# Conservative reasoning about the probability of failure on demand of a 1-out-of-2 software-based system in which one channel is "possibly perfect"

Bev Littlewood, Andrey Povyakalo

Centre for Software Reliability, City University London

*Abstract*

In earlier work, (Littlewood and Rushby 2012) (henceforth LR), an analysis was presented of a 1-out-of-2 software-based system in which one channel was "possibly perfect". It was shown that, at the aleatory level, the system *pfd* (probability of failure on demand) could be bounded above by the product of the *pfd* of channel *A* and the *pnp* (probability of non-perfection) of channel *B*. This result was presented as a way of avoiding the well-known difficulty that for two certainly-fallible channels, failures of the two will be dependent, i.e. the system *pfd* cannot be expressed simply as a product of the channel *pfd*s. A price paid in this new approach for avoiding the issue of failure dependence is that the result is conservative. Furthermore, a complete analysis requires that account be taken of *epistemic uncertainty* – here concerning the numeric values of the two parameters $pfd_A$ and $pnp_B$. Unfortunately this introduces a different difficult problem of dependence: estimating the dependence between an assessor's beliefs about the parameters. The work reported here avoids this problem by obtaining results that require only an assessor's *marginal* beliefs about the individual channels, i.e. they do not require knowledge of the dependence between these beliefs. The price paid is further conservatism in the results.

KEY WORDS: *Software reliability; fault tolerance; software perfection; probability of failure; epistemic uncertainty; software diversity; multi-version software*

## 1 Introduction

Intellectual *diversity* has been used from time immemorial to improve the dependability of human activities. Most people believe that, for many activities, "two heads are better than one": e.g. it is often better to have another person check your work than to do it yourself. The use of diversity to build reliable systems long pre-dates the use of computers. For example, the use of diverse multi-channel safety protection systems based on physically different variables (temperatures, pressures, flow-rates…) has for a long time been an attractive design approach.

Design diversity of this kind has been applied to software-based systems for several decades, and there are reports of apparently successful industrial applications to critical systems, see, e.g., (Littlewood, Popov et al. 2002; Wood, Belles et al. 2010). For example, the safety-critical flight control systems of Airbus fleets (Rouquet and Traverse

1986) have experienced massive operational exposure (Boeing 2012) with apparently no critical failure. We might conclude, *after the fact*, that these systems are very reliable.

However, there are serious difficulties in assessing the reliability of such systems *before* operational use. The stringent dependability requirements for safety-critical systems usually mean that black-box operational testing would require infeasible times on test (Butler and Finelli 1993; Littlewood and Strigini 1993). Furthermore, it is well-known that it is not possible to claim, with certainty, independence between the failures of multiple software-based channels of a system: see (Knight and Leveson 1986; Eckhardt, Caglayan et al. 1991) for experimental evidence, and (Eckhardt and Lee 1985; Littlewood and Miller 1989) for theoretical reasons for this assertion. So, for a 1-out-of-2 demand-based system, the system *pfd* will not be a simple product of the channel *pfd*s: it must be assumed that

$$pfd_{sys} > pfd_A \times pfd_B \quad (1)$$

because there will usually be *positive* association between the failures of channel *A* and those of channel *B*.[1] In fact, statistical independence is probably a rather rare phenomenon in the world: see (Kruskal 1988) for an amusing but serious discussion of inappropriate assumptions of independence. He says: "If I have a moral, it is this: Do not multiply lightly."

If independence cannot be assumed between channel failures, the problem of assessing the reliability of the system becomes difficult: we need to know *how* dependent the failures of the channels are. Assessing this dependence directly seems as hard as treating the system as a black box and measuring its *pfd* directly, and as we have noted above this is known to be generally infeasible.

In recent work, a way around this difficulty has been proposed for certain special architectures (Littlewood and Rushby 2012), henceforth LR. The idea here is that in some 1-out-of-2 systems, one channel (say *A*) may be highly functional and complex, and so (effectively certainly) failure-prone, but the other channel (*B*) may be very simple and thus *possibly perfect*. By "perfect" we mean that this channel cannot fail in its entire life, no matter how much exposure it receives, i.e. its *pfd* is zero. By "possibly perfect" we mean that such perfection will not be known with certainty. Claims about *A* will be expressed as a probability of failure on a randomly selected demand ($pfd_A$); claims about *B* will be expressed as a probability that it is not perfect ($pnp_B$). See (Littlewood and Rushby 2012) for extensive discussion, and examples of the kinds of systems for which this kind of architecture may be appropriate.

The key idea in LR is that, at the aleatory level, it can be shown that there is conditional independence between the events "*A* fails on a randomly selected demand" and "*B* is not perfect," given that the probabilities of these events, respectively $pfd_A$ and $pnp_B$, are known. It is then shown that a conservative bound for the system's (conditional) probability of failure on demand is simply the product of the probabilities of these two events, i.e.

---

[1] Whilst negative association is theoretically possible (Littlewood and Miller 1989) – thus reversing the inequality in (1) – we are not aware of any means of claiming this with high confidence in a particular instance.

$$pfd_{sys} \leq pfd_A \times pnp_B \tag{2}$$

where the conservatism arises by assuming that, if $B$ is imperfect, it always fails when $A$ does: see (Littlewood and Rushby 2012) for proof. An assessor can then use the right hand side of (2) for the probability of failure on demand of the system, and be confident that this is conservative.

The new result is useful because it provides a conservative numerical bound for the system *pfd* which is simply the product of two (hopefully small) numbers, and is thus (hopefully) *a very small number*. In other words, we have a result that is similar in nature to the one we would use if we could assume channel failures to be independent (the product of two small channel *pfd*s).

In reality, of course, the assessor will be uncertain about the values of the parameters in the discussion above: such uncertainty is called *epistemic*, and arises from the imperfect knowledge of the assessor. In LR the assessor beliefs are represented formally by a Bayesian posterior distribution,

$$F(p_A, p_B) = P(pfd_A < p_A, pnp_B < p_B) \tag{3}$$

that incorporates all the evidence that the assessor has about the unknown parameters.

The assessor's probability of system failure on a randomly selected demand is then bounded by the posterior mean of the product, from (2):

$$\int_{\substack{0 \leq p_A \leq 1 \\ 0 \leq p_B \leq 1}} p_A \times p_B \, dF(p_A, p_B) \tag{4}$$

If $F$ factorised, i.e. the assessor's beliefs about the two parameters were independent, then (4) would simplify into the product of the means of the posterior marginal distributions of the parameters. Unfortunately, assessors' beliefs are unlikely to be independent in this way, and this *epistemic dependence* poses a serious problem.

In the current paper we propose ways around this difficulty. These new results rely solely upon assessors' *marginal* beliefs about the individual channel parameters – $pfd_A$, $pnp_B$ – and do not require epistemic dependence between them to be estimated.

In summary, the results of LR and the results of the present paper address two difficult problems of dependence. In the case of LR, the basic result overcomes the problem that the system *pfd* cannot simply be assumed to be a product of the channel *pfd*s. The results in this paper address the problem that an assessor's beliefs about the parameters ($pfd_A$, $pnp_B$) of LR will not be dependent.

There is a price paid, not surprisingly, for this latter simplification: further conservatism is introduced into the claims that can be made about the system *pfd*, over and above that arising from LR.

## 2 Conservative bounds on mean system *pfd*

We begin with the result (2). Instead of dealing with the complete bivariate distribution, (3), representing the assessor's posterior beliefs about the parameters $pfd_A$ and $pnp_B$, we shall assume only that the assessor can tell us something about their separate *marginal*

distributions for these parameters, which we shall call $F(p_A)$ and $F(p_B)$ in an obvious notation. Clearly this places upon the assessor a much less onerous requirement in describing their epistemic uncertainty, inasmuch as they do not need to say anything about the *dependence* in their beliefs about the parameters.

Initially, we assume that the assessor is able to give us only a single percentile for each distribution, i.e. they are able to express their subjective confidence in a single bound (e.g. a bound that may arise from the requirements of a wider safety case):

$$\begin{aligned} P(pfd_A < p_A) &= 1-\alpha_A \\ P(pnp_B < p_B) &= 1-\alpha_B \end{aligned} \tag{5}$$

So $p_A$ is their 100(1-$\alpha_A$)% upper confidence bound for the parameter $pfd_A$; equivalently, $\alpha_A$ can be thought of as their *doubt* that $pfd_A$ is smaller than $p_A$, etc.

We have the following:

## *Theorem 1*

If

$P(pfd_A < p_A) = 1-\alpha_A$ and $P(pnp_B < p_B) = 1-\alpha_B$

represent the assessor's marginal posterior beliefs about the parameters, and without loss of generality

$\alpha_A \le \alpha_B$,

then

$$E(pfd_{sys}) \le p_A \times p_B \times (1-\alpha_B) + p_A \times \alpha_B + (1-p_A) \times \alpha_A \tag{6}$$

## *Proof*

Denote the unknown joint probability, $P(pfd_A > \alpha_A, pnp_B > \alpha_B)$, i.e. of lying in BCFE in Figure 1, by $z$. Now

$$\begin{aligned} pfd_{sys} &\le E(pfd_A \times pnp_B) \\ &= p_A \times p_B \times (1-\alpha_A-\alpha_B+z) + p_A \times (\alpha_B - z) + p_B \times (\alpha_A - z) + z \\ &= p_A \times p_B \times (1-\alpha_A-\alpha_B) + \alpha_A \times p_B + \alpha_B \times p_A + z \times (1-p_A-p_B+p_A \times p_B) \\ &\le p_A \times p_B \times (1-\alpha_A-\alpha_B) + \alpha_A \times p_B + \alpha_B \times p_A + \min(\alpha_A,\alpha_B) \times (1-p_A-p_B+p_A \times p_B) \\ &= p_A p_B (1-\alpha_B) + p_A \alpha_B + (1-p_A)\alpha_A \end{aligned} \tag{7}$$

because

$0 \le z \le \min(\alpha_A,\alpha_B) = \alpha_A$

and

$1 - p_A - p_B + p_A \times p_B \ge 0$

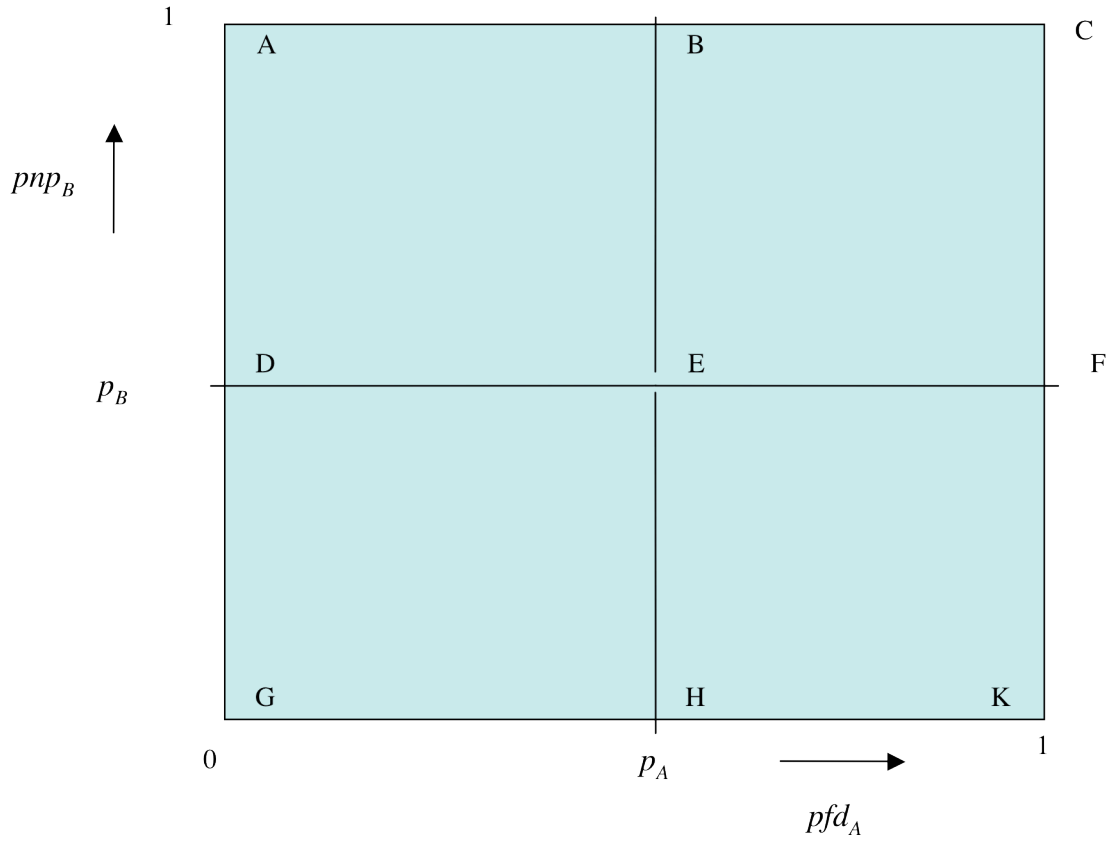


Figure 1. The random variable ($pfd_A$,$pnp_B$) is defined on the unit square. Note that this figure has been exaggerated for clarity: in reality E would be very close to the origin.

The result (7) can be seen as follows. Consider the four rectangles in Figure 1: DEHG, ABED, EFHK, BCFE. The product $pfd_A \times pnp_B$ is a random variable which is everywhere smaller than $p_A \times p_B$ within DEHG. The probability associated with DEHG is $(1-\alpha_A-\alpha_B+z)$. Thus the contribution to $pfd_{sys} = E(pfd_A \times pnp_B)$ associated with DEHG is bounded above by the product $p_A \times p_B \times (1-\alpha_A-\alpha_B+z)$. Hence the first term in (6). Similarly, within the rectangle ABED, the product $pfd_A \times pnp_B$ is a random variable which is everywhere smaller than $p_A$ (which value it takes at the point B); and the probability associated with this rectangle is $(\alpha_B - z)$; so the contribution to the mean of this rectangle is bounded by the product of these. Hence the second term in (7). Similar reasoning about EFHK, BCFE give the third and fourth terms of (7), respectively.

This completes the proof.

## *Example 1*

If the assessor can provide a single percentile (i.e. a bound with associated confidence level) for each marginal posterior distribution (i.e. for $pfd_A$ and for $pnp_B$) then the theorem provides a means of computing a conservative posterior mean of $pfd_{sys}$.

So, if the assessor is 95% confident that $pfd_A$ is smaller than $10^{-5}$, and 95% confident that $pnp_B$ is smaller than $10^{-2}$ we have, from (6):

$$E(pfd_{sys}) \leq 10^{-5} \times 10^{-2} \times (1-0.05) + 10^{-5} \times 0.05 + (1-0.05) \times 0.05 \approx 0.05 \quad (8)$$

which of course is *very* conservative.

If the assessor is 99% confident that $pfd_A$ is smaller than $10^{-3}$, and 99.9% confident that $pnp_B$ is smaller than $10^{-1}$, the bound on his posterior mean for the system *pfd* is about $1.1 \times 10^{-3}$.

In fact, since "doubts" will usually be considerably greater than "claims", this way of bounding the assessor's posterior *pfd* for the system will give a result that is approximately the same as the smallest of the two doubts.

So these results are very conservative. One reason for this is that it is assumed that there is probability mass over the whole unit square: that is, the assessor cannot rule out the possibility of the parameters taking *any* value. This probability mass is assigned most pessimistically in each of the rectangles making up the unit square, e.g for the random variable ($pfd_A$, $pnp_B$) lying in the upper right rectangle, *all* probability is assigned to the point (1,1), i.e. it is assumed with this probability that channel *A* fails, and channel *B* is imperfect, so that the system fails with certainty. This is similar to the LR reasoning.

Such beliefs may be too pessimistic for real assessors. We have heard safety assessors reason as follows: "I have confidence ($1-\alpha_A$) that channel *A*'s *pfd* is smaller than $p_A$, but I am *certain* that it is smaller than $p_A^U$, where $p_A << p_A^U$", with similar certainty that $pnp_B$ is smaller than $p_B^U$. This is illustrated in Figure 2, where now there is non-zero probability mass only in the rectangle QSWG: outside this rectangle the distribution (4), $F(p_A, p_B)$, takes the value 1 everywhere.

We can now obtain a tighter conservative bound as follows:

## *Theorem 2*

If

$P(pfd_A < p_A) = 1 - \alpha_A$ and $P(pnp_B < p_B) = 1 - \alpha_B$

and

$P(pfd_A < p_A^U) = 1$ and $P(pnp_B < p_B^U) = 1$

represent the assessor's marginal posterior beliefs about the parameters, and without loss of generality

$\alpha_A \leq \alpha_B$,

then

$$E(pfd_{sys}) \le p_A \times p_B \times (1-\alpha_B) + p_A \times p_B^{\,U} \times \alpha_B + p_B^{\,U} \times (p_A^{\,U} - p_A) \times \alpha_A \quad (9)$$

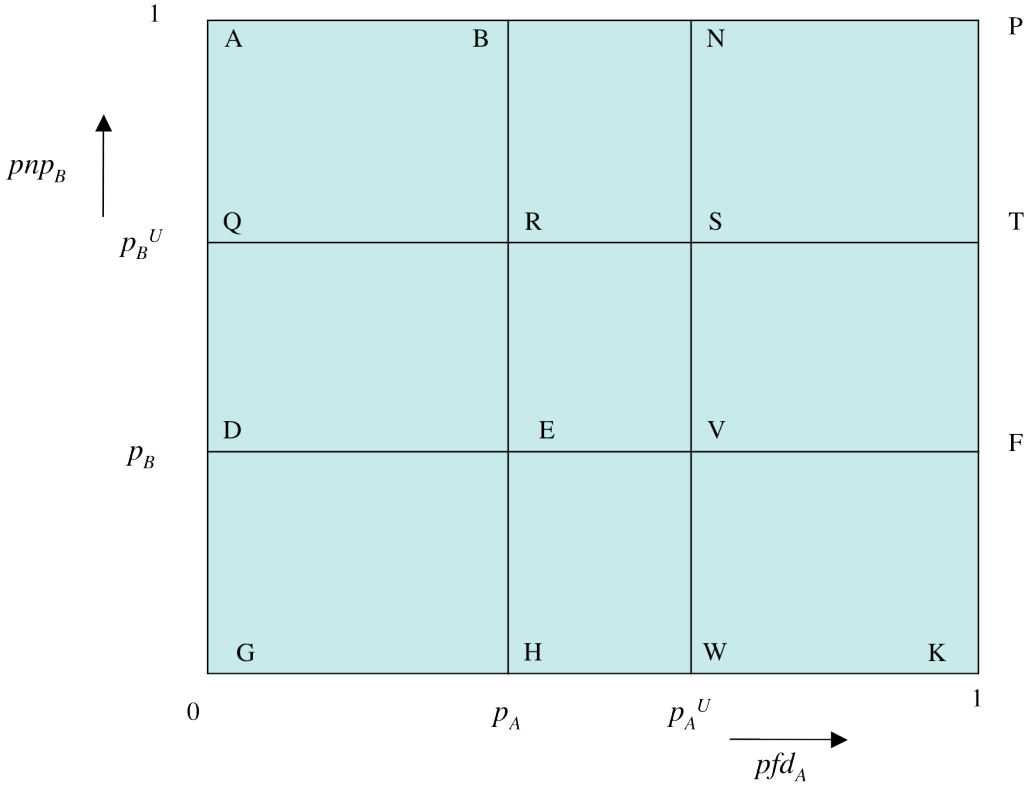


Figure 2. As Figure 1, except that now, in addition, the assessor is *certain* that $pfd_A$ does not exceed $p_A^{\,U}$ and $pnp_B$ does not exceed $p_B^{\,U}$. So there is zero probability mass outside QSWG.

## *Proof*

This is similar to the proof of the previous theorem, in terms of the four rectangles making up QSWG. Once again, denote by $z$ the unknown probability mass associated with RSVE.

In DEHG, the random variable $pfd_A pnp_B$ is bounded above by $p_A p_B$ and the probability mass here is $(1-\alpha_A-\alpha_B+z)$. So the contribution of DEHG to the posterior mean of the system *pfd* is bounded above by $p_A p_B(1-\alpha_A-\alpha_B+z)$.

By similar reasoning, the contribution from QRED is bounded by $p_A p_B^{\,U}(\alpha_B - z)$; that from EVWH by $p_A^{\,U} p_B(\alpha_A - z)$; that from RSVE by $p_A^{\,U} p_B^{\,U} z$.

Adding all these contributions together, and using the fact that $0 \le z \le \min(\alpha_A, \alpha_B) = \alpha_A$, the result follows as in the previous theorem after some rearrangement.

Notice that, as expected, (9) reduces to (6) when $p_A^U=p_B^U=1$.

## *Example 2*

As Example 1 with, additionally, $p_A^U=10^{-3}$, $p_B^U=10^{-1}$.

$$E(pfd_{sys}) \le 10^{-7} \times 0.95 + 10^{-4} \times 0.05 + 0.05 \times (10^{-3} - 10^{-5}) \times 10^{-1}$$
$$\approx 0 + 0.5 \times 10^{-5} + 0.5 \times 10^{-5}$$
$$= 1 \times 10^{-5} \quad (10)$$

Clearly this is better than the bound in Example 1. And it is an order of magnitude improvement on the very crude bound that simply multiplies the two marginal upper bounds, i.e. $10^{-4}$.

We now obtain some conservative bounds for the system *pfd* for situations in which the assessor knows the first two *moments* of their marginal distributions for the parameters, rather than percentiles as above:

## *Theorem 3*

$$E(pfd_{sys}) \le E(pfd_A \times pnp_B)$$
$$< \sqrt{\left[\left(E(pfd_A)^2 + Var(pfd_A)\right).\left(E(pnp_B)^2 + Var(pnp_B)\right)\right]} \quad (11)$$
$$< \left(E(pfd_A) + SD(pfd_A)\right).\left(E(pnp_B) + SD(pnp_B)\right) \quad (12)$$

## *Proof*

By the Cauchy-Schwarz inequality

$$\left(E(pfd_A.pnp_B)\right)^2 < E(pfd_A^2).E(pnp_B^2)$$
$$= \left(E(pfd_A)^2 + Var(pfd_A)\right).\left(E(pnp_B)^2 + Var(pnp_B)\right)$$

which gives (11). And

$$E(pnp_B)^2 + Var(pnp_B) < \left(E(pnp_B) + SD(pnp_B)\right)^2$$

with a similar expression involving $pnp_B$, so (12) follows.

## *Example 3*

The result requires knowledge of the first two moments of the marginal distributions of the two model parameters. In particular, the closeness of the bound to the "ideal" independence result (i.e. product of the marginal means of the parameters) depends on the relative sizes of the marginal standard deviations and marginal means. So, if

$SD(pfd_A)<4.E(pfd_A)$ and $SD(pnp_B)<4.E(pnp_B)$

we have

$E(pfd_{sys})<25.\ E(pfd_A).\ E(pnp_B)$

Another way in which (12) might be used is as follows. One way that we have heard assessors reason in the presence of difficult-to-assess dependence is to make a trade-off between "lack of independence" and "pessimism of channel claims". The reasoning is something like this: "I realize I cannot simply multiply my marginal beliefs about the *pfd* of channel *A* and the *pnp* of channel *B* to obtain a bound for the system *pfd*, so I will instead multiply together *pessimistic* values for these two channel beliefs. The pessimism here will counteract the optimism of the independence assumption implicit in the simple multiplication of the numbers[2]." The result (12) provides a formalism for this kind of reasoning. It shows *how much* pessimism is needed to justify such reasoning: a system claim made in this way will be a conservative one if each channel claim is conservative by an amount equal to the standard deviation of the marginal distribution.

Finally, we present conservative bounds for the situation where an assessor's beliefs about the two marginal distributions involve both *means* and *percentiles* as follows:

## *Theorem 4*

If

$P(pfd_A > p_A) = \alpha_A$ and $P(pnp_B > p_B) = \alpha_B$

and

$E(pfd_A) \le p_A$ and $E(pfd_B) \le p_B$

then

$$E(pfd_{sys}) \le E(pfd_A \times pnp_B)$$

$$\le \frac{E(pfd_A) \times E(pnp_B)}{\sqrt{\alpha_A \times \alpha_B}} \quad (13)$$

$$\le \frac{p_A \times p_B}{\sqrt{\alpha_A \times \alpha_B}} \quad (14)$$

## *Proof*

We require the following

**Lemma:**

---

[2] This kind of reasoning is more common at the aleatory level. We have seen arguments in which pessimistic claims have been made for each channel *pfd* and then these have been multiplied together to obtain a figure for the system *pfd*. The trade-off here is between *channel failure dependence* and *channel pfd pessimism*. See Bishop, P., R. Bloomfield, et al. (2011). "Towards a formalism for conservative claims about the dependability of software-based systems." IEEE Trans Software Engineering **37**(5): 708-717.

If

$0 \le X \le 1$, and $P(X>p)=\alpha$, and $E(X) \le p$,

then

$$E(X^2) \le \frac{E(X)^2}{\alpha} \le \frac{p^2}{\alpha}$$

**Proof: see appendix**

From the lemma we have:

$$E(pfd_A^2) \le \frac{E(pfd_A)^2}{\alpha_A} \le \frac{p_A^2}{\alpha_A}$$

and

$$E(pnp_B^2) \le \frac{E(pnp_B)^2}{\alpha_B} \le \frac{p_B^2}{\alpha_B}$$

And by the Cauchy-Schwarz inequality:

$$E(pfd_A \times pnp_B) \le \sqrt{E(pfd_A^2) \times E(pnp_B^2)}$$

from which the result follows.

## *Example 4*

If the assessor has a single percentile for each marginal distribution, as in Example 1:

$p_A=10^{-5}$, $\alpha_A=0.05$, and $p_B=10^{-2}$, $\alpha_B=0.05$

*and* the assessor is certain that $E(pfd_A) \le p_A$ and $E(pnp_B) \le p_B$, then

$$E(pfd_{sys}) \le \frac{10^{-5} \times 10^{-2}}{\sqrt{0.05 \times 0.05}} = 2 \times 10^{-6}$$

Obviously this is a tighter bound than in Example 1, using Theorem 1. In general, bounds (13) and (14) will be better than (6) and (9) whenever $\alpha_A >> p_A$ and $\alpha_B >> p_B$, which will generally be the case (claims will usually be much smaller numerically than doubts).

In fact it is even tighter than the bound in Example 2. At first glance this is surprising, since the latter requires the assessor to know *with certainty* upper bounds on the parameters, in addition to a percentile for each. The result here, however, similarly depends upon the assessor being *certain* that the marginal means are smaller than $p_A$, $p_B$ respectively. This is so even though the weaker bound of Theorem 4, (14), which is used in the example, does not depend on the numerical values of these marginal means.

In summary, the assessor does not need to know both the marginal means and the percentiles to use the theorem. Useful bounds on system *pfd* can be obtained by knowing either

(a) $E(pfd_A)$, $E(pnp_B)$, $\alpha_A$, $\alpha_B$ for result (13)

or

(b) $p_A$, $p_B$, $\alpha_A$, $\alpha_B$ for result (14)

but in each case they must be certain that, in addition, the marginal means are smaller than the corresponding percentiles (even if the exact values of some of these are not known to him).

Of the two options, (a) gives the tighter bound and thus can be regarded as preferable in those cases where the assessor knows each of $E(pfd_A)$, $E(pnp_B)$, $\alpha_A$, $\alpha_B$, $p_A$, $p_B$. In both cases, the bounds will be tighter for larger values of $\alpha_A$, $\alpha_B$. But of course larger values of $\alpha_A$, $\alpha_B$ are associated with smaller values of $p_A$, $p_B$, and if these are *too* small the bounds on the marginal means in Theorem 4 will be violated.

The tightest bound would occur if the assessor's percentiles ($p_A$, $p_B$) coincided exactly with their marginal means $E(pfd_A)$, $E(pnp_B)$ - in which case (a) and (b) give the same bound. Is it feasible that an assessor would be able to make them coincide in this way? In some cases an assessor may be prepared to specify a complete marginal distribution for each parameter (e.g. by accepting a parametric family, such as a 2-parameter Beta distribution, that is "fixed" by the determination of two percentiles – see Section 3). In that case the assessor will know $E(pfd_A)$, $E(pnp_B)$, they can choose $p_A$, $p_B$ to coincide with these values, and then compute the corresponding $\alpha_A$, $\alpha_B$ which will give the tightest bound.

## 3 Confidence bounds for system *pfd*

A different approach from the above obtains conservative *confidence bounds* for the system *pfd*, again without requiring estimation of the dependence of the assessor's beliefs about the unknown parameters $pfd_A$ and $pnp_B$.

As before, we assume that the expert can provide a marginal percentile for each parameter, as in (5). We again use the LR result concerning aleatory uncertainty.

Given these beliefs of the assessor concerning the individual channels of the 1-out-of-2 system, we are interested in obtaining a confidence bound for the *system pfd*. That is, we want to evaluate the probability

$$P(pfd_{sys} < p_{sys}) \tag{15}$$

for some value of $p_{sys}$.

### *Theorem 5*

Given the confidence bounds in (5), i.e.

$P(pfd_A < p_A) = 1 - \alpha_A$

$P(pnp_B < p_B) = 1 - \alpha_B$

we have

$$P(pfd_{sys} < p_A \times p_B) > 1 - (\alpha_A + \alpha_B) \tag{16}$$

*Proof:*

From (2)

$$P(pfd_{sys} < p_A \times p_B) > P(pfd_A \times pnp_B < p_A \times p_B) \tag{17}$$

Now

$$\begin{aligned} &P(pfd_A \times pnp_B > p_A \times p_B) \\ &< P(pfd_A > p_A) + P(pnp_B > p_B) - P(pfd_A > p_A, pnp_B > p_B) \end{aligned} \tag{18}$$

This is because the left hand side is the probability mass associated with the area above the hyperbola in Figure 3; this is smaller than the probability mass associated with the L-shaped region comprising rectangles ABED, BCFE, EFKH; which in turn is equal to probability masses of BCKH plus ACFD minus BCFE; these three probability masses correspond to the three terms on the RHS of (18), in the same order.

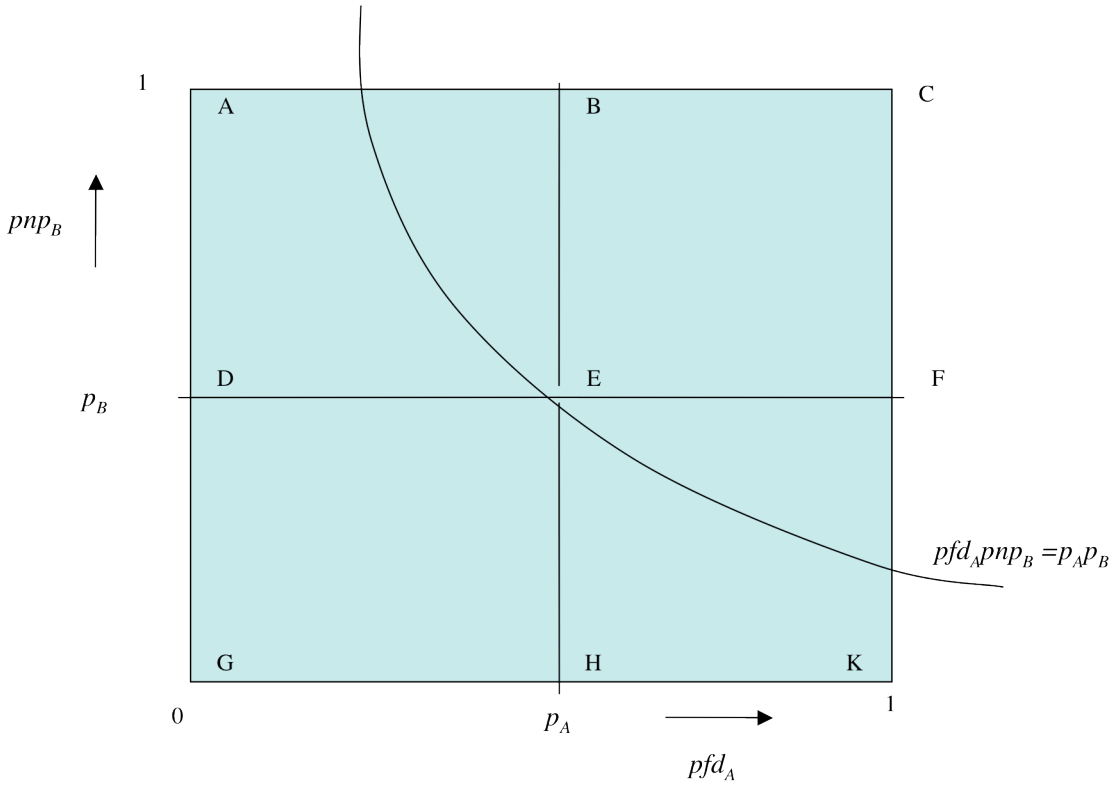


Figure 3. Essentially as Figure 1. Here the probability mass associated with the area below the hyperbola, $pfd_A pnp_B = p_A p_B$, corresponds to the probability on the right hand side of equation (13).

The last term on the right hand side of (18) is (most likely) not known – it would require the assessor to know about dependence between beliefs about parameters. So, conservatively, we have

$$P(pfd_A \times pnp_B > p_A \times p_B) < P(pfd_A > p_A) + P(pnp_B > p_B) \quad (19)$$

So finally

$$P(pfd_{sys} < p_A \times p_B) > P(pfd_A \times pnp_B < p_A \times p_B)$$

$$= 1 - P(pfd_A \times pnp_B > p_A \times p_B)$$

$$> 1 - P(pfd_A > p_A) - P(pnp_B > p_B) = 1 - (\alpha_A + \alpha_B)$$

which completes the proof.

Informally, the theorem states that the system claim is the product of the channel claims $(p_A \times p_B)$, and the doubt in this system claim is simply the sum of the channel claim doubts $(\alpha_A + \alpha_B)$.

## *Example 5*

For example, if an assessor is 95% confident (5% doubt) that $pfd_A$ is smaller than $10^{-5}$, and 95% confident (5% doubt) that $pnp_B$ is smaller than $10^{-2}$, then they are at least 90% confident (5%+5%=10% doubt) that $pfd_{sys}$ is smaller than $10^{-7}$.

## *Example 6*

If the assessor can provide two (or more) percentiles for each distribution, then multiple conservative percentiles can be generated for the distribution of $pfd_{sys}$. So if, in addition to the two percentiles above, the assessor is 99% confident that $pfd_A$ is smaller than $10^{-3}$, and 99.9% confident that $pnp_B$ is smaller than $10^{-1}$, the following conservative percentiles apply to their beliefs about the system *pfd*:

1. $Pfd_{sys}$ is smaller than $10^{-4}$ with 98.9% confidence (doubt = 1.1%)
2. $Pfd_{sys}$ is smaller than $10^{-5}$ with 94% confidence (doubt = 6%)
3. $Pfd_{sys}$ is smaller than $10^{-6}$ with 94.9% confidence (doubt = 5.1%)
4. $Pfd_{sys}$ is smaller than $10^{-7}$ with 90% confidence (doubt = 10%)

Notice that the bounding confidence in 3 above is greater than that in 2, even though the claim in 3 is a stronger one ($10^{-6}$ rather than $10^{-5}$): it should be recalled that these are conservative bounds, not exact values for confidence levels, and the "degree" of conservatism can vary. For example, an important contribution to the conservatism comes from ignoring the probability mass associated with the rectangle BCFE in Figure 1, and this will vary according to the marginal claims $p_A$, $p_B$.

This result can be generalized for the case where the assessor offers more than two percentiles for each distribution:

## *Corollary*

If the assessor offers several percentiles representing their beliefs about the parameters, as follows:

$$P(pfd_A < p_A^{(i)}) = 1 - \alpha_A^{(i)},\ \text{i} = 1, 2, ..., \text{m}$$
$$P(pnp_B < p_B^{(j)}) = 1 - \alpha_B^{(j)},\ \text{j} = 1, 2, ..., \text{n} \quad (20)$$

then all the following are conservative statements about the system *pfd*:

$$P(pfd_{sys} < p_{sys} = p_A^{(i)} \times p_B^{(j)}) > 1 - (\alpha_A^{(i)} + \alpha_B^{(j)}) \quad \forall (i, j) \quad (21)$$

Notice that different $(i,j)$ pairs may give the same "claim", $p_A^{(i)} \times p_B^{(j)}$, for different values of the "doubt", $(\alpha_A^{(i)} + \alpha_B^{(j)})$. Since all statements (25) are correct, it would be reasonable in such a case to use the smallest value of the doubt, since this will still be conservative.

In some cases, an assessor may be prepared to provide complete distributions, $F_A$, $F_B$, to represent their marginal beliefs about the two parameters $pfd_A$, $pnp_B$. Typically this might happen when the assessor is prepared to accept some parametric family of distributions (e.g. Beta distributions) that approximate to their general beliefs, and they can "fix" a particular pair by declaring one or more percentiles for each. In that case there will be a *continuous* version of the corollary above. That is, there will be an infinite number of $(\alpha_A,\alpha_B)$ pairs, each corresponding to one of an infinite number of $(p_A,p_B)$ pairs. For each statement of the kind $pfd_{sys} < p_{sys}$ there will be an infinite number of conservative doubts, as in (21) above. It is appropriate, as above, to take the least conservative in each case, so we have:

## *Theorem 6*

If, in a slightly extended notation, the functions

$$P(pfd_A > p_A) = \alpha_A(p_A)$$

and

$$P(pnp_B > p_B) = \alpha_B(p_B)$$

represent the assessor marginal doubts for all possible claims about the two parameters, there exists a *bounding distribution* for $pfd_{sys}$:

$$P(pfd_{sys} < t) = \max_{0<p_A\le 1}\left[0, \left(1 - \alpha_A(p_A) - \alpha_B(t / p_A)\right)\right]$$

## *Proof*

From (21)

$$P(pfd_{sys} < p_A \times p_B) > 1 - \alpha_A(p_A) - \alpha_B(p_B)$$

and

$$P(pfd_{sys} < t) > \max_{p_A \cdot p_B = t} \left[0, \left(1 - \alpha_A(p_A) - \alpha_B(p_B)\right)\right] = \max_{0 < p_A \le 1} \left[0, \left(1 - \alpha_A(p_A) - \alpha_B(t / p_A)\right)\right]$$

and the result follows.

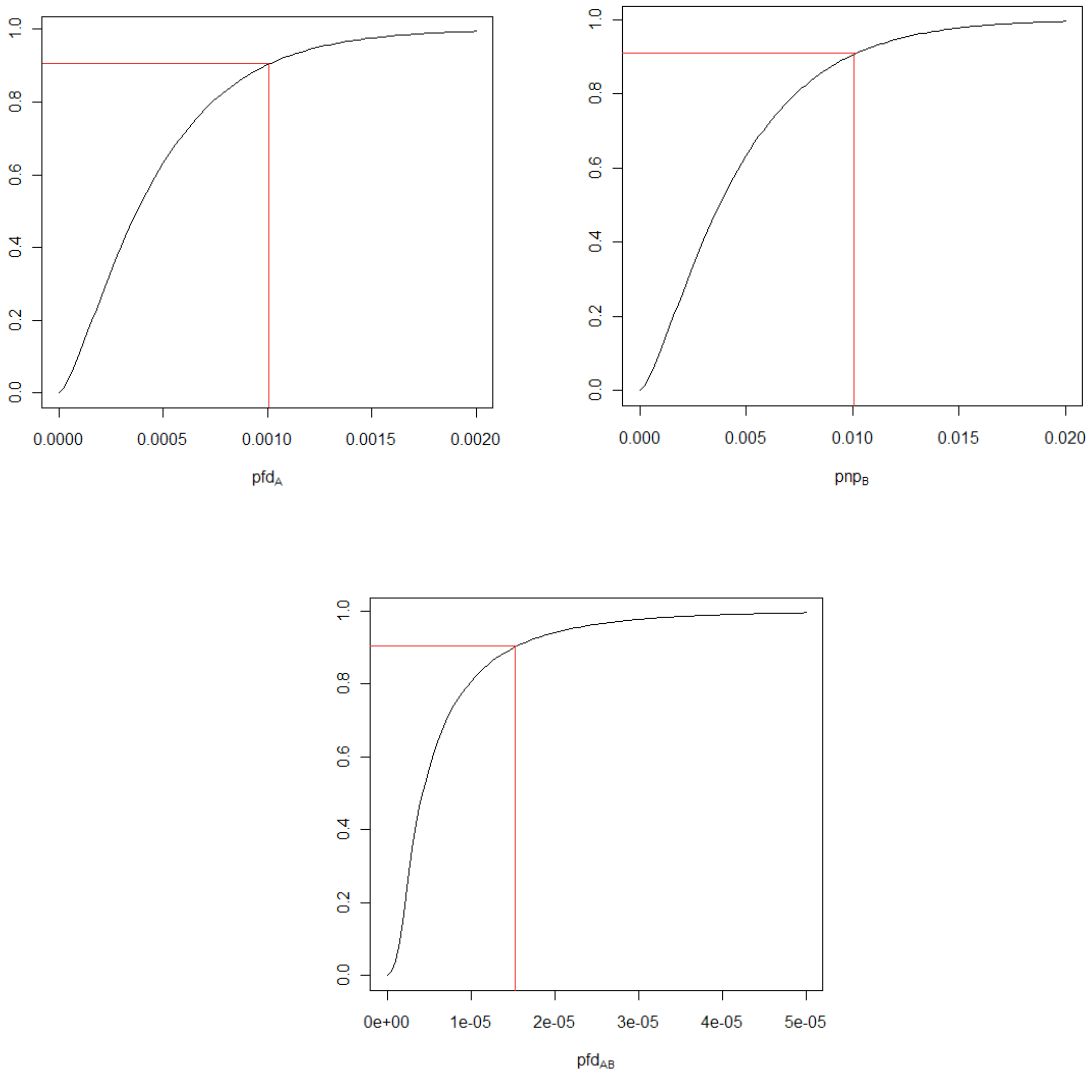


Figure 4 An example where the marginal distributions for $pfd_A$ and $pnp_B$ in the first two plots are respectively Beta(1.5, 3150) and Beta(1.5, 315). The third plot shows the resulting conservative (bounding) distribution for the system *pfd*.

In the case where the marginal distributions are continuous, the function

$$\alpha_A(p_A) + \alpha_B(t / p_A)$$

has a stationary point at $p_A = p_A^*$ satisfying the equation

$$(p_A^*)^2 \cdot \alpha'_A(p_A^*) = t \cdot \alpha'_B(t / p_B^*)$$

and using this for all $t$ we can obtain a bounding *continuous distribution* for $pfd_{sys}$. Alternatively this can be found by numerical optimization.

One way this result might be used is to take a particular level of doubt and propagate claims with this same fixed doubt throughout a case, or part of a case:

### *Example 7*

Figure 4 shows a case where both of the marginal distributions for the parameters are Betas. In the Figure we show the percentiles corresponding to a doubt of 10% for each of the three distributions (the choice of 10% is purely for illustration – it is not intended to represent a realistic figure for real cases). At this level of doubt, claims of 1.0e-03 and 1.0e-02 for $pfd_A$ and $pnp_B$ allow a claim of approximately 1.5e-05 to be made for the system *pfd*, and this is, of course, conservative. Readers may think that this near-product of the claims (1.5e-05 versus 1.0e-05), for the fixed 10% doubt, is rather a tight bound.

### *Example 8*

Because complete marginal distributions for the parameters, and the distribution of the system *pfd*, are known in this case, it follows that the corresponding means are known:

$E(pfd_A) = 0.000476, E(pnp_B) = 0.00474$

and

$E(pfd_{sys}) = 6.97\text{e-}06$

This compares favourably with the (unattainable) "perfect independence" case:

$E(pfd_A).E(pnp_B) = 2.26\text{e-}06.$

In fact, the Cauchy-Schwarz bound in this case, (11), is 3.75e-06 and is even tighter. However, the point of this approach, via a bounding *distribution*, is that it allows all bounding percentiles of the system *pfd* to be computed, not merely the mean.

## 4 Discussion

Problems concerning different kinds of dependence have dogged the assessment of design diverse multi-version systems since this approach was first proposed in the 1970s.

The LR work addressed the problem of *aleatory dependence between failures*. It showed that, for a 1-out-of-2 system in which one channel is possibly perfect, the system *pfd* is bounded above by the simple product of $pfd_A$ and $pnp_B$, so that the (presumed) dependence between failures of the channels is not required to be known. *Pace* Kruskal: "You *can* multiply", and the resulting product will be conservative.

In practice, of course, the parameters $pfd_A$ and $pnp_B$ will be unknown. This introduces another problem of dependence, namely that of *epistemic dependence between an assessor's beliefs about the two parameters*. In the present paper we have presented two new ways of avoiding this problem of epistemic dependence: each depends only upon an

assessor's *marginal* distributions for (i.e. beliefs about) the parameters. Once again, though, the price paid is further conservatism in the results.

The first approach, in Section 2, obtains bounds as in LR for an assessor's posterior mean system *pfd*. In our second approach, in Section 3, we obtain conservative confidence bounds for the system *pfd*.

The different bounds in these sections are based upon different (marginal) beliefs that the assessor may have about the two channel parameters – i.e. what they know, or are prepared to declare. Not surprisingly, the problem is generally easier for *pfd* than it is for *pnp*. For example, in (Bishop, Bloomfield et al. 2011), conservative results for a software-based channel *pfd* are obtained based on observation of extensive failure-free working; in (Littlewood and Wright 2007) results are obtained based on evidence of failure-free working *and* evidence of successful formal verification. Although both papers briefly address probability of perfection, as does (Bertolino and Strigini 1998), this is a problem that has received comparatively little attention: it is the subject of current research by the authors.

Throughout this paper, and previously in (Littlewood and Rushby 2012), it has been emphasised that the results are conservative. At the time that these kinds of assessment take place – e.g. when a decision is being made about whether a system is "good enough" to be allowed into service – it will not be possible to say *how* conservative they are (the true system *pfd* will not be available, of course, to compare with the conservative estimates obtained from the results here). Later, after massive operational exposure, such as that experienced with some aircraft types and reported in (Boeing 2012), it may be possible to obtain trustworthy estimates of the true *pfd* and so assess the degree of conservatism.

The main sources of conservatism here are two-fold. Firstly, there is that which arises from the avoidance of failure dependence issues in the LR model, represented by equation (2). Secondly, there is that which arises from the avoidance of epistemic dependence issues (so that only marginal beliefs about the LR parameters are needed), represented by the theorems of this paper.

In both cases the conservatism is intrinsic to the approach, and thus unavoidable. However, in cases where an assessor is able to express beliefs about the parameters in different ways – so making more than one bounding result available from sections 2 or 3 – it would be reasonable to choose the least conservative one(s).

The choice between the two approaches of sections 2 and 3 will depend upon how the results will be used, and in particular upon the demands of a wider safety case for which claims about the present 1-out-of-2 system (e.g. a protection system) are only a part.

The motivation in the original LR work for obtaining a bound on the assessor's posterior expected system *pfd* (as we have done here in Section 2) was that, for a Bayesian assessor, this *is* their system probability of failure on demand. It is the number they would give in answer to the question: "What is the probability that the system will fail on a randomly selected demand?" However there are some subtleties here that present pitfalls for the unwary. For example, the answer to the question "What is the probability that the system will survive the $n$ demands it will experience in its lifetime?" is not a simple function of the posterior expected system *pfd* of section 2. That is:

$$E\left((1-pfd_{sys})^n\right) \neq \left(1-E(pfd_{sys})\right)^n$$

It follows that the results of Section 2 do not provide an answer to this question, and other similar ones, directly.

A different view is that the assessor is uncertain about the system *pfd* – their uncertainty being represented by their posterior *distribution* for this – and so they should propagate *this uncertainty* through the wider plant safety case (alongside, for example, uncertainties associated with other subsystems), so that any top-level plant claim will have an associated confidence. This is more in the spirit of the results of Section 3. However, it should be noted that such propagation of "complete" uncertainty throughout a complex wider case could be very difficult.

Even if the results from this approach were to be *very* conservative, we nevertheless believe that it offers a useful rigour for reasoning about certain types of critical systems. Assuming the availability of the necessary parameters for the theorems in Sections 2 and 34, our approach allows safety cases to be based on claims about system dependability that are guaranteed to be conservative. These are surely better than ones based, for example, upon informal claims about "degrees of dependence": we know no way such claims can be justified rigorously.

Finally, it is worth emphasising that all the results here depend critically on the basic LR result concerning aleatory uncertainty: that system *pfd* can be conservatively bounded by the simple product of channel *A*'s *pfd* and channel *B*'s *pnp*. None of these results can be applied to the case of a 1-out-of-2 system in which *pfd* claims must be made about *both* channels (because each is too complex for a claim of "possibly perfect"): clearly many systems are of this kind (e.g. some complex control or guidance systems). However, we maintain our belief that this special architecture is a plausible one for some important real systems (e.g. some protection systems, e.g. some architectures in which the second channel is a simple monitor) – see (Littlewood and Rushby 2012) for further discussion and examples.

## Acknowledgements

We are grateful to the Associate Editor and four reviewers for thoughtful comments on an earlier version of this paper.

Support for the work reported here came from:

- the INDEED project, funded by EPSRC;
- the UnCoDe project, funded by the Leverhulme Trust;
- The DISPO project - funded under the C&I Nuclear Industry Forum (CINIF) Nuclear Research Programme by EDF Energy Limited, Nuclear Decommissioning Authority (Sellafield Ltd, Magnox Ltd), AWE plc, Urenco UK Ltd and Horizon Nuclear Power.The views expressed in this paper are those of the author(s) and do not necessarily represent the views of CINIF members. CINIF does not accept liability for any damage or loss incurred as a result of the information contained in this paper.

## References

Bertolino, A. and L. Strigini (1998). "Assessing the risk due to software faults: estimates of failure rate vs evidence of perfection." Journal of Software Testing, Verification and Reliability **8**(3): 155-166.

Bishop, P., R. Bloomfield, et al. (2011). "Towards a formalism for conservative claims about the dependability of software-based systems." IEEE Trans Software Engineering **37**(5): 708-717.

Boeing (2012). Statistical Summary of Commercial Airplane Accidents, Worldwide Operations, 1959-2011. Seattle, Aviation Safety, Boeing Commercial Airplanes.

Butler, R. W. and G. B. Finelli (1993). "The infeasibility of quantifying the reliability of life-critical real-time software." IEEE Trans Software Engineering **19**(1): 3-12.

Eckhardt, D. E., A. K. Caglayan, et al. (1991). "An experimental evaluation of software redundancy as a strategy for improving reliability." IEEE Trans Software Eng **17**(7): 692-702.

Eckhardt, D. E. and L. D. Lee (1985). "A Theoretical Basis of Multiversion Software Subject to Coincident Errors." IEEE Trans. on Software Engineering **11**: 1511-1517.

Knight, J. C. and N. G. Leveson (1986). "Experimental evaluation of the assumption of independence in multiversion software." IEEE Trans Software Engineering **12**(1): 96-109.

Kruskal, W. (1988). "Miracles and Statistics: The Casual Assumption of Independence." Journal of the American Statistical Association **83**(404): 929-940.

Littlewood, B. and D. R. Miller (1989). "Conceptual Modelling of Coincident Failures in Multi-Version Software." IEEE Trans on Software Engineering **15**(12): 1596-1614.

Littlewood, B., P. Popov, et al. (2002). "Modelling software design diversity - a review." ACM Computing Surveys **33**(2): 177-208.

Littlewood, B. and J. Rushby (2012). "Reasoning about the reliability of diverse two-channel systems in which one channel is 'possibly perfect'." IEEE Trans Software Engineering **38**(5): 1178-1194.

Littlewood, B. and L. Strigini (1993). "Validation of ultra-high dependability for software-based systems." CACM **36**(11): 69-80.

Littlewood, B. and D. Wright (2007). "The use of multi-legged arguments to increase confidence in safety claims for software-based systems: a study based on a BBN of an idealised example." IEEE Trans Software Engineering **33**(5): 347-365.

Rouquet, J. C. and P. J. Traverse (1986). Safe and reliable computing on board the Airbus and ATR aircraft. Safecomp: 5th IFAC Workshop on Safety of Computer Control Systems, Pergamon Press.

Wood, R. T., R. Belles, et al. (2010). Diversity Strategies for Nuclear Power Plant Instrumentation and Control Systems. Washington, DC, US Nuclear Regulatory Commission.

## Appendix

**Lemma:**

If

$0 \le X \le 1$, and $P(X>p)=\alpha$, and $E(X) \le p$,

then

$$E(X^2) \le \frac{E(X)^2}{\alpha} \le \frac{p^2}{\alpha}$$

**Proof:**

Let $E(X) = m$ and $E(X^2) = s^2$

We show that there exists a two-point discrete random variable, $Y$, as follows:

$$P(Y = y) = 1 - \alpha$$
$$P(Y = z) = \alpha$$

where

$$0 \le y \le m \le z \le 1$$

and

$$E(Y) = y \times (1-\alpha) + z \times \alpha = m$$
$$E(Y^2) = y^2 \times (1-\alpha) + z^2 \times \alpha = s^2 \qquad \text{(A1)}$$

From (A1) we have $y = \frac{m - z \times \alpha}{1-\alpha}$ and

$$E(X^2) = E(Y^2) = G(z) = \frac{(m - z \times \alpha)^2}{1-\alpha} + z^2\alpha$$

If $\alpha > 0$ the equation $G(z) = s^2$ has a positive real root:

$$z = m + \sqrt{(s^2 - m^2) \times \frac{1-\alpha}{\alpha}}$$

since $s^2 - m^2 = Var(X) \ge 0$.

It follows that the random variable $Y$ always exists.

Now, if z>0, then $G(z)$ is increasing because

$$\frac{dG(z)}{dz} = -\frac{2\alpha(m - z \times \alpha)}{1-\alpha} + 2z\alpha = \frac{-2m\alpha + 2\alpha^2 z + 2\alpha z - 2\alpha^2 z}{1-\alpha} = \frac{2\alpha(z-m)}{1-\alpha} \geq 0;$$

and

$y \geq 0$ implies $z \leq m/\alpha$

therefore

$$E(X^2) = E(Y^2) \leq m^2/\alpha = E(X)^2/\alpha \leq p^2/\alpha$$

that is

$$E(X^2) \leq \frac{E(X)^2}{\alpha} \leq \frac{p^2}{\alpha}$$

**QED**

**Bev Littlewood** has degrees in mathematics and statistics, and a PhD in statistics and computer science; he is a Chartered Engineer, and a Chartered Statistician. He has worked for more than 30 years on problems associated with the dependability of software-based systems, and has published many papers in international journals and conference proceedings and has edited several books. His technical contributions have largely focused on the application of rigorous probabilistic and statistical techniques in software systems engineering. In 1983 he founded the Centre for Software Reliability (CSR) at City University, London, and was its Director until 2003. He is currently Professor of Software Engineering in CSR. From 1990 to 2005 he was a member of the UK Nuclear Safety Advisory Committee. He is a member of IFIP Working Group 10.4 on Reliable Computing and Fault Tolerance, of the UK Computing Research Committee, and is a Fellow of the Royal Statistical Society. He is on the editorial boards of several international journals. In 2007 he was the recipient of the IEEE Computer Society's Harlan D Mills Award.

**Andrey Povyakalo** has MSc equivalent degree in computer-aided control and management from the Moscow State Institute for Physics and Engineering (MIPHE, 1985) and PhD equivalent degree in mathematical modelling in nuclear engineering from the Institute of Nuclear Power Engineering, Obninsk, Russia (INPE, 1994). In 1985-2001 he worked for INPE as lecturer, senior lecturer and associate professor contributing to a number of research projects related to probabilistic safety analysis and risk assessment (PSA/PRA). In 2001 Andrey joined the Centre for Software Reliability (CSR), City University, London as a research fellow working for the interdisciplinary research collaboration on dependability of computer-based systems (DIRC, 2000-2006) funded by the UK Engineering and Physical Science Research Council (EPSRC). In 2005 (in co-authorship with E. Alberdi, L. Strigini and P. Ayton) he was a recipient of the 2005 Herbert M. Stauffer Award for the "Best Clinical Paper" from the Association of University Radiologists. At present, he is a senior lecturer in dependability of socio-technical systems. His research interests are mostly related to probabilistic modeling of confidence in assurance cases for computer-based systems.