Prior Robustness for Bayesian implementation of the Fault Tree Analysis

Chaitanya Joshi, Fabrizio Ruggeri and Simon P. Wilson

***Abstract*—We propose a prior robustness approach for the Bayesian implementation of the fault tree analysis (FTA). FTA is often used to evaluate risk in large, safety critical systems but has limitations due to its static structure. Bayesian approaches have been proposed as a superior alternative to it, however, this involves prior elicitation, which is not straightforward. We show that minor mis-specification of priors for elementary events can result in a significant prior mis-specification for the top event. A large amount of data is required to correctly update a mis-specified prior and such data may not be available for many complex, safety critical systems. In such cases, prior mis-specification equals posterior mis-specification. Therefore, there is a need to develop a robustness approach for FTA which can quantify the effects of prior mis-specification on the posterior analysis.**

Here, we propose the first prior robustness approach specifically developed for FTA. We not only prove a few important mathematical properties of this approach, but also develop easy to use Monte Carlo sampling algorithms to implement this approach on any given fault tree with AND and/or OR gates. We then implement this Bayesian robustness approach on two real life examples: a spacecraft re-entry example and a feeding control system example. We also provide a step-by-step illustration of how this approach can be applied to a real life problem.

***Index Terms*—Fault tree analysis, prior elicitation, distorted band of priors, Bayesian robustness, Bayesian networks.**

NOMENCLATURE

AHP	Analytic Hierarchy Process.
BN	Bayesian Networks.
FT	Fault Tree.
FTA	Fault Tree Analysis.
TE	Top event.
w.r.t.	with respect to.

I. INTRODUCTION

FTA is often used to quantify the probability of occurrence of an undesirable event, namely the TE. FTs are constructed in a top-down fashion from the TE to their causes - represented by the intermediate and elementary events in the tree. The relationship between the events and the causes is represented using logical gates, most commonly AND and OR gates. The basic assumptions of the standard FTA are: (i) events are Bernoulli trials e.g. either happen or do not happen; (ii) events are statistically independent; (iii) the relationship between events are represented using logical AND and OR gates and (iv) the root of the FT is the undesirable TE to be analyzed [1].

C. Joshi is with the Department of Mathematics and Statistics, University of Waikato, Hamilton,3240 New Zealand e-mail: cjoshi@waikato.ac.nz

F. Ruggeri is with CNR IMATI, Milan, Italy and S.P. Wilson is with Discipline of Statistics, Trinity College Dublin, Dublin, Ireland.

Manuscript received April 17, 2017

A. FTA and Bayesian Networks

One of the common criticisms [2] of the FTA is that the events are considered independent of each other, which can be an unrealistic assumption. BN have been recommended as a way to incorporate local dependence between events, as well as providing a framework for both forward (prediction) and backward (inference) analyses; see for example [1], [2] and [3]. An FTA can be directly mapped into a BN and that permits the assumptions (i) - (iii) to be relaxed under the BN formalism [1]. Additionally, BN can model uncertainty in an outcome - instead of the deterministic outcomes implied by the logical gates in an FTA. In other words, BN can be seen as a natural extension of FTA [3].

However, BN have their own drawbacks: (1) the typical BN approach requires specifying exact prior probabilities (not just the prior distributions) of the elementary events and (2) it also requires specification of all the conditional probabilities in the system. In many cases, especially when designing new systems, sufficient prior information is not available and hence it may not be possible to even specify exact prior probabilities let alone all the conditional probabilities accurately. As the number of nodes and the dependency structures increase, specifying the information needed by BN accurately becomes increasingly unrealistic. FTA is relatively easier to implement. Also, in many cases (including the Examples IV-B and IV-C in this paper), the flexibility offered by BN is not needed. Unrelated events may be independent and hence dependence structures may not be required. Similarly, when the TE occurs for sure (deterministically) once the required elementary/intermediate events have occurred, modeling uncertainty in the outcomes is not necessary. Another disadvantage of BN is that incorporating a previously unknown part of the network becomes computationally very expensive as the size and the complexity of the network grows (see, for example, [4]). On the other hand, this is relatively easily done for FTA.

B. Fully Bayesian implementation of FTA

For these reasons, when the system satisfies the assumptions (i) - (iv) for FTA, it may be better to analyze it using a fully Bayesian treatment of the FTA. This way, one does not need to specify the exact prior probabilities and conditional probabilities, but at the same time, can incorporate the

uncertainty in the prior knowledge which can not be done in the typical FTA. Another advantage of the Bayesian generalization of FTA is that it permits uncertainty in and learning about the event probabilities to be incorporated. Prior distributions on the probabilities of elementary events are elicited and then the logic of the fault tree can be used to derive the prior of any intermediate event and the TE. These distributions are updated on observation of the TE or other events through Bayes law. Further, the posterior uncertainties can be easily integrated out by finding the posterior summary statistics such as a posterior mean or a posterior credible interval for the probability of the TE.

This approach was applied in [5], who described a process to elicit Beta distribution priors on the elementary event probabilities, and used simulations to derive the prior distribution for the probability of the TE from those of the elementary events. [5] also described a method for updating these distributions in light of observation of any event or combination of events in the fault tree and applied this approach to the spacecraft re-entry problem (see Section IV-B).

C. Prior elicitation for FTA

Eliciting prior distributions for the elementary events by quantifying expert opinion is usually not straightforward. Unfortunately, this area has received very little attention relative to modeling and the practicalities of implementing Bayesian inference. The few papers that do propose Bayesian elicitation methods for FTA such as [5], [6] and [7] propose *moment matching* approaches. The moment matching approach is one of the most common prior elicitation approaches (see, for e.g. [8]) and relies on elicitation of a small number of statistics. A prior distribution is then elicited by finding the parameter values for which the population moments match very closely with the elicited moments. For an accurate elicitation of the prior distribution, the number of statistics must be at least equal to the number of unknown parameters of the distribution. A typical choice of statistics includes measures of the central tendency as well as measures of the variability/spread. The exact choice of statistics often depends on the information available (i.e what can be elicited) as well as the distribution concerned. Since we are concerned with FTA where, typically, each event is a binary event, a standard choice for the prior distribution would be the Beta distribution. In Examples IV-B and IV-C we illustrate how Beta distribution priors can be elicited in different ways. In Example IV-D we illustrate how the moment matching approach can be used to elicit Gamma distribution priors.

Another important approach to prior elicitation that has been used (see, for example [5]) in risk assessment is the use of pairwise comparisons. The prior for the probability of one event is elicited. For other events, their relative chance in comparison to this event is then elicited, and that is used to construct their prior distributions. Other pairwise comparisons may be made to validate the elicitation. The idea is that comparisons are the easiest judgment to make and so the process is simplified, see [9], [10].

D. Bayesian robustness for FTA

Both Bayesian approaches - the BN as well as the fully Bayesian FTA require specification of priors. As [5] points out for the spacecraft re-entry example though, for complex systems with very little data, getting experts to elicit even a mean value (for the probability of occurrence for an elementary event) could be an uphill task simply because of the lack of information. As a result, it is likely that the elicited priors are not entirely accurate. The prior probabilities can be updated using the Bayes theorem. However, if they are incorrectly specified then it can take a lot of data to update the prior probabilities in any meaningful way. Such a data may not be available firstly, since the TE often being an undesirable event, is designed to be highly unlikely to occur and secondly, because the data collection may be very complex and expensive. As we illustrate in Section IV, even the relatively small errors in prior elicitation of the elementary events can snowball into large distortions to the prior distribution of the TE, especially for fault trees with a series of OR gates or with multiple elementary events feeding into an OR gate. Erroneous priors can lead to an erroneous analysis which is particularly undesirable considering the applications for which FTA is typically used.

Therefore, it is important to develop Bayesian robustness approaches for BN and for the fully Bayesian implementation of the FTA. Bayesian robustness approaches can quantify the changes to the posterior distribution (or its functionals) given the changes to the prior distributions or to the data. Thus, it enables the computation of the distortion to a posterior distribution (or its functionals) as a result of any distortions to the prior distributions.

In this paper, we develop a Bayesian robustness approach for the fully Bayesian implementation of FTA. We use the new class of priors called the *distorted band of priors* developed by [11]. This class is based on stochastic ordering and distortion functions. The rest of this paper is organised as follows. In Section II, we provide a brief introduction to Bayesian robustness, describe the distorted band of priors and prove a few important properties of the distorted band obtained using power functions. We develop the robustness approach for analyzing the fault trees in Section III. In Section IV we illustrate the application of this approach to challenging real life problems. We close this paper with a summary and future work in Section V.

II. Bayesian robustness and class of priors

Bayesian robustness is the study of how sensitive the results of a Bayesian analysis are to its inputs, namely, the prior, data and the likelihood [12]. Often, a Bayesian robustness approach would be concerned with only one of the inputs. In this paper, we discuss robustness of a Bayesian implementation of the FTA w.r.t. the specification of its prior distributions. When studying the robustness w.r.t. the prior distribution, the assumption is that the subjective elicitation

of any single prior is in practice impossible and instead a class of priors is the natural alternative [13]. A class of priors can be defined in many ways. In this paper we propose to use the class of priors called the distorted band of priors [11]. We use this class because it is readily applied to the FTA, is easy to interpret and is also computationally easy to implement using the rejection sampling algorithms we develop in Section III-C. The distorted band of priors is described below. For a general introduction and details on Bayesian robustness, the reader is referred to [14].

A. The distorted band of priors

Before defining the distorted band of priors, we need to define a few related concepts. A *distortion function* h is a non-decreasing continuous function $h : [0,1] \longrightarrow [0,1]$ such that $h(0) = 0$ and $h(1) = 1$. When h is used to transform the distribution function F,

$$F_h(X) = h \circ F(x) = h[F(x)]$$

represents a perturbation of F in order to measure the uncertainty about it. Note that $F_h(X)$ is also a distribution function for a particular random variable denoted by X_h and the distorted density is given by

$$f_h(X) = h'[F(x)] \cdot f(x).$$

Random variables can be assigned ordering in different ways. For two random variables X and Y, X is said to be *smaller than* Y in the *stochastic order* sense (denoted by $X \leq_{st} Y$) if

$$F_X(t) \geq F_Y(t),\ \forall t \in \mathbb{R}.$$

For absolutely continuous [discrete] random variables X and Y with densities [discrete densities] f_X and f_Y, respectively, X is said to be *smaller than* Y in the *likelihood ratio order* sense (denoted by $X \leq_{lr} Y$) if

$$\frac{f_Y}{f_X} \text{ increases over the union of the supports of } X \text{ and } Y.$$

A desirable choice for distortion functions is to consider convex and concave functions. This is because it can be shown [11] that, if π is a specific prior belief with distribution function F_π and h is a convex (concave) distortion function in $[0,1]$, then $\pi \leq_{lr} (\geq_{lr}) \pi_h$. Thus, if the decision maker is able to represent the changes to a prior belief π by a concave distortion function h_1 and a convex distortion function h_2, then it leads him to two distorted distributions π_{h_1} and π_{h_2} such that $\pi_{h_1} \leq_{lr} \pi \leq_{lr} \pi_{h_2}$. This defines the class of priors called the distorted band of priors $\Gamma_{h_1,h_2,\pi}$ as

$$\Gamma_{h_1,h_2,\pi} = \{\pi' : \pi_{h_1} \leq_{lr} \pi' \leq_{lr} \pi_{h_2}\}. \tag{1}$$

It is evident that $\pi \in \Gamma_{h_1,h_2,\pi}$. Therefore the distorted band can be seen as a particular "neighbourhood" band of π. A popular choice for distortion functions h_1 and h_2 are *power functions* given by

$$h_1(x) = 1 - (1-x)^\alpha \text{ and } h_2(x) = x^\alpha,\ \forall \alpha > 1. \tag{2}$$

Note that if we take $\alpha = n \in \mathbb{N}$ in (2), then $F_{\pi_{h_1}}(\theta) = 1 - (1 - F_\pi(\theta))^n$ and $F_{\pi_{h_2}} = (F_\pi(\theta))^n$ which correspond to the distribution functions of the minimum and the maximum, respectively, of an i.i.d. random sample of size n from the baseline prior distribution π.

While the distorted bands obtained using (2) are centered around the prior belief π, the parameter α controls the width of the interval. Increasing α not only monotonically increases the width of the distorted bands but a distorted band obtained using a smaller α is completely contained inside the band obtained using a larger α. Here, we prove this intuitive but important property. Let $X \sim F(x)$, $X_{h_1}(\alpha_1)$ and $X_{h_2}(\alpha_1)$ be the random variables corresponding to $F_{h_1,\alpha_1}(x) = 1 - [1 - F(x)]^{\alpha_1}$ and $F_{h_2,\alpha_1}(x) = [F(x)]^{\alpha_1}$ respectively. Similarly, let $X_{h_1}(\alpha_2)$ and $X_{h_2}(\alpha_2)$ be the random variables corresponding to $F_{h_1,\alpha_2}(x) = 1 - [1 - F(x)]^{\alpha_2}$ and $F_{h_2,\alpha_2}(x) = [F(x)]^{\alpha_2}$ respectively. Also, let $f_{h_1,\alpha_1}(x), f_{h_2,\alpha_1}(x), f_{h_1,\alpha_2}(x)$ and $f_{h_2,\alpha_2}(x)$ be the corresponding densities.

Lemma II.1. *Given* $1 \leq \alpha_1 \leq \alpha_2 \Rightarrow X_{h_2}(\alpha_1) \leq_{st} X_{h_2}(\alpha_2)$ *and* $X_{h_1}(\alpha_1) \geq_{st} X_{h_1}(\alpha_2)$.

Proof. $1 \leq \alpha_1 \leq \alpha_2 \Rightarrow [F(x)]^{\alpha_1} \geq [F(x)]^{\alpha_2} \Rightarrow X_{h_2}(\alpha_1) \leq_{st} X_{h_2}(\alpha_2)$. Also, $[1-F(x)]^{\alpha_1} \geq [1-F(x)]^{\alpha_2} \Rightarrow 1 - [1 - F(x)]^{\alpha_1} \leq 1 - [1 - F(x)]^{\alpha_2} \Rightarrow X_{h_1}(\alpha_1) \geq_{st} X_{h_1}(\alpha_2)$. □

It is well known that

$$X \leq_{lr} Y \Rightarrow X \leq_{st} Y.$$

See for example [11], [15] and [16]. However, we will show here that when the distortion funtions are defined as in (2), the relationship between the likelihood ratio order and the stochastic order also holds in the other direction.

Lemma II.2. *Given* $1 \leq \alpha_1 \leq \alpha_2,\ X_{h_2}(\alpha_1) \leq_{st} X_{h_2}(\alpha_2) \Rightarrow X_{h_2}(\alpha_1) \leq_{lr} X_{h_2}(\alpha_2)$ *and* $X_{h_1}(\alpha_1) \geq_{st} X_{h_1}(\alpha_2) \Rightarrow X_{h_1}(\alpha_1) \geq_{lr} X_{h_1}(\alpha_2)$.

Proof.

$$\frac{f_{h_1,\alpha_1}(x)}{f_{h_1,\alpha_2}(x)} = \frac{\alpha_1 f(x)[1 - F(x)]^{\alpha_1 - 1}}{\alpha_2 f(x)[1 - F(x)]^{\alpha_2 - 1}} = \frac{\alpha_1}{\alpha_2}[1 - F(x)]^{\alpha_1 - \alpha_2},$$

which is an increasing function in x and hence $X_{h_1}(\alpha_1) \geq_{lr} X_{h_1}(\alpha_2)$. Similarly,

$$\frac{f_{h_2,\alpha_2}(x)}{f_{h_2,\alpha_1}(x)} = \frac{\alpha_2}{\alpha_1}[F(x)]^{\alpha_2 - \alpha_1},$$

which is an increasing function in x and hence $X_{h_2}(\alpha_1) \leq_{lr} X_{h_2}(\alpha_2)$. □

Theorem II.1. *When the distortion functions are defined as in (2), (i)*$X \leq_{st} Y \Rightarrow X \leq_{lr} Y$, *(ii)* $1 \leq \alpha_1 \leq \alpha_2 \Rightarrow \Gamma_{\alpha_1} \subset \Gamma_{\alpha_2}$ *and (iii)* $\Gamma_\alpha \to F(x)$ *as* $\alpha \downarrow 1$.

Proof. (i) From Lemma II.2. (ii) From Lemma II.1 and Lemma II.2.

(iii) By extending Theorem II.1 (ii) as a continuous function of α and that

$$\lim_{\alpha\to 1^+} 1-(1-F(x))^\alpha = F(x), \lim_{\alpha\to 1^+} F(x)^\alpha = F(x).$$

□

A very useful consequence of the likelihood ratio definition is that, if the two prior distributions are ordered in the $\leq_{lr}$ sense, then the corresponding posterior distributions are also ordered in the same sense, see [17]. That is, given two prior distributions π_1 and π_2 such that $\pi_1 \leq_{lr} \pi_2$, the corresponding posterior distributions also satisfy $\pi_{1x} \leq_{lr} \pi_{2x}$. Similarly, for all $\pi' \in \Gamma_{h_1,h_2,\pi}$, we obtain that $\pi_{h_1 x} \leq_{lr} \pi'_x \leq_{lr} \pi_{h_2 x}$. That is, the posteriors of the lower and upper bound of the prior distortion band are also the lower and upper bounds for the family of all posterior distributions, Γ_x, in the $\leq_{lr}$ sense. Γ_x can be considered as the distortion band of the posterior belief for some particular concave and convex functions.

B. Measuring uncertainty using a metric

Various distance measures can be used to compare the original prior and its distortions and also the original posterior and its distortions. The preference towards the use of a particular metric is usually based on their mathematical tractability and interpretation. It is for these reasons, that here, we prefer to use the Kolmogorov metric. It has a very practical interpretation and a concise mathematical formula that enables elicitation of α (see Section III-A).

The Kolmogorov metric measures the maximum absolute difference between the two distribution functions and is defined by

$$K(X,Y) = \sup_{x\in\mathbb{R}} |F_X(x) - F_Y(x)|. \tag{3}$$

It can be shown that if h is a differentiable (concave or convex) distortion function, then the Kolmogorov distance between π and π_h is given by

$$\begin{aligned} K(\pi,\pi_h) &= \sup_{x\in\mathbb{R}} |F_\pi(x) - F_{\pi_h}(x)| \\ &= \begin{cases} p_0 - h(p_0), & \text{if } h \text{ is convex,} \\ h(p_0) - p_0, & \text{if } h \text{ is concave,} \end{cases} \end{aligned} \tag{4}$$

where p_0 satisfies $h'(p_0) = 1$ and the argument of the maximum is achieved at $\theta_0 = F_\pi^{-1}(p_0)$. Further, it can also be shown that if the distortion functions are defined as in (2) then, the Kolmogorov metric is given by the following expression:

$$K(\pi,\pi_{h_1}) = K(\pi,\pi_{h_2}) = \frac{\alpha-1}{\sqrt[\alpha-1]{\alpha^\alpha}}. \tag{5}$$

III. Bayesian robustness approach for fault trees

Given that a prior distribution has been elicited for each of the elementary events (for the probability of the occurrence of that event/ probability of failure of that component), we propose the following approach to implement Bayesian robustness methods on fault trees. The details of how some of these steps can be achieved are discussed in the remainder of this section.

Bayesian robustness for FTA - an outline

- Build a distorted band of priors for each event - see III-A.
- Simulate through the FT using algorithms A1 - A4 (see III-B and III-C) to find the prior distribution and the distorted band of priors for the intermediate events and the TE.
- Find the posterior distribution for the TE given the prior distribution and the data.
- Find the lower and the upper distortion bands for the posterior distribution of the TE given the distorted bands for the prior and the data.

A. Use of the Kolmogorov metric

We propose to build the distorted band of distributions using the power functions described in (2). In practical applications, the main question is how to correctly choose the value of α. The Kolmogorov metric between the original distribution and the upper and the lower bounds of the distorted band obtained by the power function is given by (5). The value of α can be elicited by inverting this relationship. The Kolmogorov metric measures the maximum absolute discrepancy between any given pair of distributions. After having elicited the prior distribution, the expert may be prepared to guess how far off this prior distribution could be from the truth. For example, the expert may say that she expects the prior to be no more than 10% off the mark. This implies that $K(\pi,\pi_{h_1}) = K(\pi,\pi_{h_2}) \leq 0.1$ (since $0 \leq F(x) \leq 1$). Thus, 0.1 would be a conservative estimate of K. Using a computer program, it is now possible to find the value of α that yields $K = 0.1$ using (5).

Alternatively, a rough estimate of α can be obtained using the following approximation. Let $K(\pi,\pi_{h_1}) = K(\pi,\pi_{h_2}) = K$. Then,

$$\begin{aligned} K &= \frac{\alpha-1}{\sqrt[\alpha-1]{\alpha^\alpha}} \\ &\approx \frac{\alpha-1}{\alpha} \text{ assuming } \sqrt[\alpha-1]{\alpha^\alpha} \approx \alpha \\ \Rightarrow \alpha &\approx \frac{1}{1-K}. \end{aligned} \tag{6}$$

Note that this estimate of α obtained using (6) will be an optimistic estimate for a given value of K. In our case though, since the value of K itself was conservatively chosen, the estimate of α should be reasonably accurate. For example, suppose the maximum 10% discrepancy occurs when $F(x) = 0.6$ (in reality this would not be known beforehand), meaning that the true $K = 0.06$. The corresponding true value of α, obtained using a computer code is 1.175, whereas by using the conservative estimate $K = 0.1$ in (6) gives $\alpha = 1.1111$.

K represents how far off the elicited prior is likely to be in the worst case and it is difficult to imagine how this could be more than a 100%. Therefore K will take values < 1. In most cases, however, K is likely to take smaller values, say, between 0.05 to 0.3 corresponding to between 5% and

30% deviation from the true prior. (6) can be used as long as $0 \leq K < 1$. and (5) is valid for any $K > 0$ in principal.

B. Distortion bands for intermediate and top events

One important question that we want to possibly answer is that given the distortion bands for the elementary events what can we say about the distortion bands for the intermediate and the top events? Consider two simple fault trees illustrated in Fig. 1. Each one has three elementary events X_1, X_2 and X_3 and a top event Y. The difference is that in tree [a], the elementary events are linked to the top event using an OR operation whereas in [b] they are linked using an AND operation.

Assuming that each $X_i \sim$ Bernoulli(θ_i) and $Y \sim$Bernoulli(θ), where $\theta_i = P(X_i = 1)$, and $\theta = P(Y = 1)$, then $\theta = 1 - \prod_{i=1}^{3}(1 - \theta_i)$ for [a] and $\theta = \prod_{i=1}^{3} \theta_i$ for [b]. Let $\pi_i(\cdot)$ be the prior distribution on θ_i, $i = 1, 2, 3$, and $\pi(\cdot)$ be the prior distribution on θ. Let $\Gamma_i = (\pi_{h_{1i}}, \pi_{h_{2i}})$ be the distortion band for π_i and $\Gamma = (\pi_{h_1}, \pi_{h_2})$, the distortion band for π.

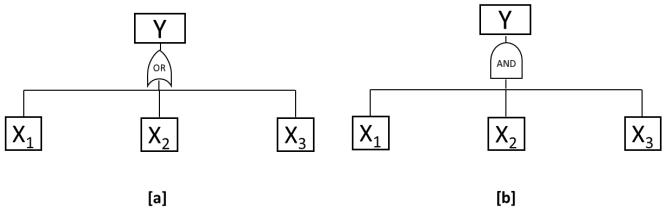

Fig. 1. Two simple fault trees (a) using an OR operation and (b) using an AND operation.

Then we have that

$$\pi(\theta = \tau) = \int_{\Omega_\tau^O} \pi_1(\tau_1)\pi_2(\tau_2)\pi_3(\tau_3)\, d\tau_1 d\tau_2 d\tau_3, \quad (7)$$

where $\Omega_\tau^O = \{\tau_i \in [0,1],\, i = 1,2,3 : 1 - \prod_{i=1}^{3}(1 - \tau_i) = \tau\}$ for the OR operation, and

$$\pi(\theta = \tau) = \int_{\Omega_\tau^A} \pi_1(\tau_1)\pi_2(\tau_2)\pi_3(\tau_3)\, d\tau_1 d\tau_2 d\tau_3, \quad (8)$$

where $\Omega_\tau^A = \{\tau_i \in [0,1],\, i = 1,2,3 : \prod_{i=1}^{3} \tau_i = \tau\}$ for the AND operation. Extending the logic, we would like to assume that for the distorted band of priors the following holds:

$$\pi_{h_1}(\theta = \tau) = \int_{\Omega_\tau^O} \pi_{h_{11}}(\tau_1)\pi_{h_{12}}(\tau_2)\pi_{h_{13}}(\tau_3)\, d\tau_1 d\tau_2 d\tau_3, \quad (9)$$

and

$$\pi_{h_2}(\theta = \tau) = \int_{\Omega_\tau^O} \pi_{h_{21}}(\tau_1)\pi_{h_{22}}(\tau_2)\pi_{h_{23}}(\tau_3)\, d\tau_1 d\tau_2 d\tau_3, \quad (10)$$

From Equations 7 and 8 it is clear that the prior $\pi(\cdot)$ for an intermediate or a top event will seldom be available in closed form. Mostly, the only way to obtain π will be by simulation. Similarly, the only way to obtain the distortion bands for the intermediate and the top events will be by simulation. Below we present algorithms to do these. For the purposes of the following algorithms, let θ denote the probability of failure. Also, we assume that there are K elementary events in the fault tree.

Algorithm $A1$: to simulate prior distributions for intermediate and top events

1) Sample $\theta_{ij} \sim \pi_i(\theta_i),\ j = 1, \ldots, N,\ i = 1, \ldots, K$.
2) For each j, run through the whole tree to find the probability of failure θ at the intermediate and the top events.
3) This generates the empirical prior distributions for the intermediate and the top events.

Algorithm $A2$: to simulate distortion bands for the prior distributions for intermediate and top events

1) Sample $\theta_{ij} \sim \pi_{h_{1i}}(\theta_i),\ j = 1, \ldots, N,\ i = 1, \ldots, K$.
2) For each j, run through the whole tree to find the probability of failure θ at the intermediate and the top events.
3) This generates the empirical prior distributions π_{h_1} for the intermediate and the top events.
4) Repeat steps $1 - 3$ for $\pi_{h_{2i}},\ i = 1, \ldots, K$ to obtain the empirical prior distribution π_{h_2}.

Note that $A2$ assumes that it is sufficient to sample from $\pi_{h_{1i}}$'s to obtain π_{h_1} and to sample from $\pi_{h_{2i}}$'s to obtain π_{h_2}. We'll now prove that this assumption is indeed valid. To do this, we define a $(1 - \delta)$ probability interval for each θ_i as $\mathcal{I}_i(\delta) = (\theta_{i_L}, \theta_{i_U})$ such that $P(\theta_{i_L} \leq \theta_i \leq \theta_{i_U}) = 1 - \delta$, for $0 \leq \delta \leq 1$. $\mathcal{I}_i(\delta)$ represents the interval we can sample from with probability $(1 - \delta)$. Let events $E_1, \ldots, E_K$ be connected to an event E through an OR gate, that is $E = \cup_{i=1}^{K} E_i$ and let $\mathcal{I}(\delta) = (\theta_L, \theta_U)$ represent the $(1 - \delta)$ probability interval for E. Then, we can prove the following.

Theorem III.1. *$\theta_L = 1 - \prod_{i=1}^{K}(1 - \theta_{i_L})$, $\theta_U = 1 - \prod_{i=1}^{K}(1 - \theta_{i_U})$ and $\theta_L < 1 - \prod_{i=1}^{K}(1 - \theta_i') < \theta_U$ for any $\theta_{i_L} < \theta_i' < \theta_{i_U}$.*

Proof. Suppose, on the contrary that, for every i, $\exists\, \theta_i' > \theta_{i_L}$ such that $\theta_L = 1 - \prod_{i=1}^{K}(1 - \theta_i')$. But

$$\begin{aligned} \theta_{i_L} < \theta_i' \quad &\Rightarrow \quad \prod_i^K (1 - \theta_{i_L}) > \prod_i^K (1 - \theta_i')\ (0 \leq \theta_{i_L} < \theta_i' < 1) \\ &\Rightarrow \quad 1 - \prod_i^K (1 - \theta_{i_L}) < 1 - \prod_i^K (1 - \theta_i') \Rightarrow \theta_L \\ &< \quad 1 - \prod_{i=1}^K (1 - \theta_i'). \end{aligned}$$

Similarly, we can show that $1 - \prod_{i=1}^{K}(1 - \theta_i') < \theta_U$. □

Now, suppose that events $E_1, \ldots, E_K$ are connected to an event E through an AND gate, that is $E = \cap_{i=1}^{K} E_i$. Then, we can prove the following.

Theorem III.2. *$\theta_L = \prod_{i=1}^{K} \theta_{i_L}$, $\theta_U = \prod_{i=1}^{K} \theta_{i_U}$ and $\theta_L < \prod_{i=1}^{K} \theta_i' < \theta_U$ for any $\theta_{i_L} < \theta_i' < \theta_{i_U}$.*

Proof. Similar as above. □

Note that in general, $F_{h_1} \leq_{st} F \leq_{st} F_{h_2} \Rightarrow \theta_{h_1 L} \leq \theta_L \leq \theta_{h_2 L}$ and $\theta_{h_1 U} \leq \theta_U \leq \theta_{h_2 U}$. When defining distortion functions using (2), this ordering is also implied by $\pi_{h_1} \leq_{lr} \pi \leq_{lr} \pi_{h_2}$ because of Theorem II.1. In fact, one can see that $\theta_L(\theta_U)$ is a decreasing (increasing) function of α and that for $1 \leq \alpha_1 < \alpha_2$, $\mathcal{I}_{\alpha_1}(\delta) \subset \mathcal{I}_{\alpha_2}(\delta)$. Theorems III.1 and III.2 imply that in order to obtain the distorted lower (upper) bands for the intermediate/top event by sampling from them, it is necessary and sufficient to sample only from the respective lower (upper) bands of the elementary events. In Section IV-A we use simulation to illustrate this property. Further we also show that sampling through the distorted bands of the elementary events obtained using a smaller α will generate the distorted bands for the intermediate/top events that are contained within the distorted bands corresponding to a larger α.

C. Sampling from the distorted priors

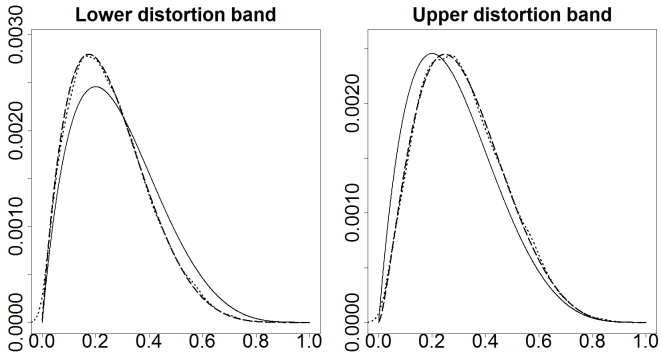

Fig. 2. Beta$(2,5)$ (*continuous*). **Left** The true distorted lower band for $\alpha = \mathbf{1.25}$ obtained using the concave function h_1 (*long-dashed*) against the lower band obtained by sampling using Algorithm $A3$ (*dotted*). **Right** The true distorted upper band using the convex function h_2 (*long-dashed*) against the upper band obtained by sampling using Algorithm $A4$ (*dotted*).

Algorithm $A2$ requires sampling $\theta_{ij} \sim \pi_{h_{1i}}$ and $\theta_{ij} \sim \pi_{h_{2i}}$ respectively. However, $\pi_{h_{1i}}$ and $\pi_{h_{2i}}$ may not be available in closed form and hence sampling from them may not be straightforward. We show that it is possible to sample from $\pi_{h_{1i}}$ and $\pi_{h_{2i}}$ using a rejection sampling scheme.

Note that

$$\pi_{h_{1i}} = \frac{d}{d\theta_i} F_{h_{1i}}(\theta_i) = \frac{d}{d\theta_i} h_1[F_i(\theta_i)] = \pi_i(\theta_i) h_1'[F_i(\theta_i)] \leq \pi_i(\theta_i) h_1'[0], \quad (11)$$

since h_1 is concave and therefore h_1' is decreasing (Arias-Nicolás et al. (2015)). This implies that

$$\frac{\pi_{h_{1i}}}{h_1'[0]\pi_i} = \frac{\pi_i(\theta_i) h_1'[F_i(\theta_i)]}{\pi_i(\theta_i) h_1'[0]} \leq 1.$$

Similarly, since h_2 is convex and therefore h_2' is increasing, it can be shown that

$$\pi_{h_{2i}} \leq \pi_i(\theta_i) h_2'[1]. \quad (12)$$

Using these inequalities, and based on Arias-Nicolás et al. (2015), we propose the following rejection sampling[1] algorithms to sample from $\pi_{h_{1i}}$ and $\pi_{h_{2i}}$, respectively.

Algorithm $A3$: to simulate from $\pi_{h_{1i}}$ for h_{1i} concave

1) Sample $\theta_{ij} \sim \pi_i(\theta_i),\ j = 1, \ldots, N,\ i = 1, 2, 3$ and $u_j \sim U(0,1)$ independently.
2) For each j, check if $u_j \leq \frac{h_1'[F_i(\theta_{ij})]}{h_1'[0]}$
   - If this holds, accept θ_j as a realisation of $\pi_{h_{1i}}$.
   - If not, reject the value θ_{ij}.

Algorithm $A4$: to simulate from $\pi_{h_{2i}}$ for h_{2i} convex

1) Sample $\theta_{ij} \sim \pi_i(\theta_i),\ j = 1, \ldots, N,\ i = 1, 2, 3$ and $u_j \sim U(0,1)$ independently.
2) For each j, check if $u_j \leq \frac{h_2'[F_i(\theta_{ij})]}{h_2'[1]}$
   - If this holds, accept θ_j as a realization of $\pi_{h_{2i}}$.
   - If not, reject the value θ_{ij}.

Note that when using power functions described in Equation (2),

$$\frac{\pi_{h_{1i}}}{h_1'[0]\pi_i} = \frac{\pi\alpha[1-F(\cdot)]^{\alpha-1}}{\pi\alpha} = [1-F(\cdot)]^{\alpha-1},$$
$$\frac{\pi_{h_{2i}}}{h_2'[1]\pi_i} = \frac{\pi\alpha[F(\cdot)]^{\alpha-1}}{\pi\alpha} = [F(\cdot)]^{\alpha-1}. \quad (13)$$

Therefore, for $u_j \sim U(0,1)$, we accept θ_j:

$$\text{in Algorithm } A3 \text{ if: } u_j \leq [1-F(\cdot)]^{\alpha-1},$$
$$\text{in Algorithm } A4 \text{ if: } u_j \leq [F(\cdot)]^{\alpha-1}. \quad (14)$$

Fig. 2 compares the true (mathematically obtained) distortion bands for a $Beta(2,5)$ distribution against the bands obtained by sampling using algorithms $A3$ and $A4$. It illustrates that algorithms $A3$ and $A4$ can sample from the distorted distributions quite accurately.

IV. Applications

A. Example: distortion bands for simple fault trees

Consider the two simple fault trees in Fig. 1. We now illustrate how the prior distributions and the distorted bands for the prior distributions can be obtained for the intermediate/top events. Assume that the prior distributions for the probabilities θ_1, θ_2 and θ_3 of the elementary events X_1, X_2, X_3 are Beta(1,10), Beta(2,10) and Beta(3,20) respectively for both the trees. The prior distribution for probability θ of the TE Y is obtained by algorithm $A1$. See Fig. 3.

Algorithms $A3$ and $A4$ were used to sample from the distorted priors (lower and upper, respectively) for θ_1, θ_2 and θ_3. Then the distorted bands for the prior distribution for θ were obtained using algorithm $A2$ for fault trees [a] and [b]. These are shown in Fig. 4 along with the original priors. Fig. 4 also illustrates that $\Gamma_{\alpha_1} \subseteq \Gamma_{\alpha_2}$, when $1 \leq \alpha_1 \leq \alpha_2$ holds even for the distorted bands of the intermediate/top events.

[1]Rejection sampling is a Monte Carlo sampling method to enable sampling from distributions which are not available in closed form or are otherwise difficult to sample from. See, for example, [18] for details.

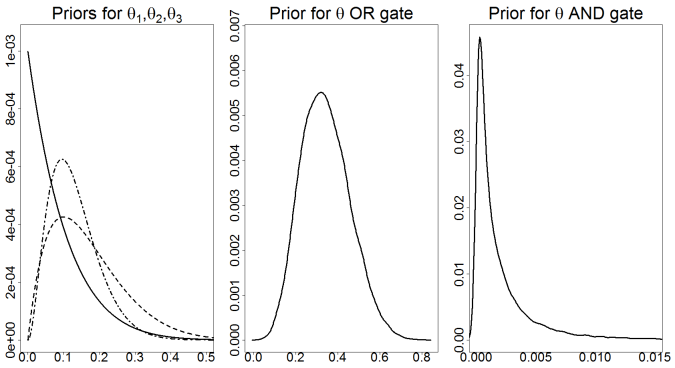

Fig. 3. **Left** Priors for θ_1 (*continuous*), θ_2 (*dashed*) and θ_3 (*dot-dashed*). **Centre** Prior distribution for θ for the fault tree [a] and **Right** for fault tree [b].

Event	Description
TE	Explosion of the spacecraft
E_{21}	Chemical reaction of propellant and air
E_{22}	Burst of pressure vessel
E_{23}	Chemical reaction between hypergolic propellants
E_{24}	Burst of battery cell
E_{11}	Sudden release of propellant (due to burst of pressure vessel E_{22})
E_{12}	Slow release of propellant
E_{01}	Valve leakage
E_{02}	Tank destruction
E_{03}	Pipe rupture
E_{13}	Chemical reactions
E_{14}	Over pressure
E_{15}	Short circuit
E_{16}	Corrosion
E_{17}	Over charge
E_{18}	Over discharge
E_{19}	Over temperature
E_{110}	Cell degradation

TABLE I
DESCRIPTION OF EVENTS FOR THE SPACECRAFT RE-ENTRY EXAMPLE

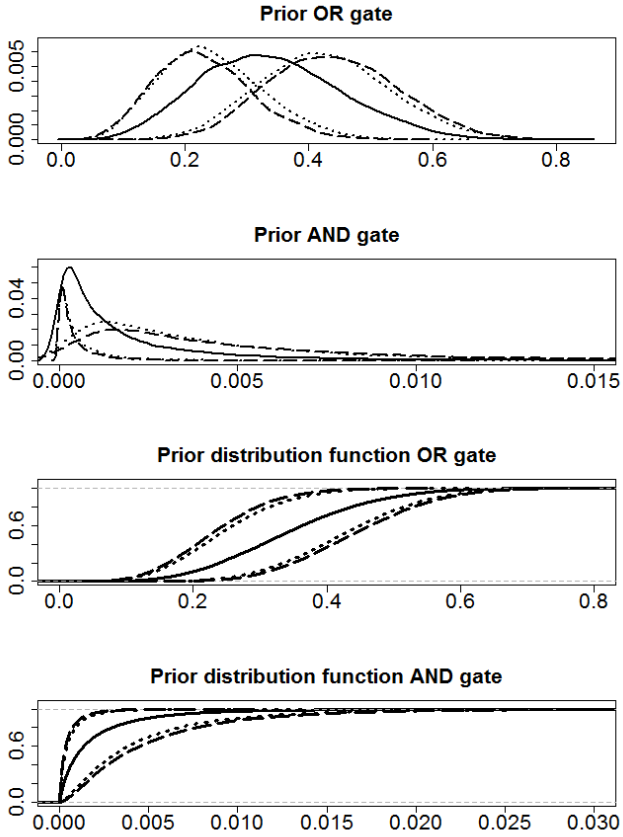

Fig. 4. Prior for the probability θ of the TE Y (*continuous*), the distorted bands using $\alpha = \mathbf{2}$ in long-dashed lines for the fault tree [a] (**top most**) and fault tree [b] (**second from the top**). Distorted bands obtained using $\alpha = \mathbf{1.8}$ are shown in dotted lines. Corresponding distribution functions are shown for the fault tree [a] (**third from the top**) and fault tree [b] (**bottom most**).

B. Spacecraft re-entry example

Risk elicitation in complex systems is challenging, especially when there are multiple elementary events and the observed data is rare/sparse. One such example is that of a spacecraft re-entering the earth's atmosphere. A Bayesian approach to elicit the risks during a spacecraft re-entry was recently developed by [5]. Their approach uses a fault tree to model the probability of an explosive break-up of the spacecraft during the re-entry process. The elementary events in this case are the (assumed independent) causes that can eventually lead to the explosive break-up. Eliciting priors required interviewing several experts because the causes are varied and no one individual was an expert on all of them. Since access to the experts was time limited, typically, information was sought from an expert to elicit prior for one of the elementary events he/she was an expert on. The expert was then asked to make a pairwise comparison between the events under their expertise in terms of which of the events are more likely. The Analytic Hierarchy Process (AHP) [19] was used to derive weights to the remaining events to elicit a prior distribution for these events. These priors were then used to determine the prior distribution of the top event (explosive break-up) and eventually to find the posterior distribution of the top event given observation of elementary or intermediate events.

For applications such as the spacecraft re-entry example, we want to highlight the following important points.

1) Prior elicitation is prone to multiple errors
2) Erroneous prior $\Rightarrow$ erroneous posterior analysis.

We elaborate on these points below.

Eliciting prior probabilities on elementary events is subject to errors from multiple sources. Firstly, the elicitation is subject to higher uncertainty in the absence of enough prior data/knowledge. Secondly, it is subject to the errors made in eliciting a prior distribution based on the information provided. This could either be because the expert was not able to provide information on the minimum number of parameters necessary to elicit a unique distribution or because the information provided was not accurate. For example, this could be either because the expert was only able to elicit a mean value, which does not lead a to unique Beta prior or because the values cited by the expert were not accurate. Then, it is also subject to the errors introduced by the methods used to elicit a prior distribution. For example,

the errors introduced by the AHP used in the spacecraft example and also by the accuracy of the computer code used to match the Beta distributions to the parameters elicited by the expert. Finally, it is influenced by the subjectivity/bias of the experts. This is especially true in situations where only one expert is consulted for eliciting a particular prior - which was the case for the spacecraft re-entry example. Seeking opinion from multiple independent and impartial experts can reduce this error - but this may not be possible in many cases due to the time and resource constraints.

In Bayesian analysis it is well known that, if only a small amount of data is available then the posterior distribution is likely to be dominated by the prior distribution. For a complex and highly expensive system such as a spacecraft used for re-entry, the data available is sparse at best and no more than a very few observations are available. This means that, in practice, the posterior will be nearly identical to the prior distribution. This is illustrated in Fig. 5. The prior distribution indicates the prior uncertainty around a spacecraft break-up event. In this case, the prior distribution has a mode just below 0.3. It takes the data on 10 identical re-entries with 1 break-up event (indicating the true probability of break up to be around 0.1,) for the posterior to be noticeably different to the prior distribution. Even then, the posterior is only slightly different than the prior. Note that for a spacecraft re-entry application, one is most likely to have only one or two observations on identical systems, not ten. For this reason, the prior and the posterior distributions will be nearly identical, as can be seen in Fig. 8. In other words, any errors made while eliciting the prior information will be directly transferred to the posterior distribution.

Therefore, it is vitally important to assess the robustness of the Bayesian procedure with respect to the mis-specification of the prior. We will now apply the robustness approach developed for the fault trees in Section III to determine the robustness of the Bayesian approach proposed by [5].

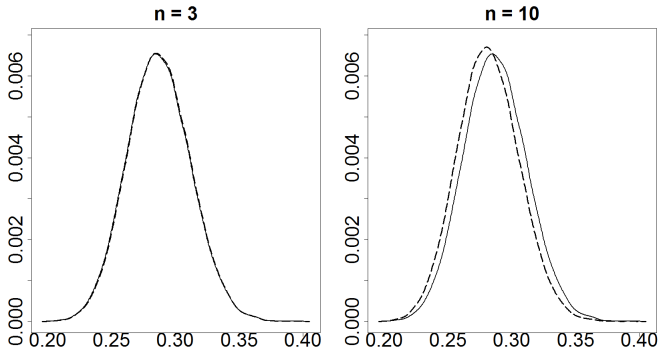

Fig. 5. Prior distribution for θ_{TE} (*continuous*) contrasted against its posterior (*long-dashed*) distribution. **Left** $n = 3$ and one break-up is observed. **Right** $n = 10$ and one break-up is observed. Note that for $n = 3$, the prior and the posterior are almost identical.

The fault tree used to model the spacecraft re-entry problem is shown in Fig. 6 [a]. The description of the events is detailed

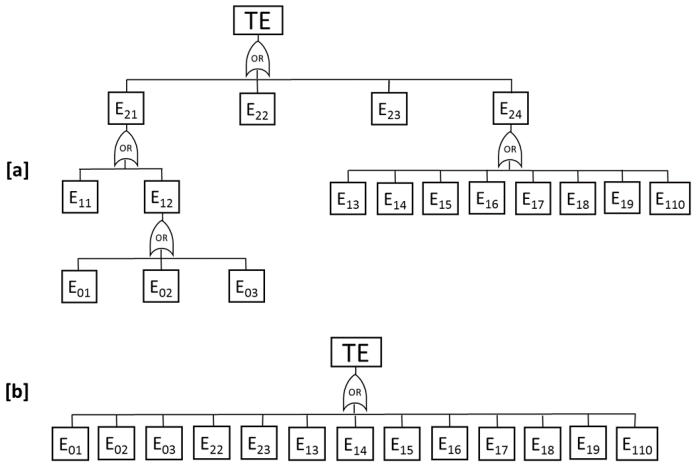

Fig. 6. [a] The fault tree used to model the spacecraft re-entry problem and [b] the simplified fault tree in minimum cut-set representation.

in Table I. Since the event E_{11} is equivalent to the event E_{22} and all the gates in the tree are OR gates, there are 13 minimal cut sets each consisting of exactly one elementary event. Therefore, the simplified version of the tree consists of only thirteen elementary events all connected by a single OR gate as shown in Fig. 6 [b]. Note that the representation [a] is still relevant since, in practice, we may observe one or more of these intermediate events while other elementary events or the top event are unobservable (see [5] for details). An event E_j can only take two values: 1 (event occurs) or 0 (event does not occur) and hence the events were modeled as $E_j \sim Bernoulli(\theta_j)$ and the priors as $g(\theta_j) \sim Beta(\alpha_j, \beta_j)$ for $\forall j$. For prior elicitation, the events were divided into three groups:

Group $G1 = \{E_{22}, E_{23}\}$,
Group $G2 = \{E_{01}, E_{02}, E_{03}\}$ and
Group $G3 = \{E_{13}, E_{14}, E_{15}, E_{16}, E_{17}, E_{18}, E_{19}, E_{110}\}$.

We assume that the priors were elicited for the event E_{22} in group $G1$, event E_{01} in group $G2$ and event E_{13} in group $G3$. This was done by matching the central 95% probability interval specified by the expert with the 2.5^{th} and the 97.5^{th} percentiles of a Beta distribution and identifying the values of the parameters α and β (using a computer code) that provide the closest match. The AHP was then used to quantify the expert opinion on how likely the remaining events in each group were compared to the one already elicited. The AHP matrices thus obtained are as follows:

$$G1 = \begin{bmatrix} 1 & 5 \\ 1/5 & 1 \end{bmatrix}, \quad G2 = \begin{bmatrix} 1 & 3 & 1 \\ 1/3 & 1 & 1/3 \\ 1 & 3 & 1 \end{bmatrix} \quad \text{and}$$

$$G3 = \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix}.$$

Note that all events in group $G3$ were deemed equally likely and hence matrix $G3$ is a matrix of $1's$. The weights derived using these AHP matrices were then used to elicit priors for the remaining events using a method described in [5]. The weights and the elicited priors are listed in Table II.

Event	Weight	Range	Elicited prior
E_{22}	0.83333	(0.01, 0.05)	Beta(6.3,233) *
E_{23}	0.16667	(0.002,0.01)	Beta(6.4,1214)
E_{01}	0.42857	(0.01, 0.04)	Beta (8.3,360)*
E_{02}	0.1428	(0.0033,0.0133)	Beta(8.3,1104)
E_{03}	0.42857	(0.01,0.04)	Beta(8.3,360)
E_{13}	0.125	(0.014, 0.055)	Beta (8.4,261)*
E_{14}	0.125	(0.014, 0.055)	Beta (8.4,261)
E_{15}	0.125	(0.014, 0.055)	Beta (8.4,261)
E_{16}	0.125	(0.014, 0.055)	Beta (8.4,261)
E_{17}	0.125	(0.014, 0.055)	Beta (8.4,261)
E_{18}	0.125	(0.014, 0.055)	Beta (8.4,261)
E_{19}	0.125	(0.014, 0.055)	Beta (8.4,261)
E_{110}	0.125	(0.014, 0.055)	Beta (8.4,261)

TABLE II
ELICITED PRIORS OBTAINED USING THE AHP PROCESS. * INDICATES THAT THE PRIOR WAS ELICITED USING THE RANGE PROVIDED BY THE EXPERT

The top event (TE) corresponds to whether the spacecraft exploded or not during the re-entry. Thus, TE can only take two values: 1 (exploded) or 0 (did not explode). It can be seen that $TE \sim Bernoulli(\theta_{TE})$, where $\theta_{TE} = 1 - \prod_j (1 - \theta_j)$. Thus, if the data was obtained from n identical spacecraft re-entries then $TE \sim Binomial(n, \theta_{TE})$. We assume that only the top event is observed and that none of the elementary events are directly observed.

Our synthetic data is obtained from the re-entry of 3 spacecrafts under identical conditions. The three observations are considered independent. Without loss of generality, we assume that the spacecraft exploded only during the second of the three re-entries. Thus, the observed data for TE is $\{0, 1, 0\}$. The likelihood is therefore given by:

$$L(TE|\theta_{TE}) = \binom{3}{1} \theta_{TE}(1 - \theta_{TE})^2. \tag{15}$$

The prior distribution for θ_{TE} is obtained by sampling from the prior distributions of the elementary events and simulating through the fault tree (algorithm $A1$). The posterior distribution is obtained using the importance sampling[2] approach described in [5]. The prior and posterior distributions for TE are plotted in Fig. 8.

Distortion bands were constructed for the priors of all the elementary events using the power functions defined in Equation 2. We assume that $K = 0.15$, that is, each elicited prior distribution could be at most 15% *off the mark* and obtain $\alpha = 1.51$ using Equation 5. The priors and the distortion bands obtained for each of them are shown in Fig. 7. The distortion bands (lower/upper) for the prior distribution of θ_{TE} were obtained by first sampling from the distorted priors (lower/upper) using Algorithms $A3/A4$ and then using algorithm $A2$. The posterior distribution of θ_{TE} were obtained for both the lower distortion band as well as the upper distortion band of the prior distribution of θ_{TE}. The posterior distributions were obtained using the importance sampling method described in [5]. These are plotted in Fig. 8.

[2]Importance sampling is a weighted sampling method of Monte Carlo integration which samples from a candidate density for the sake of efficiency and computes a modified weighted integral which converges to the given integral. See, for example, [18] for details.

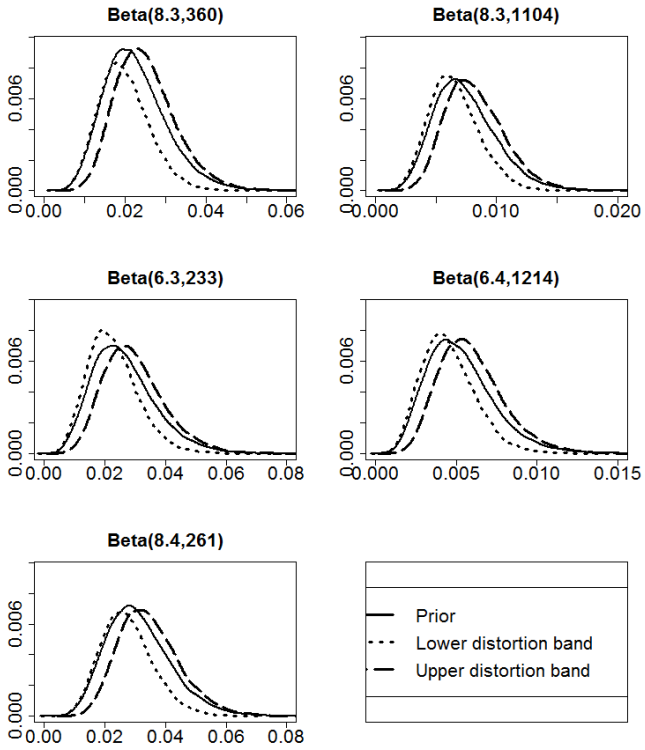

Fig. 7. Each of the unique priors (*continuous*) for the spacecraft example and the distorted bands obtained - lower band in *dotted* and upper band in *long-dashed*.

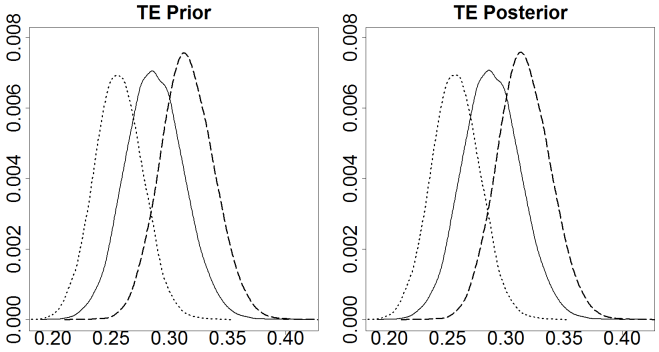

Fig. 8. **Left** The prior distribution of θ_{TE} and its distortion bands. **Right** The posterior distribution of θ_{TE} and its distortion bands: lower band (*dotted*) and upper band (*long-dashed*).

Fig. 8 illustrates two very important points.

1) Even minor mis-specifications in the prior distributions of the elementary events can result in a significant mis-specification of the prior distribution for the TE. Observing the distortion bands for the individual priors (Fig. 7) may lead one to conclude that such minor mis-specification in prior for each of the elementary events

will have a minor effect on the prior distribution of θ_{TE}. However, Fig. 8 (left) illustrates that this is not the case and that the distortions in the prior distribution for the TE are quite significant - the mode shifting by about 0.05 in each direction.

2) When the data is sparse, prior mis-specification $\approx$ posterior mis-specification. As seen in Fig. 5, when the data is sparse ($n = 3$), the posterior distribution is almost entirely governed by the prior. Fig. 8 (right) highlights that in this case the mis-specification in prior distribution results in an almost identical mis-specification in the posterior distribution. So much so that both the figures look identical.

Thus, Fig. 8 vividly illustrates why studying the robustness of the prior distribution in the Bayesian implementation of the FTA is essential. The methods developed in this paper enable us to do that.

C. Feed Control System

We now consider a safety analysis example discussed by [2]. This concerns the performance of a feeding control system used to transfer propane from a propane evaporator to a scrubbing column. Here, an improper control of the feeding system is the TE and all components are assumed binary (*work/fail*). The list of components, their description and the prior probabilities (of their failure) are listed in Table III. The fault tree used to model this problem is shown in Fig. 9 [a]. Since there is an AND gate at the top of the fault tree, there are 10 minimal cut sets each consisting of one of the components E_{21}, E_{22} and one of the components $E_{01}, E_{02}, E_{12}, E_{24}, E_{25}$. The simplified fault tree is shown in Fig. 9 [b].

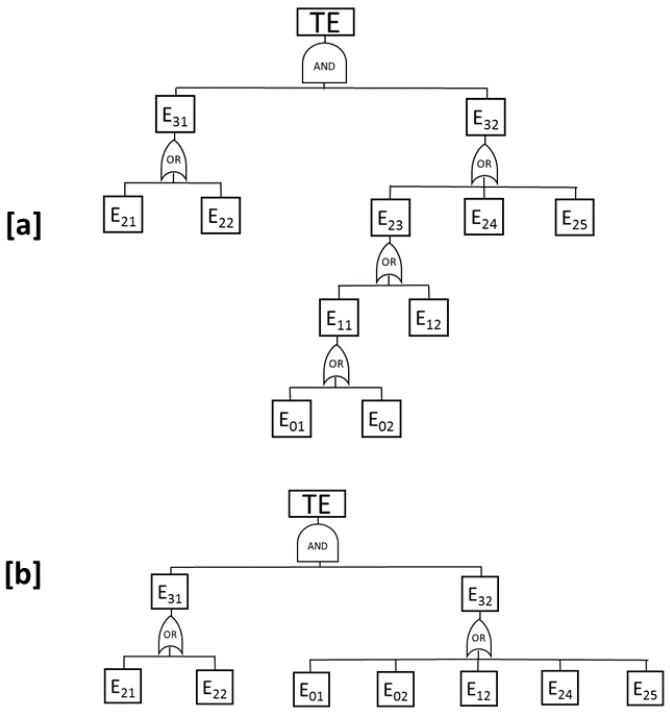

Fig. 9. [a] The fault tree used to model the feeding control system and [b] the simplified fault tree.

The first step in implementing FTA using a fully Bayesian approach is to specify priors for each of the elementary events. For this example, we know the prior probabilities as reported in [2]. We can consider these values as the mean values for the corresponding probability parameters θ_i and try to find the parameters α_i and β_i of $Beta(\alpha_i, \beta_i)$ whose mean is closest to the given probability value for each event E_i. Note that for a Beta distribution, the mean is $\mu = \alpha/(\alpha + \beta)$, which implies $\beta = \alpha(1 - \mu)/\mu$. That is, for a given μ there are infinitely many solutions of α and β. Therefore, we need to match one more statistic to find a unique combination of parameters. Here, we assume that the variance for each event is 0.03 and use a computer code to find the unique Beta distribution which matches the given mean and the variance. The prior distributions thus elicited are also listed in Table III.

We use algorithm $A1$ to sample $N = 20,000$ from the prior distribution of the TE. The mean of the sampled values is 0.2720, which matches the prior probability of the TE obtained by [2] using the standard FTA and the BN approaches. The distorted band of priors were obtained by assuming that $K = 0.2$ (which yields $\alpha = 1.75$ using Equation 5) and by sampling using algorithms $A3$ and $A4$. The original priors and the distorted band of priors obtained are shown in Fig. 10. The distorted band of priors for the TE were obtained using algorithm $A2$ and are shown in Fig. 11. The mean of the sampled values for the distorted lower and the upper bands of distributions are 0.13 and 0.4 respectively. Here too, one can see that small distortions to the prior distributions of the elementary events have resulted in significant distortion to the prior distribution of the TE.

Here, we have improved the analysis performed by [2] in two distinct ways. Firstly, we implement the fully Bayesian implementation of FTA as proposed by [5] and are therefore able to quantify the uncertainty around the prior probability of TE. For example, we can now say that while the mean prior probability of TE is 0.2720, the median is 0.2480 and the 95% probability interval is $(0.0365, 0.6247)$. This immediately highlights the fact that the prior probability of failure could be as high as more than 0.6. A simple BN or a simple FTA can not provide this information. Secondly, we have implemented the prior robustness approach developed in this paper to determine the upper and lower distortion bands of the prior distribution of the TE. These bands account for the errors in eliciting prior distributions for elementary events and provide the possible worst case scenarios. While the mean prior probability of TE is only 0.1147 for the lower distortion band, it is as high as 0.4070 for the upper distortion band. It may be possible that the mean prior probability of failure to be as high as 0.4 is considered unacceptable. Recall that this robustness analysis was performed assuming that the priors elicited for each of the elementary events can each be at most 20% *off the mark*. This could easily be the case for many complex systems when very little to no prior information is available. The 95% probability

Event	Description	Probability	Elicited prior
TE	Feed system improper control	AND gate	NA
E_{31}	Manual valve improper control	OR gate	NA
E_{32}	Automatic valve improper control	OR gate	NA
E_{21}	Manual valve mechanical failure	0.1393	Beta (0.4,2.4715)
E_{22}	Human failure in operating manual valve	0.2696	Beta (1.5,4.0638)
E_{23}	No signal received by actuator	OR gate	NA
E_{24}	Actuator mechanical failure	0.2015	Beta (0.9,3.5665)
E_{25}	Automatic valve mechanical failure	0.3403	Beta (2.2,4.2648)
E_{11}	No signal received by pressure controller	OR gate	NA
E_{12}	Pressure relay failure	0.1538	Beta (0.5,2.7509)
E_{01}	Pressure transmitter failure	0.1647	Beta (0.6,3.0429)
E_{02}	Pressure controller failure	0.2818	Beta (1.6,4.0777)

TABLE III
DESCRIPTION OF EVENTS FOR FEED CONTROL SYSTEM EXAMPLE ALONG WITH THE PRIOR PROBABILITY AND THE CORRESPONDING ELICITED PRIOR DISTRIBUTION

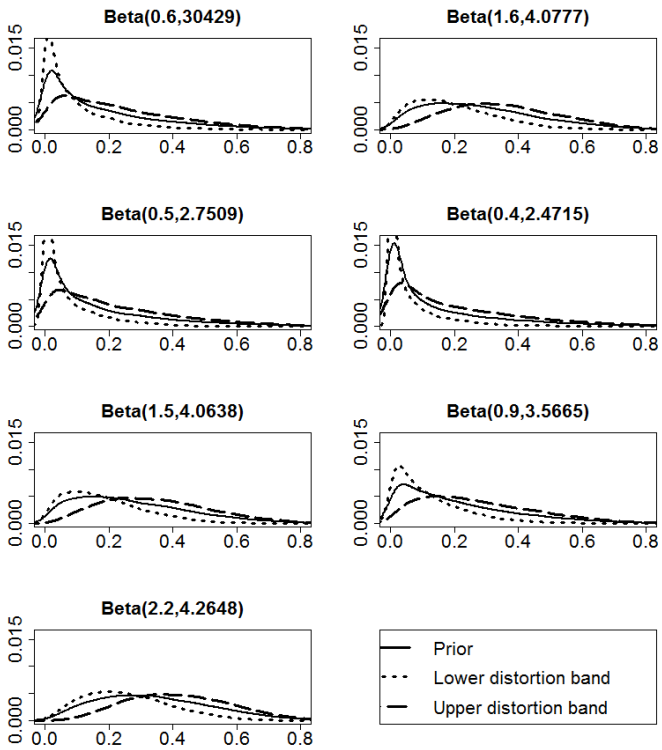

Fig. 10. Each of the unique priors (*continuous*) for the feed control system example and the distorted bands obtained - lower band in *dotted* and upper band in *long-dashed*.

interval for the upper distortion band is $(0.1160, 0.7430)$. Only the prior robustness analysis proposed in this paper allows one to discover that in fact that the mean prior probability of failure could be as high as 0.4 and that the prior probability of failure can be as high as 0.74. This information can have significant implications on the perceived reliability of a system.

D. Triple Modular Redundancy

The methodology proposed in this paper is applicable to events following virtually any distribution. Here, we illustrate how this works in a more general case. We do so by providing a step-by-step guide to implementing our method.

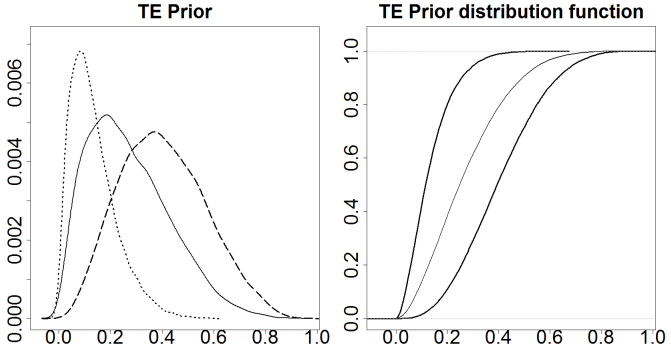

Fig. 11. **Left** The prior distribution of θ_{TE} and its distortion bands obtained using $\alpha = 1.75$ lower band (*dotted*) and upper band (*long-dashed*). **Right** Corresponding distribution functions are also shown.

Consider the simple FT in Fig. 1 [a], but now assume that variables $X_i, i = 1, 2, 3$, denote time to failure and are therefore modeled as $Exp(1/\lambda_i)$, λ_i being the survival parameter. This FT represents a triple modular redundancy system where the system will fail only if all three of the modules fail.

Step I: Elicit prior distributions for elementary events:

Let $g_i(\lambda_i)$ be the prior distribution for λ_i and $G_i(\lambda_i)$ be the corresponding distribution function. Since λ_i's are survival parameters, a natural choice for $g_i(\lambda_i)$ could be the Gamma distribution. The first task then is to elicit prior distributions. As described in Section I-C, this is often done using the *moment matching* approach. Examples IV-B and IV-C illustrate how this can be done in two different ways for the Beta distribution priors. Here we illustrate this for Gamma distribution priors.

We'll use a $Gamma(a_i, b_i)$ prior for λ_i. The mean μ_i and the variance σ_i^2 for the Gamma random variables are $\mu_i = a_i/b_i$ and $\sigma_i^2 = a_i/b_i^2$. This gives

$$a_i = \frac{\mu_i^2}{\sigma_i^2} \text{ and } b_i = \frac{\mu_i}{\sigma_i^2}. \tag{16}$$

Thus, prior distributions can be elicited using Equation (16) by eliciting the expected time to failure and the variance for each X_i. Suppose that each of the modules X_i have

an expected lifespan of 10 but because they are each made using different designs or materials, the expert believes that their variances are 1, 2 and 4 respectively. That is, the expert considers X_1 to be the most consistent and X_3 the least. Using Equation (16) gives us the following priors: $\lambda_1 \sim Gamma(100, 10)$, $\lambda_2 \sim Gamma(50, 5)$ and $\lambda_3 \sim Gamma(25, 2.5)$.

Step II: Priors for intermediate events & TE:

Given the prior distributions for the elementary events, we now need to compute the prior distributions for the intermediate events and the TE. As explained using Equations 7 and 8, the priors for intermediate events or the TE will seldom be available in closed form. We'll obtain these priors by simulation. This involves sampling from the prior distribution of each of the elementary events and simulating through the FT using algorithm $A1$.

We first sample $N = 20,000$ $\lambda_i \sim g_i(.)$ for $i = 1, 2, 3$, and then simulate a failure time for each of them. We now have N instances of the failure times for each of the three modules. This allows us to compute the failure time for the TE, namely the maximum of the three failure times. This gives us a sample of size N for the failure times for the TE from which the prior distribution of the TE can be calculated using kernel density estimation.

Step III: Quantify & account for errors in prior elicitation:

Once the priors for the elementary events have been elicited, we need to quantify the errors made in eliciting them and obtain a distorted band of prior distributions for each of them to account for these errors. The width of the distortion band is governed by the power function parameter α as described in Equation (2). Higher value of α leads to wider distortion bands indicating greater uncertainty/error in the elicited prior. As described in Section III-A, the expert is asked to quantify *how far off* can their prior be from the truth in percentage terms. This quantification gives the Kolmogorov distance K. α can then be elicited using Equations (5) or (6).

This gives us the distorted band of priors $(G_{ih_1}(\lambda_i), G_{ih_2}(\lambda_i))$ around the elicited prior $G_i(\lambda_i)$, where $G_{ih_1}(\lambda_i) = 1 - (1 - G_i(\lambda_i))^{\alpha_i}$ and $G_{ih_2}(\lambda_i) = (G_i(\lambda_i))^{\alpha_i}$. Note that the corresponding distorted probability densities are $g_{ih_1}(\lambda_i) = \alpha_i g_i(\lambda_i)(1 - G_i(\lambda_i))^{\alpha_i - 1}$ and $g_{ih_2}(\lambda_i) = \alpha_i g_i(\lambda_i)(G_i(\lambda_i))^{\alpha_i - 1}$. In most cases, the distortion bands for the elementary events, g_{ih_1} and g_{ih_2} won't be in the form of standard statistical distributions. As a result, drawing samples or computing probabilities won't be straightforward. Therefore, we use Algorithms $A3$ and $A4$ to draw samples from these distorted distributions. For $u_j \sim U(0, 1)$, we accept θ_j using Equation (14):

$$\begin{aligned} \text{in Algorithm } A3 \text{ if: } u_j &\leq [1 - G_i(.)]^{\alpha_i - 1}, \\ \text{in Algorithm } A4 \text{ if: } u_j &\leq [G_i(.)]^{\alpha_i - 1}. \end{aligned} \quad (17)$$

For standard distributions, this can be easily done using any statistical programming package.

Continuing the example, suppose that the expert believes that the elicited priors for the elementary events are each at most 20% off the mark. Using Equation (5) and a computer code, this gives $\alpha_1 = \alpha_2 = \alpha_3 = 1.735$. We pass the N samples simulated in Step II through Algorithms $A3$ and $A4$ to generate samples from g_{ih_1} and g_{ih_2}. In this case, the acceptance rate was about 57%, so we end up having about $11,500$ samples from the distorted band of each of the prior distributions. The priors and their distorted bands are shown in Fig. 12.

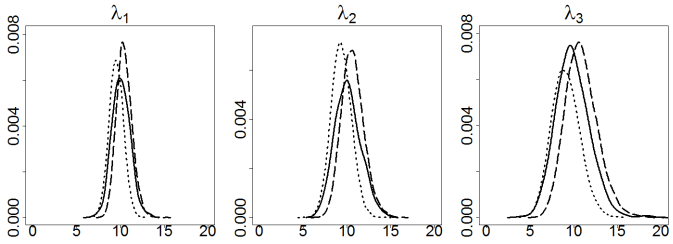

Fig. 12. **Left** The elicited prior and the distortion band for λ_1 using $\alpha_1 = 1.735$. **Centre** For λ_2 using $\alpha_2 = 1.735$. **Right** For λ_3 using $\alpha_3 = 1.735$.

Step IV: Distortion bands for intermediate events & TE:

Using the accepted samples, and similar to Step II, we can find the distortion bands for the intermediate events and the TE by simulating through the FT using Algorithm $A2$.

We have (about $11,500$) samples of λ_i from each of the distorted bands g_{ih_1} and g_{ih_2} for each of the priors g_i. As we did in Step II, we now simulate the failure times for each instance of λ_i and then compute the resulting failure time for the TE. This gives us a sample (of about $11,500$) from the lower and the upper distortion bands for the prior distribution of the TE. The distortion bands can now be computed using kernel density estimation as in Step II. These are plotted in Fig. 13.

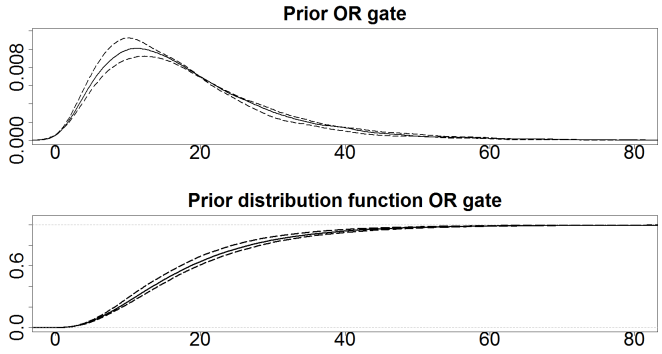

Fig. 13. **Top** The resulting prior density and the distortion bands for the TE **Bottom** Corresponding distribution functions.

Step V: Posterior distribution & its distortion bands:

The previous steps compute the prior distributions. These are helpful in the risk assessment before conducting trials/experiments to collect data. Once the data is collected,

the final step is to calculate the posterior distribution for the TE and the distortion bands of the posterior distribution for the TE. This allows one to update the risk assessment post observing the data. Posterior distributions and their distortion bands can also be obtained for any of the intermediate events or indeed the elementary events. Sampling based methods such as importance sampling or Approximate Bayesian Computation (ABC) can be used to compute the posterior distributions.

This was demonstrated in Example IV-B, where importance sampling was used to calculate posterior distribution of the TE. See [5] for details on this approach.

V. Summary and Future Work

A. Why Bayesian prior robustness?

In order to improve on the static nature of the FTA, Bayesian approaches have been proposed. This includes the BN approach as well as a fully Bayesian implementation of the FTA. Bayesian approaches enable updating the prior beliefs based on the observed data using the Bayes theorem. While BN allows for dependency between events, it has some important drawbacks as highlighted in Section I-A. One of the main drawbacks of BN is that it requires specification of exact prior probabilities and can not account for any uncertainty around the specified values. This is an unrealistic requirement since often the prior information may be subject to uncertainty. A fully Bayesian implementation of FTA has an advantage over the BN in that it accounts for the uncertainty in the prior value by specifying a prior distribution for each of the elementary events.

In many complex applications though, eliciting priors on elementary/intermediate events is not straightforward for a number of reasons. Firstly, eliciting prior distributions can be subject to errors from multiple sources as discussed in Example IV-B. Secondly, for complex systems with very little prior data, getting experts to accurately elicit even a mean value can be challenging simply because of the lack of information.

Further in some applications, actual observed data is scarce at best. This implies that the posterior beliefs are nearly identical to the prior beliefs. This also implies that the mis-specifications to the prior distributions are also the mis-specifications to the posterior. Therefore, prior robustness analysis is essential.

One of the goals of this paper was to warn engineers about the potential pitfalls around the unique specification of prior distributions when employing a fully Bayesian approach to the FTA - namely, the various types of errors that can be made and the resulting uncertainty around the true priors. Here, we provide a way to model such uncertainty and evaluate the consequences.

B. Methodology proposed in this paper

In this paper, we develop a prior robustness approach for the fully Bayesian implementation of FTA. This now provides a way to model the uncertainties around the prior specifications and to quantify the effect these uncertainties have on the perceived reliability of the system. A prior robustness approach to Bayesian FTA has not been developed before.

Our approach is based on a new class of priors called the distorted band of priors [11] which are obtained by using convex and concave power functions. We develop the computational algorithms needed to implement this approach on Bayesian FTA. We also mathematically prove that these algorithms will always work and will accurately generate the distortion bands needed. Our mathematical contribution is detailed below.

We prove some important properties related to the distorted band class of priors. In particular we have shown that the lower/upper bounds of the distorted bands for the intermediate/top events can only be obtained by using the lower/upper bounds of the distorted bands for the elementary events. This property is important since it implies that it is sufficient and also necessary to sample only from the lower/upper bounds of the distorted bands for each of the elementary events in order to obtain the distorted bands for the intermediate/top events. If this property was not true, then one would have to sample from all the (infinite number of) distributions that are members of the distorted band of priors for an elementary event in order to obtain the accurate distorted bands for the intermediate/top events.

An important problem while working with the distorted band of priors is to be able to correctly elicit the power function parameter α. The value of α determines the width of the distortion band and thus quantifies the uncertainty around the specified prior distribution. The correct elicitation of α is therefore vital to be able to accurately estimate the uncertainty around the reliability of the system. Yet, there was no guidance on how to do this. We show how the power function parameter α can be elicited through the expert knowledge. We provide both an exact equation that can be solved using a computer code and also an approximation that does not require using a computer code.

Finally, we provide a step-by-step illustration of how this methodology can be implemented on a real life example.

C. What difference does it make?

We consider two real life examples: a spacecraft re-entry example discussed by [5] and a feeding control system example discussed by [2]. Since eliciting prior distribution is so critical, we illustrate how prior distributions can be elicited in different ways using these two examples. We also illustrate how the distorted prior and posterior bands can be obtained

for the top event. We use these examples to illustrate that minor distortions in the prior/posterior distributions of the elementary events can result in significant distortions to the prior/posterior distributions of the top event. These distortions can change the perceived reliability of the system highlighting the need for this kind of analysis to be implemented when using the fault tree approach to model complex events.

The two key messages here are: 1. A practical warning to the engineers that, wherever possible an extra effort should be placed on eliciting prior distributions with more precision so as to reduce the uncertainty in prior specification. 2. A prior robustness approach such as the one developed here should be employed to quantify the uncertainty around the Bayesian analysis so as to get a more realistic model of the reliability of the system.

We also show that this methodology can be just as easily applied for non-Bernoulli data by illustrating its use to model time to failure in a triple modular redundancy system. Indeed, the distortion bands can be obtained for any distribution using the algorithms described in this paper.

An important point to note is that the use of the distorted band class of priors actually relaxes the distributional assumptions made when eliciting a prior distribution (whether it is Beta or Exponential or any other). This is because the class will contain (infinitely many) distributions which will not follow the form of the original prior. Therefore, the resulting analysis can be considered to be largely independent of any distributional assumptions made.

D. Future work

The spacecraft re-entry example discussed also uses the AHP to elicit priors for some of the elementary events. The AHP itself is also subject to errors due to mis-specification. The approach proposed here assumes that the distortions due to the mis-specification in the AHP are also taken into account while eliciting the power function parameter α. A robustness approach to explicitly quantify the distortions to AHP will be an important future contribution.

BN have been shown to be a superior alternative to FTA, especially when the independence assumption is not met. However, BN not only require specification of exact prior probabilities but also the conditional probabilities to define dependencies between events. As the number of nodes and dependencies increase, the BN require an increasingly large amount of prior specification. The BN outcome is therefore subject to the mis-specification of not only the priors but also of conditional probabilities. An important future contribution would be to develop a Bayesian prior robustness approach for BN.

ACKNOWLEDGMENT

The authors would like to thank the three anonymous referees and the Associate Editor for their valuable feedback which has helped in greatly improving this paper.

REFERENCES

[1] A. Bobbio, L. Portinale, M. Minichino, and E. Ciancamerla, "Improving the analysis of dependable systems by mapping fault trees into bayesian networks," *Reliability Engineering and System Safety*, vol. 71, pp. 249 – 260, 2001.

[2] N. Khakzad, F. Khan, and P. Amyotte, "Safety analysis in process facilities: Comparison of fault tree and bayesian network approaches," *Reliability Engineering and System Safety*, vol. 96, pp. 925 – 932, 2011.

[3] P. Trucco, E. Cagno, F. Ruggeri, and O. Grande, "A Bayesian belief network modelling of organisational factors in risk analysis: A case study in maritime transportation," *Reliability Engineering and System Safety*, vol. 93, pp. 823 – 834, 2008.

[4] D. Neidermayer, "An introduction to bayesian networks and their contemporary applications," *in Innovations in Bayesian Networks*, 2008, ed. D. E. Holmes and L.C. Jain.

[5] C. DePersis, *A risk assessment tool for highly energetic break-up events during the atmospheric re-entry*. Trinity College Dublin, Ireland, 2016, Ph.D. Thesis.

[6] D. Lindley and N. Singpurwalla, "Reliability (and fault tree) analysis using expert opinions," *Journal of the American Statistical Association*, vol. 81, no. 393, pp. 87 – 90, 1986.

[7] A. Irving, "Fault tree uncertainty analysis using a Monte Carlo method," *in* 10^{th} *Advances in Reliability Technology Symposium*, 1988, ed. G.P. Libberton.

[8] A. Hagan and M. West, *The Oxford Handbook of Applied Bayesian Analysis*, ser. Oxford Handbooks in Mathematics. OUP Oxford, 2010.

[9] H.-H. Por and D. V. Budescu, "Eliciting subjective probabilities through pair-wise comparisons," *Journal of Behavioral Decision Making*, vol. 30, no. 2, 2017.

[10] T. A. Mazzuchi, W. G. Linzey, and A. Bruning, "A paired comparison experiment for gathering expert judgment for an aircraft wiring risk assessment," *Reliability Engineering & System Safety*, vol. 93, no. 5, pp. 722 – 731, 2008.

[11] J. P. Arias-Nicolás, F. Ruggeri, and A. Suárez-Lorens, "New classes of priors based on stochastic orders and distortion function," *Bayesian Analysis*, vol. 11, no. 4, pp. 1107 – 1136, 2016.

[12] J. Berger, D. Ríos Insua, and F. Ruggeri, "Bayesian robustness," *in Robust Bayesian Analysis*, 2000, ed. D. Ríos Insua and F. Ruggeri.

[13] E. Moreno, "Global bayesian robustness for some classes of prior distributions," *in Robust Bayesian Analysis*, 2000, ed. D. Ríos Insua and F. Ruggeri.

[14] D. Ríos Insua and F. Ruggeri, *Robust Bayesian Analysis*. Springer, New York, 2000.

[15] M. Shaked and G. Shanthikumar, *Stochastic Orders*, ser. Springer Series in Statistics. Springer-Verlag New York, 2007.

[16] A. Müller and D. Stoyan, *Comparison Methods for Stochastic Models and Risk*. John Wiley & Sons Inc., Chichester, 2002.

[17] F. Spizzichino, *Subjective probability models for lifetimes*. Chapman and Hall, CRC, Boca Raton, 2001.

[18] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*, 2nd ed. Springer, New York, 2004.

[19] T. Saaty, *The Analytic Hierarchy Process*. McGraw Hill, New York, 1980.

Chaitanya Joshi Chaitanya Joshi is a Senior Lecturer at the Department of Mathematics and Statistics, University of Waikato, Hamilton, New Zealand. He received his B.Sc. degree in Statistics from University of Mumbai, his M.Sc. degree in Statistice from the Indian Institute of Technology Kanpur and his Ph.D. degree in Statistics from Trinity College Dublin, Ireland.

His main research interests include computational Bayesian methods, Bayesian robustness and statistical modeling of complex systems.

Fabrizio Ruggeri Fabrizio Ruggeri received the B.Sc. degree in Mathematics from the University of Milano, Italy; the M.Sc. degree in Statistics from Carnegie Mellon University, Pittsburgh, PA, USA; and the Ph.D. degree in Statistics from Duke University, Durham, NC, USA.

He is Research Director at the Italian National Research Council in Milano, and Adjunct Faculty at Queensland University of Technology. His interests are mostly in Bayesian and industrial statistics, especially in robustness, decision analysis, reliability, and stochastic processes; recently, he got involved in biostatistics.

Dr. Ruggeri is ASA and ISBA Fellow, Zellner Medal recipient, ISI Elected Member, former ISBA and ENBIS President, and current ISI Vice President and ISBIS President-Elect, besides Chair of the ISBA Industrial Section. He is also Editor-in-Chief of Applied Stochastic Models in Business and Industry, Encyclopedia of Statistics in Quality and Reliability, and Wiley StatsRef Online.

Simon P. Wilson Simon Wilson is professor in statistics at Trinity College Dublin where he is head of the Statistics Discipline. He has a Ph.D. from George Washington University, Washington DC. His research interests are in statistical reliability with a focus on the use of Bayesian methods as well as more generally applications of Bayesian methods across diverse fields from astronomy to zoology.