Abstract
This contribution focuses on the design of optimal maintenance schedules for metallic structures prone to develop fatigue cracks. The crack propagation phenomenon is addressed using a fracture mechanics approach. The problem of maintenance scheduling is addressed within the framework of reliability-based optimization (RBO). Thus, it is possible to minimize the costs associated with maintenance and eventual failure while explicitly considering uncertainties in the crack propagation phenomenon and inspection activities. The underlying RBO problem is solved using an efficient method recently developed by the authors. A numerical example demonstrating the application of the proposed approach is presented.

Key words: maintenance scheduling, fatigue cracks, advanced simulation techniques, reliability-based optimization, reliability sensitivity

1. Introduction
Metallic structural components operating under cyclic loading are prone to develop fatigue cracks during their respective life span. These cracks can lead to loss of serviceability, partial failure or even collapse, depending on the particulars of the structural system. An effective means for mitigating the negative effects of crack propagation is the scheduling of inspection and repair activities (see, e.g. [26, 43, 56, 59, 63, 64, 68, 76]). The scheduling of maintenance activities usually involves the definition of the inspection frequency (e.g. monthly, annually), the type of inspection technique (e.g. visual inspection) and the definition of a particular repair strategy. Although maintenance activities constitute an effective means for coping with deterioration, their scheduling is extremely challenging, as the propagation of fatigue cracks is a highly uncertain phenomenon (see, e.g. [24, 37, 69]). Moreover, inspection activities are also uncertain, e.g. inspection activities may assess the damage incorrectly or may not detect any damage at all.

In the involved decision-making scenario described above, optimization methods which explicitly model the effects of uncertainty have received increased attention (see, e.g. [8, 67]). In particular, reliability-based optimization (RBO) has shown to be an adequate tool to seek the best tradeoff between the
overall costs (i.e. construction, operation, maintenance and possible failure costs) and an adequate level of reliability (see, e.g. [27, 40, 61]).

This contribution presents an approach for scheduling maintenance activities such that the cost of operation associated with a particular fatigue-prone mechanical component is minimized. The damage due to fatigue is addressed using a fracture mechanics approach. This allows simulating the crack propagation phenomenon by integrating appropriate laws that describe the crack growth. Uncertainties are considered explicitly by modeling physical parameters related to crack propagation as random variables. Moreover, it is considered that a mechanical component may develop several fatigue cracks, i.e. the component may undergo multi-site damage. In order to mitigate the effects of crack growth in the mechanical component, inspection and repair activities are scheduled. Inspection activities attempt detecting cracks in the mechanical component; as inspection is – in most cases – imperfect, the probability of detecting a crack is modeled explicitly. Once a crack is detected, it may be repaired (the decision to repair is taken according to some prescribed rules). In order to schedule the inspection and repair activities, two variables must be identified: the quality of inspection ($q$) – which influences the probability that a crack is detected – and the time at which inspection is performed ($T_I$); for the sake of simplicity, a single inspection activity and perfect repair are considered in this contribution. These two variables are selected based on the cost associated with the operation of the mechanical component, i.e. the summation of maintenance cost (the costs associated with inspection and repair for a particular value of the variables $q$ and $T_I$) and eventual failure cost (the cost associated with failure due to downtime, damages, etc.).

The study of the crack propagation phenomenon considering uncertainties and the scheduling of maintenance activities has received considerable attention in the literature (see, e.g. [57, 58, 78]). A brief survey on some of the approaches developed within this context is presented in Appendix A. Compared with these approaches, this contribution contains several innovative concepts. On one side, Subset Simulation – which was introduced in [4, 7] – is applied for evaluating expected costs; the major advantage of applying this method is that it allows studying several different failure modes while being numerically efficient. On the other side, the RBO problem is solved using an efficient algorithm based on line search recently developed by the authors [31]; this algorithm combines standard optimization techniques with an efficient algorithm for estimating reliability sensitivity; the latter algorithm is extended in this contribution for estimating the sensitivity associated with expected costs.

The structure of this paper is as follows. Section 2 provides a general overview on the problem being studied and the relevance of maintenance activities. Then, the model for crack propagation considered in this contribution is described in Section 3. Section 4 discusses the different relevant events that may take place during the lifetime of a mechanical component and the economical costs associated with these events. Section 5 presents an approach for calculating the sensitivity of the aforementioned costs with respect to the variables defining the maintenance schedule. The strategy for solving the
2. Importance of Maintenance and the Effect of Uncertainties

2.1. General Overview

The crack propagation phenomenon in metallic components (operating under cyclic loading) can be characterized in three different stages: the crack initiation (CI) stage, where micro cracks nucleate (e.g. at points of stress concentration); the stable crack growth (SCG) stage, where the crack propagates in a stable manner; finally, the unstable crack growth (UCG) stage, at which the fast growth of the crack leads to the collapse of the component. For design purposes, the duration of the UCG stage is commonly ignored. On the other hand, the duration of CI stage will vary according to the system being analyzed: for weldments, the duration of CI phase may be neglected [14]; in aerospace components, the duration of the CI phase may account for the vast majority of the lifetime due to high quality standards of fabrication [37] and, hence, it must be explicitly modeled. A schematic representation of the crack propagation stages is shown in fig. 1, where \( a \) represents the crack length and \( N \), the number of load cycles.

![Figure 1: Stages of fatigue crack propagation phenomenon](image)

It should be noted that in a particular mechanical component, one or more cracks may develop. For example, in an aircraft, fatigue cracks may develop at multiple sites of a critical component, i.e. multi-site damage, causing drastic reduction of the lifetime [71].

Maintenance activities can be an extremely cost-effective means for controlling the damage accumulation due to crack propagation. That is, inspection activities during the lifetime of a mechanical component allow identifying the presence of cracks. Based on the information provided by inspection, a decision on repairing the mechanical component may be taken, e.g. in case the detected cracks exceed a prescribed threshold level, they can be removed using a special procedure. In this way, the chances that the mechanical component does not fail during the design lifetime are highly increased. It should be noted that the costs associated with inspection and repair of a component are – usually
much cheaper than assuming the costs associated with the failure (collapse). In this context, it is important to remark that inspection and repair activities attempt at mitigating partial damage states, i.e. although damage is present in the component, it does not cause failure at the time being considered. The relevance of considering partial damage states in practical design situations has been widely acknowledged in the literature [25, 40].

Within the scope of this contribution, the definition of a maintenance schedule involves the selection of two design variables. The first one refers to the quality of inspection \((q)\). This variable characterizes the ability of a particular non-destructive inspection (NDI) method for detecting a crack. For example, while a low quality inspection method may fail in detecting a certain crack, a high quality method will (most likely) detect the same crack. For simplicity, it is assumed that \(q\) can be modeled by means of a positive, real number [68]. The second design variable is the time of inspection \((T_I)\), which determines at which instant of the lifetime of a component inspection and repair are performed. This variable plays a major role, as maintenance activities scheduled at an inappropriate time may be completely ineffective [77]. In this contribution, \(T_I\) is modeled as a positive, real number, just as in the case of \(q\). However, it should be noted that depending on the scope of the specific application being studied, a more suitable model for either \(q\) or \(T_I\) could be a discrete variable; however, for the sake of simplicity, this last option is not pursued further in this contribution.

2.2. Uncertainties and Their Effect

One of the major challenges in the scheduling of maintenance activities is that both the crack propagation phenomenon and the inspection activities are affected by uncertainties. In the case of the crack propagation, the time to appearance and length of a crack at a particular instant cannot be characterized as deterministic variables (see, e.g. [37, 69]). Therefore, the uncertainty in the crack propagation phenomenon should be explicitly modeled, e.g. by means of random variables.

As in the case of the crack propagation, the inspection activities are also subject to uncertainties. This is due to the fact that inspection is not infallible, i.e. a certain inspection technique may fail in detecting a particular crack. The uncertainties related with the detection of crack can be modeled by means of the probability of detection (POD), i.e. the probability that a crack of length \(a\) is detected by inspection. In addition, the length of a crack that has been detected may be measured incorrectly, thus affecting the decision of repair.

The effect of the aforementioned uncertainties is reflected in the performance of the mechanical component; more specifically, these uncertainties prevent from predicting the behavior of the component during its target lifetime in a deterministic way. Instead, a spectrum of possible events may occur. For a better understanding of these different events and their interaction, consider the schematic representation in fig. 2, which corresponds to an event tree (see, e.g. [16, 23, 43]). This figure illustrates the possible events that may take place during the life span of a metallic stripe subject to cyclic loading that develops a single edge crack. The occurrence of any of these events will be defined by the physics
of the problem (in this case, the crack propagation phenomenon), the particulars of the maintenance schedule and how the uncertainties affect the crack propagation and inspection, respectively.

At the beginning of the life span, the stripe in fig. 2 is undamaged. As load cycles are applied over the stripe, a crack is developed. It may occur that this crack grows until causing failure ($F$) or the component may survive until inspection. Upon the outcome of the inspection, repair may or may not take place ($R$ and $\overline{R}$, respectively). In case the component is repaired, it should survive ($\overline{F}$) until the end of the life span because repair is assumed to be perfect (however, in a component with several cracks, this might not be the case). In case no repair takes place, the component may fail ($F$) or it may still survive ($\overline{F}$) until the end of the life span in spite of the presence of a crack.

2.3. Scheduling of Maintenance Activities

The schematic representation of the possible events that can occur during the life span of a mechanical component in fig. 2 serves for illustrating the fact that in the scheduling of maintenance activities, the effects of uncertainty cannot be ignored. In particular, the task of determining an appropriate maintenance schedule is rather challenging, as the costs of inspection, repair and failure associated with a particular scheme are not fixed but vary randomly, due to inherent uncertainties.

In this challenging scenario, reliability-based optimization (RBO) offers a rational approach for taking into account the effects of uncertainty. In particular the problem of optimal design of a maintenance schedule can be interpreted as the minimization of the total cost ($C_T$) associated with the operation of the mechanical component during its lifetime, i.e. the summation of the cost associated with inspection activities ($C_I$), repair ($C_R$) and eventual failure ($C_F$) [21, 40]. The cost function is a random variable itself, as it depends on the occurrence of repair and failure events. Then, the problem of minimization of $C_T$ can be addressed by defining an appropriate deterministic substitute problem.
Thus, the substitute optimization problem is defined as follows.

$$\min_{(q,T)} E[C_T(q,T_i,\theta)]$$

subject to

$$h_j(q,T_i) \leq 0, \quad j = 1, \ldots, n_{DC}$$

(1)

In eq. (1), $E[\cdot]$ denotes expected value; $\theta$ denotes a $n_\theta$-dimensional vector of random variables (which is characterized by means of a joint probability density function $f(\theta)$) representing the uncertainties in the crack propagation phenomenon and the inspection activities; $h_j(\cdot,\cdot)$, $j = 1, \ldots, n_{DC}$ are constraints of the problem (e.g. side constraints on $(q,T_i)$). In addition, the expected total cost $E[C_T(\cdot,\cdot,\cdot)]$ is defined as:

$$E[C_T(q,T_i,\theta)] = E[C_I(q,T_i,\theta)] + E[C_R(q,T_i,\theta)] + E[C_F(q,T_i,\theta)]$$

(2)

The solution of the optimization problem in eq. (1) yields a maintenance schedule that is optimal with respect to the costs of operation while explicitly taking into account the effects of uncertainties in crack propagation and inspection. In the remaining part of this contribution, the details on how to solve this optimization problem are examined in depth.

3. Crack Propagation

3.1. Crack Growth Model

A number of different models have been developed for studying the crack propagation phenomenon [2]. In this contribution, the crack propagation phenomenon is characterized as one-dimensional (governed by Mode I loading) and it is described using the Paris-Erdogan law [50], i.e.:

$$\frac{da}{dN} = C(\Delta K)^m, \quad a(N = 0) = a_0$$

(3)

In eq. (3), $C$ and $m$ are material properties, $a_0$ is the initial crack length and $\Delta K$ is the stress intensity factor (SIF) range, which is calculated according to the principles of linear fracture mechanics by means of:

$$\Delta K = Y(a)\sqrt{\pi a} \Delta \sigma$$

(4)

where $Y(a)$ is a geometry-dependent function and where $\Delta \sigma$ is the range of the far-field stress, i.e. $\Delta \sigma = \sigma_{max} - \sigma_{min}$. For simulating the propagation of $n_C$ cracks in a component, a system of $n_C$ differential equations of the type shown in eq. (3) is solved using an appropriate numerical integration scheme until the target life time of the component is reached; in case failure occurs before reaching the target life time, the integration is stopped. The presence of several cracks in the component is
reflected in the integration of the Paris-Erdogan law in the SIF’s, i.e. the SIF at the tip of the $i$-th crack is a function of the length of the different cracks present in the component. In this context, it is important to note that analytical expressions for calculating SIF’s exist for very simple cases only, usually involving a single crack or several symmetric cracks. In more general cases, it is necessary to resort to numerical methods for calculating SIF’s. In this contribution, the so-called Finite Element Alternating Method (FEAM) [72, 73] is applied for estimating SIF’s. A brief description of this method is presented in Appendix B of this contribution. It should be noted that the FEAM is not the only alternative for estimating stress intensity factors: other methods available are, e.g. Boundary Element Method (see, e.g. [17, 36]), Meshless methods (see, e.g. [20, 48]), Extended Finite Element Method (see, e.g. [47]), etc.

Within the framework of linear elastic fracture mechanics, the failure of a mechanical component occurs when the stress intensity factor $K_I(a) = Y(a)\sqrt{\pi a} \sigma_{\text{max}}$ exceeds the critical value of the fracture toughness, $K_{Ic}$. But in most cases, this criterion by itself may not be representative, as failure can occur due to a combined effect, e.g. interaction between brittle fracture (i.e. $K_I(a) > K_{Ic}$) and ductile failure (i.e. applied force over mechanical component exceeding its capacity). A feasible means for taking into account this interaction is the application of the so-called R6 curve criterion (see, e.g. [2, 34]). Thus, failure will occur whenever the following inequality holds:

$$\frac{K_I(a)}{K_{Ic}} \geq C_f(a), \quad C_f(a) = \frac{\sigma_{\text{max}}}{\sigma_C(a)} \left( \frac{8}{\pi^2} \log \left( \sec \left( \frac{\pi}{2} \frac{\sigma_{\text{max}}}{\sigma_C(a)} \right) \right) \right)^{-0.5}$$

(5)

where $\sigma_C(\cdot)$ is the collapse stress of the mechanical component; in the case of a structure under axial tension, collapse occurs when the stress in the net section reaches the flow stress ($\sigma_y$) of the material of the component.

The Paris-Erdogan law has been thoroughly used by the engineering community in order to model the crack propagation phenomenon due to its simplicity and straightforward implementation. In particular, this law is quite successful in characterizing crack growth under constant amplitude cyclic loading, small scale yielding (i.e. yielding ahead of the crack tip) and long cracks. For those cases where these conditions are not met, the Paris-Erdogan law may not be appropriate. However, several extensions and alternative models have been proposed in the literature in order to overcome this drawback. For example, the so-called cohesive model [49] allows modeling crack initiation, small cracks and the effects of plasticity. It should be noted that the approach for optimal maintenance planning proposed in this contribution can be applied – in principle – in conjunction with any appropriate crack propagation model. Therefore, the use of the Paris-Erdogan law does not imply a limitation.

3.2. Uncertainties in Crack Propagation

In this contribution, the uncertainty in the crack propagation phenomenon is addressed by direct integration of the equation governing the crack growth (cf. eq. (3)). Such an approach is one of several
alternatives for considering uncertainties in crack propagation [75]; in other approaches, this issue is addressed, e.g. by incorporating a factor that models a random process [76].

The approach for considering the uncertainties based on direct integration consists in the following. The different parameters that govern the Paris-Erdogan law and the failure criterion are modeled as random variables. Then, for simulating the crack propagation, samples (realizations) of the vector of uncertain parameters are drawn and then, the corresponding differential equation is integrated. In this contribution, the parameters that are modeled using random variables are the initial crack length, the parameter $C$ of the Paris-Erdogan law, the fracture toughness and the yield stress.

The crack length is a random variable itself, as it depends on a number of basic random variables, as shown in eq. (6).

$$a_i = a_i(q, T_I, \theta, t), \quad i = 1, \ldots, n_C$$

In the equation above, $a_i$ is the length of the $i$-th crack; $\theta$ is an $n_\theta$-dimensional vector of random variables; $t$ denotes time and it should be such that $0 \leq t \leq T$, where $T$ is the target lifetime. In order to simplify the notation, the crack length is denoted as $a_i = a_i(t), \quad i = 1, \ldots, n_C$ in the remaining part this contribution.

4. Lifetime Events and their Expected Costs

This Section provides a detailed description on how the different relevant events that may occur during the lifetime of a fatigue-prone mechanical component are defined. Moreover, the economical costs associated with each of these events and the approach for estimating their expected values are discussed as well.

4.1. Inspection Activities

4.1.1. Description

Inspection activities take place at time $T_I$. The aim of inspection is detecting the presence of possible cracks in a mechanical component and also measuring their length. As discussed in Section 2, inspection activities are not perfect, as they may not be able detect and measure cracks correctly. In particular, the uncertainty in the detection of a crack is modeled by means of the probability of detection (POD). Different inspection techniques have associated different POD [39]. In this contribution, the POD is modeled by means of the following function [79]:

$$POD(q, a(t)) = 1 - e^{\lambda a(t)}, \quad a \geq 0$$

where $q$ denotes the quality of inspection and where $\lambda$ is a constant depending on the specific inspection technique that is applied. In addition, the error in the measurement of the length of a crack is modeled
by means of a random variable $\theta$, $\theta \in \theta$ [79]. Thus, the measured crack length $a_{\text{meas},i}$ is defined as indicated in eq. (8).

$$a_{\text{meas},i}(T_I) = a_i(T_I) + \theta_{i,i}, \quad i = 1, \ldots, n_C$$

(8)

4.1.2. Expected Inspection Cost

The expected inspection cost is calculated as the product between the unit inspection cost (corrected by the discount rate) and the probability that inspection takes place [68]. Mathematically, this expected cost is defined as:

$$E[C_I(q, T_I, \theta)] = \frac{c_I(q)}{(1 + r)T_I} \left(1 - p_{T_I}^F\right)$$

(9)

where $c_I(q)$ is the unit cost of performing inspection of quality $q$, $r$ is the discount rate (this term is included for calculating capitalized costs) and $p_{T_I}^F$ is the probability that failure occurs before $T_I$; usually, $p_{T_I}^F << 1$. Thus, the expected inspection cost can be approximated as:

$$E[C_I(q, T_I, \theta)] \approx \frac{c_I(q)}{(1 + r)T_I}$$

(10)

This simplification is most convenient, as $E[C_I(q, T_I, \theta)]$ is an explicit function of the variables defining the maintenance schedule. Thus, the value of the expected inspection cost can be computed straightforwardly; also the derivatives of $E[C_I(q, T_I, \theta)]$ can be computed analytically.

4.2. Repair Activities

4.2.1. Description

The repair event $R$ consists in the removal of one or more cracks. This event takes place based on the outcome of the inspection. In order to define the repair event formally, consider the following concepts. The event in which the $i$-th crack is repaired is denoted as $R_i$, $i = 1, \ldots, n_C$. In case the repair event $R_i$ takes place, it is performed immediately after inspection and it is assumed that the crack is fully removed (perfect repair), i.e. $a_i = 0$ from the time of inspection ($T_I$) until the end of the life time ($T$). In mathematical terms, $R_i$ is defined as follows.

$$R_i = \{\theta \in \Omega_\theta : D_{R_i}(q, T_I, \theta) \geq 1\}, \quad i = 1, \ldots, n_C$$

(11)

In eq. (11), $D_{R_i}(\cdot, \cdot, \cdot), \quad i = 1, \ldots, n_C$ is the so-called normalized demand function associated with $R_i$. In order to define this function, it should be considered that repair of the $i$-th crack takes place upon the occurrence of two basic events, i.e.:

1. The crack should be detected. The chance of detecting a particular crack is modeled by means of the probability of detection (POD), as indicated in Section 4.1. Therefore, in order to model the detection of the crack within the context of simulation, the condition $\text{POD}(q, a(T_I))/\theta_U \geq 1$
(where $\theta_U$ is a uniformly distributed random variable, $\theta_U \in \Theta$) is employed. It should be noted that this condition is fulfilled with probability $POD(q,a(T_I))$, which is precisely the target probability that must be simulated.

2. The measured crack length ($a_{meas}$) should be larger than a certain threshold $a_{cr}$. Again, this event can be modeled within the context of simulation using the condition $a_{meas}/a_{cr} \geq 1$.

The two events described above must occur simultaneously in order to ensure that repair takes place. Therefore, by taking the minimum between the two and comparing it to unity, it is possible to simulate the target event. Thus, $D_{R_i}(\cdot,\cdot,\cdot)$ is defined as (see, e.g. [30]):

$$D_{R_i}(q,T_I,\theta) = \min \left( \frac{a_{meas,i}(T_I)}{a_{cr}}, \frac{POD(q,a_i(T_I))}{\theta_U,i} \right), \quad \theta_U \in \Theta, \quad i = 1, \ldots, n_C$$

(12)

Considering the definitions in eqs. (11) and (12), it is possible to define in mathematical terms the repair event $R$ and the associated normalized demand function $D_{R}(\cdot,\cdot,\cdot)$ as shown below.

$$R = \{ \theta \in \Omega_{\theta} : \; D_{R}(q,T_I,\theta) \geq 1 \}$$

(13)

$$D_{R}(q,T_I,\theta) = \max_{i=1,\ldots,n_C} (D_{R_i}(q,T_I,\theta))$$

(14)

It should be noted that the normalized demand function associated with the repair event can be non differentiable as it includes $\min(\cdot)/\max(\cdot)$ functions.

4.2.2. Expected Repair Cost

The expected repair cost is defined as a multi-dimensional integral of the cost function associated with repair over the region of the uncertain parameters associated with the repair event $R$ defined above. In mathematical terms, this integral is defined as:

$$E[C_{R}(q,T_I,\theta)] = \int_{D_{R}(q,T_I,\theta) \geq 1} C_{R}(q,T_I,\theta)f(\theta)d\theta$$

(15)

where $C_{R}(\cdot,\cdot,\cdot)$ is the cost function associated with repair and is defined as shown below.

$$C_{R}(q,T_I,\theta) = \sum_{i=1}^{n_C} I_{R_i}(q,T_I,\theta) \frac{c_R}{(1+r)^{T_I}}$$

(16)

In eq. (16), $c_R$ represents the cost of repairing one crack and $I_{R_i}(\cdot,\cdot,\cdot)$ is an indicator function which is equal to one in case $D_{R_i}(q,T_I,\theta) \geq 1$ and zero, otherwise. Thus, $C_{R}(\cdot,\cdot,\cdot)$ can be interpreted as the product between the unit capitalized repair cost times the number of cracks being repaired.

A quantity that is closely related with the calculation of the expected repair cost is the probability of
repair ($p_R$), which is defined as follows.

$$p_R(q, T_I) = P[D_R(q, T_I, \theta) \geq 1] = \int_{D_R(q, T_I, \theta) \geq 1} f(\theta) d\theta$$

$$= \int_{\theta \in \Omega} I_R(q, T_I, \theta) f(\theta) d\theta$$  \hspace{1cm} (17)

In the last equation, $I_R(\cdot, \cdot, \cdot)$ is an indicator function which is equal to one in case $D_R(q, T_I, \theta) \geq 1$ and zero, otherwise.

### 4.2.3. Evaluation of Expected Repair Cost

The evaluation of the expected cost associated with repair is quite challenging, as it involves the computation of a multi-dimensional integral. The evaluation of this integral is closely related with the evaluation of reliability; thus, methods of structural reliability have been applied in the literature in order to evaluate expected costs. For example, in [43], the estimation of the expected costs and reliability is performed using the so-called First Order Reliability Method (FORM, see, e.g. [19]). In this case, each of the branches of the event tree (see fig. (2)) must be analyzed in order to estimate reliability; this can be quite involved. Another possibility for estimating expected cost is the application of Monte Carlo Simulation (MCS, see, e.g. [45]); the advantage of MCS is that one simulation run allows assessing the reliability associated with the full event tree [74], e.g. considering several structural components and failure modes. However, the main disadvantage of MCS is that it is numerically demanding, i.e. a large number of simulations may be required, especially for analyzing rare occurrence events (such as the failure event). Therefore, in this contribution, Subset Simulation (SS) is applied for assessing reliability and evaluating expected costs. This method – which was introduced in [4, 7] – is an advanced simulation method: it retains the generality of MCS but possesses a higher efficiency.

For describing the application of SS for evaluating the multi-dimensional integral in eq. (15), note that the expected cost can be expressed as the product of the expected repair cost conditioned on the repair event ($E[C_R(q, T_I, \theta)]/R$) times the probability of repair ($p_R(q, T_I)$, see eq. (17)) by means of the Bayes’ theorem, as shown in the equation below.

$$E[C_R(q, T_I, \theta)] = E[C_R(q, T_I, \theta)]/R \cdot p_R(q, T_I)$$  \hspace{1cm} (18)

In SS, the repair domain $R$ ($R = \{\theta \in \Omega_\theta : D_R(q, T_I, \theta) \geq 1\}$) is defined as a sequence of subsets (or intermediate failure events) $R_p$, $p = 1, \ldots, P$ such that $R_1 \supset F_2 \supset \cdots \supset R_P = R$. Thus, the failure probability can be expressed as:

$$p_R(q, T_I) = P[D_R(q, T_I, \theta) \geq 1] = P[R] = P[R_1] \prod_{p=1}^{P-1} P[R_{p+1} | R_p]$$  \hspace{1cm} (19)
where \( P[R_{p+1} | R_p] \) is the probability of occurrence of the event \( R_{p+1} \) conditioned on the event \( R_p \). The key issue in SS is selecting the intermediate failure events such that the probabilities \( P[R_{p+1} | R_p] \), \( p = 1, \ldots, P - 1 \) and \( P[R_1] \) are sufficiently large (e.g. \( P[.] \approx 0.1 \)) in order to estimate them using direct MCS. Thus, a small failure probability can be calculated as the product of larger, conditional probabilities. This allows SS estimating \( p_R(q, T_I) \) most efficiently. Moreover, it should be noted that at the last step of SS, samples of the uncertain parameters conditioned on the failure event are generated. These samples can be used to estimate the expected failure cost conditioned on the repair event, i.e.:

\[
E[C_R(q, T_I, \theta)] \approx \left( \frac{1}{N_R} \sum_{j=1}^{N_R} C_R(q, T_I, \theta^{(j)}) \right) p_R(q, T_I), \quad \theta^{(j)} \sim f(\theta/R), \quad j = 1, \ldots, N_R 
\]

(20)

where \( N_R \) is the number of samples of the uncertain parameters conditioned on the repair event. For a better understanding of how the expected repair cost is evaluated using SS, consider the schematic representation in fig. 3 in the space of the uncertain parameters. As it can be noticed from the figure, the repair event takes place only for a fraction of the possible values of the uncertain parameters, i.e. \( D_R(q, T_I, \theta) \geq 1 \). Using SS, it is possible to draw samples of the uncertain parameters in this region, as shown with the dots in the figure. It should be noted that the value of the function associated with cost repair is – in general – different for each point, e.g. while the point \( \theta^{(j_1)} \) in the figure can be associated with the repair of a single crack, the point \( \theta^{(j_2)} \) can be associated with the repair of \( n \) cracks. Thus, the value expected repair cost conditioned on the repair event can be calculated as the average of the value of the cost function for the different samples generated.

\[ \bullet \theta^{(j)} \sim f(\theta/R), \quad j = 1, \ldots, N_R \]

Figure 3: Schematic representation of multi-dimensional integral associated with the evaluation of the expected repair cost
As a side remark, it should be considered that an interesting feature of SS is that it explores the space of uncertain parameters by means of successive subsets $R_p$, $p = 1, \ldots, P$, each of which is of rarer occurrence than the previous one, i.e. $P[R_{p-1}] > P[R_p]$. This implies that a single run of SS provides the probability of occurrence of the event $P[D_R(q, T_I, \theta) \geq b]$, where $b \in [0, 1 + \delta]$ and $\delta$ is a small constant, e.g. $\delta = 0.05$.

4.3. Failure Event

4.3.1. Description

Failure corresponds to the occurrence of the event where the mechanical component does no longer fulfill its intended design purpose. In this contribution, the failure event $F$ is defined considering two possible failure criteria (see, e.g. [51]). In the first one, the stress intensity factor at the tip of one (or more) crack(s) exceeds the critical value of the material toughness (considering the effects of plasticity, see eq. (5)) during the life time of the mechanical component. In the second one, failure occurs whenever two or more cracks linkup. In this context, it should be noted that the linkup of cracks does not necessarily imply the collapse of a mechanical component. However, the application of such a stringent failure criterion ensures a conservative modeling of the failure event within the context of this contribution.

In mathematical terms, the failure event $F$ is defined as follows.

$$F = \{ \theta \in \Omega_\theta : D_F(q, T_I, \theta) \geq 1 \} \quad (21)$$

In eq. (21), $D_F(\cdot, \cdot, \cdot)$ is the normalized demand function associated with $F$. In turn, this function is defined as the maximum between two individual normalized demand functions, i.e.:

$$D_F(q, T_I, \theta) = \max (D_{F,S}(q, T_I, \theta), D_{F,L}(q, T_I, \theta)) \quad (22)$$

where $D_{F,S}(\cdot, \cdot, \cdot)$ and $D_{F,L}(\cdot, \cdot, \cdot)$ are the normalized demand functions associated with the first and second failure criterion described above, respectively. The function $D_{F,S}(\cdot, \cdot, \cdot)$ measures the ratio between the stress intensity factor at the crack of a tip and the fracture toughness. It is defined as:

$$D_{F,S}(q, T_I, \theta) = \max_{i=1,\ldots,n_C} \left( \max_{t \in [0, T]} \left( \frac{K_{I,i}(a(t))}{C_f(\theta, a(t))K_{Ic}(\theta)} \right) \right) \quad (23)$$

where $K_{I,i}(\cdot)$ is the SIF associated with the $i$-th crack and $a(t) = \langle a_1(t), a_2(t), \ldots, a_{n_C}(t) \rangle^T$. The function $D_{F,L}(\cdot, \cdot, \cdot)$ is associated with the occurrence of linkup of cracks. This event is defined using Swift’s criterion (see, e.g. [9]), i.e. linkup occurs whenever the plastic zones from adjacent crack tips merge. This criterion is shown schematically in fig. (4). In this sketch, cracks emanating from rivet holes are depicted; $d_{i_1, i_2}$ denotes the separation between the holes, $s_{i_1, i_2}$ refers to the separation between the plastic zones and $r_{pl,j}$, $j = i_1, i_2$ denotes the size of the plastic zones ahead of the crack.
The size of the plastic zone is calculated as \( r_{pl,j} = \frac{K_{I,j}^2}{(2\pi \sigma_y^2)} \).

\[
\text{Figure 4: Schematic representation of Swift’s criterion}
\]

In mathematical terms, \( D_{F,L}(\cdot,\cdot,\cdot) \) is defined as shown below.

\[
D_{F,L}(q,T,\theta) = \max_{l=1,\ldots,n_L} \left( \max_{t\in[0,T]} \left( \frac{d_{l1,l2} - s_{l1,l2}(a(t))}{d_{l1,l2}} \right) \right)
\]  

(24)

In eq. (24), it is considered that linkup could potentially occur at \( n_L \) locations, involving the cracks \( a_{l1} \) and \( a_{l2} \), where \( (l_1,l_2) \in [1,\ldots,n_C] \); additionally, the separation between the plastic zones is a function of time, i.e. \( s_{l1,l2} = s_{l1,l2}(a(t)) \).

As in the case of \( D_R(\cdot,\cdot,\cdot) \), the normalized demand function associated with the failure event can be non differentiable as it includes max(\cdot) functions.

4.3.2. Expected Failure Cost

The expected failure cost is defined as a multi-dimensional integral of the cost function associated with failure over the region of the uncertain parameters associated with the repair event \( F \) defined above. In mathematical terms, this integral is defined as:

\[
E[C_F(q,T_l,\theta)] = \int_{D_F(q,T_l,\theta) \geq 1} C_F(q,T_l,\theta) f(\theta) d\theta
\]  

(25)

where \( C_F(\cdot,\cdot,\cdot) \) is the cost function associated with failure; it is defined as follows.

\[
C_F(q,T_l,\theta) = \frac{c_F}{(1+r)T_F(q,T_l,\theta)}
\]  

(26)

In eq. (26), \( c_F \) is the cost associated with failure (i.e. \( c_F \) accounts for the consequences or losses associated with failure) and \( T_F(\cdot,\cdot,\cdot) \) is the time at which failure takes place. This time to failure is a byproduct of the simulation of the crack propagation phenomenon. More specifically, the crack propagation equation is integrated using a numerical scheme until the normalized demand function associated with failure exceeds one; the time at which the normalized demand exceeds one is characterized as the time to failure.

As in the case of the expected repair cost, a quantity that is closely related with the calculation of
\[ E[C_F(q, T_I, \theta)] \] is the probability of failure \( (p_F) \), which is defined below.

\[
p_F(q, T_I) = P[D_F(q, T_I, \theta) \geq 1] = \int_{D_F(q, T_I, \theta) \geq 1} f(\theta) d\theta = \int_{\theta \in \Omega} I_F(q, T_I, \theta) f(\theta) d\theta
\]  

(27)

In the last equation, \( I_F(\cdot, \cdot, \cdot) \) is an indicator function which is equal to one in case \( D_F(q, T_I, \theta) \geq 1 \) and zero, otherwise.

4.3.3. Evaluation of Expected Failure Cost

The expected failure cost is evaluated in a similar way as the expected repair cost. Repeating the steps presented in eqs. (18), (19) and (20), it is also possible to estimate the expected cost associated with failure, as shown below in eq. (28). In this equation, \( N_F \) indicates the number of samples of the uncertain parameters conditioned on the repair event and \( p_F(q, T_I) \) is the probability of failure (see eq. (27)).

\[
E[C_F(q, T_I, \theta)] \approx p_F(q, T_I) \frac{1}{N_F} \sum_{j=1}^{N_F} \frac{c_F}{(1+r)^{T_F(q, T_I, \theta^{(j)})}}, \quad \theta^{(j)} \sim f(\theta/F), \quad j = 1, \ldots, N_F
\]  

(28)

For a better understanding of how the expected failure cost is evaluated using SS, consider the schematic representation in fig. 5 in the space of the uncertain parameters. Using SS, it is possible to draw samples of the uncertain parameters in the failure region, as shown with the dots in the figure. Each of these samples has associated a different time to failure \( T_F(\cdot, \cdot, \cdot) \). Thus, the value expected failure cost conditioned on the failure event can be calculated as the average of the value of the cost function for the different samples generated.

5. Sensitivity Estimation of Expected Costs

While the previous Section focused on the definition of the relevant events during the lifetime of a mechanical component and the evaluation of the expected costs associated with these events, this Section is concerned with the sensitivity (derivative) of these costs with respect to the variables defining the maintenance schedule. That is, the objective of this Section is developing efficient estimators that quantify the change in the expected cost functions due to perturbations in the quality of inspection or the time of inspection. Such information is most useful when solving the optimal maintenance scheduling problem, as discussed in detail in Section 6.1.

For simplifying the presentation of this Section, the following notation is considered: the design variables defining the maintenance scheme are grouped in the vector \( y \), i.e. \( y = \langle q, T_I \rangle^T \).
5.1. Sensitivity of Expected Inspection Cost

The estimation of the sensitivity of the expected inspection cost with respect to the quality of inspection and the time of inspection is straightforward. This is due to the fact that there is an expression for estimating this cost (see eq. (10)) which is explicit with respect to the variables defining the maintenance schedule. Therefore, the sought derivatives can be calculated analytically.

5.2. Sensitivity of Expected Repair Cost

The evaluation of the sensitivity of the expected repair cost is a challenging task, as it comprises analyzing the multi-dimensional integral in eq. (15) with respect to variations in the variables defining the maintenance schedule. This problem has been previously studied in the literature, see e.g. [23]; however, the approach proposed in that contribution (developed within the framework of FORM) is not applicable within the scope of this contribution. Therefore, an approach presented in [31] is applied and extended in the following in order to compute the sought sensitivity.

For describing the approach followed for computing the sensitivity, it should be noted that the expected repair cost function may be non differentiable with respect to the variables that define the maintenance strategy. Therefore, the sought sensitivity is calculated with respect to an approximate representation of this expected cost that is differentiable, i.e.: 

$$ \frac{\partial E[C_R(y, \theta)]}{\partial y_l} = \lim_{\Delta y_l \to 0} \frac{E[C_R(y + \psi(l)\Delta y_l, \theta)] - E[C_R(y, \theta)]}{\Delta y_l} $$

(29)

where $\psi(l)$ is a vector with all entries equal to zero, except the $l$-th entry, which is equal to one.

For implementing this approximate representation of the expected repair cost and the evaluating its sensitivity, three approximation concepts are introduced, namely:
1. The probability that the normalized demand $D_R(y, \theta)$ exceeds a threshold $b$ is approximated by means of an exponential function [27, 31, 33, 31], i.e.:

$$P \left[ D_R(y^k, \theta) \geq b \right] \approx P \left[ D_R(y^k, \theta) \geq 1 \right] e^{\psi(b-1)}, \quad b \in [1 - \delta, 1 + \delta] \quad (30)$$

where $\psi$ is a constant and $\delta$ is a small number, e.g. $\delta = 0.05$. It should be noted that the constant $\psi$ can be readily determined using a least square procedure, as the relation between $P \left[ D_R(y^k, \theta) \geq b \right]$ and $b$ is a byproduct of SS (see Section 4).

2. The normalized demand function at a perturbed design $y^*$ is approximated as the normalized demand function at a nominal design $y$ plus a linear combination of the vector $(y^* - y)$, i.e.:

$$D_R(y^*, \theta) \approx \overline{D}_R(y^*, \theta) = D_R(y, \theta) + \sum_{l=1}^{2} a_l (y^*_l - y_l) \quad (31)$$

where $a_l, l = 1, 2$ are real coefficients calculated using a least squares approach. Details on how to compute these coefficients are described in Appendix C. The expression in eq. (31) could be interpreted as an incomplete expansion of $D_R(y^*, \theta)$ because the possible interactions between $y$ and $\theta$ are ignored as well as the influence of higher order terms. Thus, in most cases, this expansion will not be exact, as changes in $D_R(\cdot, \cdot)$ due to perturbations in the design variables are non linear, implicit functions. Nonetheless, as long as the coefficients $a_l, l = 1, 2$ are calibrated appropriately, it could be expected that $\overline{D}_R(y^*, \theta)$ can approximate $D_R(y^*, \theta)$ reasonably well.

3. The cost function associated with repair (cf. eq. (16)) at a perturbed design $y^*$ is approximated also as the cost function at a nominal design $y$ plus a linear combination of the vector $(y^* - y)$, i.e.:

$$C_R(y^*, \theta) \approx \overline{C}_R(y^*, \theta) = E[C_R(y, \theta)/R] + \sum_{l=1}^{2} c_l (y^*_l - y_l) \quad (32)$$

In eq. (32), $E[C_R(y^k, \theta)/R]$ is the expected repair cost conditioned on the repair event, which can be calculated using SS (see eq. (20)), and $c_l, l = 1, 2$ are real coefficients; details on how to calculate them are described in Appendix C. It should be noted that this linear approximation of the cost function associated with repair is similar to the one introduced for the normalized demand function, i.e. it is an incomplete expansion.

While the two first approximation concepts described above were introduced in [31], the last approximation of the function associated with cost of repair constitutes a novel concept. The advantage of introducing the representations shown in eqs. (30), (31) and (32) is that they allow estimating an approximation of the sensitivity of the approximate expected repair cost explicitly, as they do not depend on $\theta$. For demonstrating this last point, consider the introduction of eqs. (15), (31) and (32)
in the limit of eq. (29), i.e.

\[
\frac{\partial E[C_R(y, \theta)]}{\partial y_l} \approx \lim_{\Delta y_l \to 0} \left( \int_{D_R(y, \theta) \geq 1} (E[C_R(y, \theta)/R] + c_l \Delta y_l) f(\theta)d\theta \right) - \int_{D_R(y, \theta) \geq 1} E[C_R(y, \theta)/R] f(\theta)d\theta \]

(33)

As the relation \( P[D_R(y, \theta) \geq b] \) has been approximated using an exponential function (cf. eq. (30)), the sought sensitivity can be further simplified.

\[
\frac{\partial E[C_R(y, \theta)]}{\partial y_l} \approx -a_l \psi P[D_R(y, \theta) \geq 1] E[C_R(y, \theta)/R] + c_l P[D_R(y, \theta) \geq 1]
\]

(34)

Thus, eq. (34) provides an explicit approximation of the sensitivity of the (approximate) expected repair cost with respect to the design variables. It is most interesting to note that this expression for sensitivity depends on the coefficient \( \psi \), which is obtained from a standard reliability analysis using SS, and on \( a_l, c_l \), \( l = 1, 2 \), which are determined using a least squares procedure. That is, a single reliability analysis plus some additional computations of the normalized demand and cost functions (for perturbed values of the design variable vector) suffice for generating an approximate estimate of the sensitivity of \( E[C_R(y, \theta)] \). Within the scope of this contribution, this is a most remarkable feature, as the total number of simulations of the crack propagation phenomenon can be drastically reduced, thus increasing the efficiency of the proposed approach.

Finally, it must be pointed out that according to the numerical experience of the authors, the approximation introduced in the eq. (32) has very little influence in estimating the sought sensitivity. That is, the expected repair cost conditioned on the repair event does not vary significantly due to small changes in the values of the design variables (quality of inspection and time of inspection). However, the approximation is still presented in the paper for the sake of completeness, as it is not possible to ensure that in every specific situation being analyzed, the aforementioned observation is fulfilled.

5.3. Sensitivity of Expected Failure Cost

The calculation of the sensitivity of the expected failure cost can be considerably more involved than its counterpart related with the repair event. This is due to the fact that the normalized demand function associated with the failure event may present considerable discontinuities due to the repair activities. For a better understanding of this idea, consider the following example: for a certain realization of the design variables and uncertain parameters, a mechanical component fails (i.e. \( D_F(y, \theta) > 1 \)) and the repair event does not take place; however, \( D_R(y, \theta) \) is quite close to 1. Then, in case the maintenance schedule is slightly altered (e.g. a higher inspection quality is used) but all the other parameters remain constant, it may occur that repair takes place and that the value of the normalized demand function associated with failure is much smaller than 1. The fact that \( D_F(y, \theta) \)
may present considerable discontinuities implies that the linear approximation introduced in eq. (31) for the normalized demand function associated with repair can not be applied in the case of failure. Thus, the scheme for calculating sensitivities developed for the expected repair cost is not directly applicable in the case of the failure event.

In order to estimate the sensitivity of the expected failure cost, a number of approximation concepts are introduced. These approximation concepts allow constructing an approximate representation of the expected failure cost that is differentiable. First, consider the definition of an auxiliary failure event $F_0$. This event refers to the failure of the mechanical component without considering any inspection and repair, i.e. $F_0$ is independent of the value of $y$. In mathematical terms, $F_0$ can be defined as follows.

\[ F_0 = \{ \theta \in \Omega_\theta : D_F(0, \theta) \geq 1 \} \]  

In the last equation, the null vector $0$ denotes that no maintenance activities are scheduled. Considering this auxiliary event, it is possible to define a rough, auxiliary approximation of the expected failure cost; this approximation is denoted as $E \left[ C_A^F(y, \theta) \right]$ and is defined in the equation below.

\[ E \left[ C_A^F(y, \theta) \right] = \int_{D_R(y, \theta) < 1, \theta \in F_0} C_F(\theta) f(\theta) d\theta \]  

In eq. (36), the cost function $C_F(\cdot)$ does not depend on $y$, as repair never takes place. The definition of $E \left[ C_A^F(y, \theta) \right]$ includes all those realization of the uncertain parameters that lead to failure and that do not imply repair. It should be noted that these realizations constitute a subset of the realizations of the uncertain parameters that lead to the failure event $F$. The difference between these two sets are those samples of the uncertain parameters for which repair takes place but failure still occurs. From a physical point of view, the difference between these sets arises in those cases where some of the cracks present in a mechanical component are repaired but the remaining ones still lead to failure. Although the auxiliary approximation of the expected failure cost $E \left[ C_A^F(y, \theta) \right]$ is not equal $E \left[ C_F(y, \theta) \right]$ (due to the difference discussed above), it should be noted that still these two expected costs correlate well. Therefore, the sensitivity measure of $E \left[ C_A^F(y, \theta) \right]$ is used to approximate the sensitivity of $E \left[ C_F(y, \theta) \right]$. Numerical validation has indicated that such approximation is appropriate. Moreover, the sensitivity of the auxiliary expected failure cost can be calculated using the approach developed in Section 5.2. Thus, for computing an approximation of the sensitivity of the expected failure cost, the following approximations are introduced.

- An auxiliary demand function $D_R^*(y, \theta)$ is defined, such that $D_R^*(y, \theta) = 2 - D_R(y, \theta)$.
- The probability that $D_F(y, \theta)$ exceeds a threshold $b$ is approximated by means of an exponential function, i.e.:

\[ P \left[ D_F(y, \theta) \geq b \right] \approx P \left[ D_F(y, \theta) \geq 1 \right] e^{\psi^*(b-1)} \]  

(37)
where $\psi^*$ is a constant which is determined using the samples generated at the last stage of SS (see Section 4).

- The auxiliary demand function at a perturbed design $y^*$ is approximated as the auxiliary demand function at a nominal design $y$ plus a linear combination of the vector $(y^* - y)$, i.e.;

$$D_R^*(y^*, \theta) \approx D_R^*(y, \theta) + \sum_{l=1}^2 a_l^*(y_l^* - y_l) \tag{38}$$

where $a_l^*$, $l = 1, 2$ are real coefficients calculated using a least squares approach, as indicated in Appendix D.

- As the cost function associated with failure in eq. (36) does no depend on the design variables, it is set equal to $E[C_F(y, \theta)]/F$. Note that this value is already available from the simulation step using SS.

Using an argumentation similar to the one described in Section 5.2, it is possible to derive the following approximate expression for estimating the sensitivity of the expected failure cost.

$$\frac{\partial E[C_F(y, \theta)]}{\partial y_l} \approx \frac{\partial E[C_A^*(y, \theta)]}{\partial y_l} \approx -a_l^* \psi P[D_F(y, \theta) \geq 1] E[C_F(y, \theta)/F] \tag{39}$$

6. Solution Strategy

This Section presents the algorithm applied for solving the problem posed in eq. (1) that is related with the optimal design of a maintenance schedule considering uncertainties. In addition, some issues on the practical implementation of the solution scheme are discussed as well.

6.1. Strategy for Solving the RBO problem

For solving the problem posed in eq. (1), an optimization algorithm based on descent feasible directions is applied. This class of algorithm has been widely applied in the field of deterministic structural optimization (see, e.g. [3, 29]) and its application in RBO problems has been just recently investigated in [31]. This algorithm is – by construction – monotonically convergent. This feature is most advantageous when solving challenging problems, as the optimization process can be discontinued at any stage, still leading to feasible designs which are better than the initial one. The optimization algorithm can be summarized in two steps, namely the identification of a descent feasible direction and a line search step (along the descent direction). Each of these steps is briefly described below in Sections 6.2 and 6.3, respectively. For more details on the optimization algorithm and the implementation of its two steps, it is referred to [31].

20
6.2. Selection of Descent Feasible Direction

Given a feasible design \( y^k \) (i.e. a design that fulfills the constraints of the optimization problem), a descent feasible direction \( d \) is defined such that for a sufficiently small step size \( \alpha \), it is ensured that \( y^k + \alpha d \) is feasible and that \( E[C_T(y^k + \alpha d, \theta)] < E[C_T(y^k, \theta)] \). The issue on how to select a descent feasible direction depends on the particulars of \( y^k \), i.e. in case the current feasible design \( y^k \) is an interior point (i.e. no constraints are active), a possible choice for \( d \) is the steepest descent direction \(-\nabla E[C_T(y^k, \theta)]\) (i.e. the negative of the gradient of the expected total cost function); in case one or more constraints are active at the current feasible design, \( d \) can be determined by solving a linear programming (LP) problem involving the gradients of the expected total cost function and the active constraints [3]. In any of the two aforementioned cases, the sensitivities of the different functions involved in eq. (1) are required. The sensitivities of the expected cost functions of inspection, repair and failure can be calculated using the formulations presented in Section 5, while the sensitivity of the deterministic constraint functions can be calculated using any appropriate procedure, e.g. finite differences.

In order to visualize how a descent feasible direction \( d \) is determined, consider the schematic representation in fig. (6). At the initial design \( y^1 \), no constraints are active; therefore, \( d \) is set equal to \(-\nabla E[C_T(y^1, \theta)]\). At \( y^2 \), a nonlinear constraint is active and \( d \) is selected such that the angle between \( d \) and \(-\nabla E[C_T(y^2, \theta)]\) is acute and the angle between \( d \) and \( \nabla h_1(y^2) \) is obtuse. Finally, at \( y^3 \), a linear constraint is active; in this case, \( d \) is obtained directly as the projection of \(-\nabla E[C_T(y^3, \theta)]\) over the plane orthogonal to \( \nabla h_2(y^3) \).
6.3. Line Search

Once a descent feasible direction has been determined, the next step is to perform a line search in the single-dimensional space defined by $d$ in order to solve the following optimization problem.

$$\min_{\alpha} \ E[C_T(y^k + \alpha d, \theta)]$$

subject to

$$h_j(y^k + \alpha d) \leq 0, \quad j = 1, \ldots, n_{DC}$$

(40)

First, the step $\alpha$ to the nearest intersecting boundary defined by the constraints of the problem is determined using any appropriate search scheme [3], e.g. by means of bisection. In this context, it is assumed that the evaluation of $h_j(\cdot), \ j = 1, \ldots, n_{DC}$ is numerically inexpensive and, thus, the determination of $\alpha$ is straightforward. Then, for determining the optimal step $\alpha^{kopt}$ ($\alpha^{kopt} \in [0, \bar{\alpha}]$) that solves the optimization problem in eq. (40), it is necessary to compute the expected total cost function at several trial points; however, the evaluation of this function is numerically involved, as it is necessary to compute multi-dimensional integrals associated with expected cost, as discussed in Section 4. In order to avoid computing $E[C_T(y^k + \alpha d, \theta)]$ repeatedly for different values of $\alpha$, the expected total cost function along the search direction is replaced with a polynomial approximation, i.e.:

$$E[C_T(y^k + \alpha d, \theta)] \approx c_0 + c_1 \alpha + c_2 \alpha^2, \quad \alpha \in [0, \bar{\alpha}]$$

(41)

where $c_k, \ k = 0, 1, 2$ are real coefficients. These coefficients are determined using a least square procedure, i.e. the expected total cost function is evaluated using the approaches described in Section 4 in a grid of $n_p$ points ($n_p \geq 3$) in order to determine the sought coefficients; in this contribution, $n_p$ is selected equal to 4. Using the approximation in eq. (41), the value of the optimal step $\alpha^{kopt}$ can be determined using, e.g. a bisection scheme.

Validation calculations have indicated that the application of the approximate representation in eq. (41) can be very effective. However, it has been observed that generating an approximation valid over the range $[0, \bar{\alpha}]$ can be challenging in case $\bar{\alpha}$ is too large, i.e. there is a considerable distance between the current optimal design and the active constraint(s) along the search direction. A possibility for solving this issue is reducing the search space to a value smaller than smaller than $\bar{\alpha}$. For determining such value, e.g. a percentage of the euclidean norm of the current optimal design could be selected. This approach has been found a simple yet effective strategy, e.g. in the first steps of optimization, percentages around 80% have provided satisfactory results.

6.4. Additional Implementation Issues

As already discussed in Section 4, the expected costs associated with repair and failure are calculated using Subset Simulation (SS). Although SS is a very efficient and general simulation technique,
one of its major disadvantages is that the variability of the estimates produced can be considerable [5]. Within the context of RBO, this variability may even prevent determining an optimal design solution. For example, large variability of the estimates of expected cost can render a line search strategy useless.

A possible means for coping with the variability of the estimates of SS is the application of common random numbers (CRN) and smoothing of indicator functions. The effectiveness of both of these strategies was investigated within the context of RBO in [65]. However, it should be noted that these strategies do not eliminate the variability of the estimates. Instead, they introduce a consistent estimation error. That is, although the estimates (of, e.g. reliability) for two different sets of the design variables still contain error, these estimates are still comparable between them.

Similar strategies to the ones proposed in [65] are adopted in this contribution in order to cope with the variability of the estimates of SS. The application of CRN implies that the same stream of random numbers is used for evaluating expected costs associated with two different maintenance schedules $y^{k_1}$ and $y^{k_2}$. The use of a smooth indicator function implies that the expressions for estimating probability in eqs. (17) and (27), respectively, are modified. More specifically, the indicator function is replaced by the cumulative density function of a normal distribution ($\Phi(\cdot)$) with mean $\mu = 1$ and a small standard deviation, e.g. $\sigma = 0.01$; the argument of $\Phi(\cdot)$ is the associated normalized demand function. This formulation is shown below.

$$p_R(y^k) = \int_{\theta \in \Omega_\theta} \Phi(D_R(y^k, \theta)) f(\theta) d\theta$$

$$p_F(y^k) = \int_{\theta \in \Omega_\theta} \Phi(D_F(y^k, \theta)) f(\theta) d\theta$$

7. Example

7.1. Description

In this section, the design of an optimal maintenance schedule for a fatigue-prone structural element is considered. The model involves a symmetric plate including four identical rivet holes under axial tension, as shown in fig. (7).

The load over the element is such that the range of the far-field stress is $\Delta \sigma = 90$ [MPa] and the far-field maximum stress, $\sigma_{\text{max}} = 100$ [MPa]; during one year of operation, a total of $1.2 \times 10^5$ load cycles are applied.

The plate must endure a life period of 10 years. It is assumed that at the beginning of this period, the element has already developed fatigue cracks; more specifically, there are a total of 8 cracks emanating from the rivet holes. A schematic representation of these cracks emanating from two rivet holes is shown in fig. (4). The uncertainty in the length of these initial cracks ($a_{0,i}$, $i = 1, \ldots, 8$) is characterized by means of a log-normal distribution with expected value 1 [mm] and standard
deviation 0.5 [mm], i.e. $a_{0,i} \sim \text{LN}(1, 0.5)$ [mm]. For the sake of simplicity, the random variables modeling the initial crack length are assumed to be uncorrelated; however, it has been observed that there is a considerable correlation between the length of cracks emanating from the same hole (see, e.g. [60]). The parameters of the Paris-Erdogan law are modeled such that $m = 2$ and $C \sim \text{LN}(2.5 \times 10^{-23}, 2.5 \times 10^{-24})$ [Pa$^{-2}$cycles$^{-1}$]. The stress intensity factors are calculated using the FEAM (see Appendix B).

Concerning the parameters involved in the failure criterion (cf. eq. (5)), the fracture toughness is characterized by a log-normal distribution, i.e. $K_{IC} \sim \text{LN}(80, 8)$ [MPa m$^{-0.5}$]. The collapse stress is given by the relation $\sigma_C = \sigma_Y (1 - (0.016 + \sum_{i=1}^{8} a_i)/128)$, where $\sigma_Y$ is the yield stress ($\sigma_Y \sim \text{LN}(325, 32.5)$ [MPa]).

The parameter modeling the probability of detection (cf. eq. (7)) is set equal to $\lambda = 100$ [1/m]. The error in the measurement of the crack length is modeled as a random variable, i.e. $\theta_e \sim \text{LN}(0.5, 0.1)$ [mm]. A crack is repaired provided that its measured length exceeds $a_{cr} = 5$ [mm].

Finally, the costs of inspection, repair and failure are set equal to $c_I(q) = 10q$ [MU] (where [MU] denotes monetary units), $c_R = 300$ [MU] and $c_F = 2 \times 10^5$ [MU], respectively; the discount rate is assumed to be $r = 2\%$.

The RBO problem associated with the design of a maintenance schedule is formulated as:

$$\min_{(q, T_I)} E \left[ C_T(q, T_I, \theta) \right]$$

subject to

$$1 \leq q \leq 5, \ 1 \leq T_I \leq 9$$  \hspace{1cm} (44)
where $T_I$ is measured in years. The expected cost associated with repair and failure is estimated using SS; a total of 1000 samples are used at each stage of this advanced simulation method.

7.2. Results and Discussion

Prior to the design of the optimal maintenance strategy, the dependence of $p_R$ and $p_F$ on the design parameters is analyzed for the particular problem under study. For this purpose, the value of these probabilities is estimated over a suitable grid, as shown in figs. (8) and (9), respectively. Concerning the repair activities (cf. fig. (8)), it is noted that $p_R$ increases with a higher quality of inspection, as chances of detecting a particular crack become larger; $p_R$ also increases with the time of inspection, as the average crack size increases with time (note that generally a larger crack has a higher probability of being detected).

![Figure 8: Probability of repair ($p_R$) as a function of quality of inspection ($q$) and the time of inspection ($T_I$)](image)

Regarding the failure event (cf. fig. 9), it can be observed that $p_F$ decreases with an increase of the quality of inspection because there is a higher probability of repair; with respect to the time of inspection, $p_F$ increases for both an early inspection time (because cracks are too small to be detected) and a late inspection time (because failure occurs before inspection). These observations are in agreement with other contributions in the literature (see, e.g. [28, 77]), i.e. the effectiveness of early inspection is quite limited, as cracks are too small and thus, difficult to detect; therefore, the beneficial effect of maintenance activities is lost.

In a next step, the RBO problem formulated in eq. (44) is solved using the approach described in Section 6 (see fig. (10)). Starting from an initial, feasible design $\mathbf{y} = (q, T_I)^T = (1.0, 2.0)^T$, it is possible to find a maintenance schedule which is in the vicinity of the optimum after only two iterations (i.e. two cycles of determination of a descent direction followed by a line search step) at $\mathbf{y}^{opt} = (4.0, 5.6)^T$. In this context, it is important to note that the region where the optimal solution lies is quite flat (see also, e.g. [11, 23]). Thus, although the exact location of the optimum is not determined, this is
not detrimental for the proposed approach for solving RBO problems, as the value of the objective function can not be improved substantially.

Based on the results presented in fig. (10), it can be noted that the time at which inspection is performed has a large impact in $E[C_T]$; in case inspections take place either too early or too late, the effectiveness of maintenance activities is completely lost. In addition, it can be observed that the proposed approach for solving RBO problems is most efficient, as the algorithm arrives to the vicinity of the optimum after two iterations. In this context, the monotonic convergence properties of the optimization algorithm are most advantageous, as the optimization algorithm may be interrupted at any iteration, still leading to an improved design.

As it can be also noted from fig. (10), the optimal solution is a compromise between costs of inspection, repair and failure. That is, in the optimal solution, the quality of inspection is such that the cost of
inspection is neither a maximum nor a minimum while the cost associated with repair and failure are kept at an intermediate value. Moreover – for the example under study – the optimal solution does not coincide with the maintenance schedule that minimizes the probability of failure, highlighting the importance of modeling repair activities explicitly, which correspond to partial damage states [40]. In order to illustrate this last point, fig. (11) presents a comparison between the expected total costs for different maintenance schemes. In this figure, each bar represents the expected total cost, while the colors within each bar indicate the portion of the expected total cost due to inspection, repair and failure, respectively. When no maintenance activities are performed (see bar 1), the expected total cost equals the expected failure cost. Moreover, this expected cost largely exceeds largely the cost when maintenance activities are performed. The expected total costs in bar 2 show the case when the expected failure cost is minimized individually. Although in this case the total cost is much lower than in the case of no maintenance, it is still non optimal, as the interaction with the costs of inspection and repair are not considered. On the contrary, in bar 3, the total cost is minimized (i.e. the summation of the cost of inspection, repair and failure) yielding a solution which is optimal.

![Figure 11: Expected total cost and individual costs for different maintenance strategies: (1) no maintenance (2) minimization of expected failure cost (3) minimization of expected total cost](image)

8. Conclusions and Outlook

In this contribution, an approach for optimal maintenance scheduling of fatigue prone components and structures considering uncertainties has been presented. Starting from an initial, feasible maintenance schedule, it is possible to determine a series of steadily improved schedules. The key ingredients of the proposed strategy are the application of an advanced simulation algorithm for estimating ex-
pected costs, the determination of a feasible descent direction (by means of an efficient algorithm for estimating sensitivity of the expected cost) and the application of a line search step. In this way, the number total number of simulations of the crack propagation phenomenon – which is a critical issue from the point of view of efficiency – can be considerably reduced.

The application of an advanced simulation technique showed that it is possible not only to account for the effects of uncertainty in the crack propagation phenomenon but also to consider several failure criteria. This is of paramount relevance, as in real structures, different failure modes may occur. Moreover, the simulation scheme can be applied independently of the model used to quantify the crack propagation phenomenon, i.e. in principle, the proposed approach is not limited to elastic fracture mechanics and the Paris-Erdogan law.

The scheme for assessing the sensitivity of the expected cost is particularly well suited for the problem at hand. Its major advantage is that no additional runs of the simulation algorithm are required. That is, due to the approximations introduced, it is necessary to compute only a few additional realizations of the normalized demand functions associated with the repair and failure event.

The numerical example is most instrumental in highlighting the importance of optimization tools in context with the design of an optimal maintenance schedule. The results indicate that the optimal maintenance strategy is not necessarily the one that minimizes the costs associated with failure; in fact, the optimal solution is a tradeoff between several different economical factors. Moreover, the time at which an inspection is performed plays a key role in the overall effectiveness of the maintenance strategy. Thus, inspection activities - even those of highest quality - can be totally useless (although expensive) if scheduled at an inappropriate time.

Future research directions include the extension of the proposed approach for more general cases. This includes, e.g. the possibility of modeling several cycles of inspection and repair activities throughout the life time of a mechanical component; consideration of imperfect repair activities and inclusion of the crack initiation phase when characterizing fatigue life.

Acknowledgment

This research was partially supported by the Austrian Research Council (FWF) under Project No. P20251-N13 which is gratefully acknowledged by the authors. The help of Mr. P. Beaurepaire in carrying out part of the calculations of the numerical example is deeply appreciated.

A. Brief Literature Survey: Crack Propagation and Maintenance Activities

Stochastic fatigue growth has been studied using different approaches, e.g. Markov-chain models [12, 62], using time-dependent noise terms [80] or random coefficients [42], bilinear crack growth models [6], random field models [54], particle filtering techniques (see, e.g. [13]). Uncertainty in inspection of cracks, its influence in reliability and the beneficial aspects of repair activities have also received
considerable attention in the literature (see, e.g. [18, 30, 46, 55, 60, 70, 79]). Reliability assessment of mechanical components in view of crack propagation has been addressed by means of Monte Carlo Simulation (see, e.g. [10, 35, 53, 60]), First- and Second-order reliability method (see, e.g. [41, 42]), Importance Sampling (see, e.g. [51, 52]), etc. Finally, the problem of maintenance scheduling has been solved using, e.g. Markov chain models [61], First-order reliability method in combination with a non-linear programming algorithm (see, e.g. [1, 15, 32, 43, 68]), Monte Carlo simulation (see, e.g. [28, 74]), diffusive crack growth models [66], introducing appropriate simplifications in inspection intervals and reliability estimation (see, e.g. [11, 22]), analytical or semi-analytical models [38, 77], computing sensitivities efficiently [23], application of Importance Sampling in combination with the response surface method [56], etc.

B. Finite Element Alternating Method

The FEAM allows calculating the SIF’s most efficiently by linear superposition of two solutions. The SIF’s are estimated in the cracked component by iterating between the analytical solution for an infinite body with a crack and the finite element solution for the uncracked component. The iteration involves the residual tractions closing the crack in the uncracked body and the far field stresses in the boundary of the infinite body. At the end of the iteration, the subtraction of the stresses in the uncracked component and the infinite body allows obtaining the solution for the target cracked component. This idea is represented schematically in fig. (12), for the case of a plate with a single crack.

![Figure 12: Basic concept of the Finite Element Alternating Method](image)

A salient feature of the FEAM is that the stiffness matrix (associated with the finite element solution of the uncracked body) must be computed only once. For details on the implementation of the FEAM, it is referred to [72, 73].
C. Linear Approximations for Estimating Sensitivity of Expected Repair Function

The linear expansion of the normalized demand provides an approximation of the function $D_R(y^*, \theta)$ (see eq. (31)). The coefficients $a_l$, $l = 1, 2$ of this expansion can be determined, e.g. in a least square sense. This implies the minimization of square error $S^2$ between the exact and approximate models, which are evaluated at a series of realizations of the vector of design variables and uncertain parameters, i.e.:

$$S^2 = \sum_{s=1}^{N_S} \left( D_R \left( y^{(s)}, \theta^{(s)} \right) - D_R \left( y, \theta \right) - \sum_{l=1}^{2} a_l (y_l^{(s)} - y_l^k) \right)^2$$  \hspace{1cm} (45)

In eq. (45), the pair of vectors $(y^{(s)}, \theta^{(s)})$, $s = 1, \ldots, N_S$ are realizations of the vector of design variables and of the uncertain parameters, respectively; moreover, $y_l$, $l = 1, 2$ denotes the $l$-th component of the vector $y$.

The details of the algorithm for calibration of the coefficients $a_l$, $l = 1, 2$ are described below (for a detailed description of this algorithm, it is referred to [31]).

1. Create a set of $N_S$ samples of the uncertain parameters such that $D_R(y, \theta^{(s)}) \approx 1$, $s = 1, \ldots, N_S$. Note that these samples can be selected among those samples generated at the last stage of SS (see Section 4) used to calculate the expected repair cost. Numerical experience indicates that a suitable number of samples is $N_S = 100$ [31]. Set $l = 1$.

2. Generate a set of $N_S$ realizations of the vector of design variables, i.e. $y^{(s)}$, $s = 1, \ldots, N_S$; each vector $y^{(s)}$ must be equal to $y$, except the $l$-th component $y_l^{(s)}$. This component is defined by introducing a small perturbation, i.e. $y_l^{(s)} = (1 + \xi_l^{(s)}) y_l$, $s = 1, \ldots, N_S$, where $\xi_l^{(s)}$, $s = 1, \ldots, N_S$ are independent Gaussian random variables with mean $\mu = 0$ and a small standard deviation $\sigma$, e.g. $\sigma = 0.01$.

3. Evaluate $D_R \left( y^{(s)}, \theta^{(s)} \right)$, $s = 1, \ldots, N_S$.

4. Estimate the coefficient $a_l$ in a least squares sense, i.e.:

$$a_l = \frac{\sum_{s=1}^{N_S} (y_l^{(s)} - y_l) \left( D_R \left( y^{(s)}, \theta^{(s)} \right) - D_R \left( y, \theta^{(s)} \right) \right)}{\sum_{s=1}^{N_S} (y_l^{(s)} - y_l)^2}$$ \hspace{1cm} (46)

5. In case $l = n_y$, stop the algorithm. Otherwise, set $l = l + 1$ and return to step 2.

For generating an approximation of the function associated with the cost of repair (cf. eqs. (16) and (32)), a procedure similar to the one described above is followed, except that the residual is formulated in terms of $C_R(\cdot, \cdot)$ instead of $D_R(\cdot, \cdot)$. In fact, the same samples used to calibrate $a_l$, $l = 1, 2$ can be used to determine $c_l$, $l = 1, 2$. However, the approximation is calibrated using samples of the design variables and uncertain parameters that belong to the repair event (cf. (20)). Therefore, samples
which are not included in the repair event must be excluded from steps 1 and 4 of the algorithm described above.

D. Linear Approximation of Auxiliary Demand Function

The procedure for constructing an approximation of the auxiliary demand function (cf. eq. (38)) is similar to the one outlined in Appendix C, except by the following details.

- The residual $S$ is formulated in terms of $D_{R^*}((,\cdot))$ instead of $D_R((,\cdot))$.
- In step 1, samples of the uncertain parameters are selected such that $D_F(y,\theta^{(s)}) \geq 1$ and $D_R(y,\theta^{(s)}) < 1$, $s = 1,\ldots,N_S$; these samples are chosen from the last stage of SS.
- In step 4, the samples of the design variables and uncertain parameters which do not fulfill the conditions $D_F(y^{(s)},\theta^{(s)}) \geq 1$ and $D_R(y^{(s)},\theta^{(s)}) < 1$ should be discarded.

References


Munich, Germany, EU, 1999.


