can be calculated accordingly for the parallel system. P_ts is also calculated using Eq. 42. The only other difference is that the life quantile has to be calculated with respect to the parallel system structure as root of the following equation:

$$q-1=\prod_{l=1}^{K}\left(1-e^{-\left(\frac{t_q}{T_l}\right)^{b_l}}\right) \tag{46}$$

For systems which are composed of both series and parallel substructures, the system equation is to be set up according to the system structure in analogy to Eq. 31 and corresponding partial derivatives are to be formed. Thus, only the terms in P (of Eq. 30) are to be replaced by the valid ones, because the parameters of the failure modes are also determined independently.

B. CALCULATION OF P_{ts} FOR SUCCESS RUN TESTS CONSIDERING THE SYSTEM STRUCTURE

If the SR tests are performed with the entire system, the equations for calculation are identical to those used with only one prevailing failure mode. Thus, Eq. 19 to 23 can be used for the calculation. Only the value of reliability from prior knowledge R_p must be calculated according to the system structure from prior knowledge of the failure distributions of the failure modes. For a strictly serial system with Weibull distributed failure modes R_p is calculated via:

$$R_\mathrm{p} = e^{\sum_{l=1}^{K}\left(\frac{t_\mathrm{req}}{T_l}\right)^{b_l}} \tag{47}$$

However, if subsystems or components are tested in separate SR tests, the calculation of the P_ts must be adapted. The system structure, which determines the logic connection of the respective failure modes of the tested subsystems or components to the overall system, is irrelevant for the calculation of the P_ts, because reliability demonstration at system level can only be provided if all SR tests are successful. If the number of allowed failures is already exceeded in one of the SR tests, system reliability demonstration cannot be provided. According to the product law of probabilities, this relationship corresponds to a multiplication of all survival probabilities of the individual SR tests. Since the P_ts describes exactly these probabilities in SR tests, Eq. 19 to 23 can be used to calculate the probabilities $P_{ts,1}, \ldots, P_{ts,l}, \ldots, P_{ts,K}$ of the K SR tests, i.e.

$$P_\mathrm{ts} = \prod_{l=1}^{K} P_{ts,l} \tag{48}$$

However, the system structure is required for determining the necessary sample sizes of the individual SR tests. These must be determined in such a way that the resulting reliability distribution at system level can meet the reliability target. For this purpose, the beta distributions of the individual SR tests according to the sample size can be linked, for example, by the method of moments for the system reliability distribution and the fulfillment of the reliability target can be checked [21], [22], [23]. It is not necessary to divide the reliability target of the system among the subsystems or components of the system (reliability partitioning), because only the fulfillment of the reliability target of the system is relevant, a fixed partitioning of the reliability requirement for demonstration purposes only limits the solution space of test configurations, see also Chapter VII.

VI. CONSIDERING UNCERTAINTY IN PRIOR KNOWLEDGE FOR P_{ts} CALCULATION

The information about failure times is usually coming from one or more observations. Due to the fact, that the amount of observations is always limited, the information is subject to epistemic uncertainty. Thus, the prior knowledge about the reliability and the failure distribution in particular are also subject to uncertainty. In order to be able to consider these in the considerations of the reliability demonstration planning the following methods and procedures are presented and a corresponding calculation of the P_ts is introduced.

A. EFFECT OF UNCERTAINTY IN PRIOR KNOWLEDGE

The reliability is the aleatoric uncertainty of the products lifetime. The lifetime is varying from product to product and is therefore described by a probability. The estimation of the actual underlying probability of the lifetime can only be made inadequately, i.e., not exactly due to the sampling error. This inadequacy of observation is called epistemic uncertainty.

Prior knowledge in the context of reliability corresponds to information about the reliability itself (type SR test: reliability distribution) or information about the failure distribution (type EoL test: failure distribution or entire sample). Thus, it is the aleatory uncertainty - the uncertainty about the failure times, or reliability. The uncertainty about the correct determination of reliability (epistemic uncertainty) must be additionally considered in the test planning process, because prior knowledge is always subject to epistemic uncertainty.

The two types of prior knowledge in certain (fixes values) and uncertain form are shown in Fig. 4. Thus, the epistemic uncertainty in prior knowledge can be determined by specifying the reliability distribution, for example, by a beta distribution for a SR Test. The uncertainty in prior knowledge within an EoL test can be determined, e.g., by the original sample size of the failure distribution [21], [22].

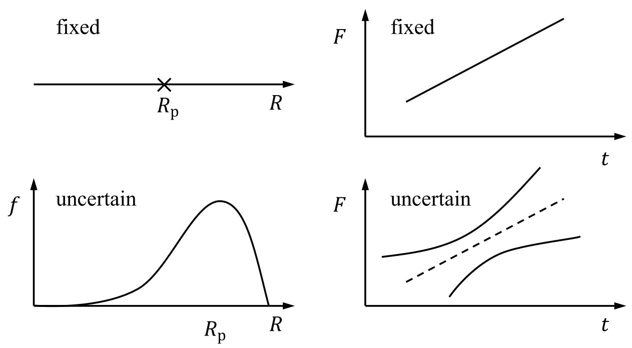

FIGURE 4. Types of prior knowledge in its certain (fixed values) and uncertain form for type SR test (left) and type EoL test (right).

B. CALCULATION OF P_{ts} WITH UNCERTAINTY FOR EOL TESTS

Prerequisite for the calculation of P_ts in EoL Tests is a prior knowledge of the type EoL Test. That is a failure distribution with specification of the original sample size or the specification of the original sample in the form of failure and suspension times itself. The procedure is formulated for the two-parameter Weibull distribution. However, the procedure can also be applied to other distributions, as long as the appropriate likelihood function, the variance-covariance matrix, the quantile function and the respective partial derivatives for calculating the asymptotic variance of the quantile exist, see Chapter III. Because those variances are also required here.

The main idea in the analytical-approximative approach is to use the asymptotic distribution of the lifetime quantile in order to calculate the P_ts in an analytical way [24]. Here, the asymptotic standard deviations σ_{H_0} and σ_{H_1} are determined under the validity of the two hypotheses using the Central Limit Theorem and a Taylor series approximation. The distributions of the lifetime quantiles are then obtained as normal distributions.

The scale parameters μ_{H_0} and μ_{H_1} of these distributions are predetermined via prior knowledge, because they are specified by the defined hypotheses and are independent of the Central Limit Theorem considerations and the sample size. However, because of the uncertainty in prior knowledge, it is not possible to further specify a single value for the parameters of the failure distribution, which means that the scale of the lifetime quantile is also subjected to variance. However, with the same approach, the variance can also be determined as a normal distribution via the Central Limit Theorem and Taylor series approximation. For this purpose, only the sample size n_0 on which the prior knowledge was based, has to be used. Thus, the normal distribution of the scale of the lifetime quantile under validity of the alternative hypothesis μ_{H_1} results as:

$$\mu_{H_1} \sim \mathcal{N}\left(T_{H_1}\left(-\ln\left(R_{\text{req}}\right)\right)^{1/b}, \sigma_{H_1,n_0}\right) \tag{49}$$

However, it is important to note that the variance in this equation is determined by means of the synthetic failure times corresponding to n_0, instead of n. The relationships between the likelihood, its derivatives and the variances and covariances are still valid, see [24].

The scale parameter of the asymptotic distribution of the scale of the lifetime quantile is identical with the scale of the lifetime quantile in case no uncertainty would be considered. Using the transformation of Eq. 15, the normal distribution of the scale of the lifetime quantile for the validity of H_0 results in:

$$\begin{aligned}\mu_{H_0} &\sim \mathcal{N}\left(t_{\text{req}}, (1-s)\cdot\sigma_{H_1,n_0}\right)\\ &= \mathcal{N}\left(t_{\text{req}}, \frac{t_{\text{req}}}{T_{H_1}}\left(-\ln\left(R_{\text{req}}\right)\right)^{-1/b}\cdot\sigma_{H_1,n_0}\right)\end{aligned} \tag{50}$$

with:

$$\begin{aligned}\sigma_{H_0,n_0} &= (1-s)\cdot\sigma_{H_1,n_0}\\ &= \frac{t_{\text{req}}}{T_{H_1}}\left(-\ln\left(R_{\text{req}}\right)\right)^{-1/b}\cdot\sigma_{H_1,n_0}\end{aligned} \tag{51}$$

In such a case, when the scale parameter of a normal distribution is again normally distributed, it is also called a mixture distribution and the resulting distribution is a normal distribution again [33]. The scale parameter of the resulting distribution corresponds to that of the distribution which describes the scattering scale parameter. The variances are summed up, which means that the resulting standard deviation σ_Σ is the geometric sum of the standard deviations of the two distributions. Thus, the asymptotic normal distribution of the lifetime quantile under the validity of H_1 with consideration of the uncertainty of the prior knowledge is determined as follows:

$$\begin{aligned}t_{R,H_1} &\sim \mathcal{N}\left(t_{\text{p}}, \sigma_{\Sigma,H_1}\right)\\ &= \mathcal{N}\left(T_{H_1}\left(-\ln\left(R_{\text{req}}\right)\right)^{1/b}, \sqrt{\sigma_{H_1,n_0}^2+\sigma_{H_1}^2}\right)\end{aligned} \tag{52}$$

Accordingly, the following distribution applies analogously to the validity of the null hypothesis:

$$t_{R,H_0} \sim \mathcal{N}\left(t_{\text{req}}, \sigma_{\Sigma,H_0}\right) \tag{53}$$

with:

$$\begin{aligned}\sigma_{\Sigma,H_0} &= \sqrt{\sigma_{H_0,n_0}^2+\sigma_{H_0}^2}\\ &= \frac{t_{\text{req}}}{T_{H_1}}\left(-\ln\left(R_{\text{req}}\right)\right)^{-1/b}\cdot\sqrt{\sigma_{H_1,n_0}^2+\sigma_{H_1}^2}\\ &= (1-s)\cdot\sigma_{\Sigma,H_1}\end{aligned} \tag{54}$$

Finally, P_{ts} resulting in:

$$\begin{aligned}P_{\text{ts}} = 1-\Phi\Bigg(&\Phi^{-1}\left(C_{\text{req}}; 0, \frac{t_{\text{req}}}{T_{H_1}}\left(-\ln\left(R_{\text{req}}\right)\right)^{-1/b}, \sigma_{\Sigma,H_1}\right);\\ &T_{H_1}\left(-\ln\left(R_{\text{req}}\right)\right)^{1/b}-t_{\text{req}}, \sigma_{\Sigma,H_1}\Bigg)\end{aligned} \tag{55}$$

Based on the equations presented, it can be seen that the P_{ts} of EoL tests while considering the uncertainty in prior knowledge can reach at most the value that the sample size of the prior knowledge is reaching in a calculation without uncertainty. This is due to the summation of the variances. The variance of the lifetime quantile can therefore never be smaller than the variance that results from the sample size of the prior knowledge alone. For practical purposes, this means, that the values of P_{ts} will always be smaller if the uncertainty in the prior knowledge is considered. Furthermore, a sample size larger than that of the prior knowledge $n > n_0$ would not lead to an increase in the P_{ts} beyond the corresponding value of n_0.

C. CALCULATION OF P_{ts} WITH UNCERTAINTY FOR SUCCESS RUN TESTS

If the uncertainty in the prior knowledge is to be considered for SR Tests, the prior knowledge must first be translated into a suitable form if necessary. If the prior knowledge is available in the form of a beta distribution, it can be used directly.

The approach to calculate the P_{ts} of Eq. 21 using the binomial distribution cannot be further used here, because the parameter of the success probability, which in this context is the probability of failure, is not further a single, fixed value, but scatters according to the beta distribution of prior knowledge. However, this corresponds exactly to the information given by the beta-binomial distribution. Accordingly, the P_{ts} with prior knowledge can be calculated using a beta distribution as follows to account for the uncertainty in the prior knowledge:

$$P_{\text{ts}} = \sum_{i=0}^{k}\binom{n}{i}\cdot\frac{\beta\left(B+i, A+n-i\right)}{\beta\left(B, A\right)} \tag{56}$$

The beta distribution with the parameters A and B is describing the reliability distribution for the required service life. In the planned SR Test with sample size n a maximum of k failures are allowed.

VII. STUDY OF THE INFLUENCES ON THE P_{ts}

With the calculation methods developed, the influences on the P_{ts} can now be studied. Initially, only the individual main effects will be discussed.

A. INFLUENCE OF CONSIDERING PRIOR KNOWLEDGE WITH BAYES' THEOREM

Fig. 5 shows the results for a system with one failure mode as weibull distributed with the parameters $T = 1$ and $b = 3$ ($W(1; 3)$) and the reliability target $R_{\text{req}} = 92$ %, $C_{\text{req}} = 90$ % with $t_{\text{req}} = 0.28$, the change in P_{ts} when varying the sample sizes of prior knowledge n_0 and of an uncensored EoL test n. A symmetry with respect to the n and n_0 axis can be seen, which implies that the two sample sizes have the same effect on the P_{ts}. That is, the increase in P_{ts} by increasing the sample size of test n, will likewise be achieved by an equally increased sample size n_0 of prior knowledge. Thus, the sample of prior knowledge is equal to the sample of the test. This confirms what was stated in Chapter IV and the findings from Eq. 27 and 28 of the applied MAP.

In Fig. 6, the lifetime requirement t_{req} was modulated using the safety distance s. Two cases of the sample size of prior knowledge and the test were calculated and, in addition, the P_{ts} is shown without taking prior knowledge into account. It can be seen that the additional consideration results in a higher P_{ts} here.

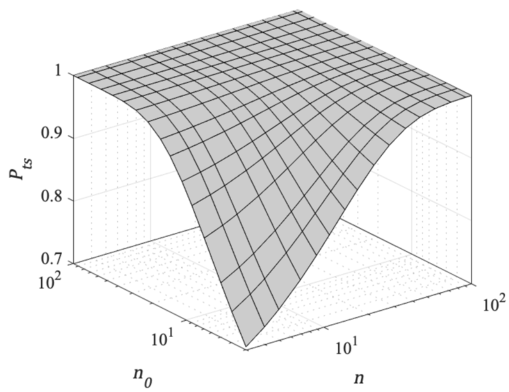

FIGURE 5. P_{ts} **of an uncensored EoL test with additional consideration of prior knowledge** $W(1; 3)$ **with Bayes' theorem.**

For an SR test, the additional consideration of prior knowledge directly reduces the required sample size and the P_{ts} is only dependent on the failure distribution and the sample size, see Eq. 21. The reliability requirement is involved by the sample size required by this, but the actual dependence is with respect to the sample size itself. Fig. 7 shows the possible prior knowledge for an example with the reliability target $R_{\text{req}} = 92$ %, $C_{\text{req}} = 90$ % with $t_{\text{req}} = 0.28$. The prior knowledge here is a beta distribution $\mathcal{B}(A_0; B_0)$. According to the basic idea, it can be seen that P_{ts} increases sharply as A_0 of the prior knowledge beta distribution increases. Similarly, the P_{ts} decreases as B_0 increases. This is mainly due to the associated change in the required sample size,

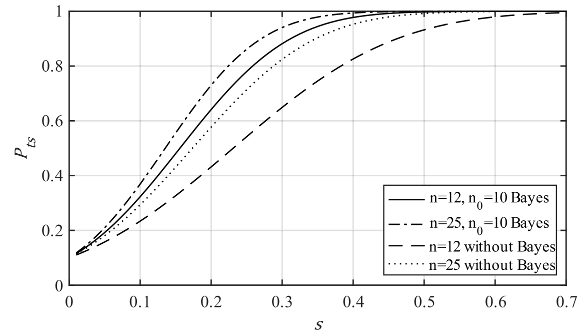

FIGURE 6. P_{ts} **of an uncensored EoL test as a function of the product safety distance** s **with additional consideration of prior knowledge** $W(1; 3)$ **using Bayes' theorem.**

because by increasing the "good parts" of prior knowledge with the parameter A_0, fewer specimens need to survive the test time, thus increasing the probability of passing the test – P_{ts} increases. The opposite happens when the parameter B_0 of prior knowledge increases, because the "failures" of prior knowledge have to be compensated by additional survivors in the test, the probability that all these additional specimens actually survive is decreasing – P_{ts} decreases. It has to be noted here, that prior knowledge was considered in the form of $W(1; 3)$ for calculation of P_{ts}. The additional consideration of prior knowledge in the form of the beta distribution with parameters A_0 and B_0 using Bayes' theorem, does alter the required sample size and thus the P_{ts} changes.

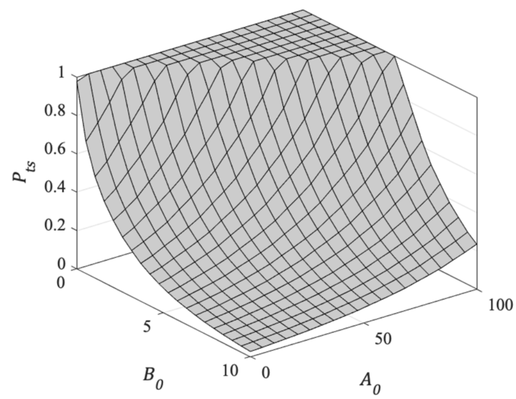

FIGURE 7. P_{ts} **of a Success Run test as a function of the prior knowledge** A_0 **and** B_0 **with additional consideration of prior knowledge** $W(1; 3)$ **with Bayes' theorem.**

B. INFLUENCE OF THE SYSTEM STRUCTURE

The system structure must be considered in the system reliability demonstration anyway - except in the case of a sole SR test of the whole system level. This is because the reliability information obtained for the system must be evaluated separately according to failure modes.

In the following, the influences on the system reliability demonstration are examined.

The test configuration has the greatest influence on the P_{ts} when considering a fixed reliability target, fixed system structure, and failure distributions of the failure modes. Unlike the usual practice of dividing the system requirements among individual components [12], [13], by examining the P_{ts}, it is possible to choose sample sizes for subsystems, components, or parts that maximize the probability of successfully demonstrating the system reliability. In this case, a partitioning of the requirement is not necessary as the focus solely lies on the reliability target of the overall system. Since theoretically infinite system structures are possible, the effects of the apportionment of specimen on the P_{ts} of the two basic forms, namely the series system and the parallel system, will be presented below.

In a strictly serial system, the system failure is predominantly influenced by the failure mode with the lowest reliability at t_{req} due to the multiplicative combination of the reliability of the failure modes. In contrast, in strictly parallel systems, the system failure is mainly influenced by the failure mode with the highest reliability due to the multiplicative combination of the failure probabilities of the failure modes, as the system only fails when the last failure occurs.

To identify the allocation of sample sizes to different tests for maximum P_{ts}, the P_{ts} can be calculated as a function of the sample sizes. The maximum can be found using a suitable optimization algorithm, for example, by setting the maximum possible number of test specimens overall or the resulting total costs as an upper limit. Alternatively, for simple systems, the P_{ts} can be plotted against the allocation of test specimens. This is exemplified in Fig. 8 for a series system and a parallel system for EoL tests at the component level with a total sample size n_{tot} of 50. In this example, both systems consist of only two failure modes, which are directly addressed in the component tests. The sample sizes of these failure modes are n_1 and n_2, and the failure modes in both systems are determined by the Weibull distributions $W_1(0.6; 3)$ and $W_2(1; 3)$. The reliability target is $R_{req} = 92$ % with $C_{req} = 90$ %, and the life cycle requirements are chosen as $t_{req,Series} = 0.175$ for the series system and $t_{req,Parallel} = 0.45$ for the parallel system to achieve similar values of maximum P_{ts} in this example. The considerations regarding the influence on system failure in series and parallel systems are reflected as follows: To achieve maximum P_{ts}, in the series system, the failure mode with lower reliability – and thus more critical for the system – needs to be tested more, i.e., $n_1 > n_2$, and for the optimum, $n_1 = 38$ and $n_2 = 12$. Similarly, in the parallel system, the more critical failure mode in the system, namely failure mode 2, needs to be tested more. The optimum is achieved with $n_2 > n_1$, specifically with sample sizes of $n_1 = 15$ and $n_2 = 35$. Therefore, when conducting EoL tests at the component level, a configuration that guarantees maximum P_{ts} will also provide a better estimation for the more critical failure mode in the system.

Since the P_{ts} is calculated using Eq. 23 for both series and parallel systems in SR tests, the qualitative behavior of the P_{ts} with respect to the apportionment of sample sizes is identical. This is exemplified in Fig. 9.

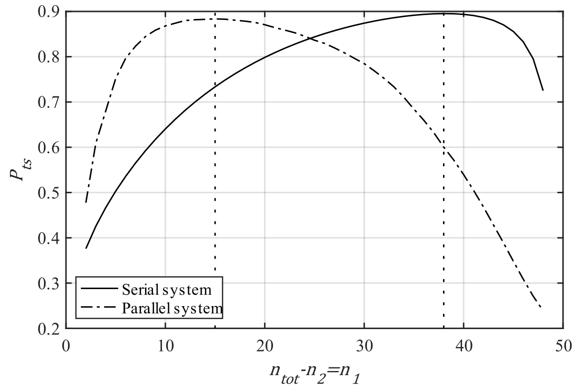

FIGURE 8. P_{ts} **of EoL tests of components from a series and parallel system with two failure modes each and a total sample size of 50.**

The sample sizes were chosen such that the system reliability target can be met through the SR tests. The systems considered are the same as those in the EoL scenario with the same reliability requirements. Due to the significant differences in the logical relationships between the series and parallel systems, the SR tests for the series system require significantly larger sample sizes, namely $n_{tot} = 94$, compared to the parallel system with $n_{tot} = 10$. This is primarily due to the fact that the sample sizes for the SR tests are solely determined by the required reliability and the desired level of confidence when no lifetime ratio is used.

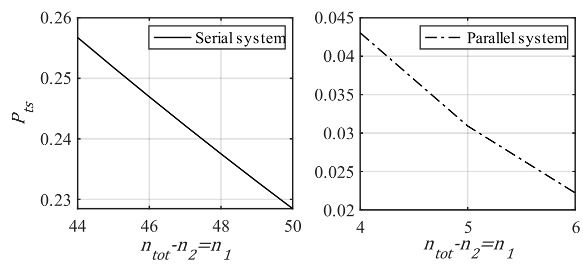

FIGURE 9. P_{ts} **of Success Run tests of components from a series and parallel system with two failure modes each.**

In SR tests, only very small deviations from the equal apportionment of specimen are possible, while still satisfying the system reliability requirement. For the parallel system, the variation in the sample sizes in the component tests is limited to only 4 to 6 whereas for the series system, it can vary between 44 and 50 test objects per component test, while maintaining the same total sample size n_{tot}.

Unlike EoL component tests, which yield maximum P_{ts} when the most critical failure mode incurs the highest test effort (sample size), SR component tests yield maximum P_{ts} when the failure mode with the lowest reliability incurs the highest test effort. This may not be desirable in the case of parallel systems, as it does not allow for simultaneously achieving maximum P_{ts} and accurate identification of the most critical failure mode. This is always possible in EoL testing however.

In addition to testing individual components, the entire system can also be tested. Fig. 10 and 11 show the values of P_{ts} for EoL component-level tests and system-level tests plotted against the total sample size n_{tot}, for two additional example systems.

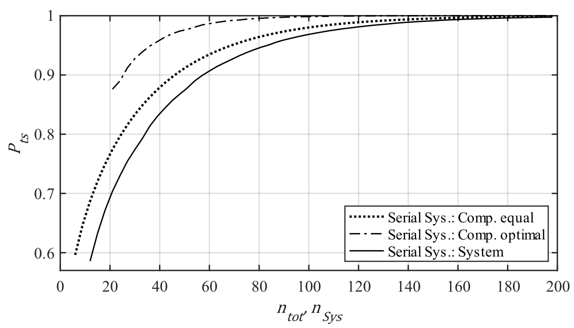

FIGURE 10. P_{ts} **comparison as a function of the sample size for EoL test of the series system.**

These systems consist of a strictly serial system and a strictly parallel system, both including the same failure modes: $W(1.5; 1.2)$, $W(2; 2)$, and $W(1; 3)$. The reliability requirement is $R_{\mathrm{req}} = 92$ % and $C_{\mathrm{req}} = 90$ %, with the required lifetimes of $t_{\mathrm{req}} = 0.05$ for the series system and $t_{\mathrm{req}} = 0.9$ for the parallel system. Both the equal apportionment of specimen among the three components and the optimal apportionment are shown alongside the P_{ts} for system-level tests.

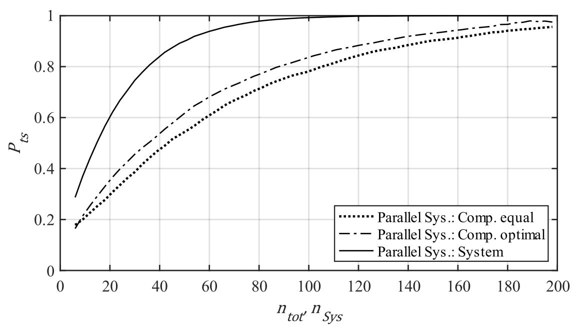

FIGURE 11. P_{ts} **comparison as a function of the sample size for EoL tests of the parallel system.**

It can be observed that the optimal apportionment can significantly increase the P_{ts}. The advantage of the system-level test is that a single specimen can be used to evaluate all failure modes, whereas component-level tests can only assess the dominant failure modes in the tested component. In contrast, the specimens used in the system-level test to evaluate the failure modes are censored according to the system structure. In component-level tests, the number of failures for evaluating a specific failure mode can be precisely controlled, which can result in the P_{ts} of a series system's component-level tests exceeding those of the system-level test, despite the fact that the test objects in the system-level test can be essentially used three times, as shown in Fig. 10. For the parallel system, the system-level test consistently yields significantly higher P_{ts} values. This is because system failure occurs only when all components have failed. Therefore, a test object from the system-level test always provides information about the failure times of all failure modes (in this example of the strictly parallel system). This effect is higher the more failure modes are in the system.

Since the sample sizes in the SR test are determined by the reliability requirement and in this example, no lifetime ratios or failures are allowed, the values of P_{ts} cannot be plotted against the sample size. Instead, the values are presented as a bar chart for the three cases: overall system test, component test with equal specimen allocation, and optimal allocation, as shown in Fig. 12 for the strictly serial system and strictly parallel system.

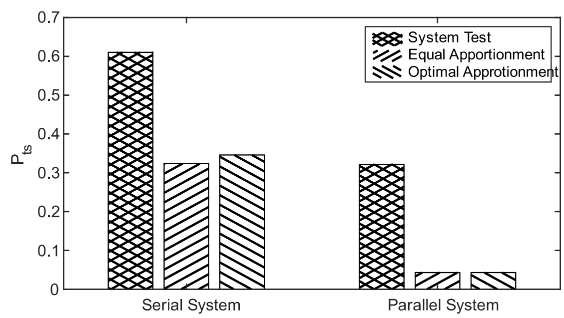

FIGURE 12. P_{ts} **of the SR tests.**

The three cases require significantly different sample sizes. The sample size required for the parallel system does not allow for any apportionment other than the equal one, as only three specimens per component are necessary to achieve the requirement. Changing to two specimens in a component test would yield such poor results that no matter how many specimens are tested in another component, it would not compensate for in the resulting system reliability. Both systems generally show small values of P_{ts}, with the component tests yielding even lower values than the system tests. This is due to the significantly higher sample sizes, with each specimen needing to survive for a successful demonstration. However, none of the SR tests can be recommended for system demonstration in these example systems, as all of them yield maximum P_{ts} values $< 50\%$. Additionally, the component tests of the series system require a total of $n_{tot} = 192$ specimens, whereas the component tests of the parallel system only require $n_{tot} = 7$ specimens in total. Since no equal apportionment is possible with this sample size, only the optimal is calculated, which allocates 3 specimen to second component and 2 each for the first and third component. The system SR test does require $n_{\mathrm{Sys}} = 28$ specimen.

C. INFLUENCE OF UNCERTAINTY IN PRIOR KNOWLEDGE

The calculation of P_{ts} taking into account the uncertainty in prior knowledge holds potential, as it indirectly assesses the quality of prior knowledge. Generally, P_{ts} decreases when uncertainty is considered in EoL tests, as the lifetime quantiles exhibit larger variations. This is illustrated using a simple system with a failure mode of $W(1; 3)$ and $n_0 = 12$ and $n_0 = 35$ in Fig. 13. A significant reduction occurs compared to the calculation without considering uncertainty. The values of P_{ts}, taking into account uncertainty in prior knowledge, tend to approach a maximum.

This maximum can be observed in Fig. 13, where it precisely matches the P_{ts} value calculated without considering prior knowledge for a sample size of $n = 12$ and $n = 35$, respectively. The sample size of the prior knowledge limits the maximum achievable P_{ts}. This is quite evident for simple systems with only a single failure mode, but becomes less apparent for more complex systems. In practical terms, this means that uncertainty should be taken into account. However, this calculation should always be accompanied by a calculation without considering uncertainty, as it represents the expected P_{ts}.

In contrast to the EoL tests, the SR test does not exhibit a consistent trend of decreasing or increasing P_{ts} when considering uncertainty in prior knowledge. Fig. 14 demonstrates that for high P_{ts} values, there is a decrease, while for low P_{ts} values, there is an increase. This phenomenon can be attributed to the increased variability of reliability, which follows a beta distribution in the context of prior knowledge.

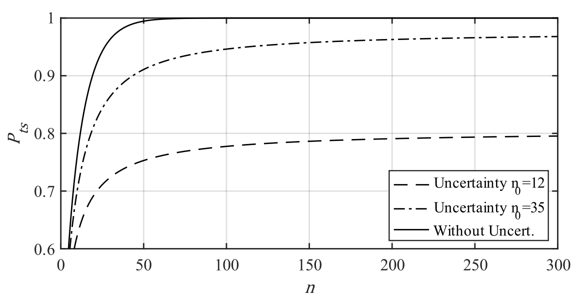

FIGURE 13. P_{ts} **as a function of the sample size for an EoL test considering uncertainty in prior knowledge.**

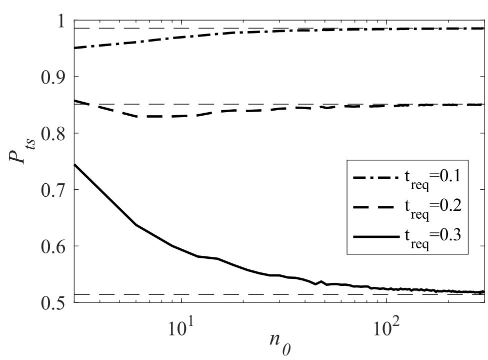

FIGURE 14. P_{ts} **as a function of the prior knowledge sample size for a Success Run test considering uncertainty in prior knowledge.**

When the reliabilities are already large values, the increased variance leads to a higher occurrence of smaller reliabilities compared to larger ones, relative to the reference value used to calculate P_{ts} without uncertainty. This is because when reliability values approach the maximum value of 1, the probability of encountering even larger values diminishes significantly. Conversely, when uncertainty is taken into account, small P_{ts} values exhibit an upward trend. This can be attributed to the decreased probability of obtaining even smaller reliability values as they approach the minimum value of 0. The distinctive behavior of P_{ts} when considering uncertainty in prior knowledge between the EoL test and the SR test arises from the inherent domain constraint of reliability itself. Reliability is confined to the interval $R \in [0, 1]$, while the test statistic τ, which is relevant in the calculation of P_{ts} for EoL tests, has no such restriction. The only restriction in this case is the requirement of strictly positive lifetimes. However, this does not significantly limit the behavior of P_{ts}.

VIII. INFLUENCE OF IMPORTANT INTERACTIONS

In the second stage of the analysis, critical interactions when using more than one of the presented approaches will be presented.

A. INFLUENCE OF THE SYSTEM STRUCTURE WHEN CONSIDERING PRIOR KNOWLEDGE USING BAYES' THEOREM

The additional consideration of prior knowledge using Bayes' theorem also leads to higher P_{ts} values in both EoL and Success Run tests, as well as in system and component tests. The apportionment of specimen during component testing, as discussed in Chapter VII, is dependent on the incorporation of prior knowledge through Bayes' theorem. This corresponds to the idea that the samples of prior knowledge and those of the test are of equal importance. Therefore, when prior knowledge is taken into account, it replaces a portion of the sample sizes required for optimization. As a result, the required sample sizes for maximizing P_{ts} change accordingly.

This effect also implies that the optimal apportionment is no longer independent of the sample size of the test when prior knowledge is considered. The proportion of prior knowledge samples replacing the test samples changes with the sample size of the test. To illustrate this fact, Fig. 15 depicts the optimal apportionment of specimen for a series system with two failure modes of identical distribution, $W(1; 3)$, based on the total sample size of the uncensored EoL test. It can be observed that for failure mode 1, a larger sample size must initially be assigned due to the smaller sample size of the prior knowledge in order to achieve the maximum P_{ts}. Since both failure modes are identical, the optimal apportionment without additional consideration using Bayes' theorem would be an equal apportionment. This can also be observed for very small and very large total sample sizes.

B. INFLUENCE OF UNCERTAINTY WITH ADDITIONAL CONSIDERATION OF PRIOR KNOWLEDGE USING BAYES' THEOREM

As discussed in Chapter IV, the additional consideration of prior knowledge using Bayes' theorem results in the ability to replace test units if the prior knowledge is positive enough.

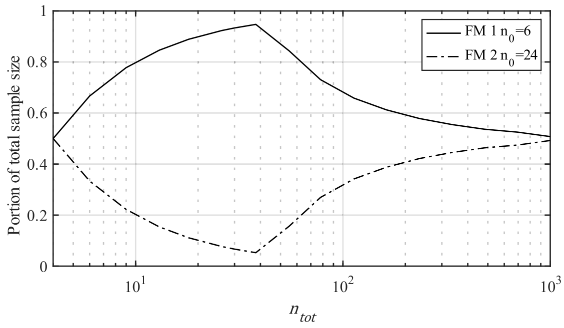

FIGURE 15. Change of the optimal sample allocation of an EoL component test of a series system with two failure modes $W(1;3)$.

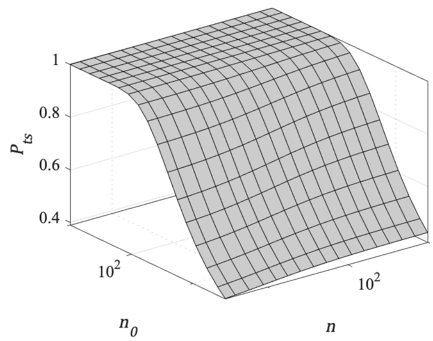

FIGURE 16. P_{ts} as a function of n and n_0 considering uncertainty and prior knowledge with Bayes' Theorem in an EoL test.

Thus, the additionally considered prior knowledge is equivalent to the information obtained from the test, as illustrated in Chapter VII. Furthermore, it was shown that incorporating the uncertainty in prior knowledge in EoL tests leads to a decrease in P_{ts}, which is bounded by the sample size n_0. Therefore, the maximum achievable P_{ts} is the one that would theoretically result without considering uncertainty at $n = n_0$.

The combined consideration of uncertainty in prior knowledge and prior knowledge using Bayes' theorem has the consequence that the maximum P_{ts} is determined by n_0 (the effect of uncertainty) and that this maximum is always achieved because n and n_0 are equivalent (the effect of Bayes' theorem). This implies that the sample size of prior knowledge n_0, solely determines the value of P_{ts}. For $n > n_0$, the effect of uncertainty limits P_{ts} to the maximum corresponding to n_0. On the other hand, for $n < n_0$, the effect of additional consideration using Bayes' theorem ensures that the maximum P_{ts} is still achieved because n and n_0 are equivalent. In this case, the effective sample size is always large enough to reach the maximum value, at least n_0. Therefore, P_{ts} is practically not further dependent on the sample size of the test n.

To illustrate this, Fig. 16 exemplarily calculates the P_{ts} of the uncensored EoL test as a function of n and n_0, considering the prior knowledge with $W(1;3)$ using Bayes' theorem for the demonstration, as well as the uncertainty in prior knowledge. It is clearly observed that P_{ts} does not change with n.

When prior knowledge is used in the reliability assessment through Bayes' theorem, it means that a higher reliability, longer lifetime, or a higher level of confidence can be demonstrated compared to not considering prior knowledge. Thus, there is always a gain in any case. However, the uncertainty in prior knowledge becomes irrelevant once the test has been conducted and the evidence has been provided.

For practical application in test planning with P_{ts}, it is important to consider the uncertainty but also calculate P_{ts} without considering uncertainty. When additionally considering prior knowledge using Bayes' theorem, the value of P_{ts} obtained without considering uncertainty should be used, while the one with uncertainty can be considered as a lower bound or worst-case scenario for P_{ts}.

In the case of SR tests, the relationships are somewhat simpler. The test configuration is already determined by the reliability requirement. Therefore, the calculation of P_{ts} is initially independent of the reliability demonstration. As discussed in Chapter VII, considering uncertainty can result in both a decrease and an increase in P_{ts}, while incorporating prior knowledge using the Bayes' theorem leads to a smaller required sample size, resulting in an increase in P_{ts}.

The increases and decreases in P_{ts} are still present but are overshadowed by the effects of Bayes' theorem. To illustrate this, the same example from Fig. 14, with the additional consideration of prior knowledge using the Bayes theorem for $t_{req} = 0.35$ ($s = 5.8$ %), is depicted in Fig. 17. It can be observed that when uncertainty and the Bayes' theorem are taken into account, P_{ts} approaches the values obtained when considering uncertainty alone, which in turn approximate the case without consideration of uncertainty or the Bayes' theorem.

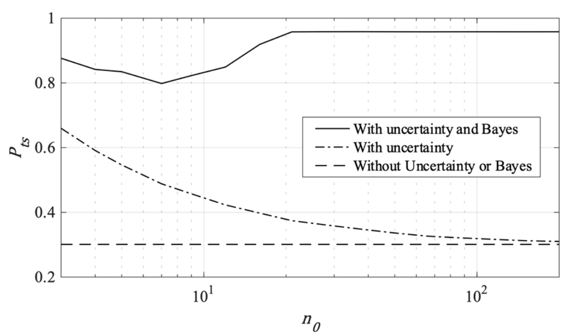

FIGURE 17. P_{ts} as a function of n_0 considering uncertainty and prior knowledge with Bayes' Theorem in a Success Run test.

However, incorporating prior knowledge using Bayes' theorem allows for a significant reduction in the required sample size, leading to an additional increase in P_{ts}. Starting from a sample size of prior knowledge $n_0 = 70$, P_{ts} further increases because a large portion of the demonstration is already achieved through the additional consideration of prior knowledge. Theoretically, with $n_0 = 200$, no SR test is necessary as the demonstration is already fully accomplished based on prior knowledge alone. However, this is irrelevant for practical applications since a demonstration should never rely solely on prior knowledge. For this reason, a required sample size approaching zero is not allowed for here. This is why the P_{ts} is not reaching 100 %, even for very large sample sizes.

IX. SUMMARY AND CONCLUSION

The development of technical products must ensure that their functionality is guaranteed over the desired service life. This means that the reliability requirement imposed on them must be fulfilled. Reliability demonstration tests are conducted to verify the stated requirements. The challenges involved are diverse: on the one hand, the demonstration must ensure that it has the necessary statistical power to even provide the evidence, i.e., it must be capable of achieving the desired result, and on the other hand, it must be feasible with the available resources.

The Probability of Test Success is capable of evaluating the suitability of different test configurations and types in terms of their probability of achieving a successful reliability demonstration. This enables the selection of the most suitable test. The present work utilizes the Probability of Test Success and develops methods to evaluate tests of real systems with multiple subsystems, components, parts, and failure modes. Additionally, the uncertainty of the required prior knowledge is taken into account, as well as the additional consideration of prior knowledge using the Bayes theorem. The analytical calculation method allows leveraging the asymptotic properties of quantile estimation methods through the variance-covariance matrix and Taylor series approximation, enabling a simple and fast analytical calculation based on the normal distribution. The additional consideration of prior knowledge using Bayes' theorem results in an extension of the sample size similar to the maximum a posteriori estimation for failure-based tests. To account for the reduced variance of the lifetime quantiles due to the fixed prior knowledge, the variances are reduced using a correction factor calculated with the Bayes' theorem. Complex systems with multiple failure modes and system levels require a specific approach in the reliability demonstration, namely, the evaluation of test data must be performed separately for each failure mode in failure-based tests. As a result, the calculation of the Probability of Test Success reflects this requirement. To incorporate this in the evaluation using the Probability of Test Success, the calculation procedure was extended to include the variance of the epistemic uncertainty using a mixture distribution.

The evaluation of main effects and interactions has shown that there are clearly favored test strategies in several scenarios. The interaction analysis has ultimately revealed that dealing with prior knowledge is of particular importance.

The consideration of the Boolean system structure of the product does allow for an optimal apportionment of specimen onto the respective component tests in order to demonstrate system reliability in a most efficient way.

REFERENCES

[1] M. Dazer, D. Brautigam, T. Leopold, and B. Bertsche, "Optimal planning of reliability life tests considering prior knowledge," in *Proc. Annu. Rel. Maintainability Symp. (RAMS)*, Jan. 2018, pp. 1–7, doi: 10.1109/RAM.2018.8463098.

[2] M. Dazer, "Zuverlässigkeitstestplanung mit Berücksichtigung von Vorwissen aus stochastischen Lebensdauerberechnungen," Ph.D. dissertation, Fac. 7—Eng. Des., Prod. Eng. Automot. Eng., Univ. Stuttgart, Stuttgart, Germany, 2019. [Online]. Available: https://elib.uni-stuttgart.de/bitstream/11682/10518/3/190_Dissertation_Martin_Dazer.pdf

[3] M. Dazer, T. Herzig, A. Grundler, and B. Bertsche, "R-OPTIMA: Optimal planning of reliability tets," in *Proc. 7th Int. Conf. Integrity-Rel.-Failure*, 2020, pp. 695–702.

[4] A. Grundler, M. Dazer, and T. Herzig, "Statistical power analysis in reliability demonstration testing: The probability of test success," *Appl. Sci.*, vol. 12, no. 12, p. 6190, Jun. 2022, doi: 10.3390/app12126190.

[5] B. Bertsche and M. Dazer, *Zuverlässigkeit im Fahrzeug-und Maschinenbau*, 4th ed. Berlin, Germany: Springer, 2022.

[6] A. Albers, C. Mandel, S. Yan, and M. Behrendt, "System of systems approach for the description and characterization of validation environments," in *Proc. 15th Int. Design Conf.*, 2018, pp. 2799–2810.

[7] J. Neyman and E. S. Pearson, "IX. On the problem of the most efficient tests of statistical hypotheses," *Philos. Trans. Roy. Soc. A, Math., Phys. Eng. Sci.*, vol. 231, pp. 289–337, Feb. 1933.

[8] A. Grundler, M. Dazer, T. Herzig, and B. Bertsche, "Efficient system reliability demonstration tests using the probability of test success," in *Proc. 31st Eur. Saf. Rel. Conf. (ESREL)*, 2021, pp. 1654–1661.

[9] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*. New York, NY, USA: Lawrence Erlbaum Associates, 1988.

[10] J. O. Berger, *Statistical Decision Theory and Bayesian Analysis*. New York, NY, USA: Springer, 1985.

[11] A. Benard and E. C. Bos-Levenbach, "Het uitzetten van waarnemingen op waarschijnlijkheids-papier1," *Statistica Neerlandica*, vol. 7, no. 3, pp. 163–173, Sep. 1953.

[12] X. Larrucea, F. Belmonte, A. Welc, and T. Xie, "Reliability engineering," *IEEE Softw.*, vol. 34, no. 4, pp. 26–29, Feb. 2017.

[13] *Reliability of Military Electronic Equipment*, AGREE, Washington, DC, USA, 1957.

[14] W. H. von Alven, *Reliability Engineering*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1964.

[15] D. G. Anirban, *Asymptotic Theory of Statistics and Probability*. New York, NY, USA: Springer, 2006.

[16] H. A. David and H. N. Nagaraja, *Order Statistics*, 3rd ed. Hoboken, NJ, USA: Wiley, 2003.

[17] M. Dazer, M. Stohrer, S. Kemmler, and B. Bertsche, "Planning of reliability life tests within the accuracy, time and cost triangle," in *Proc. IEEE Accelerated Stress Test. Rel. Conf. (ASTR)*, Sep. 2016, pp. 1–9, doi: 10.1109/ASTR.2016.7762270.

[18] R. Fisher, "Frequency distribution of the values of the correlation coefficients in samples from an indefinitely large population," *Biometrika*, vol. 10, no. 4, pp. 21–507, 1915.

[19] A. Grundler, M. Bartholdt, and B. Bertsche, "Statistical test planning using prior knowledge-advancing the approach of Beyer and Lauster," in *Safety and Reliability—Safe Societies in a Changing World*. London, U.K.: CRC Press, 2018, doi: 10.1201/9781351174664-102.

[20] A. Grundler, M. Bollmann, M. Obermayr, and B. Bertsche, "Berücksichtigung von Lebensdauerberechnungen als Vorkenntnis im Zuverlässigkeitsnachweis," in *Tagung Technische Zuverlässigkeit 2019*. VDI Verlag GmbH, 2019.

[21] A. Grundler, M. Dazer, and B. Bertsche, "Reliability-test planning considering multiple failure mechanisms and system levels—An approach for identifying the optimal system-test level, type, and configuration with regard to individual cost and time constraints," in *Proc. Annu. Rel. Maintainability Symp. (RAMS)*, Jan. 2020, pp. 1–7, doi: 10.1109/RAMS48030.2020.9153678.

[22] A. Grundler, M. Dazer, and B. Bertsche, "Effect of uncertainty in prior knowledge on test planning for a brake caliper using the probability of test success," in *Proc. Annu. Rel. Maintainability Symp. (RAMS)*, May 2021, pp. 1–9, doi: 10.1109/RAMS48097.2021.9605727.

[23] A. Grundler, M. Dazer, T. Herzig, and B. Bertsche, "Efficient reliability demonstration using the probability of test success and Bayes theorem," in *Proc. Probabilistic Saf. Assessment Manag. Conf.*, 2021. [Online]. Available: https://www.iapsam.org/PSAM16/papers/AL3-PSAM16.pdf

[24] A. Grundler, M. Dazer, T. Herzig, and B. Bertsche, "Considering multiple failure mechanisms in optimal test design," in *Proc. 7th Int. Conf. Integrity-Rel.-Failure*, J. F. S. Gomes and S. A. Meguid, Eds., 2020, p. 673. [Online]. Available: https://paginas.fe.up.pt/~irf/Proceedings_IRF2020/data/papers/16469.pdf

[25] A. Grundler, M. Göldenboth, F. Stoffers, M. Dazer, and B. Bertsche, "Effiziente zuverlässigkeitsabsicherung durch berücksichtigung von simulationsergebnissen am beispiel einer hochvolt-batterie," in *VDI-Fachtagung Technische Zuverlässigkeit 2021*. VDI Verlag GmbH, 2021.

[26] M. S. Hamada, G. Alyson, C. Wilson, R. Shane, and H. F. Martz. *Bayesian Reliability* (Springer Series in Statistics). New York, NY, USA: Springer, 2008, doi: 10.1007/978-0-387-77950-8.

[27] M. Maisch, "Zuverlässigkeitsorientiertes Erprobungskonzept für Nutzfahrzeuggetriebe unter berücksichtigung von betriebsdaten," Ph.D. dissertation, Univ. Stuttgart, Stuttgart, Germany, 2017.

[28] W. Q. Meeker, L. A. Escobar, and F. G. Pascual, *Statistical Methods for Reliability Data*, 2nd ed. Hoboken, NJ, USA: Wiley, 2021.

[29] J. Raphson, *Analysis Aequationum Universalis Seu ad Aequationes Algebraicas Resolvendas Methodus Generalis, & Expedita, Ex Nova Infinitarum Serierum Methodo, Deducta AC Demonstrata*. London, U.K.: Typis Tho. Braddyll, Prostant Venales Apud Johannem Taylor, 1967.

[30] T.-R. Tsai, Y.-T. Lu, and S.-J. Wu, "Reliability sampling plans for Weibull distribution with limited capacity of test facility," *Comput. Ind. Eng.*, vol. 55, no. 3, pp. 721–728, Oct. 2008, doi: 10.1016/j.cie.2008.02.010.

[31] K. J. Wilson and M. Farrow, "Assurance for sample size determination in reliability demonstration testing," *Technometrics*, vol. 63, no. 4, pp. 523–535, Oct. 2021, doi: 10.1080/00401706.2020.1867646.

[32] O. P. Yadav, N. Singh, and P. S. Goel, "Reliability demonstration test planning: A three dimensional consideration," *Rel. Eng. Syst. Saf.*, vol. 91, no. 8, pp. 882–893, Aug. 2006, doi: 10.1016/j.ress.2005.09.001.

[33] T. Gneiting, "Normal scale mixtures and dual probability densities," *J. Stat. Comput. Simul.*, vol. 59, no. 4, pp. 375–384, 1997.

ALEXANDER GRUNDLER was born in Stuttgart, Germany, in 1989. He received the B.Sc. degree in mechanical engineering from the University of Stuttgart, Stuttgart, in 2014, the M.Sc. degree in mechanical engineering, in 2017, and the Dr.-Ing. degree in reliability engineering from the University of Stuttgart, in January 2024.

From 2017 to 2022, he was a Research Assistant with the Institute of Machine Components, University of Stuttgart, working on several reliability topics, such as reliability demonstration and optimal test planning. He is currently a Reliability Engineer with Robert Bosch GmbH, Schwieberdingen, Germany. He has published over 20 articles in the field of reliability engineering. His current research interests include efficient reliability demonstration, design for reliability, and risk assessment.

MARTIN DAZER was born in Villingen, Stuttgart, Germany, in 1989. He received the B.Sc. degree in mechanical engineering from Baden-Württemberg Cooperative State University (DHBW), in 2011, the M.Sc. degree in mechanical engineering, in 2014, and the Dr.-Ing. degree in reliability engineering from the University of Stuttgart, in 2019.

From 2015 to 2018, he was a Research Assistant with the Institute of Machine Components working on stochastic fatigue calculations and optimization methods in reliability test planning. He is currently the Head of the Reliability and Drive Technology Department, Institute of Machine Components. He is also the Founder and a Consultant of RelTest-Solutions GmbH, Stuttgart, offering highly advanced reliability consulting, coaching, and training for industry. His current research interest includes the multitude of aspects of reliability engineering with its main focus on life testing.

Dr. Dazer is a member of the Advisory Board of Safety and Reliability of VDI. He is also a member of the Program Committee of the Technical Reliability Conference, Germany. He is the Head of the Technical Committee of Reliability Management of the Association of German Engineers (VDI). He is working as a reviewer of several reliability journals.