A Robust Data Driven Approach to Quantifying Common-Cause Failure in Power Networks

Matthias C. M. Troffaes
Durham University, UK
matthias.troffaes@gmail.com

Simon Blake
Newcastle University, UK
simon.blake@ncl.ac.uk

Abstract

The standard alpha-factor model for common cause failure assumes symmetry, in that all components must have identical failure rates. In this paper, we generalise the alpha-factor model to deal with asymmetry, in order to apply the model to power networks, which are typically asymmetric. For parameter estimation, we propose a set of conjugate Dirichlet-Gamma priors, and we discuss how posterior bounds can be obtained. Finally, we demonstrate our methodology on a simple yet realistic example.

Keywords. robust, alpha-factor, failure, reliability, Gamma, Dirichlet

1 Introduction

When modelling power networks, typically, the basic event we are interested in are loss of so-called *security zones*. A security zone makes up a collection of components, so that if one component in the zone fails, power in the whole zone is lost. Security zones are typically bounded by circuit breakers, which allow isolating consequences of faults.

An interesting problem occurs when faults in different zones do not occur independently. For example, power lines in adjacent zones often share transmission towers. A landslide, for instance, can cause the tower to collapse, affecting both zones simultaneously. It is important that the frequency of such events is taken into account, as otherwise the actual risk to the network might be underestimated.

The standard literature for common cause failure modelling assumes symmetry [5, 9], however, clearly, for our purpose, security zones will typically not exhibit symmetry, due to differences in layout, composition, and age of constituents. In this paper, we adapt the approach of Troffaes et al. [9] to allow for asymmetry.

In doing so, as opposed to existing methods [3, 4], we enable a more data driven approach to network reliability analysis. Specifically, we allow actual failure data on the network—which is most informative, but typically also very sparse—to be combined with say national average failure rates—such data is typically far more abundant, but also not necessarily as applicable to the specific network at hand due to specific local conditions which may be hard to model, let alone to be quantified.

A key feature of our approach is built-in sensitivity analysis against ill-known parameters, following [10, 7, 8, 9]. Following recent work on prior-data conflict [11, 12, 9], in this paper, we will focus on sensitivity analysis in the so-called learning parameters of the model, which essentially tells us how much we should weigh network specific data against our prior expectations informed by say national averages.

The paper is structured as follows. Section 2 derives the mathematical model for dealing with common cause failures in asymmetric two component systems. Section 3 discusses the statistical problem of how to estimate the parameters of the model. We construct a likelihood for typical kinds of data available. We then propose a conjugate prior, which is an independent product of a Dirichlet (or beta) prior and two Gamma priors. Finally, we discuss how sensitivity analysis can be performed to obtain posterior bounds. Section 4 works through an actual example. Section 5 concludes the paper.

2 Modelling Common Cause Failure for Asymmetric Components

2.1 Two Component Model

In this discourse, a 'component' denotes any subsystem, which, for the purpose of common cause analysis, we do not subdivide any further. In particular, it does not need to denote a separate electrical component of

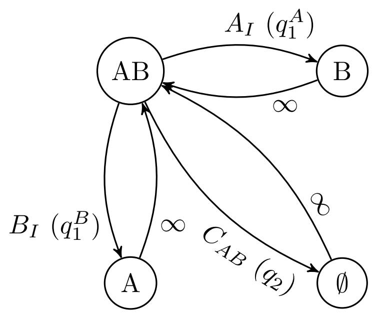

Figure 1: Markov chain for failure with instant repair. The nodes show non-faulty zones.

the power network. For example, if we are merely interested in the loss of security zones, a component could be taken to be such security zone.

Let us call these components A and B. Now, following the basic parameter model of Mosleh et al. [5] (also see [9]), one traditional way to model common cause failures is to attribute all failures to any of the following three events:

- A_I: independent failure of A
- B_I: independent failure of B
- C_{AB}: common cause failure of both A and B

These three events are assumed to be generated by independent Poisson processes. For simplicity, in this exposition, we assume that repair is immediate.[1] Figure 1 depicts the corresponding continuous time Markov chain, along with rates for all transitions.

Following standard notation in common cause failure modelling, by q_1^A we denote the rate of A_I, by q_1^B we denote the rate of B_I, and by q_2 we denote the rate of C_{AB}. The subscript of the q denotes the number of components involved (or is t for 'total', as in the next paragraph). The superscript denotes the particular component, and is required due to lack of symmetry. For comparison, in the standard basic parameter model, we would have $q_1^A = q_1^B = q_1$.

A key challenge is that we do not observe these events directly. Indeed, often, we have a good idea of the rate at which each component fails, that is, we know

$$q_t^A = q_1^A + q_2 \tag{1}$$
$$q_t^B = q_1^B + q_2 \tag{2}$$

Additionally, we may have a fairly good idea of what fraction α_2 of faults is due to a common cause. The fraction of faults not due to a common cause is $\alpha_1 := 1 - \alpha_2$.

For example, say that we have a sequence of 100 independent observations in which a fault occurs, and say that in exactly 18 of those observations, both components failed. Then, to a good approximation, α_2 would simply be 0.18. The parameters α_1 and α_2 are called *alpha-factors*.

So, we have three observable quantities: q_t^A, q_t^B, and α_1—note that $\alpha_2 = 1 - \alpha_1$. From these, we need to derive three model parameters: q_1^A, q_1^B, and q_2. Here, the only difference with the standard basic parameter model in the literature is that we do not assume $q_1^A = q_1^B$ (and whence, also not that $q_t^A = q_t^B$). This difference may seem only very subtle, particularly for the case where only two components are involved, however, the consequent mathematical treatment is notably different to merit a careful consideration, as follows.

We can easily express α_1 and α_2 in terms of the above parameters, once noted that a fraction of faults can be written as a ratio of fault rates:

$$\alpha_1 = \frac{q_1^A + q_1^B}{q_1^A + q_1^B + q_2} \tag{3}$$
$$\alpha_2 = \frac{q_2}{q_1^A + q_1^B + q_2} \tag{4}$$

Now, consider the combination $\alpha_1 + 2\alpha_2$:

$$\alpha_1 + 2\alpha_2 = \frac{q_1^A + q_1^B + 2q_2}{q_1^A + q_1^B + q_2} = \frac{q_t^A + q_t^B}{q_1^A + q_1^B + q_2} \tag{5}$$

Consequently,

$$q_1^A + q_1^B + q_2 = \frac{q_t^A + q_t^B}{\alpha_1 + 2\alpha_2} \tag{6}$$

where the right hand side now consists of observable quantities. Plugging this expression into our earlier expression for α_2, we find:

$$\alpha_2 = \frac{q_2(\alpha_1 + 2\alpha_2)}{q_t^A + q_t^B} \tag{7}$$

so, consequently:

$$q_2 = \frac{\alpha_2}{\alpha_1 + 2\alpha_2}(q_t^A + q_t^B) \tag{8}$$

We recovered one of the model parameters. For the other ones, simply use:

$$q_1^A = q_t^A - q_2 \tag{9}$$
$$q_1^B = q_t^B - q_2 \tag{10}$$

[1] Note that, consequently, simultaneous failures due to independent causes of the two components have probability zero.

2.2 Preliminary Example

To demonstrate the theory so far developed, we apply it on a simple example. Athough in the following, the probabilities are entirely made up, they are representative of typical power networks.

Suppose we have a collection of customers supplied from two security zones, named A and B, where loss of power in both zones will result in customer interruption. Suppose, for the sake of argument, that, on average, per year, we observe 3 faults in zone A, and 5 faults in zone B. We also know that, from historical data, 15% of all faults in these zones results in customer interruption. What is the rate at which we lose customers?

Following the above model, we have:

$$q_t^A = 3 \tag{11}$$

$$q_t^B = 5 \tag{12}$$

assuming rates are expressed per year, and

$$\alpha_1 = 0.85 \tag{13}$$

$$\alpha_2 = 0.15 \tag{14}$$

Then, following the earlier analysis, we find that:

$$q_2 = \frac{\alpha_2}{\alpha_1 + 2\alpha_2}(q_t^A + q_t^B) \tag{15}$$

$$= \frac{0.15}{0.85 + 2 \times 0.15}(3 + 5) = 1.043 \tag{16}$$

The rate at which customer interruption occurs is exactly $q_2 = 1.043$, or about one per year. Note that we can also derive the rate at which independent failures occur:

$$q_1^A = q_t^A - q_2 = 3 - 1.043 = 1.957 \tag{17}$$

$$q_1^B = q_t^B - q_2 = 5 - 1.043 = 3.957 \tag{18}$$

3 Parameter Estimation from Data

An obvious challenge with our statistical model is that we need to estimate the failure rates of each component (or, security zone), as well as the fraction of double failures. Information relating to these probabilities can come from a variety of sources.

Two options present themselves:

1. Use historical failure data of single and double failures in the network under study to estimate the parameters q_1^A, q_1^B, and q_2, directly, say using maximum likelihood. A problem here is that, typically, for one specific network, not very much data may be available.
2. Use average nationwide failure rates q_t^A and q_t^B, along with average nationwide double failure fraction α_2. The methodology of Section 2.1 then applies to find q_1^A, q_1^B, and q_2. As there is far more nationwide data available, one would hope that this leads to more accurate estimates for q_t^A, q_t^B, and α_2. A key problem here is that it is not clear to what extent nationwide averages will also apply to the specific network under study.

In this treatment, we use both sources of information: aggregated nationwide failure probabilities for components, obtained by averaging, as well as local data specific to the location of interest. As already mentioned, the latter sort of data is typically very sparse, but at the same time also more informative, as it can incorporate known information about individual asset condition, age, location (e.g. exposure to extreme weather, marine corrosion or industrial pollution), level of utilisation and actual fault history.

We now propose a conjugate Bayesian model for dealing with both types of data. Specifically, we use the aggregated data to construct a prior, and then use the likelihood of the local data to update this prior to a posterior. From a likelihood perspective, the prior simply represents pseudo counts, so effectively, we are really simply adding local data to the nationwide data to obtain a local prediction.

A key question is: how strong should the nationwide data be weighed in comparison to the local data? Or, phrased differently: how relevant do we believe is the nationwide data for making predictions about the local situation? In conjugate analysis [1], there is a natural parameter which represents this subjective judgement. What we will do is perform a sensitivity analysis against this parameter, very similar to what is done in for instance the imprecise Dirichlet model, or more generally, in the exponential family [10, 6, 11, 12].

Observe that the expression for q_2 in terms of the alpha-factors α_1, α_2 and total failure rates q_1^A, q_1^B (Eq. (8)) can be written as a function of just the alpha-factors, times a function of just the total failure rates. So, inspired by [9], for a joint prior, we use an independent product of two Gamma distributions, one on q_t^A and one on q_t^B, and a Dirichlet (or, beta) distribution jointly on α_1 and α_2. We now elaborate on this in the following sections.

3.1 Dirichlet Prior for Alpha-Factors

A natural way to estimate alpha-factors goes via a sequence of N observations, where n_1 of those involved single failures of either A or B (but not both), and the remaining n_2 involved double failures of both A and B. The corresponding likelihood is:

$$\Pr(n_1, n_2 \mid \alpha_1, \alpha_2) = \binom{N}{n_1} \alpha_1^{n_1} \alpha_2^{n_2}. \tag{19}$$

A conjugate prior for the above likelihood is the Dirichlet density (or, beta density, as we have only two categories):

$$f(\alpha_1, \alpha_2 \mid s, t_1, t_2) \propto \alpha_1^{st_1-1} \alpha_2^{st_2-1} \tag{20}$$

with hyperparameters $s > 0$ and t_1, $t_2 \in (0,1)$ such that $t_1 + t_2 = 1$. The posterior density is simply:

$$f(\alpha_1, \alpha_2 \mid n_1, n_2, s, t_1, t_2) \propto \alpha_1^{st_1+n_1-1} \alpha_2^{st_2+n_2-1} \tag{21}$$

By Eq. (8), we will need to find the posterior expectation of

$$\frac{\alpha_2}{\alpha_1 + 2\alpha_2}, \tag{22}$$

where we remind the reader that $\alpha_1 + \alpha_2 = 1$, and typically, α_2 is expected to be small. As discussed in great detail in [9], we can do so via Taylor expansion. For example, with second order expansion:

$$E\left(\frac{\alpha_2}{\alpha_1 + 2\alpha_2} \middle| n_1, n_2, s, t_1, t_2\right) \tag{23}$$

$$\approx E(\alpha_2 - \alpha_2^2 | n_1, n_2, s, t_1, t_2) \tag{24}$$

$$= \frac{n_2 + st_2}{N + s} \left(1 - \frac{n_2 + st_2 + 1}{N + s + 1}\right) \tag{25}$$

using the well-known properties of the Dirichlet distribution (for example, see [9, Eq. (10)]); we remind the reader that $N = n_1 + n_2$. For this approximation, the absolute error is less than:

$$\frac{n_2 + st_2}{N + s} \frac{n_2 + st_2 + 1}{N + s + 1} \frac{n_2 + st_2 + 2}{N + s + 2}. \tag{26}$$

3.2 Gamma Prior for Total Failure Rates

To estimate total failure rates, assume we have observed a component (A or B) for time T, during which this component failed M times. The likelihood for the failure rate q_t of this component is then:

$$\Pr(M \mid q_t, T) = \frac{(q_t T)^M \exp(-q_t T)}{M!} \tag{27}$$

as we assumed a Poisson process. A conjugate prior for this likelihood is the Gamma density:

$$f(q_t \mid u, v) \propto q_t^{uv-1} \exp(-q_t u) \tag{28}$$

with hyperparameters $u > 0$ and $v > 0$. The posterior density is:

$$f(q_t \mid M, T, u, v) \propto q_t^{uv+M-1} \exp(-q_t(u + T)) \tag{29}$$

By Eq. (8), of interest is the posterior expectation of q_t, which is simply:

$$E(q_t \mid M, T, u, v) = \frac{T}{u+T}\frac{M}{T} + \frac{u}{u+T}v. \tag{30}$$

Considering this posterior expectation when $T = 0$, we see that v represents a prior expectation for q_t, and considering this posterior expectation when $T = u$, we see that u represents the time T needed before the posterior starts to move away from this prior [9, Sec. 3.2].

3.3 Full Analysis

Let us put everything together.

For the alpha-factors, suppose our prior expected fraction of single failures is t_1, and our prior fraction of double failures is t_2. Moreover, we observed n_1 single failures, and n_2 double failures. We are rather unsure about how much weight to assign to the prior, that is, we are unsure about the hyperparameter s. Remember, in a likelihood interpretation of Bayesian inference, s can be thought of the total pseudo count assigned to the prior. Say, $s \in [\underline{s}, \overline{s}]$; for example, with $\underline{s} = 0$ and $\overline{s} = 5$, we count the prior for no more than five observations. As discussed in [9], it seems quite sensible to perform a sensitivity analysis over s, to properly cope with prior-data conflict.

For the total failure rates, suppose our prior expected failure rates are v^A and v^B. Moreover, we observed M^A failures of component A during a time span of T, and M^B failures of component B during a time span of T. For simplicity we take the observed time spans for both components to be identical, as this is the case for our application, but it could be relaxed easily. Again, we are rather unsure about the hyperparameters u^A and u^B—for simplicity, we will also take these to be equal: $u := u^A = u^B$ (again this could be relaxed easily). Here, u can be thought of a pseudo observation time assigned to the prior. Say, $u \in [\underline{u}, \overline{u}]$; for example, with $\underline{u} = 0$ and $\overline{u} = 3$, we count the prior failure rates for no more than 3 years.

Consequently, by Eqs. (8), (25), and (30),

$$\underline{E}(q_2 \mid D) = \inf_{\substack{s \in [\underline{s}, \overline{s}] \\ u \in [\underline{u}, \overline{u}]}} E(q_2 \mid D, s, u), \tag{31}$$

$$\overline{E}(q_2 \mid D) = \sup_{\substack{s \in [\underline{s}, \overline{s}] \\ u \in [\underline{u}, \overline{u}]}} E(q_2 \mid D, s, u). \tag{32}$$

When we assume independence between the alpha-factors and the total failure rates, the expectation decomposes into a product:

$$E(q_2 \mid D, s, u) = \frac{n_2 + st_2}{N + s}\left(1 - \frac{n_2 + st_2 + 1}{N + s + 1}\right) \times \frac{u(v^A + v^B) + M^A + M^B}{u + T} \tag{33}$$

and

$$D \coloneqq (n_1, n_2, M^A, M^B, T, t_1, t_2, v^A, v^B). \tag{34}$$

Note that the optimization problem for s and u can be solved through two independent optimisation problems, one in just s, and one in just u. For the optimisation in u, due to the monotonicity of the objective function, it suffices look at just $\underline{u}$ and $\overline{u}$. The objective function in s is not always monotone (although it often will be), but nevertheless numerical optimisation is still quite easy. The example provides more detail.

Note that bounds for the lower and upper posterior expectations of q_1^A and q_1^B can be derived in a very similar way, through Eqs. (9) and (10)—we leave this to the reader.

4 Network Risk Example

4.1 Problem Description

Following is a generic double circuit reliability problem, based on an actual case study in the North-East of England.

There are two unequal circuits. Circuit A has an expected failure rate of 0.3856 per year, based on 2 transformers and 24.1 km of line and cable. Circuit B has an expected failure rate of 0.3279 per year, based on 1 transformer and 21.5 km of line and cable. No adjustments have been made for asset condition. In the past 12 years, circuit A has experienced 7 failures in 12 years, and circuit B has experienced 4 failures in 12 years. Of these failures, 3 were double failures. For a group of 11 neighbouring (and similar) circuits, there have been 38 failures, of which 24 were single failures, and 14 were double failures—these 38 include the circuit we are studying. On average, for a much larger group of circuits at that voltage, but not necessarily similar to the double circuit under study, about 18% of all failures are double failures.

4.2 Prior and Data

As global prior for the alpha-factors, we use the global average: $t_1 = 0.82$ and $t_2 = 0.18$. It seems reasonable to use neighbouring circuits to correct our prior information about the alpha-factors of our circuit: so $n_1 = 24$ and $n_2 = 14$.

For the total failure rates, an expert provided us with some prior expectations based on global averages of failures for the particular components that make up the circuits: $v^A = 0.3856$ and $v^B = 0.3279$. We have $M^A = 7$ failures during $T^A = 12$ years of circuit A, and $M^B = 4$ failures during $T^B = 12$ years of circuit B.

All we need in addition is some assessment about s (number of total failures needed before we start to move away from the prior in the direction of the data for alpha-factors) and u (time needed before starting to move away from prior in direction of the data for total failure rates). As discussed in Section 3.3, we will perform a sensitivity analysis over intervals for both s and u. Let us take $s = [0, 15]$ and $u = [0, 10]$, which seem conservative yet reasonable given their interpretation discussed earlier.

4.3 Posterior Bounds

We must solve the optimisation problems in Eqs. (31) and (32), using Eq. (33). Let

$$f(s) \coloneqq \frac{n_2 + st_2}{N + s}\left(1 - \frac{n_2 + st_2 + 1}{N + s + 1}\right), \tag{35}$$

$$e(s) \coloneqq \frac{n_2 + st_2}{N + s}\frac{n_2 + st_2 + 1}{N + s + 1}\frac{n_2 + st_2 + 2}{N + s + 2}, \tag{36}$$

$$g(u) \coloneqq \frac{u(v^A + v^B) + M^A + M^B}{u + T}. \tag{37}$$

where $e(s)$ represents a bound on the absolute error, as $f(s)$ is only an approximation (see Eq. (26)). With

$$\underline{f} \coloneqq \inf_{s \in [0,15]} f(s) \qquad \overline{f} \coloneqq \sup_{s \in [0,15]} f(s) \tag{38}$$

$$\overline{e} \coloneqq \sup_{s \in [0,15]} e(s) \tag{39}$$

and

$$\underline{g} \coloneqq \inf_{u \in [0,10]} g(u) = \min_{u \in \{0,10\}} g(u) \tag{40}$$

$$\overline{g} \coloneqq \sup_{u \in [0,10]} g(u) = \max_{u \in \{0,10\}} g(u), \tag{41}$$

where we used that g is a monotone function, we then have that,[2]

$$\underline{E}(q_2 \mid D) \geq (\underline{f} - \overline{e})\underline{g} \tag{42}$$

$$\overline{E}(q_2 \mid D) \leq (\overline{f} + \overline{e})\overline{g} \tag{43}$$

[2] Instead of $\overline{e}$, we could use $e(\arg\inf_{s \in [0,15]} f(s))$ and $e(\arg\sup_{s \in [0,15]} f(s))$ to arrive at slightly better error bounds, but in practice it makes little difference.

By numerical optimisation, we find

$$\underline{f} = 0.212 \qquad \overline{f} = 0.227 \qquad \overline{e} = 0.057 \tag{44}$$
$$\underline{g} = 0.824 \qquad \overline{g} = 0.917 \tag{45}$$

Concluding,

$$\underline{E}(q_2 \mid D) \geq 0.128 \tag{46}$$
$$\overline{E}(q_2 \mid D) \leq 0.260 \tag{47}$$

Note that the absolute error $\overline{e}$ is rather large in comparison to $\underline{f}$ and $\overline{f}$—this is due to the fact that the data reflects a rather high value for α_2, and low order approximations only work well when α_2 is less than 0.1. Using instead a sixth order approximation (the equations are very easy to compute, but rather long to write down, see [9] for details; also note that the approximation scheme is designed for ease of computation at the expense of requiring the use of higher order terms, and that more sophisticated techniques might achieve this accuracy with fewer terms), we find:

$$\underline{f} = 0.237 \qquad \overline{f} = 0.266 \tag{48}$$
$$\overline{e} = 0.002 \tag{49}$$

so, because $\overline{e}$ is quite small,

$$\underline{E}(q_2 \mid D) \approx 0.194 \tag{50}$$
$$\overline{E}(q_2 \mid D) \approx 0.245, \tag{51}$$

or in other words, we expect a double failure every four or five years.

5 Conclusions

We have explored a model for dealing with common cause failures in simple power networks, allowing data from various sources to be merged into a meaningful number, or range of numbers when robustness is at stake.

We assumed immediate repair, which is clearly not realistic. Non-immediate repair is typically modelled through continuous time Markov chains [2, Chapters 7–13], which have not yet received that much attention in the imprecise literature. The other unrealistic assumption is the Markov assumption itself, although that assumption seems still pervasive in the standard literature. In practice, failure rates are rarely independent of the history of the system, so the ability to build some level of non-stationarity into the model would be desirable. Moreover, it is not entirely clear how the typical simulation techniques that deal with these issues can be made to work to achieve a robust analysis over a range of parameters.

For more complex power networks, the model would need to be extended to handle multiple components. Although this is mathematically quite easy, difficulties are to be expected with estimating parameters that relate to common cause events, because there can be many more ways in which multiple failures occur when three or more components are involved. Some level of symmetry between common cause events would likely need to be accepted.

Another interesting question would be to investigate how the analysis impacts decisions, say on asset replacement.

Acknowledgements

The first author gratefully acknowledges funding from the Durham Energy Institute (EPSRC Grant EP/J501323/1). We thank all reviewers for their comments which have helped us improving various aspects of the paper. We are also indebted to Chris Dent for many useful discussions and comments.

References

[1] Jose M. Bernado and Adrian F. M. Smith. *Bayesian Theory*. John Wiley and Sons, 1994.

[2] R. Billinton and R. N. Allan. *Reliability Evaluation of Power Systems*. Plenum Press, 2nd edition, 1996.

[3] Simon Blake and Philip Taylor. *Handbook of Power Systems II*, chapter Aspects of Risk Assessment in Distribution System Asset Management: Case Studies, pages 449–480. Springer, 2010.

[4] Simon Blake, Philip Taylor, and David Miller. A composite methodology for evaluating network risk. In *CIRED 21st International Conference on Electricity Distribution*, Frankfurt, Germany, June 2011.

[5] A. Mosleh, K. N. Fleming, G. W. Parry, H. M. Paula, D. H. Worledge, and D. M. Rasmuson. Procedures for treating common cause failures in safety and reliability studies: Procedural framework and examples. Technical Report NUREG/CR-4780, PLG Inc., Newport Beach, CA (USA), January 1988.

[6] Erik Quaeghebeur and Gert de Cooman. Imprecise probability models for inference in exponential families. In Fabio G. Cozman, Robert Nau, and Teddy Seidenfeld, editors, *ISIPTA'05:*

Proceedings of the Fourth International Symposium on Imprecise Probabilities and Their Applications, pages 287–296, Pittsburgh, USA, July 2005.

[7] Matthias C. M. Troffaes, Dana L. Kelly, and Gero Walter. Elicitation and inference for the imprecise Dirichlet model with arbitrary sets of hyperparameters. In *Programme and Abstracts: 5th CSDA International Conference on Computational and Financial Econometrics (CFE 2011) and 4th International Conference of the ERCIM (European Research Consortium for Informatics and Mathematics) Working Group on Computing & Statistics (ERCIM 2011)*, page 38, 2011.

[8] Matthias C. M. Troffaes, Dana L. Kelly, and Gero Walter. Imprecise Dirichlet model for common-cause failure. In *Proceedings of PSAM 11 & ESREL 2012*, June 2012.

[9] Matthias C. M. Troffaes, Gero Walter, and Dana Kelly. A robust Bayesian approach to modelling epistemic uncertainty in common-cause failure models. Submitted. `arXiv:1301.0533`.

[10] Peter Walley. Inferences from multinomial data: Learning about a bag of marbles. *Journal of the Royal Statistical Society, Series B*, 58(1):3–34, 1996.

[11] Gero Walter and Thomas Augustin. Imprecision and prior-data conflict in generalized Bayesian inference. *Journal of Statistical Theory and Practice*, 3:255–271, 2009.

[12] Gero Walter, Thomas Augustin, and Frank P. A. Coolen. On prior-data conflict in predictive Bernoulli inferences. In F. Coolen, G. de Cooman, T. Fetz, and M. Oberguggenberger, editors, *ISIPTA'11: Proceedings of the Seventh International Symposium on Imprecise Probability: Theories and Applications*, pages 391–400, Innsbruck, 2011. SIPTA.