

1 Manuscript, October 2015

2

3

4 **A discretization procedure for rare events in Bayesian networks**

5 Kilian Zwirgmaier & Daniel Straub

6 Engineering Risk Analysis Group, Technische Universität München

7 (kilian.zwirgmaier@tum.de, straub@tum.de, www.era.bgu.tum.de)

8 **Abstract**

9 Discrete Bayesian networks (BNs) can be effective for risk- and reliability assessments, in
10 which probability estimates of (rare) failure events are frequently updated with new
11 information. To solve such reliability problems accurately in BNs, the discretization of
12 continuous random variables must be performed carefully. To this end, we develop an
13 efficient discretization scheme, which is based on finding an optimal discretization for the
14 linear approximation of the reliability problem obtained from the First-Order Reliability
15 Method (FORM). Because the probability estimate should be accurate under all possible
16 future information scenarios, the discretization scheme is optimized with respect to the
17 expected posterior error. To simplify application of the method, we establish parametric
18 formulations for efficient discretization of random variables in BNs for reliability problems
19 based on numerical investigations. The procedure is implemented into a software prototype.
20 Finally, it is applied to a verification example and an application example, the prediction of
21 runway overrun of a landing aircraft.

22 **Keywords**

23 Bayesian networks; discretization; near-real-time; structural reliability; updating

24 **1 Introduction**

25 For operational risk and reliability management, it is often desirable to compute the
26 probability of a rare event F under potentially evolving information. Examples include
27 warning systems for natural and technical hazards, or the planning of inspection and
28 intervention actions in infrastructure systems. Ideally, this is achieved through Bayesian
29 updating of $\Pr(F)$ with the new information Z to the posterior probability $\Pr(F|Z)$. When
30 physically-based or empirical models for predicting the rare event exist, such updating is
31 possible with structural reliability methods (SRM) (Sindel and Rackwitz, 1998, Straub, 2011,
32 Straub et al., 2016). However, it is often difficult to perform the required computations in
33 near-real-time, due to a lack of efficiency or robustness. A modeling and computational
34 framework that does facilitate efficient Bayesian updating is the discrete Bayesian network
35 (BN). Hence it was proposed to combine SRMs with discrete Bayesian networks for near-
36 real-time computations (Friis-Hansen, 2000, Straub and Der Kiureghian, 2010a, Straub and
37 Der Kiureghian, 2010b).

38 BNs are based on directed acyclic graphs (DAGs), to efficiently define a joint probability
39 distribution $p(\mathbf{Y})$ over a random vector \mathbf{Y} (Jensen and Nielsen, 2007, Kjaerulff and Madsen,
40 2013). The DAG of a BN, which is often referred to as the qualitative part of a BN, consists
41 of a node for each variable in \mathbf{Y} and a set of directed links among nodes representing
42 dependence among the variables. In the case of discrete BNs, conditional probability tables
43 (CPTs) quantitatively define the type and strength of the dependence among the variables.
44 The entries of the CPT of a variable Y_i are the probabilities for each state of Y_i conditional on
45 all possible combinations of states of its parents.

46 For hybrid BNs, which include both discrete and continuous variables, exact inference is
47 available only for two special cases, which are BNs with Gaussian nodes, whose means are
48 linear functions of their parents, and BNs, whose nodes are defined as a mixture of truncated
49 basic functions (MoTBFs) (Langseth et al., 2009, Langseth et al., 2012). Otherwise,
50 approximate inference algorithms are available for hybrid BNs based on sampling techniques,
51 e.g. (Lerner, 2002, Hanea et al., 2006). However, these are computationally demanding and
52 not generally suitable for near-real-time decision support (Hanea et al., 2015). As an
53 alternative, the continuous random variables can be discretized, which enables the use of
54 exact inference algorithms that exist for general discrete BNs. These include the variable
55 elimination algorithm (Zhang and Poole, 1994) and the junction tree algorithm (Lauritzen and
56 Spiegelhalter, 1988, Jensen et al., 1990).

57 The size of discrete BNs, and the associated computational effort, increases approximately
58 exponentially with the number of discrete states of its nodes, which motivates the
59 development of efficient discretization algorithms. While efficient discretization in the
60 context of machine learning and BNs in general has been investigated by multiple researchers
61 (Dougherty et al., 1995, Kotsiantis and Kanellopoulos, 2006), research on efficient
62 discretization in the context of engineering risk analysis or structural reliability has been
63 limited. In general, it is to be distinguished between static and dynamic discretization. While

64 the former discretizes the BN a-priori before entering evidence (offline), the latter is based on
65 an iterative scheme that updates the discretization scheme in function of the evidence (online).

66 Dynamic discretization for risk analysis applications has been developed mainly by (Neil et
67 al., 2008), based on the work by (Kozlov and Koller, 1997). The procedure starts with an
68 initial discretization of a hybrid BN, for which an approximate entropy error is calculated. If
69 the error complies with a convergence criterion, the current discretization is accepted.
70 Otherwise the discretization is iteratively altered, by splitting the intervals with the highest
71 entropy error, until the convergence criterion is fulfilled. The approach is implemented in the
72 software AgenaRisk (Agena, 2005). Other dynamic discretization algorithms for reliability
73 analysis have been proposed, e.g. in (Zhu and Collette, 2015) for dynamic BNs. The
74 advantage of dynamic discretization is its flexibility when evidence is entered in the BN, i.e.
75 when the model is updated with new observation.

76 Static discretization has the advantage of being computationally faster and simple to
77 implement. Some considerations for static discretization of BNs in reliability applications
78 have been presented in (Friis-Hansen, 2000, Straub, 2009, Straub and Der Kiureghian, 2010a).
79 As pointed out by (Friis-Hansen, 2000), for applications in which extreme events are
80 important, discretization of the distribution tails should be performed with care. Static
81 discretization facilitates a careful representation of these tails. However, the accuracy of the
82 static discretization varies with the available evidence. The difficulty is thus to find a
83 discretization scheme that is optimal under a wide variety of posterior distributions.

84 In this paper we derive a procedure for efficiently performing static discretization of
85 continuous reliability problems. An optimal discretization scheme is sought, which minimizes
86 the expected approximation error with respect to possible future observations (evidence). To
87 solve this optimization problem, we propose to approximate the reliability problem by the
88 First-Order Reliability Method (FORM). Section 2 of the paper describes the proposed
89 methodology. Section 3 presents numerical parameter studies, and simple parametric relations
90 for defining an efficient discretization scheme are derived. In Section 4, the procedure is
91 applied to a set of verification examples and to the computation of the probability of runway
92 overrun of a landing aircraft. While the theory is introduced for problems with only one
93 design point, considerations regarding problems with multiple design points are given in the
94 last verification example and in the discussion.

95 **2 Methodology**

96 **2.1 Structural reliability**

97 Since the 1970s structural reliability methods have been developed and applied in the
98 engineering community to estimate failure probabilities $\Pr(F)$ of components or systems,
99 based on physical or empirical models. The performance of engineering components is

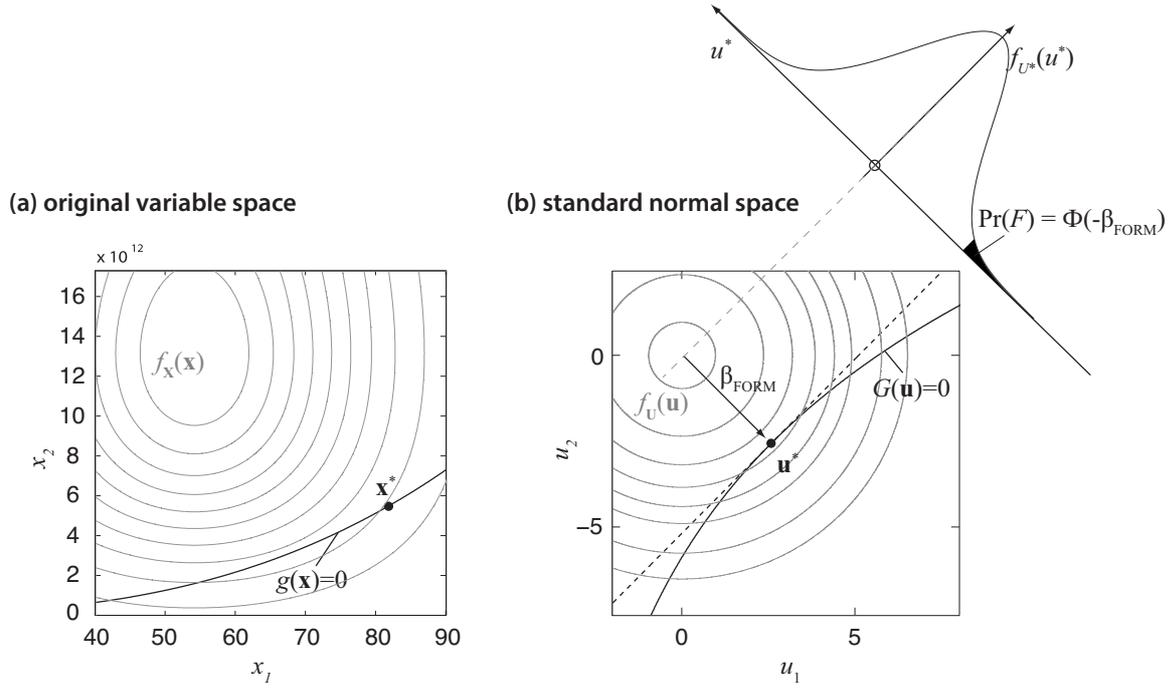
100 described by a limit state function (LSF) $g(\mathbf{x})$, where $\mathbf{X} = [X_1; \dots; X_n]$ is a vector of basic
 101 random variables influencing the performance of the component. By definition, failure
 102 corresponds to $g(\mathbf{x})$ taking non-positive values, i.e. the failure event is $F = \{g(\mathbf{X}) \leq 0\}$. $g(\mathbf{x})$
 103 includes the physical or engineering model, which is often computationally demanding. The
 104 probability of failure is calculated by integrating the probability density function (PDF) of \mathbf{X} ,
 105 $f_{\mathbf{X}}(\mathbf{x})$, over the failure domain:

$$\Pr(F) = \int_{g(\mathbf{x}) \leq 0} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} \quad (2)$$

106 The formulation can be extended to the reliability of general systems by defining the failure
 107 domain as a combination of series and parallel systems (Ditlevsen and Madsen, 2007). In the
 108 general case, there is no analytical solution to Eq. 2 and the integral is potentially high-
 109 dimensional. For this reason, structural reliability methods (SRMs) are applied to approximate
 110 it. These include the first- and the second order reliability method (FORM and SORM) as
 111 well as a large variety of sampling methods, including importance sampling methods such as
 112 directional importance sampling, and sequential sampling methods such as subset simulation.
 113 These methods are well-documented in the literature (Au and Beck, 2001, Rackwitz, 2001,
 114 Der Kiureghian, 2005, Ditlevsen and Madsen, 2007).

115 **2.2 First order reliability method (FORM)**

116 To obtain an approximation of the probability of failure through FORM, the LSF $g(\mathbf{X})$ is
 117 transformed to an equivalent LSF $G(\mathbf{U})$ in the space of uncorrelated standard normal random
 118 variables $\mathbf{U} = [U_1; \dots; U_n]$ (Fig. 1). The transformation is probability conserving, so that
 119 $\Pr[g(\mathbf{X}) \leq 0] = \Pr[G(\mathbf{U}) \leq 0] = \Pr(F)$. A suitable transformation for this purpose, which is
 120 consistent with the BN, is the Rosenblatt transformation (Hohenbichler and Rackwitz, 1981).
 121 In case all basic random variables are independent, this transformation reduces to the
 122 marginal transformations: $U_i = \Phi^{-1}[F_{X_i}(X_i)]$, with Φ^{-1} being the inverse standard normal
 123 CDF.



124
 125 Figure 1. Design point and linear approximation of the limit state surface. Left side: original random
 126 variable space; right side: standard normal space (from (Straub, 2014a)).

127
 128 The FORM approximation of $\Pr(F)$ is obtained by substituting the LSF in U-space $G(\mathbf{U})$ by a
 129 linear function $G_L(\mathbf{U})$, i.e. a first-order Taylor expansion of $G(\mathbf{U})$. The key idea of FORM is
 130 to choose as the expansion point the so-called design point \mathbf{u}^* , which is the point that
 131 minimizes $\|\mathbf{u}^*\|$ subject to $G_L(\mathbf{U}) \leq 0$. \mathbf{u}^* also known as the most likely failure point, as it is
 132 the point in the failure domain with the highest probability density. Since all marginal
 133 distributions of the standard uncorrelated multinormal distribution are standard normal, it can
 134 be shown that the FORM probability of failure $\Pr[G_L(\mathbf{U}) \leq 0]$ is:

$$\Pr[G_L(\mathbf{U}) \leq 0] = \Phi(-\beta_{FORM}) \quad (3)$$

135 where Φ is the standard normal CDF and β_{FORM} is the distance from the origin to the design
 136 point, i.e. $\beta_{FORM} = \|\mathbf{u}^*\|$. The problem thus reduces to finding the design point \mathbf{u}^* . If $G(\mathbf{U})$ is
 137 linear, the FORM solution of the probability of failure is exact, otherwise it is an
 138 approximation, which however is sufficiently accurate in most practical applications with
 139 limited numbers of random variables (Rackwitz, 2001).

140 The linearized LSF $G_L(\mathbf{U})$ can be written as:

$$G_L(\mathbf{U}) = \beta_{FORM} - \boldsymbol{\alpha}^T \mathbf{U} \quad (4)$$

141 where $\boldsymbol{\alpha} = [\alpha_1, \dots, \alpha_n]$ is the vector of FORM importance measures. These importance
 142 measures are defined as:

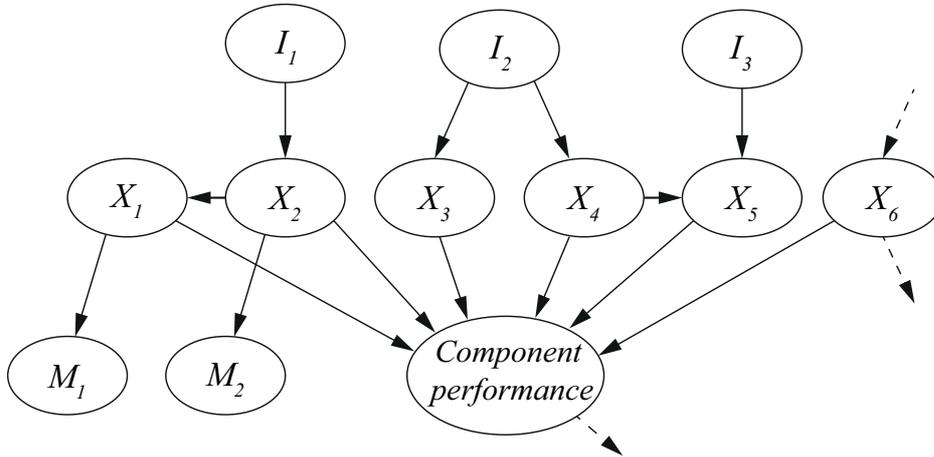
$$\alpha_i = \frac{u_i^*}{\beta_{FORM}} \quad (5)$$

143 where u_i^* is the i -th component of the design point coordinates. The α_i 's take values between
 144 -1 and 1, and it is $\|\boldsymbol{\alpha}\| = 1$. α_i is 0, if the uncertainty on U_i has no influence on $\Pr(G_L(\mathbf{U}) \leq$
 145 0), and it is 1 or -1, if U_i is the only random variable affecting $\Pr(g_L(\mathbf{U}) \leq 0)$. When the
 146 original random variables X_i are mutually independent, the α_i 's are readily applicable also in
 147 the original space, otherwise the α_i 's can be transformed as described in (Der Kiureghian,
 148 2005).

149 **2.3 Treatment of a reliability problem in a BN**

150 We combine discrete BNs and structural reliability concepts to facilitate updating of rare
 151 event (failure) probabilities under new observations. The general problem setting is illustrated
 152 in the BN of Fig. 2. We here limit the presentation to component reliability problems; system
 153 problems are considered later. The binary random variable ‘Component performance’ is
 154 described by the LSF $g(\mathbf{X})$.

155



156

157 Figure 2. A general BN including a component reliability problem.

158 The basic random variables \mathbf{X} of the model are included in the BN as parents of ‘Component
 159 performance’. The nodes M_i represent measurements of individual random variables X_i , and
 160 nodes I_j represent factors influencing the basic random variables. Dependence between the
 161 variables in \mathbf{X} is modeled either directly by links among them (here $X_2 \rightarrow X_1$ and $X_4 \rightarrow X_5$) or
 162 through common influencing factors (here $I_2 \rightarrow X_3$ and $I_2 \rightarrow X_4$). The component
 163 performance node can have (multiple) child nodes, which, however, does not impact the
 164 discretization of the reliability problem.

165 Ultimately, the goal is to predict the component performance, i.e. $\Pr(F)$, conditional on other
 166 model parameters, such as measurements M_i or influencing variables I_j . Whenever new
 167 evidence on these variables is available, the BN should be evaluated (in near-real time)
 168 utilizing exact BN inference algorithms.

169 To enable exact inference algorithms, all continuous random variables are discretized. These
 170 include the \mathbf{X} , and possibly the M_i and I_j . In the general case, the computational effort for
 171 solving the BN is a direct function of the CPT size of ‘Component performance’. The size of

172 this CPT is $2 \prod_{i=1}^n n_i$, where n is the number of random variables in \mathbf{X} , and n_i is the number
 173 of states used for discretizing X_i . In this paper we do not describe the discretization of random
 174 variables M_i and I_j , since it is typically straightforward and does not contribute significantly
 175 to computational performance. The key parameter for computational efficiency and accuracy
 176 is the discretization scheme for \mathbf{X} , which is described in sections 2.5 and 2.6.

177 2.4 Simplification of BNs through node removal

178 Removing random variables from a BN is one possibility to reduce the computational effort
 179 associated with a model. A formal approach for removing nodes from a BN is described in
 180 (Straub and Der Kiureghian, 2010b). In order to decide which nodes to remove from the BN
 181 the following questions should be considered:

- 182 • Which random variables are relevant for prediction? (These include ‘Component
 183 performance’.)
- 184 • Which random variables can potentially be observed? (These include the measurement
 185 variables.)
- 186 • Which random variables simplify the modeling of dependencies? (These are e.g.
 187 common influencing factors such as I_2 in Fig. 2.)
- 188 • For which random variables is it desirable to explicitly show their influence on
 189 component performance?

190 If a random variable does not belong to any of these categories, the corresponding node in the
 191 BN can be removed. Since the computational efficiency of the model is governed by the size
 192 of the CPT of the ‘Component state’ node, the primary interest is in removing basic random
 193 variables X_i from the BN. As a measure for the relevance of a basic random variable,
 194 importance measures α_i from a FORM analysis may be used. To better understand the
 195 relation between α_i and X_i ’s relevance for prediction, consider a linearized LSF $G_L(\mathbf{U})$.
 196 Following (Der Kiureghian, 2005), the variance of $G_L(\mathbf{U})$ can be decomposed as:

$$\sigma_{G_L}^2 = \|\nabla G\|^2 (\alpha_1^2 + \alpha_2^2 + \dots + \alpha_n^2) \quad (6)$$

197 where ∇G denotes the gradient vector of the non-linearized LSF $G(\mathbf{U})$. From Eq. 6 it is seen
 198 that a random variable X_i with corresponding α_i accounts for $\alpha_i^2 \cdot 100\%$ of the variance $\sigma_{G_L}^2$.
 199 Therefore, observing a random variable X_i with $\alpha_i = 0.1$ will reduce the variance $\sigma_{G_L}^2$ by 1%,
 200 whereas observing X_j with $\alpha_j = 0.5$ will reduce $\sigma_{G_L}^2$ by 25%.

201 2.5 Discretization of basic random variables

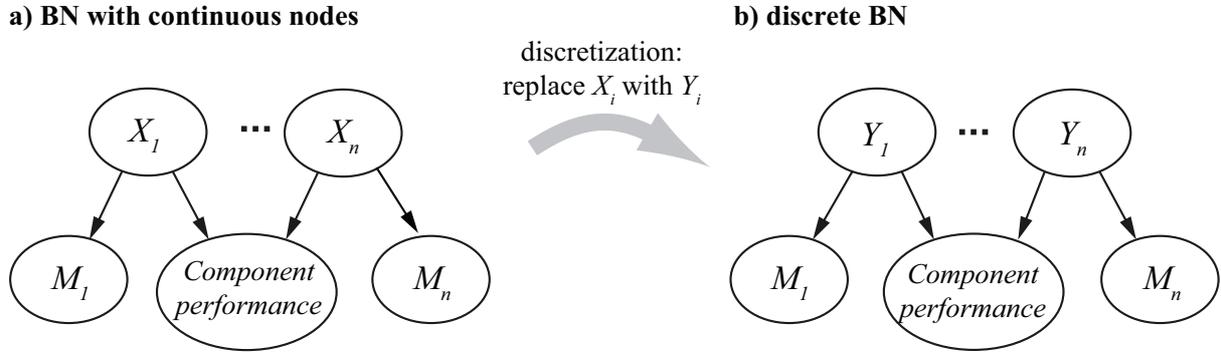
202 For ease of presentation, we first consider discretization of statistically independent basic
 203 random variables \mathbf{X} , i.e. the special case of the BN in Fig. 2 in which the X_i ’s have no parents¹.

¹ In a BN, basic random variables \mathbf{X} are independent if they are not connected through links, if they have no common (unknown) ancestors and if no evidence is available on any of their joint descendants.

204 The proposed procedure is extended to the general case of dependent basic random variables
 205 thereafter.

206 2.5.1 Independent basic random variables

207 The situation is illustrated in Fig. 3. The performance of the component depends on n
 208 statistically independent random variables and is described by a LSF $g(\mathbf{X}) = g(X_1, \dots, X_n)$.
 209 For all basic random variables X_i , corresponding measurements M_i can be performed. To
 210 obtain an equivalent discrete BN, the continuous X_i are replaced by the discrete random
 211 variables Y_i , and the LSF is replaced by the CPT of component performance conditional on
 212 $\mathbf{Y} = [Y_1; \dots; Y_n]$. For each discrete random variable Y_i with n_i states $1, 2, \dots, n_i$, we define a
 213 discretization scheme $D_i = [d_0, d_1, \dots, d_{n_i-1}, d_{n_i}]$ consisting of $n_i + 1$ interval boundaries.
 214 The first and the last interval boundaries are given by the boundaries of X_i 's outcome space.
 215



216
 217 Figure 3. Representation of a basic reliability problem with n independent basic random variables in a BN. Left:
 218 original problem with continuous basic random variables X_i , right: discrete BN, in which X_i 's are substituted with
 219 discrete nodes Y_i .

220 Since here the X_i , and thus the Y_i , have no parents, the PMF of Y_i is defined as:

$$p_{Y_i}(j) = F_{X_i}(d_j) - F_{X_i}(d_{j-1}); \quad \text{with } j \in [1, \dots, n_i] \quad (7)$$

221 where F_{X_i} denotes the cumulative distribution function (CDF) of X_i . The probability of failure
 222 corresponding to the discrete BN in Fig. 3b can be calculated as:

$$\Pr(F) = \sum_{y_1=1}^{n_1} \dots \sum_{y_n=1}^{n_n} p_{Y_1}(y_1) \cdot \dots \cdot p_{Y_n}(y_n) \cdot \Pr(F|Y_1 = y_1 \cap \dots \cap Y_n = y_n) \quad (8)$$

223 Note that the discretization does not introduce any approximation error here, as long as the
 224 conditional $\Pr(F|Y_1 = y_1 \cap \dots \cap Y_n = y_n)$ is computed exactly.

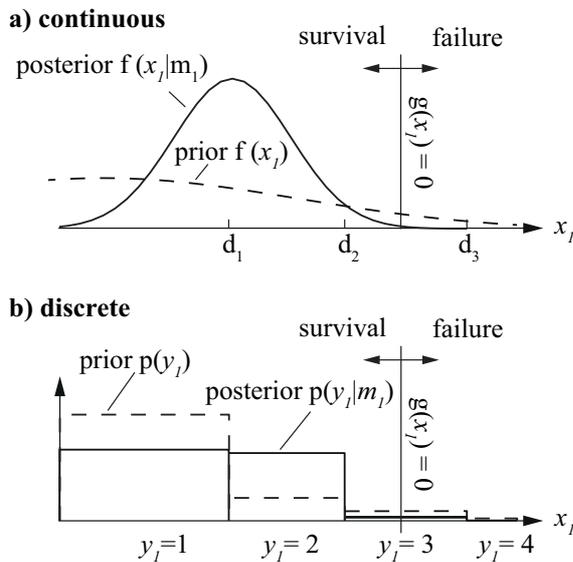
225 Once measurements from the nodes $\mathbf{M} = [M_1; \dots; M_n]$ are available, the conditional failure
 226 probability can be calculated as:

$$\Pr(F|\mathbf{M} = \mathbf{m}) \approx \frac{1}{p_{\mathbf{M}}(\mathbf{m})} \sum_{y_1=1}^{n_1} \dots \sum_{y_n=1}^{n_n} p_{Y_1}(y_1) \cdot p_{M_1|Y_1}(m_1|y_1) \cdot \dots \cdot p_{Y_n}(y_n) \cdot p_{M_n|Y_n}(m_n|y_n) \cdot \Pr(F|Y_1 = y_1 \cap \dots \cap Y_n = y_n) \quad (9)$$

227 where $\Pr(F|Y_1 = y_1, \dots, Y_n = y_n)$ is the conditional probability of component failure given
 228 y_1, \dots, y_n . If no measurements are available for some of the basic random variables, the
 229 corresponding likelihood terms $p_{M_j|Y_j}(m_j|y_j)$ are simply omitted in Eq. 9.

230 While the computation of the unconditional failure probability following Eq. 8 is exact, the
 231 computation of the conditional failure probability through Eq. 9 is only an approximation.
 232 The reason is that the dependence between the measurement variable M_i and the ‘Component
 233 performance’ variable is not fully captured in the discrete BN (see also Straub and Der
 234 Kiureghian, 2010b). In Fig. 4, this is illustrated for a reliability problem with one basic
 235 random variable X_i . Both the continuous distribution (Fig. 4a) and the corresponding
 236 discretized distribution (Fig. 4b) are updated correctly after observing M_1 . However, for Eq. 9
 237 to be exact, also the conditional failure probabilities $\Pr(F|Y_1 = y_1)$ would need to be updated.
 238 This can be observed in Fig. 4a: in interval $Y_1 = 3$, which is the one cut by the limit state
 239 surface, the ratio of the probability mass in the failure domain to that in the safe domain
 240 changes from the prior to the posterior case, i.e. $\Pr(F|Y_1 = 3) \neq \Pr(F|Y_1 = 3, M_1 = m_1)$.
 241 Since the computation of the conditional failure probability following Eq. 9 is based on the
 242 prior probability $\Pr(F|Y_1 = 3)$, the discretization introduces an approximation in this case.
 243 The error occurs only in the intervals that are cut by the limit state surface.

244

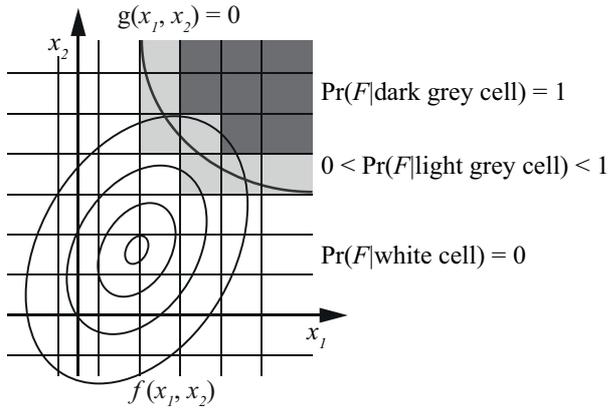


245
 246

Figure 4. Discretization error in 1D.

247 In the simple one-dimensional case of Fig. 4, an optimal discretization approach would be to
 248 discretize the whole outcome space in two intervals, one capturing the survival and one the
 249 failure domain. This discretization would have zero approximation error. However, already in
 250 a two-dimensional case, such a solution is not possible. This is illustrated in Fig. 5, where the
 251 cells cut by the limit state surface are indicated in grey. The failure probability conditional on
 252 measurements calculated according to Eq. 9 will necessarily be an approximation. The
 253 approximation error will be small, if the contribution of the cells cut by the limit state surface

254 (the grey cells in Fig. 5) to the total failure probability is small. An efficient discretization will
 255 thus limit this contribution with as few intervals as possible.



256
 257 Figure 5. Discretization error in 2D.

258 2.5.2 Dependent basic random variables

259 Eqs. (7-9) must be adjusted when dependence among the X_i 's is present, in accordance with
 260 the case-specific BN structure. However, the principles outlined above for independent
 261 X_1, \dots, X_n hold equally for dependent basic random variables: The discretization error is a
 262 function of the cells cut by the limit state function.

263 When determining an optimal discretization, we propose in the following to find the FORM
 264 approximation of the reliability problem, which can readily account for the dependence
 265 among the random variables. Hence, there is no need to distinguish between cases with
 266 independent or dependent random variables.

267 2.6 Efficient discretization

268 2.6.1 Optimal discretization of linear problems in standard normal space

269 To find an efficient discretization of \mathbf{X} , we consider the FORM solution to the reliability
 270 problem. Evaluating the linearized FORM LSF $G_L(\mathbf{U})$ is computationally inexpensive once
 271 the design point \mathbf{u}^* is known. Therefore, it is feasible to find a discretization of \mathbf{U} that is
 272 optimal for the event $\{G_L(\mathbf{U}) \leq 0\}$ through optimization. If $G(\mathbf{U})$ is not strongly non-linear,
 273 this solution will be an efficient discretization for $\{G(\mathbf{U}) \leq 0\}$ and, after a transformation to
 274 the original space, also for $\{g(\mathbf{X}) \leq 0\}$.

275 As discussed in Section 2.5.1, the approximation error of the discretization is associated with
 276 the change from the prior to the posterior distribution of the basic random variables. A
 277 measure of optimality must thus consider possible measurements of \mathbf{X} or \mathbf{U} . We consider
 278 hypothetical measurements \tilde{M}_i (this notation is used to distinguish hypothetical measurements
 279 from actual measurements $M_i = m_i$) of all standard normal random variables U_i with additive
 280 measurement error $\varepsilon_i \sim N(0, \sigma_{\varepsilon_i})$. Since both the prior distribution and the measurement error
 281 are normal distributed, the likelihood is also normal distributed:

$$\tilde{M}_i | \{U_i = u_i\} \sim N(u_i, \sigma_{\varepsilon_i}) \quad (10)$$

282 The posterior, i.e. the conditional distribution of U_i given a measurement outcome $\tilde{M}_i = \tilde{m}_i$,
 283 is the normal distribution with mean $\frac{1}{1+\sigma_{\varepsilon_i}^2} \tilde{m}_i$ and standard deviation $\sqrt{\left(1 - \frac{1}{1+\sigma_{\varepsilon_i}^2}\right)}$.

284 We define an error measure based on comparing the true posterior probability of failure
 285 $P_{F|\tilde{M}}(\tilde{\mathbf{m}})$ with the posterior probability of failure calculated from the discretized \mathbf{U} , denoted
 286 by $\hat{P}_{F|\tilde{M}}(\mathbf{d}; \tilde{\mathbf{m}})$. Here, \mathbf{d} are the parameters defining the discretization. The proposed error
 287 measure is:

$$e(\mathbf{d}, \tilde{\mathbf{m}}) = \left| \frac{\log_{10} \hat{P}_{F|\tilde{M}}(\mathbf{d}; \tilde{\mathbf{m}}) - \log_{10} P_{F|\tilde{M}}(\tilde{\mathbf{m}})}{\log_{10} P_{F|\tilde{M}}(\tilde{\mathbf{m}})} \right| \quad (11)$$

288 The error measure of Eq. 11 represents a tradeoff between the absolute and the relative error.
 289 It weights the (logarithmic) relative error by the magnitude of the posterior failure probability.
 290 This ensures that the same relative error is considered worse at a higher probability level
 291 compared to an error at a lower probability level.

292 A-priori, the measurement outcomes are not known. Hence we define the optimal
 293 discretization as the one that minimizes the expected preposterior error $E_{\tilde{M}}[e(\mathbf{d}, \tilde{M})]$:

$$\mathbf{d}^{opt} = \arg \min_{\mathbf{d}} E_{\tilde{M}}[e(\mathbf{d}, \tilde{M})] = \arg \min_{\mathbf{d}} \int_{\tilde{M}} e(\mathbf{d}, \tilde{\mathbf{m}}) f_{\tilde{M}}(\tilde{\mathbf{m}}) d\tilde{\mathbf{m}} \quad (12)$$

294 The optimization is thus based on the computation of an expected value with respect to the
 295 possible measurements outcomes \tilde{M} before having taken any measurements. This is
 296 analogous to a preposterior analysis (Raiffa and Schlaifer, 1961, Straub, 2014b). However,
 297 unlike in traditional preposterior analysis, the objective is not to identify an optimal action
 298 under future available information, but to find the optimal discretization parameters \mathbf{d}^{opt} . The
 299 integral in Eq. 12 is evaluated through a simple Monte Carlo approach. All \tilde{M}_i have the
 300 normal distribution with zero mean and variance $1 + \sigma_{\varepsilon}^2$.

301 The parameters in \mathbf{d} describing the discretization scheme are:

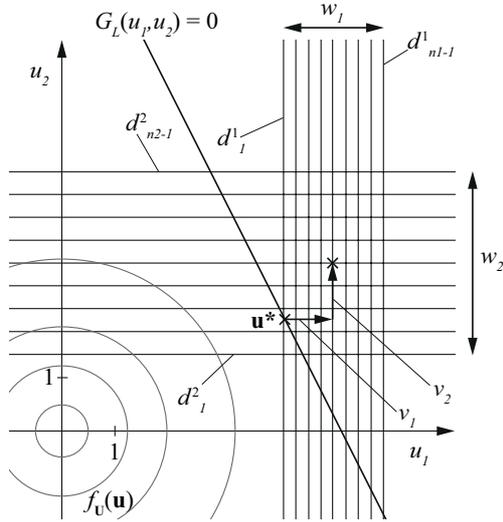
- 302 – n_i : number of intervals used to discretize each random variable U_i ,
- 303 – w_i : width of the discretization frame in the dimension of U_i , and
- 304 – v_i : position of the midpoint of the discretization frame relative to the design point

305 These parameters are illustrated in Fig. 6. For a problem with n basic random variables, the
 306 full set of optimization parameters is $\mathbf{d} = [w_1, \dots, w_n, n_1, \dots, n_{n-1}, v_1, \dots, v_n]$.

307 Clearly, the discretization error reduces with increasing n_i . Because the computational
 308 efficiency of the final BN is a direct function of the size of the CPT associated with
 309 component performance, which is $\prod_{i=1}^n n_i$, we constrain its size. To this end, we define c_{up} as

310 the maximum allowed number of parameters of the CPT of the component state node. This
 311 puts a constraint on the optimization of Eq. 12:

$$\prod_{i=1}^n n_i \leq c_{up} \quad (13)$$



312

313 Figure 6. Schematic representation of a discretization of a linear 2D reliability problem. w_i is the distance
 314 between interval boundaries d_1^i and $d_{n_i-1}^i$. All intervals between these boundaries are equi-spaced. v_i is the
 315 position of the midpoint of the discretization frame relative to the design point \mathbf{u}^* in dimension i .

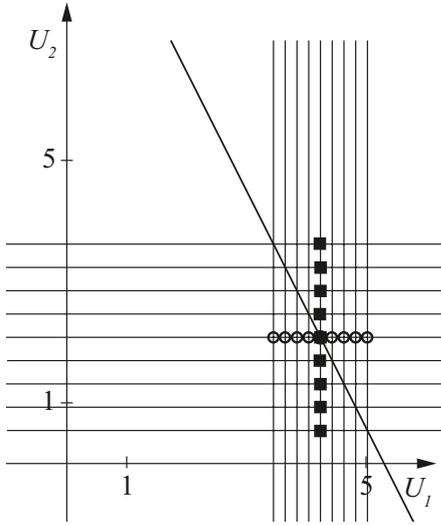
316 The optimization is implemented through a two-level approach. The optimization of the
 317 continuous parameters width w_i and position of the discretization frame v_i for all $i = 1, \dots, n$
 318 is carried out using unconstrained nonlinear optimization for fixed values of n_i . The
 319 optimization of the discrete n_i is performed through a local search algorithm. Note that the
 320 optimization is performed offline, i.e. prior to running the BN, hence it does not affect the
 321 goal of near-real time performance of the BN. Furthermore, in section 3 a heuristic is derived
 322 that can replace the time-consuming solution of the optimization problem.

323 2.6.2 Efficient discretization of the original random variables \mathbf{X}

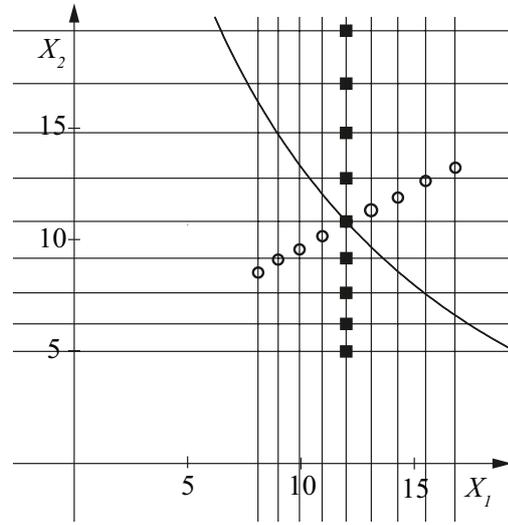
324 Since the nodes in the BN represent random variables \mathbf{X} in their original outcome space, the
 325 discretization schemes, which are derived for the corresponding standard normal random
 326 variables \mathbf{U} , need to be transformed to the \mathbf{X} -space. In the case of mutually independent
 327 random variables X_i , any point on the i -th interval boundary in \mathbf{U} -space – if transformed – will
 328 result in the same corresponding i -th interval boundary in \mathbf{X} -space. This is not the case for
 329 dependent random variables X_i , where a mapping of the interval boundaries in \mathbf{U} -space to \mathbf{X} -
 330 space will not lead to an orthogonal discretization scheme in \mathbf{X} -space. To preserve
 331 orthogonality throughout the transformation, we propose to represent each interval boundary
 332 through a characteristic point and determine the boundary in \mathbf{X} -space through a
 333 transformation of this point. For transforming the interval boundary of X_i , the characteristic

334 point is selected as the design point u^* , where the i -th element is substituted by the coordinate
 335 of the interval boundary. In Fig. 7 this is shown for an example with $n = 2$ random variables.

a) Discretization in U-space



b) Transformed discretization



336
 337 Figure 7. Transformation of a discretization scheme from U-space to X-space. To preserve orthogonality
 338 each interval boundary in U-space is represented by a characteristic point. The random variables X_1 and
 339 X_2 are Weibull distributed with scale and shape parameter 1 and their correlation is 0.5.

340 3 Development of an efficient discretization procedure

341 3.1 Optimization of the FORM approximation

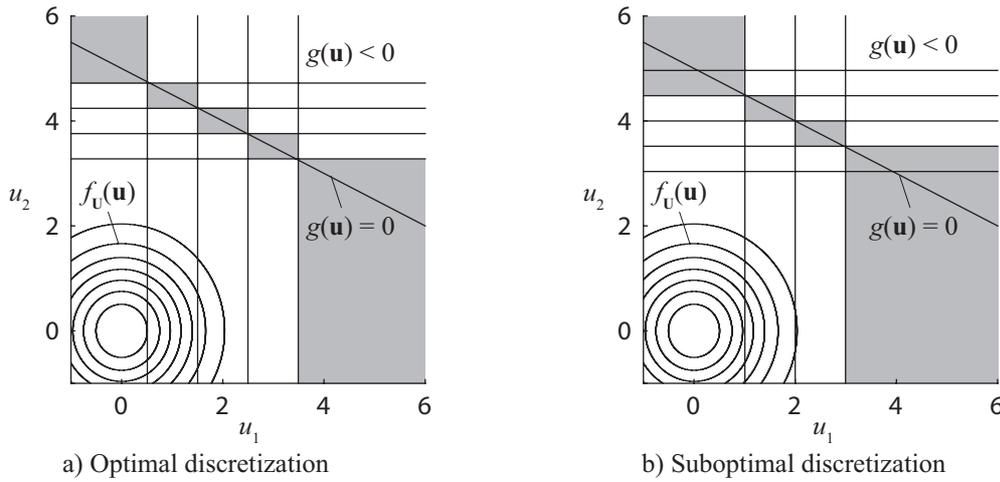
342 We present the optimal discretization for the FORM approximation $G_L(\mathbf{U})$ for $n = 2$ and
 343 $n = 3$ dimensions. Extension to higher numbers of random variables is discussed. Because
 344 the linear LSF employed in FORM is described only by the reliability index β_{FORM} and the
 345 vector α of FORM sensitives (Eq. 4), it facilitates parametric studies.

346 Initially, we consider a reliability index $\beta_{FORM} = 4.26$, corresponding to a probability of
 347 failure of 10^{-5} . The standard deviation of the additive measurement error is set to either
 348 $\sigma_\varepsilon = 0.5$ or $\sigma_\varepsilon = 1.0$. Different combinations of FORM sensitivity values α_i are selected, to
 349 investigate their effect on the optimal discretization. In all investigated cases, we find that the
 350 position of the midpoint of the optimal discretization frame coincides with the design point,
 351 i.e. $v_i^{opt} = 0$. The optimal discretization widths w_i^{opt} vary significantly with the importance
 352 measures α_i , as shown in section 3.1.2. However, the optimal number of intervals n_i^{opt} is
 353 approximately the same for all random variables in all investigated cases, independent of the
 354 α_i values, i.e. $n_i^{opt} = c_{up}^{1/n}$.

355 3.1.1 On why the optimal number of intervals n_i^{opt} is independent of α_i

356 To better understand why the number of intervals does not depend on $|\alpha_i|$ (for $|\alpha_i|$ that are
 357 significantly larger than 0), recall that an efficient discretization scheme should focus on the
 358 area around the limit state surface. More precisely, the discretization error in the posterior

359 case is induced by the cells that are cut by the limit state surface. Exemplarily, Fig. 8 shows a
 360 linear problem in standard normal space with two basic random variables U_1 and U_2 , where
 361 $\alpha_1 = 0.45$ and $\alpha_2 = 0.89$. While Fig. 8a shows an optimal discretization with 5 intervals per
 362 dimension, Fig. 8b shows a discretization scheme, where the more important random variable
 363 U_2 is discretized with 6 intervals and U_1 with 4 intervals. In both cases, the discretization
 364 frame is centered at the design point. It is observed that the probability mass of the (grey)
 365 cells, associated with the discretization error in the posterior case, is higher in Fig. 8b than for
 366 the optimal discretion scheme in 8a. In the example shown here it is $1.4 \cdot 10^{-3}$ compared to
 367 $2.5 \cdot 10^{-4}$.



368 a) Optimal discretization
 369 b) Suboptimal discretization
 369 Figure 8. Linear problem in standard normal space, with two random variables U_1 and U_2 , where $\alpha_1 = 0.45$ and
 370 $\alpha_2 = 0.89$. The intervals cut by the limit state surface, i.e. those which potentially lead to a posterior
 371 discretization error are marked in grey.

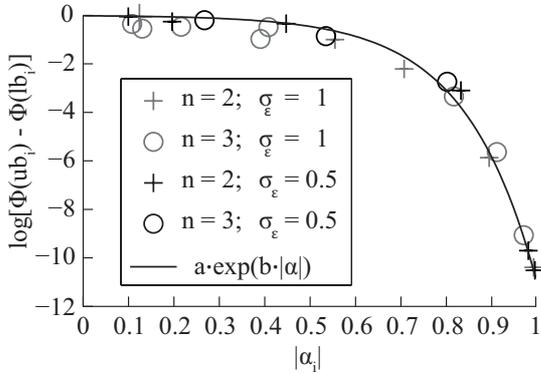
372 For input random variables X_i with a value of $|\alpha_i|$ close to zero, the above observations do
 373 not hold. Following Section 2.4, these variables should be removed from the BN prior to
 374 discretization.

375 3.1.2 Dependence of the optimal discretization width on α_i

376 In the optimization, it is found that the optimal discretization width w_i^{opt} varies strongly with
 377 the random variable's importance, expressed through α_i . It is reminded that the width w_i
 378 describes the domain in which a fine discretization mesh is applied (Fig. 6). In general, w_i^{opt}
 379 decreases with increasing α_i . This effect can be observed in Fig. 8a, where U_2 is the more
 380 important input random variable and it is $w_2^{opt} < w_1^{opt}$.

381 A clear relation between w_i^{opt} and α_i can be observed by plotting the probability mass
 382 enclosed by w_i^{opt} against α_i , as shown in Fig. 9. The results of Fig. 9 indicate that the
 383 probability mass contained within this interval should be a direct function of α_i . The more
 384 important the variable, the finer the discretization around the design point should become.
 385 The observed relationship between this probability mass and α_i follows a clear trend, and a
 386 function can be fitted (Fig. 9). Neither the dimensionality of the problem nor the standard

387 deviation of the measurement error appear to have an influence on this relation. However, as
 388 shown in the following section, it is found that the relation does depend on the prior failure
 389 probability of the problem (i.e. on β_{FORM}) and on the number of intervals n_i used to discretize
 390 the domain.



391
 392 Figure 9. Logarithm of the probability mass enclosed by the discretization frame plotted against α_i . Φ
 393 denotes the standard normal CDF and ub_i respectively lb_i the last (upper) and the first (lower) interval
 394 bound in dimension i .

395 To facilitate the application in practice and extending the results to larger numbers of random
 396 variables, in section 3.2 parametric functions are fitted to the optimization results to capture
 397 the dependency between the optimal discretization width w_i^{opt} and the FORM importance
 398 measures α_i .

399 3.1.3 Dependence of the optimal discretization on the reliability index β and the 400 number of discretization cells c_{up}

401 The influence of the prior failure probability and the maximum size of the CPT, c_{up} , on the
 402 optimal discretization is investigated through 10 problems with $n = 2$ random variables in
 403 standard normal space. The FORM importance measures of the random variables are selected
 404 between 0.1 to 0.995 and the standard deviation of the measurement error is fixed to $\sigma_\epsilon = 1.0$.
 405 We find that the optimal discretization frame is generally centered at the design point, i.e.
 406 $v_i^{opt} = 0$, and that the intervals are distributed uniformly among the dimensions.

407 Firstly, we vary the maximum CPT size c_{up} , i.e. the total number of discretization cells. The
 408 reliability index is $\beta_{FORM} = 5.2$. Fig. 10 shows the influence of c_{up} on the resulting width of
 409 the discretization frame w_i . Three cases are considered: $c_{up} = 25$, $c_{up} = 100$ and $c_{up} = 400$.
 410 These choices correspond to 5, 10 and 20 intervals for each random variable. The left side of
 411 Fig. 10 shows the relation between the optimal w_i and $|\alpha_i|$. The right side of Fig. 10 shows the
 412 same relation, where the w_i 's are scaled as in Fig. 9, i.e. the logarithm of the probability mass
 413 enclosed by the outer interval boundaries is depicted. As in Fig. 9, there is a clear dependence
 414 between the scaled w_i values and the $|\alpha_i|$'s. The interval frames increase with increasing
 415 number of random variables.

416 Secondly, we vary the prior failure probability from 10^{-3} ($\beta = 3.1$) to 10^{-7} ($\beta = 5.2$). The
 417 results are shown in Fig. 11. Again, a distinct dependence between the scaled w_i values and

418 the $|\alpha_i|$'s is found. The interval frames decrease with increasing reliability index (with
 419 decreasing failure probability).

420 3.2 Parametric function of optimal discretion frame

421 As evident from Fig. 10 and Fig. 11, there is a clear dependence of the probability mass
 422 enclosed by the optimal discretization frame (with width w_i) on the FORM sensitivity values
 423 $|\alpha_i|$. The following parameteric function captures this dependence:

$$\log(\Phi(ub_i) - \Phi(lb_i)) = a \cdot \exp(b \cdot |\alpha_i|) \quad (14)$$

424 ub_i is the upper and lb_i the lower interval boundary in dimension i , such that $w_i = ub_i - lb_i$.
 425 a and b are the parameters of the exponential function. This function is depicted in Figs. 10
 426 and 11. Tab. 1 shows the parameter values a and b for the different combinations of the prior
 427 reliability index β and number of intervals per dimension

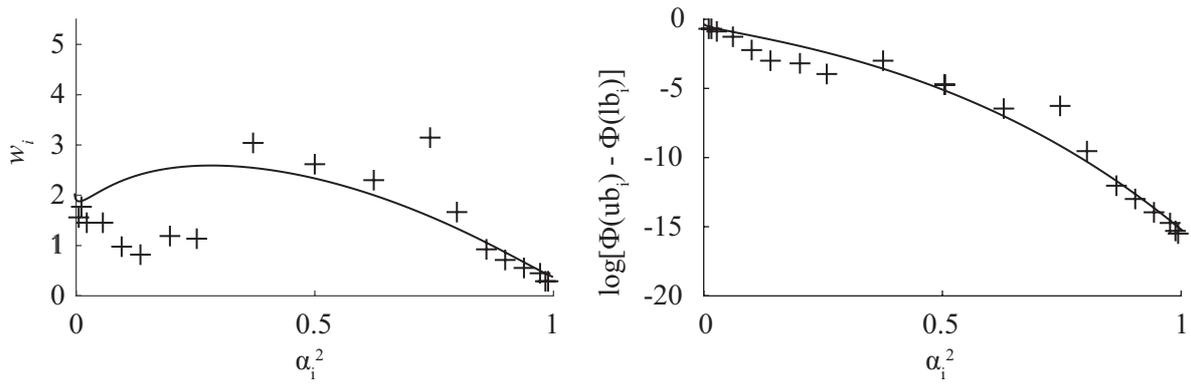
428 Table 1. Parameters a and b of Eq. 14 for $\beta = 3.1$, $\beta = 4.3$ and $\beta = 5.2$ as well as 5, 10 and 20 intervals per
 429 dimension.

$\begin{bmatrix} a, \\ b \end{bmatrix}$	$n_i = 5$	$n_i = 10$	$n_i = 20$
$\beta = 3.1$	$\begin{bmatrix} -0.28, \\ 2.9 \end{bmatrix}$	$\begin{bmatrix} -1.6 \cdot 10^{-2}, \\ 5.8 \end{bmatrix}$	$\begin{bmatrix} -9.8 \cdot 10^{-4}, \\ 8.7 \end{bmatrix}$
$\beta = 4.3$	$\begin{bmatrix} -0.15, \\ 4.3 \end{bmatrix}$	$\begin{bmatrix} -2.4 \cdot 10^{-2}, \\ 6.1 \end{bmatrix}$	$\begin{bmatrix} -2.1 \cdot 10^{-2}, \\ 6.2 \end{bmatrix}$
$\beta = 5.2$	$\begin{bmatrix} -0.36, \\ 3.7 \end{bmatrix}$	$\begin{bmatrix} -0.11, \\ 5.0 \end{bmatrix}$	$\begin{bmatrix} -3.7 \cdot 10^{-2}, \\ 6.0 \end{bmatrix}$

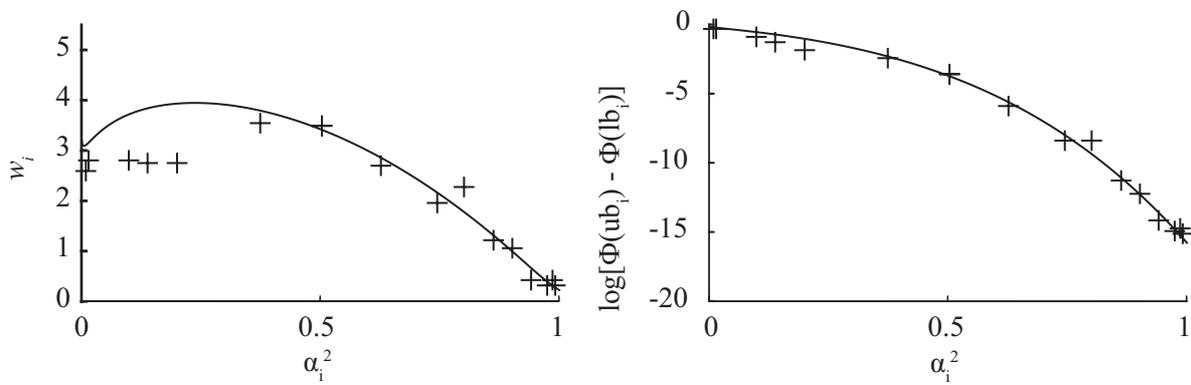
430 From the left sides of Fig. 10 and Fig. 11, it can be observed that the relation between α_i^2 and
 431 the optimal w_i is fairly diffuse for random variables with $|\alpha_i| < 0.6$. Here, the parametric
 432 relationship of Eq. 14 is less accurate. However, these random variables by definition have
 433 lower importance on the reliability estimate. Hence, the inaccuracy of Eq. 14 for random
 434 variables with $|\alpha_i| < 0.6$ is not critical, as is confirmed by the numerical investigations
 435 performed in the remainder of the paper.

436 The parameter values of Tab. 1 are derived from two-dimensional problems. In Fig. 9 it is
 437 shown that there are no notable differences between two and three dimensions. On this basis,
 438 it is hypothesized that the heuristics are applicable also to problems with higher dimensions.
 439 This assumption is furthermore supported by the verification examples presented in chapter 4,
 440 where the heuristics are applied also to four-dimensional problems without any notable
 441 deterioration in the results.

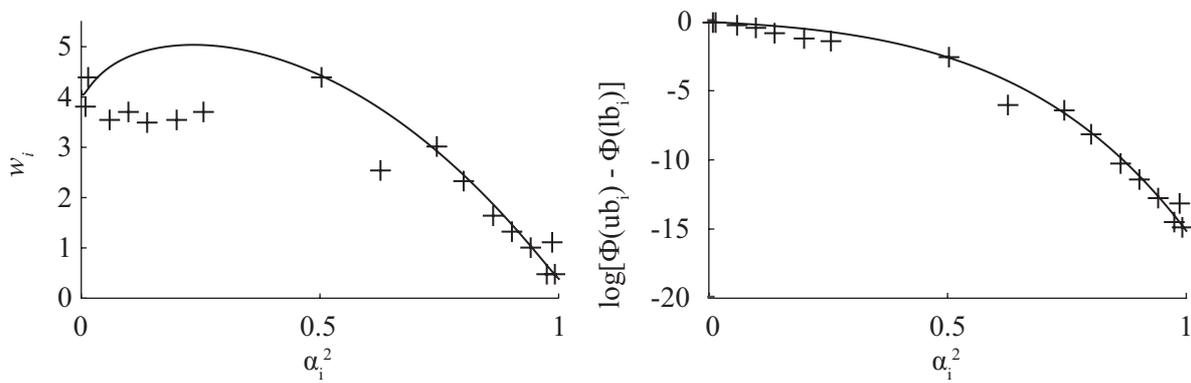
a) 5 intervals per random variable



b) 10 intervals per random variable

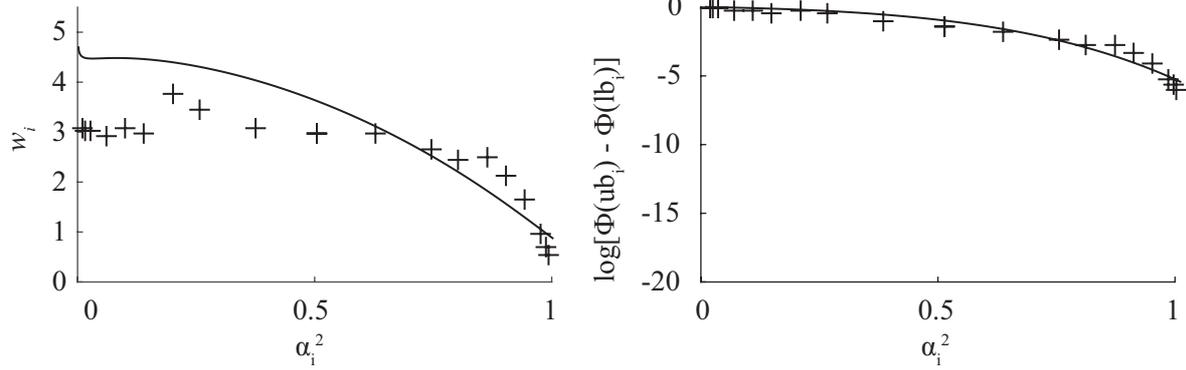


c) 20 intervals per random variable

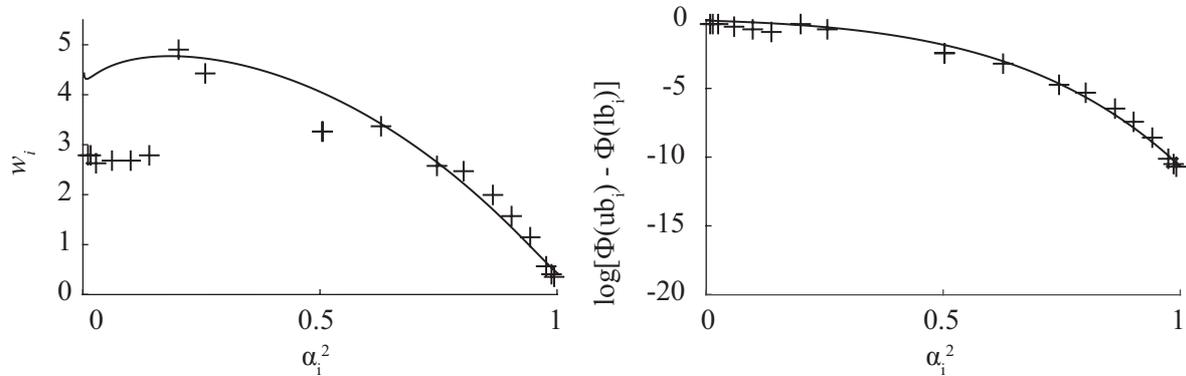


442 Figure 10. Optimization results for 10 two-dimensional, linear problems in standard normal space, which are
 443 discretized with 5, 10 and 20 intervals per dimension. In all cases the prior failure probability is 10^{-7} ($\beta = 5.2$).
 444 The crosses represent the optimization results. The solid lines are the fitted parametric functions (Eq. 14). The
 445 left-hand side shows the relation between the width of a discretization frame w_i and $|\alpha_i|$ and the right-hand side
 446 shows the relation between the probability mass enclosed by the discretization frame with width w_i and $|\alpha_i|$.

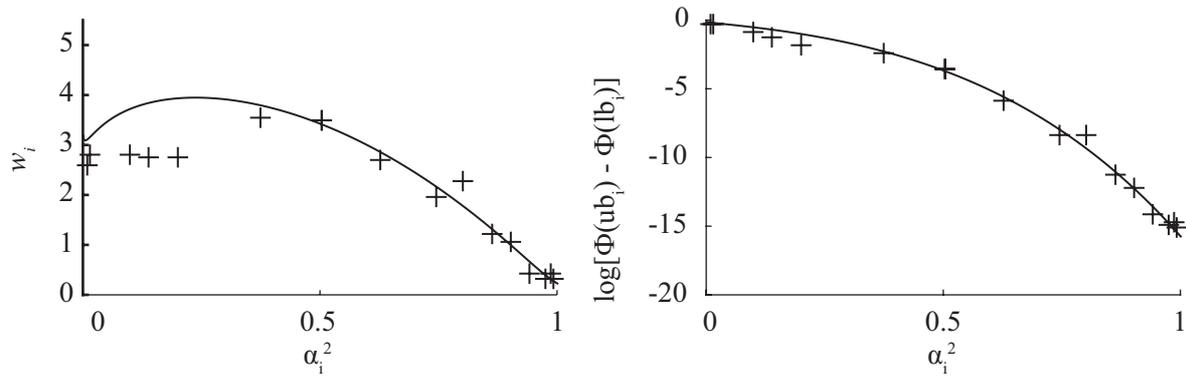
a) $\beta = 3.1, \Pr(F) = 10^{-3}$



b) $\beta = 4.3, \Pr(F) = 10^{-5}$



c) $\beta = 5.2, \Pr(F) = 10^{-7}$



447
448

449 Figure 11. Optimization results for 10 two-dimensional, linear problems in standard normal space, which are
450 discretized with 10 intervals per dimension. The prior failure probabilities are 10^{-3} ($\beta = 3.1$), 10^{-5} ($\beta = 4.3$)
451 and 10^{-7} ($\beta = 5.2$). The crosses represent the optimization results. The solid lines are the fitted parametric
452 functions (Eq. 14). The left-hand side shows the relation between the width of a discretization frame w_i and $|\alpha_i|$
453 and the right-hand side shows the relation between the probability mass enclosed by the discretization frame
454 with width w_i and $|\alpha_i|$.

455 3.3 Summary of the proposed procedure

456 The steps of the proposed procedure are:

- 457 1. Formulate the reliability problem

- 458 2. Set up the corresponding BN
459 3. Perform a FORM analysis for the reliability problem
460 4. Simplify the BN by removing nodes based on:
461 a. their importance for prediction
462 b. their observability
463 c. whether or not a node simplifies modeling of dependencies
464 d. whether or not it is desired to explicitly show a node in the BN for
465 communication purposes
466 5. Find the discretization scheme in U-space based on the proposed heuristics i.e.:
467 a. the discretization scheme is centered at the design point from the FORM
468 analysis
469 b. the same number of intervals is used for each random variable
470 c. the width of the discretization frame follows Eq. 14
471 6. Transform the discretization scheme to X-space
472 7. Compute the CPTs of the component state node and the basic random variables using
473 Monte Carlo simulation or Latin hypercube sampling
- 474 A MATLAB-based software tool performing these steps is available for download under
475 www.era.bgu.tum.de/software.

476 **4 Applications**

477 **4.1 Verification example I**

478 For verification purposes, we apply the proposed methodology to the discretization of a
479 general limit state with non-normal dependent random variables. The approximation error
480 made by this discretization is investigated for different measurement outcomes.

481 Failure is defined through the LSF $g(\mathbf{x})$:

$$g(\mathbf{x}) = a - \prod_{i=1}^n X_i \quad (15)$$

482 i.e., failure corresponds to the event $\{\prod_{i=1}^n X_i \geq a\}$.

483 The basic random variables are distributed as $X_1 \sim LN(0,0.5)$ and $X_2, \dots, X_n \sim LN(1,0.3)$
484 (values in parenthesis are the parameters of the lognormal distribution). The statistical
485 dependence among the X_i is described through a Gaussian copula model, with pairwise
486 correlation coefficients ρ_{ij} . The parameters a and ρ_{ij} determine the prior failure probability
487 P_F . Measurements $M_i = m_i$ are available for all basic random variables; they are associated
488 with multiplicative measurement errors $\varepsilon_i \sim LN(0,0.71)$. In Tabs. 2 and 3, different cases
489 with 3 and 4 random variables are shown. These cases differ with respect to the prior failure

490 probability P_F , the correlation between the random variables ρ_{ij} and the observed
 491 measurements \mathbf{m} . For each case, a reference solutions $P_{F|\mathbf{M}}$ is calculated analytically.

492
 493 Table 2. Evaluation of the discretization error for different measurement outcomes \mathbf{m} , for problems with $n = 3$
 494 random variables. a is the constant in the LSF, Eq. 15; ρ_{ij} is the correlation coefficient between X_i and X_j for all
 495 $i \neq j$; P_F and $P_{F|\mathbf{M}}$ denote the analytically calculated prior and posterior failure probabilities; $\hat{P}_{F|\mathbf{M}}$ is the
 496 conditional failure probability calculated with the discrete BN.

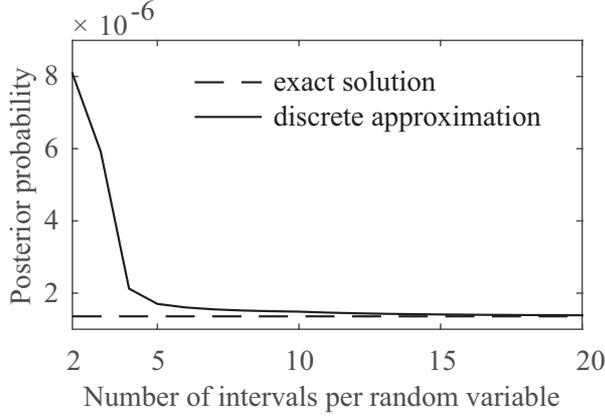
a	c_{up}	ρ_{ij}	P_F	\mathbf{m}	$P_{F \mathbf{M}}$	$\hat{P}_{F \mathbf{M}}$	Absolute error	Relative error [%]
100	10^3	0	$3.6E - 5$	[3.0,2.9,2.9]	$4.3E - 5$	$4.5E - 5$	$3E - 6$	6
100	10^3	0	$3.6E - 5$	[2.3,1.1,2.1]	$4.6E - 6$	$5.3E - 6$	$7E - 7$	14
100	10^3	0	$3.6E - 5$	[0.9,2.4,0.9]	$2.8E - 7$	$3.5E - 7$	$7E - 8$	25
200	15^3	0.5	$1.6E - 4$	[1.6,2.0,1.2]	$1.4E - 6$	$1.4E - 6$	$1E - 7$	4
400	8^3	0.5	$6.4E - 6$	[2.6,3.0,3.2]	$8.2E - 7$	$8.9E - 7$	$7E - 8$	9
400	12^3	0.5	$6.4E - 6$	[3.6,3.3,4.3]	$4.9E - 6$	$5.0E - 6$	$1E - 9$	3

497
 498 Table 3. Evaluation of the discretization error for different measurement outcomes \mathbf{m} . The number of random
 499 variables $n = 4$; a is the constant in the LSF, Eq. 15; ρ_{ij} is the correlation coefficient between X_i and X_j for all
 500 $i \neq j$; P_F and $P_{F|\mathbf{M}}$ denote the analytically calculated prior and posterior failure probabilities; $\hat{P}_{F|\mathbf{M}}$ is the
 501 conditional failure probability calculated with the discrete BN.

a	c_{up}	ρ_{ij}	P_F	\mathbf{m}	$P_{F \mathbf{M}}$	$\hat{P}_{F \mathbf{M}}$	Absolute error	Relative error [%]
400	10^4	0	$1.7E - 5$	[2.2,3.2,2.4,3.4]	$9.5E - 6$	$1.0E - 5$	$9E - 7$	9
400	10^4	0	$1.7E - 5$	[1.6,1.6,1.6,2.0]	$6.5E - 7$	$7.9E - 7$	$1E - 7$	21
400	10^4	0	$1.7E - 5$	[1.1,2.3,1.9,1.2]	$2.4E - 7$	$3.0E - 7$	$6E - 8$	26
600	10^4	0.5	$1.3E - 3$	[3.3,1.7,2.8,2.6]	$4.2E - 4$	$4.3E - 4$	$2E - 5$	4
800	8^4	0.5	$5.3E - 4$	[1.9,2.0,1.9,2.4]	$1.8E - 5$	$1.9E - 5$	$1E - 6$	8

502
 503 The results in Tables 2 and 3 show that the proposed methodology for discretization leads to
 504 errors in the posterior probability estimate that are acceptably small for most engineering
 505 applications. (It is reminded that discretization does not lead to a discretization error in the
 506 prior case.) As expected, the relative error is larger when the posterior probability is low, and
 507 the absolute error is larger when the posterior probability is high. This follows from the error
 508 measure defined in Eq. 11, which balances the relative with the absolute error. In addition,
 509 the results do not display any apparent effect of correlation on the accuracy.

510 To assess the effect of the choice of the number of discretization intervals, the failure
 511 probability $\widehat{P}_{F|M}$ was calculated for a discretization scheme with up to 20 intervals per RV for
 512 the fourth measurement case in Tab. 2. The estimated failure probabilities $\widehat{P}_{F|M}$ are plotted
 513 together with the exact solution in Fig. 12.



514
 515 Figure 12. Posterior probability $\widehat{P}_{F|M}$ as a function of the number of intervals per random variable together with
 516 the exact (analytical) solution $P_{F|M}$ for the fourth measurement case [1.6, 2.0, 1.2] in Tab 2.

517 4.2 Verification example II

518 The failure criterion applied in verification example I (Eq. 15) leads to a linear LSF in U-
 519 space. To verify the accuracy of the proposed method for problems with non-linear LSFs in
 520 U-space, we additionally investigate the following LSF:

$$g(\mathbf{x}) = a - \sum_{i=1}^n X_i \quad (16)$$

521 Again the basic random variables X_1 to X_n are distributed as $X_1 \sim LN(0,0.5)$
 522 and $X_2, \dots, X_n \sim LN(1,0.3)$. Different cases with $n = 2, 3$ and 4 random variables are
 523 investigated. Measurements $M_i = m_i$ are available for all basic random variables; associated
 524 to these measurement are multiplicative measurement errors $\varepsilon_i \sim LN(0,0.71)$. For
 525 independent random variables X_i it is possible to determine posterior distributions
 526 $f_{X_i|M_i}(x_i|m_i)$ analytically. The posterior failure probabilities $P_{F|M}$, which are used as
 527 reference solutions, are calculated through importance sampling with 10^7 samples. The
 528 results are presented in Tab. 4.

529 Table 4. Evaluation of the discretization error for different measurement outcomes \mathbf{m} . The problems have
530 $n = 2, 3$ or 4 random variables; a is the constant in the LSF, Eq. 16; ρ_{ij} is the correlation coefficient between X_i
531 and X_j for all $i \neq j$; P_F and $P_{F|\mathbf{M}}$ denote the prior respectively posterior failure probabilities, which are calculated
532 through importance sampling with 10^7 samples; $\hat{P}_{F|\mathbf{M}}$ is the conditional failure probability calculated with the
533 discrete BN. Since for correlated basic random variables there is no analytical solution, in these cases the
534 updating of the basic random variables was performed through rejection sampling with $> 5E7$ accepted samples.
535

a	c_{up}	ρ_{ij}	P_F	\mathbf{m}	$P_{F \mathbf{M}}$	$\hat{P}_{F \mathbf{M}}$	Absolute error	Rel. error [%]
<u>$n = 2$:</u>								
12	10^2	0	$1.3E - 5$	[2.8,4.5]	$1.4E - 5$	$1.2E - 5$	$2E - 6$	15
12	10^2	0	$1.3E - 5$	[2.3,2.4]	$3.3E - 6$	$3.5E - 6$	$2E - 7$	6
10	12^2	0	$1.7E - 4$	[4.0,3.2]	$4.0E - 4$	$3.7E - 4$	$3E - 5$	7
12	10^2	0.5	$1.7E - 4$	[2.3,2.4]	$4.8E - 5$	$5.0E - 5$	$2E - 6$	4
<u>$n = 3$:</u>								
15	10^3	0	$3.7E - 5$	[2.1,5.6,5.0]	$4.8E - 5$	$4.5E - 5$	$4E - 6$	7
15	10^3	0	$3.7E - 5$	[1.1,3.7,3.4]	$1.5E - 5$	$1.8E - 5$	$3E - 6$	20
13	12^3	0	$5.0E - 4$	[3.0,3.0,3.0]	$5.4E - 4$	$5.4E - 5$	$3E - 6$	1
16	10^3	0.5	$9.1E - 4$	[3.0,6.0,5.0]	$1.8E - 3$	$1.9E - 3$	$3E - 5$	2
<u>$n = 4$:</u>								
20	8^4	0	$7.4E - 6$	[2.0,4.0,3.4,3.0]	$4.9E - 6$	$5.6E - 6$	$6E - 7$	13
17	8^4	0	$3.1E - 4$	[1.0,1.4,1.2,2.0]	$4.5E - 5$	$5.2E - 5$	$7E - 6$	15
17	12^4	0	$3.1E - 4$	[3.1,2.0,3.3,2.4]	$2.5E - 4$	$2.5E - 4$	$7E - 6$	3
24	8^4	0.5	$3.7E - 4$	[1.0,1.4,1.2,2.0]	$1.1E - 5$	$1.1E - 5$	$9E - 7$	8

536

537 The results in Tab. 4 do not differ substantially from Tabs. 2 and 3. This indicates that the
538 (weak) non-linearity of the LSF function describing failure does not affect the accuracy
539 significantly.

540 4.3 Runway overrun

541 Runway overrun (RWO) of a landing aircraft is one of the most critical accidents types in
542 civil aviation (IATA, 2013). A conceptual RWO warning system is developed with the

543 proposed discretization procedure. It provides RWO probabilities conditional on observations
 544 of the landing-weight, the headwind and the approach speed for different aircraft types and
 545 different airports. For a detailed description of how this problem can be treated in BN
 546 framework we refer to (Zwirglmaier and Straub, 2015).

547 RWO is the event of the operational landing distance exceeding the available runway length
 548 (Fig. 13). Correspondingly, a LSF for runway overrun can be defined as:

$$g(\mathbf{X}) = \text{Runway length} - \text{Operational landing distance}(\mathbf{X}) \quad (17)$$

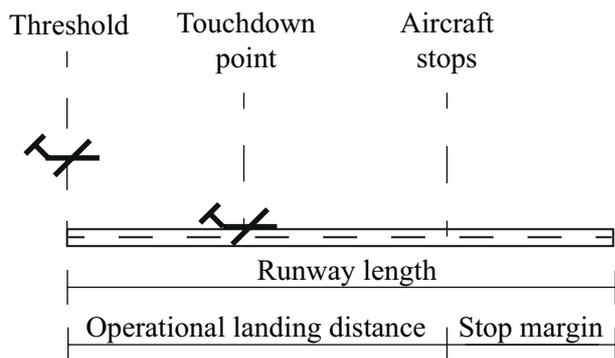
549 with \mathbf{X} representing the basic random variables of the problem.

550 (Drees and Holzapfel, 2012) proposed a model for the operational landing distance required
 551 by a landing aircraft, which is applied here. The model, as well as the basic random variables
 552 \mathbf{X} , are presented in (Zwirglmaier et al., 2014), which also includes a detailed description of
 553 the reliability and sensitivity analysis.

554 We consider two different airports (AP I and AP II) and two different aircraft types (AC A
 555 and AC B). While the aircraft type affects the landing-weight, the airport affects both the
 556 headwind and the approach speed. The distribution models for landing-weight, headwind and
 557 approach speed deviation at the different airports and with the different aircraft types are
 558 given in Tabs. 5–7. All other basic random variables of the problem are not affected by the
 559 airport and aircraft type and are as in (Zwirglmaier et al., 2014).

560 Tab. 8 summarizes the FORM importance measures of all random variables \mathbf{X} computed for
 561 the four combinations of aircrafts and airports.

562



563
 564 Figure 13. Runway definitions.

565

566

567 Table 5. Distribution models for landing weight conditional on the aircraft.
568

Landing weight [t]			
Aircraft	Distribution	Mean	Std. deviation
A	Weibull (min)	59.25	1.69
B	Weibull (min)	64.25	1.69

569 Table 6. Distribution models for head wind conditional on the airport.

Head wind [kts]			
Airport	Distribution	Mean	Std. deviation
I	Normal	5.42	5.75
II	Normal	6.51	5.75

570 Table 7. Distribution models for approach speed deviation conditional on the airport.
571
572

Approach speed deviation [kts]			
Airport	Distribution	Mean	Std. deviation
I	Gumbel (max)	4.69	4.21
II	Gumbel (max)	5.63	4.21

573 Table 8. FORM importance measures α_i for each aircraft-airport combination and every basic random variable
574 in the RWO application.
575

Random variable	α_i				Annotation
	(I/A)	(I/B)	(II/A)	(II/B)	
Landing weight [t]	0.09	0.10	0.11	0.09	Modeled
Headwind [kts]	-0.65	-0.61	-0.67	-0.60	Modeled
Temperature [°C]	0.03	-0.00	-0.03	-0.03	Not important
Air pressure [hPa]	0.01	-0.01	-0.01	-0.00	Not important
Touchdown point [m]	0.20	0.16	0.18	0.20	Modeled
Approach speed deviation [kts]	0.20	0.21	0.20	0.24	Not observable
Time of spoiler deployment [s]	-0.00	-0.00	0.01	0.01	Not important
Time of breaking initiation [s]	0.70	0.74	0.68	0.73	Not observable
Time of reverser deployment [s]	0.03	0.04	0.06	0.05	Not important
Time of breaking end [s]	-0.02	-0.01	0.01	0.02	Not important

576 4.3.1 Selection of relevant random variables

577 The applied RWO model includes 10 basic random variables. However, it is sufficient to
578 include only a selection of these explicitly in the BN. Random variables that are not relevant
579 for the prediction of RWO in the considered scenarios can be excluded. This is the case for
580 random variables with a low FORM importance, whose value does not depend significantly
581 on airport and aircraft type. Here, all random variables, whose absolute value of the FORM
582 importance measure $|\alpha_i|$'s is smaller than 0.1, are excluded (see Tab. 8). The one exception is
583 landing weight, since its mean value is substantially influenced by the aircraft type.

584 One can additionally exclude random variables that cannot be measured before the decision
585 on whether to land or not is made. This holds for Touchdown point and the time at which the
586 pilot initiates breaking. Since these basic random variables are also not needed to simplify the
587 modeling of dependencies, it is not necessary to explicitly model them in the BN, as indicated
588 in Tab. 8.

589 4.3.2 BN model

590 The resulting BN of the RWO warning system is shown in Fig. 14. During the aircraft
591 approach, measurements can be obtained for the three basic random variables included in the
592 BN.

593 The random variables are discretized separately for each aircraft-airport combination (joint
594 states of discrete parents) with 8 intervals each, following the proposed discretization
595 procedure. In a second step, the discretization schemes are merged, i.e. the regions of the
596 outcome space, which are discretized with fine intervals for at least one of the aircraft-airport
597 combinations, are discretized with the respective fine intervals also in the merged
598 discretization scheme. In the end 15 (landing-weight), 10 (headwind) and 9 (approach speed
599 deviation) intervals are used to discretize the three basic random variables.

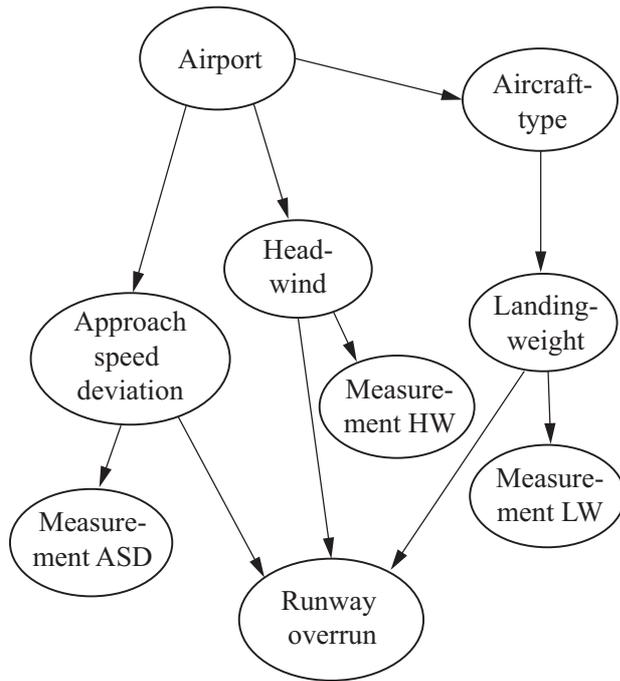
600 For all observable quantities, the measurements m_i are modeled with an additive observation
601 error:

$$m_i = x_i + \varepsilon_i \quad (18)$$

602 ε_i is modeled by a normal distribution with zero mean and standard deviation σ_{ε_i} .

603 For the random variable landing weight (at landing time), the standard deviation of the
604 measurement error is $\sigma_{\varepsilon_{LW}} = 0.34$ tons. Due to turbulences governing wind speeds, the
605 measurement of the head wind speed at the time of the measurement is only an uncertain
606 indicator for the head wind speed at landing time; we model the measurement error with a
607 standard deviation $\sigma_{\varepsilon_{HW}} = 2.88$ kts. The measurement uncertainty associated with the
608 approach speed deviation at landing has standard deviation $\sigma_{\varepsilon_{ASD}} = 4.21$ kts.

609 49 (Measurement LW), 57 (Measurement HW) and 57 (Measurement ASD) intervals are used
610 to discretize the measurement nodes.



611
612
613 Figure 14. BN structure for a RWO warning system.

614 4.3.3 Results

615 In Tab. 9, RWO probabilities for the different airports and aircrafts obtained with the discrete
616 BN are compared to solutions, which were calculated by importance sampling around the
617 design point.

618 Table 9. RWO probabilities for the different airports and aircrafts calculated with the discrete BN p_{BN} ,
619 together with solutions calculated by importance sampling around the design point p_{DS} . The latter have a
620 sampling error with coefficient of variation in the order of 10%.
621

AP/AC	p_{BN}	p_{DS}
I/A	$2.0e - 7$	$1.9e - 7$
I/B	$1.0e - 6$	$9.2e - 7$
II/A	$1.3e - 7$	$1.3e - 7$
II/B	$6.9e - 7$	$6.5e - 7$

622 In Tab. 10, results obtained with the BN for different hypothetical cases of aircrafts
623 approaching an airport are presented. In each of these cases, measurements associated with
624 landing weight, headwind and the approach speed deviation are made. A threshold on the
625 probability of RWO is used to decide, whether or not the pilot should continue landing or
626 cancel the landing attempt. Here we assume that up to a RWO probability of 10^{-6} the pilot
627 should continue landing.

628 Table 10. Probabilities of RWO and corresponding decision on landing, computed with the BN for different
629 sets of observations.

Case	Airport	Aircraft	Meas. LW	Meas. HW	Meas. ASD [kts]	Pr (RWO)	Landing
a)	I	B	63	0	10.5	$2.5e - 8$	Yes
b)	I	A	61	-10	5	$4.8e - 6$	No
c)	II	B	67	3	0	$6.5e - 10$	Yes
d)	II	A	57.5	-12	3	$1.3e - 6$	No

630 5 Discussion

631 When modeling with BNs, it is often necessary or beneficial to discretize continuous random
632 variables. When the BN includes rare events that are a function of such random variables, the
633 choice of the discretization scheme is non-trivial. In this contribution, we investigate this
634 discretization based on FORM concepts, and propose a heuristic procedure for an efficient
635 discretization in these cases. This is based on importance measures α_i obtained through a
636 FORM analysis, which represent the influence of the uncertainty associated with a random
637 variable X_i .

638 The most important finding is that discretization should focus on the area around the most
639 likely failure point (design point), identified by a FORM analysis. Furthermore, we find that
640 optimally all random variables should be discretized with approximately equal numbers of
641 intervals, independent of their importance, as long as $|\alpha_i|$ is not close to zero. The widths of
642 the intervals should be selected based on the FORM importance α_i of the random variables.
643 With increasing importance, the interval width should be reduced, leading to finer
644 discretization for larger $|\alpha_i|$. This relation is particularly evident for $|\alpha_i| \geq 0.8$. We show that
645 it is possible to fit a parametric function to approximate the relation between $|\alpha_i|$ and the
646 optimal width of the region on which the discretization should focus.

647 This parametric function is used to derive a heuristic procedure for finding an efficient
648 discretization. This allows the extrapolation of the optimization results to problems with more
649 random variables. As demonstrated by the verification examples, the heuristic procedure leads
650 to accurate results.

651 This paper is restricted to static discretization. Application of the proposed procedure within
652 dynamic discretization (e.g. (Neil et al., 2008)) should be investigated. The results of the
653 procedure can serve as an initial discretization scheme, which is iteratively adjusted within
654 dynamic discretization. This might strongly enhance the convergence performance of these
655 algorithms.

656 The gain in computational efficiency resulting from the proposed procedure over alternative
657 static discretization approaches is problem specific. Some insights can be gained from Figure
658 8. A discretization with intervals of equal width centered at the origin would require between
659 approximately 2 to 5 times more intervals per random variable to achieve the same accuracy.

660 Because of the exponential increase of computational effort with number of random variables,
661 this leads to a considerable increase in efficiency. The gain compared to discretization with
662 equal-frequency intervals is expected to be even higher, since equal frequency intervals
663 focuses the fine intervals on the region of high probability density rather than on the tails of
664 the distribution.

665 This paper focuses on component reliability problems, which are characterized by a single
666 design point. Nevertheless the heuristics derived can also be applied to system reliability
667 problems. System reliability problems can in general be treated as combinations of
668 component reliability problems. Parallel and series systems are to be distinguished. For
669 parallel systems discretization should be performed based on the joint design point of the
670 problem. For series systems, following the same line of thought as in the runway overrun
671 example, discretization can be performed separately for each component problem
672 (corresponds to the discrete cases i.e. airport- aircraft combinations in the RWO example). In
673 a second step the discretization schemes can be merged. In the same way it is possible to
674 apply the heuristic to multi state components. One can treat each limit state surface (LSF)
675 defining the boundary between two states separately and merge the discretization schemes
676 afterwards.

677 The number of basic random variables in a single LSF that can be modeled explicitly in a BN
678 is limited to around 5 to 8. This is due to the exponential growth of the target nodes CPT with
679 increasing number of parents and is independent of the discretization method. Despite this
680 limitation, BNs are applicable to many practical problems – particularly if one considers that
681 usually not all basic random variables need to be modeled explicitly as nodes, as
682 demonstrated in the presented example.

683 While in this paper the focus was on the discretization of the basic random variables, it is
684 straightforward to incorporate the BNs discussed into larger models.

685 **6 Conclusion**

686 We investigate discretization of continuous reliability problems such that they can be treated
687 in a discrete Bayesian network framework. Reliability problems with linear LSF in standard
688 normal space are considered. These can be seen as FORM approximations of reliability
689 problems. For these linear LSFs, optimal discretization schemes are found, which are optimal
690 with respect to an error measure calculated through a preposterior analysis. Since FORM is
691 known to give good approximations also for most non-linear reliability problem, the resulting
692 discretization schemes are efficient also for non-linear LSFs. The main findings presented in
693 this paper are:

- 694 • An optimal discretization scheme should discretize finely the area around the FORM
695 design point.

- 696 • The size of the sub-region of the outcome space of a random variable X_i can be
- 697 reduced significantly for random variables whose corresponding uncertainty is
- 698 dominating the reliability problem
- 699 • The number of intervals used for discretization should be approximately equal for all
- 700 basic random variables

701 On this basis, we propose a heuristic that can be used to find an efficient discretization
 702 scheme. In verification examples, this heuristic is found to give good accuracy and efficiency.

703 7 References

- 704 AGENA. 2005. Available: <http://www.agenarisk.com>.
- 705 AU, S.-K. & BECK, J. L. 2001. Estimation of small failure probabilities in high dimensions by subset
 706 simulation. *Probabilistic Engineering Mechanics*, 16, 263-277.
- 707 DER KIUREGHIAN, A. 2005. First- and Second-Order Reliability Methods. In: NIKOLAIDIS, E.,
 708 GHIOCEL, D. M. & SINGHAL, S. (eds.) *Engineering Design Reliability: Handbook*. CRC
 709 PressINC.
- 710 DITLEVSEN, O. & MADSEN, H. O. 2007. *Structural Reliability Methods*. John Wiley & Sons.
- 711 DOUGHERTY, J., KOHAVI, R. & MEHRAN, S. Supervised and Unsupervised Discretization of
 712 Continuous Features. In: PRIEDITIS, A. & RUSSEL, S., eds. *Machine Learning: Proceedings*
 713 *of the Twelfth International Conference, 1995 San Francisco*.
- 714 DREES, L. & HOLZAPFEL, F. 2012. Determining and Quantifying Hazard Chains and their
 715 Contribution to Incident Probabilities in Flight Operation. *AIAA Modeling and Simulation*
 716 *Technologies Conference*. American Institute of Aeronautics and Astronautics.
- 717 FRIIS-HANSEN, A. 2000. *Bayesian Networks as a Decision Support Tool in Marine Applications*.
 718 PhD, Technical university of denmark.
- 719 HANEA, A., NAPOLES, O. M. & ABABEI, D. 2015. Non-parametric Bayesian networks: Improving
 720 theory and reviewing applications. *Reliability Engineering & System Safety*, 144, 265-284.
- 721 HANEA, A. M., KUROWICKA, D. & COOKE, R. M. 2006. Hybrid method for quantifying and
 722 analyzing Bayesian belief nets. *Quality and Reliability Engineering International*, 22, 709-729.
- 723 HOHENBICHLER, M. & RACKWITZ, R. 1981. Nonnormal dependent vectors in structural safety.
 724 *Journal of Engineering Mechanics*, 107, 1227-1238.
- 725 IATA 2013. Safety report 2012. International Air Transport Association (IATA).
- 726 JENSEN, F. V., LAURITZEN, S. L. & OLESEN, K. G. 1990. Bayesian updating in causal
 727 probabilistic networks by local computations. *Computational statistics quarterly*, 4, 269-282.
- 728 JENSEN, F. V. & NIELSEN, T. D. 2007. *Bayesian networks and decision graphs*, New York [u.a.],
 729 Springer.
- 730 KJAERULFF, U. B. & MADSEN, A. L. 2013. *Bayesian Networks and Influence Diagrams: A Guide*
 731 *to Construction and Analysis*, Springer Publishing Company, Incorporated.
- 732 KOTSIANTIS, S. & KANELLOPOULOS, D. 2006. Discretization techniques: A recent survey.
 733 *GESTS International Transactions on Computer Science and Engineering*, 32, 47-58.
- 734 KOZLOV, A. V. & KOLLER, D. 1997. Nonuniform dynamic discretization in hybrid networks. *13th*
 735 *Annual Conference on Uncertainty in AI (UAI)*. Providence, Rhode Island.
- 736 LANGSETH, H., NIELSEN, T. D., RUMI, R. & SALMERÓN, A. 2012. Mixtures of truncated basis
 737 functions. *International Journal of Approximate Reasoning*, 53, 212-227.
- 738 LANGSETH, H., NIELSEN, T. D., RUMÍ, R. & SALMERÓN, A. 2009. Inference in hybrid Bayesian
 739 networks. *Reliability Engineering & System Safety*, 94, 1499-1509.

- 740 LAURITZEN, S. L. & SPIEGELHALTER, D. J. 1988. Local computations with probabilities on
741 graphical structures and their application to expert systems. *Journal of the Royal Statistical*
742 *Society. Series B (Methodological)*, 157-224.
- 743 LERNER, U. N. 2002. *Hybrid Bayesian networks for reasoning about complex systems*. PhD,
744 Stanford University.
- 745 NEIL, M., TAILOR, M., MARQUEZ, D., FENTON, N. & HEARTY, P. 2008. Modelling dependable
746 systems using hybrid Bayesian networks. *Reliability Engineering & System Safety*, 93, 933-
747 939.
- 748 RACKWITZ, R. 2001. Reliability analysis—a review and some perspectives. *Structural Safety*, 23,
749 365-395.
- 750 RAIFFA, H. & SCHLAIFER, R. 1961. *Applied statistical decision theory*, Boston, Division of
751 Research, Graduate School of Business Administration, Harvard University.
- 752 SINDEL, R. & RACKWITZ, R. 1998. Problems and Solution Strategies in Reliability Updating.
753 *Journal of Offshore Mechanics and Arctic Engineering*, 120, 109-114.
- 754 STRAUB, D. 2009. Stochastic Modeling of Deterioration Processes through Dynamic Bayesian
755 Networks. *Journal of Engineering Mechanics*, 135, 1089-1099.
- 756 STRAUB, D. 2011. Reliability updating with equality information. *Probabilistic Engineering*
757 *Mechanics*, 26, 254-258.
- 758 STRAUB, D. 2014a. Engineering Risk Assessment. *Risk – A Multidisciplinary Introduction*. Springer.
- 759 STRAUB, D. 2014b. Value of information analysis with structural reliability methods. *Structural*
760 *Safety*, 49, 75-85.
- 761 STRAUB, D. & DER KIUREGHIAN, A. 2010a. Bayesian Network Enhanced with Structural
762 Reliability Methods: Application. *Journal of Engineering Mechanics*, 136, 1259-1270.
- 763 STRAUB, D. & DER KIUREGHIAN, A. 2010b. Bayesian Network Enhanced with Structural
764 Reliability Methods: Methodology. *Journal of Engineering Mechanics*, 136, 1248-1258.
- 765 STRAUB, D., PAPAIOANNOU, I. & BETZ, W. 2016. Bayesian analysis of rare events. *Accepted for*
766 *publication in Journal of Computational Physics*.
- 767 ZHANG, N. L. & POOLE, D. A simple approach to Bayesian network computations. Proc. of the
768 Tenth Canadian Conference on Artificial Intelligence, 1994.
- 769 ZHU, J. & COLLETTE, M. 2015. A dynamic discretization method for reliability inference in
770 Dynamic Bayesian Networks. *Reliability Engineering & System Safety*, 138, 242-252.
- 771 ZWIRGLMAIER, K., DREES, L., HOLZAPFEL, F. & STRAUB, D. 2014. Reliability analysis for
772 runway overrun using subset simulation. *ESREL 2014*. Wroclaw, Poland.
- 773 ZWIRGLMAIER, K. & STRAUB, D. 2015. Discretization of Structural Reliability Problems: An
774 Application to Runway Overrun. *12th International Conference on Applications of Statistics*
775 *and Probability in Civil Engineering, ICASP12*. Vancouver, Canada.
- 776