Sampling Without Probabilistic Model

MICHAEL BEER
Institut für Statik und Dynamik der Tragwerke, Technische Universität Dresden, 01062 Dresden, Germany,
email: Michael.Beer@tu-dresden.de

Abstract. In this paper a novel technique for random vector sampling starting from rare data is presented. This model-free sampling technique is developed to operate without a probabilistic model. Instead of estimating a distribution function, the information contained in a given small sample is extracted directly to produce the sampling result as a second sample of considerably larger size that completely reflects the properties of the original small sample. As a further enhancement, the new sampling technique is extended to processing imprecise data.

Model-free sampling can be coupled to stochastic structural analysis and safety assessment by application to input data or to result data. In the case of limited data, for instance, due to a high numerical cost of the underlying computational model, the novel technique can be applied to generate a proper estimation of stochastic structural responses and, thanks to a sound reproduction of distribution tails, of structural reliability. In this context it can provide a basis for increasing the numerical efficiency of Monte Carlo simulations in computational stochastic mechanics.

The usefulness of the model-free sampling technique is underlined by means of numerical examples.

Keywords: Sampling; Monte Carlo simulation; Imprecise data; Fuzzy randomness; Uncertain structural analysis; Safety assessment.

1. Introduction

Simulation techniques often offer the only possibility for solving problems in which random properties must be taken into account. Indeed, Monte-Carlo simulation and further developments thereof have become versatile tools for solving a variety of problems in a wide range of engineering disciplines, see (Schuëller and Spanos, 2001).

An essential precondition for obtaining realistic results from a simulation is the availability of statistically-validated probability distributions for the input variables. The specification of these distributions thus plays an essential role, see (Schuëller, 2001b). For determining reliably parameters and forms of probability distributions, extensive data in the form of samples are required. This enables using well-developed and sophisticated methods of statistical estimation theory and test theory, which operate parametrically or non-parametrically (Mood et al., 1974). Further, the numerical procedure for processing the specified random quantities in structural analysis and safety assessment must be computationally efficient to enable the stochastic analysis of large and nonlinear systems (Schenk et al., 2005; Schenk and Schuëller, 2005; Schuëller et al., 2003). Only if a sufficient amount of structural response data is produced, their stochastic properties can be identified reliably, and failure probabilities can be estimated appropriately with the aid of statistical methods.

© 2006 *by authors. Printed in USA.*

In any case, problems may primarily occur in the following three situations. First, the available information is limited in the form of small samples (Problem I). Second, structural response data can only be produced to a limited extent due to a high computational cost in analyzing the underlying structural problem in correspondence with the simulation of random input quantities (Problem II). Third, the sample elements are characterized by uncertainty or imprecision (Problem III). As a result, probability distributions for input variables and structural responses, or probabilities of defined events such as structural failure cannot be specified to a sufficient degree of reliability.

1.1. PROBLEM I – SMALL SAMPLES

In this case statistical estimations and tests based on small samples may yield vague and ambiguous results. For an appropriate level of confidence wide intervals for the estimated values are obtained. The variety of possible probabilistic models can almost not sufficiently narrowed with the aid of tests. This applies, in particular, if the distribution type is not pure in the form of a compound or multimodal distribution, or if general, for example nonlinear, dependencies in multi-dimensional cases are present. The less information the sample contains the more subjectivity is introduced with the specification of a certain probabilistic model. On the other hand, there is no evidence that the information that is actually contained in the sample is extracted completely but only to a certain degree. The results obtained on such basis may vary dramatically. Approaches to remedy this problem aim at determining bounds for the possible range of stochastic models and prognoses. A distiction can be made here between pure probabilistic methods (Deodatis et al., 2003; Papadopoulos et al., 2005; Red-Horse and Benjamin, 2004), which are focused on finding the bounds with different externally applied search strategies, and methods based on extended uncertainty models such as p-box (Berleant and Zhang, 2004), random sets (Hall and Lawry, 2004; Tonon et al., 2000), sets of probability measures (Fetz and Oberguggenberger, 2004), or fuzzy randomness (Möller and Beer, 2004), which cover the possible range of probabilistic models at once and intrinsically contain the search for probabilistic bounds. For this intrinsic search, a generally applicable and numerically efficient optimization algorithm has been developed as modified evolution strategy (Möller et al., 2000), the usefulness of which has already been shown, for example, in safety assessments coupled to a nonlinear structural analysis (Möller et al., 2003). Moreover, the model of fuzzy randomness provides a basis for an evaluation of those problems on several levels of subjective confidence in an encapsulated manner. Further, a variety of methods based on Bayesian theory (Bernardo and Smith, 1994) can be employed if subjective information is available beyond the small sample.

1.2. PROBLEM II – HIGH COMPUTATIONAL COST

The solution to Problem II comprises a wide variety of methods to increase numerical efficiency of stochastic structural analysis and safety assessment (Schuëller, 2001a). The corresponding developments primarily concern enhancements in Stochastic Finite Element Methods and in the numerical simulation of stochastic processes, which have already reached practical relevance in solving engineering problems (Ghanem and Spanos, 1991; Schenk and Schuëller, 2005). Their practical applicability substantially hinges on an efficient representation of the random input quantities. In this context, spectral representations of stochastic processes have attracted considerable attention, which particularly refers to Karhunen-Loéve or Polynomial Chaos expansion (Du et al., 2005; Field Jr. and Grigoriu, 2004; Gutierrez and Zaldivar, 2000; Phoon et al., 2005; Schuëller et al., 2003; Spanos and Ghanem, 1989). For reliability analysis, which focuses on rare events,

further efficiency improvements of simulation techniques are pursued. Among several strategies, the class of variance reducing methods such as Importance Sampling and variants thereof (Rackwitz, 2001; Schuëller, 2001a) probably represents the most popular kind of approaches. As these developments are not directly related to the approach pursued in this paper but only act as a motivation, their consideration is not extended at this point.

1.3. PROBLEM III – IMPRECISE DATA

This problem exceeds the limits of traditional methods in uncertainty quantification and processing, and represents a reasearch field of increasing interest. Traditionally, imprecision or uncertainty of sample elements is either neglected totally or taken into account approximately by selecting "probably adverse values" with respect to structural responses and safety measures from a possible value range. However, the actual impact of such a selection of input parameters can generally not be evaluated at the pre-stage of a simulation. On the other hand, the question arises as to how to model that uncertainty or imprecision realistically. It appears, for example, in situations in which the precision of measuring devices is strongly limited, the measuring points cannot be defined precisely (rough surfaces in thickness measurements), the expert evaluations influence the value specification, the measured values are gained under dubious conditions, and linguistic assessments are accounted for. In those cases the data possess random properties and non-random properties simultaneously. A pure probabilistic solution by applying the aforementioned approaches for dealing with limited information in the form of small samples is thus somewhat critical. Only Bayesian methods (Bernardo and Smith, 1994) are capable of incorporating subjective uncertainty, but still in terms of probability, which contradicts the non-random nature of some information. For a more pertinent uncertainty modeling in the case of non-probabilistic phenomena generalized uncertainty models have been developed (Fellin et al., 2005; Helton and Oberkampf, 2004), which are related to or covered by the framework of evidence theory. A comprehensive direct modeling of the imprecision or uncertainty of the individual elements of a random sample can be realized with the aid of a fuzzy randomness approach (Möller and Beer, 2004). Statistical investigations of uncertain or imprecise data and of properties of fuzzy random variables are, to a great extend, in an initial stage of development. Related research in this regard may be found in (Bandemer and Näther, 1992), in (Viertl, 1996), and in (Körner, 1997). These developments concern the analysis of imprecise data, the definition of statistical parameters, and the investigation of statistical laws for fuzzy random variables. Publications discussing the simulation of fuzzy randomness are rare. An approach evaluating fuzzy probability distribution functions on a trajectory-by-trajectory basis is presented in (Sickert et al., 2003). Numerical investigations of statistical properties of fuzzy random variables based on simulation are discussed in (Colubi et al., 2002). However, these methods require prior knowledge about the fuzzy probability distributions or the fuzziness of the realizations to be generated. General techniques for generating fuzzy realizations of fuzzy random variables are not known at the present time. Moreover, the application of traditional sampling methods to the numerical generation of fuzzy realizations encounters considerable difficulties. For instance, the numerical effort for estimating fuzzy parameters and fuzzy probability distributions from fuzzy-valued samples (fuzzy samples) is significantly high, in particular, when interaction between the fuzzy parameters is taken into account. Further, the simulation of fuzzy realizations starting from fuzzy probability distribution functions is not unique. That is, different fuzzy samples may have identical empirical fuzzy probability distribution functions. These conflicts hinder the pursuing of traditional sampling and simulation approaches.

1.4. SOLUTION IDEA

Despite considerable developments in answering the aforementioned three problems, an overall satisfying solution does not exist. In the following an attempt is made to develop a basis for a sampling technique to improve uncertainty processing in structural analysis and safety assessment in those problematic cases. The novel sampling technique is intended to circumvent an explicit specification of a probabilistic model, which avoids an introduction of subjectivity and motivates its denotation as "model-free sampling". Further, it should be capable of attaining appropriate results starting from samples of small size that may consist of uncertain or imprecise data.

The development starts from the basic statistical assumption that all information is contained in the sample. On the basis of a small sample a second sample of considerably larger size is numerically generated that completely reflects the statistical properties and uncertainty characteristics of the original small sample. This sampling technique can be applied to rare input data as well as to rare result data of a stochastic structural analysis and might thus be helpful as a preprocessor or as a postprocessor in combination with established simulation methods in diverse cases to improve estimations of stochastic structural responses and of structural reliability, and to increase the numerical efficiency of the computations. For enhancing the model-free sampling technique to processing imprecise data the generalized uncertainty model fuzzy randomness is taken as a basis, which enables to transfer stochastic uncertainty and non-stochastic uncertainty of the input data completely and simultaneously to the results of structural analysis and safety assessment. Finally, predictions of uncertain stochastic structural responses and of uncertain structural reliability are obtained.

2. Numerical Procedure

The basic concept of the model-free sampling technique is to generate the sampling result directly from a given sample instead of estimating a probability distribution and performing the sampling according to this. The characteristics of a population are described by a sufficiently large sample. As the mathematical model of a distribution function is not employed herein, conventional statistical estimations are dispensed with. The concept of statistical estimation is applied in a generalized sense.

The starting point is the observed sample

$$\underline{S}_0 = \left\{ \underline{\mathrm{s}}_{0,\mathrm{i}},\ \mathrm{i} = 1,\ ...,\ \mathrm{n}_0 \right\} \tag{1}$$

of size n_0 as a set of realizations $\underline{\mathrm{s}}_{0,\mathrm{i}} = \underline{\mathrm{x}}_{0,\mathrm{i}}$ in $\mathbb{R}^{\mathrm{n}}$of the underlying continuous random vector $\underline{\mathrm{X}}_0$ with unknown properties. A second concrete sample

$$\underline{S}_1 = \left\{ \underline{\mathrm{s}}_{1,\mathrm{k}},\ \mathrm{k} = 1,\ ...,\ \mathrm{n}_1 \right\} \tag{2}$$

of a considerably larger size $\mathrm{n}_1 \gg \mathrm{n}_0$ is then sought that represents the original sample $\underline{S}_0$ "as well as possible". That is, the new sample $\underline{S}_1$ is expected to exhibit statistical characteristics "comparable" to $\underline{S}_0$. This is realized by the following heuristic iterative approach with the superscript $^{[.]}$ indicating the iteration step.

1. The starting point is an arbitrary estimate

$$\underline{S}_1^{[0]} = \left\{ \underline{\mathrm{s}}_{1,\mathrm{k}}^{[0]},\ \mathrm{k} = 1,\ ...,\ \mathrm{n}_1 \right\} \tag{3}$$

for the sample $\underline{S}_1$. This is broadly specified without consideration of the information contained in the observed sample $\underline{S}_0$. All sample elements $\underline{\mathrm{s}}_{1,\mathrm{k}}^{[0]} = \underline{\mathrm{x}}_{1,\mathrm{k}}^{[0]}$ of $\underline{S}_1^{[0]}$ should possess the same information content. That is, they should exhibit the same probability density $\mathrm{f}_1^{[0]}$ in their immediate surroundings,

$$\int\limits_{\|\underline{\delta}\| \le \|\underline{\varepsilon}\|} \mathrm{f}_1^{[0]} \left(\underline{\mathrm{s}}_{1,\mathrm{p}}^{[0]} + \underline{\delta} \right) \mathrm{d}\underline{\delta} \ = \int\limits_{\|\underline{\delta}\| \le \|\underline{\varepsilon}\|} \mathrm{f}_1^{[0]} \left(\underline{\mathrm{s}}_{1,\mathrm{q}}^{[0]} + \underline{\delta} \right) \mathrm{d}\underline{\delta} \tag{4}$$

$$\forall\, \underline{\mathrm{s}}_{1,\mathrm{p}}^{[0]},\, \underline{\mathrm{s}}_{1,\mathrm{q}}^{[0]} \in \underline{S}_1^{[0]},\ \|\underline{\varepsilon}\| \ll 1.$$

This leads to the specification of $\underline{S}_1^{[0]}$ by continuous uniform distribution over a sufficiently large (physically meaningful), bounded domain $\underline{D} \subset \mathbb{R}^{\mathrm{n}}$ of possible (not excludable) realizations of the random vector $\underline{\mathrm{X}}_0$ represented by $\underline{S}_0$,

$$\left(\underline{\mathrm{X}}_1 \sim \mathrm{U}(\underline{D}) \right) \rightarrow \underline{S}_1^{[0]}. \tag{5}$$

2. The sample $\underline{S}_1^{[0]}$ is compared with the observed sample $\underline{S}_0$. The purpose of this comparison is to obtain a measure $\mathrm{G}^{[0]}$ for the statistical dissimilarity between the samples $\underline{S}_1^{[0]}$ and $\underline{S}_0$. For this dissimilarity measure, a real valued function

$$\mathrm{G}^{[.]} = \mathrm{g}\left(\underline{S}_0,\, \underline{S}_1^{[.]} \right) : \quad \left(\underline{S}_0,\, \underline{S}_1^{[.]} \right) \rightarrow \mathbb{R} \tag{6}$$

is selected which yields a global minimum for $\mathrm{G}^{[.]}$ if the samples $\underline{S}_1^{[.]}$ and $\underline{S}_0$ are "as similar as possible" in a statistical sense. That is, $\mathrm{G}^{[.]}$ is intended to be minimal if $\underline{S}_1^{[.]}$ and $\underline{S}_0$ originate from the same population $\underline{\mathbf{X}}$ with probability one,

$$\mathrm{P}\left(\underline{S}_0 \subset \underline{\mathbf{X}} \wedge \underline{S}_1^{[.]} \subset \underline{\mathbf{X}} \right) \rightarrow 1 \ \Rightarrow\ \mathrm{G}^{[.]} \Rightarrow \mathrm{MIN}. \tag{7}$$

Due to the fact that intended application is for samples consisting of imprecise data, established statistical test methods cannot be implemented.

3. The sample $\underline{S}_1^{[0]}$ is modified in such a way that a subset

$$\underline{S}_1^{[0]^-} = \left\{ \underline{\mathrm{s}}_{1,\mathrm{k}_1}^{[0]^-},\ ...,\ \underline{\mathrm{s}}_{1,\mathrm{k}_{\mathrm{m}_1}}^{[0]^-} \right\} \subset \underline{S}_1^{[0]} \tag{8}$$

of m_1 elements $\underline{\mathrm{s}}_{1,\mathrm{k}}^{[0]^-}$ (stipulated number with $\mathrm{m}_1 \ll \mathrm{n}_1$) are specified by discrete uniform distribution over the indices k of the elements $\underline{\mathrm{s}}_{1,\mathrm{k}}^{[0]}$ of $\underline{S}_1^{[0]}$,

$$\left(\mathrm{X}^- \sim \mathrm{U}\left(1,\, 2,\, ...,\, \mathrm{n}_1\right) \right) \rightarrow \left\{ \mathrm{k}_1,\, \mathrm{k}_2,\, ...,\, \mathrm{k}_{\mathrm{m}_1} \right\}, \tag{9}$$

and are removed from $\underline{S}_1^{[0]}$ to obtain the reduced sample

$$\underline{S}_{1,\mathrm{red}}^{[0]} = \underline{S}_1^{[0]} \setminus \underline{S}_1^{[0]^-}. \tag{10}$$

As an replacement for the removed elements $\underline{\mathrm{s}}_{1,\mathrm{k}}^{[0]^-}$, a set

$$\underline{S}_1^{[0]^+} = \left\{ \underline{\mathrm{s}}_{1,\mathrm{k}_1}^{[0]^+}, \ldots, \underline{\mathrm{s}}_{1,\mathrm{k}_{\mathrm{m}_1}}^{[0]^+} \right\} \tag{11}$$

of m_1 new elements $\underline{\mathrm{s}}_{1,\mathrm{k}}^{[0]^+}$ are generated randomly – again with the aid of a uniform distribution over the domain $\underline{D}$ of possible realizations specified in Step 1,

$$(\underline{\mathrm{X}}_1 \sim \mathrm{U}(\underline{D})) \rightarrow \underline{S}_1^{[0]^+}. \tag{12}$$

Their union with the reduced sample $\underline{S}_{1,\mathrm{red}}^{[0]}$ then yields the modified sample

$$\underline{S}_1^{[1]} = \underline{S}_{1,\mathrm{red}}^{[0]} \cup \underline{S}_1^{[0]^+}. \tag{13}$$

Then, the measure value $\mathrm{G}^{[1]}$ is computed for the modified sample $\underline{S}_1^{[1]}$.

4. The measure values $\mathrm{G}^{[1]}$ and $\mathrm{G}^{[0]}$ are compared. If $\mathrm{G}^{[1]} \geq \mathrm{G}^{[0]}$, it is concluded that the modification in Step 3 has not led to an improved estimation for $\underline{S}_1$. The modification is then nullified,

$$\underline{S}_1^{[0]} = \left(\underline{S}_1^{[1]} \setminus \underline{S}_1^{[0]^+} \right) \cup \underline{S}_1^{[0]^-}, \tag{14}$$

and a repeat modification of $\underline{S}_1^{[0]}$ is carried out according to Step 3. If $\mathrm{G}^{[1]} < \mathrm{G}^{[0]}$, on the other hand, the modified sample $\underline{S}_1^{[1]}$ yields an improved estimation compared with $\underline{S}_1^{[0]}$. The sample $\underline{S}_1^{[1]}$ is then taken as the basis for the next iteration step and modified anew according to the rules in Step 3 to produce $\underline{S}_1^{[2]}$. Again, the result is assessed. This procedure is repeated with an iteration counter r for successful modifications,

$$\begin{aligned} \underline{S}_1^{[\mathrm{r}+1]} &= \left\{ \underline{\mathrm{s}}_{1,1}^{[\mathrm{r}+1]}, \ldots, \underline{\mathrm{s}}_{1,\mathrm{k}}^{[\mathrm{r}+1]}, \ldots, \underline{\mathrm{s}}_{1,\mathrm{n}_1}^{[\mathrm{r}+1]} \right\} \\ &= \left(\left\{ \underline{\mathrm{s}}_{1,1}^{[\mathrm{r}]}, \ldots, \underline{\mathrm{s}}_{1,\mathrm{k}}^{[\mathrm{r}]}, \ldots, \underline{\mathrm{s}}_{1,\mathrm{n}_1}^{[\mathrm{r}]} \right\} \setminus \left\{ \underline{\mathrm{s}}_{1,\mathrm{k}_1}^{[\mathrm{r}]^-}, \ldots, \underline{\mathrm{s}}_{1,\mathrm{k}_{\mathrm{m}_1}}^{[\mathrm{r}]^-} \right\} \right) \\ &\quad \cup \left\{ \underline{\mathrm{s}}_{1,\mathrm{k}_1}^{[\mathrm{r}]^+}, \ldots, \underline{\mathrm{s}}_{1,\mathrm{k}_{\mathrm{m}_1}}^{[\mathrm{r}]^+} \right\}, \end{aligned} \tag{15}$$

until it is no longer possible to obtain an improvement of $\underline{S}_1$ beyond $\underline{S}_1^{[\mathrm{r}]}$. The dissimilarity measure $\mathrm{G}^{[\mathrm{r}]}$ then attains its minimum value. As the configuration of $\underline{S}_1$ that corresponds to the minimum of $\mathrm{G}^{[.]}$ can only be realized with probability zero (continuous case), a termination limit is defined for the probability with which an improvement can be obtained. The iteration is terminated if the average success rate of modifications attains a predefined and sufficiently small value. Finally, the sample $\underline{S}_1^{[\mathrm{r}^*]}$ obtained from the last successful modification is taken as the sampling result,

$$\underline{S}_1 = \underline{S}_1^{[\mathrm{r}^*]}. \tag{16}$$

The random modifications of $\underline{S}_1^{[r]}$ within the iteration ensure that the goal sample $\underline{S}_1$ is obtained as a random sample in consistency with established sampling principles. By virtue of its general concept the model-free sampling technique is a priori not limited in its applicability.

3. Real-Valued Samples

The model-free sampling technique is developed first, to apply for processing real-valued samples. The samples are deemed real-valued in the sense that their elements are denoted by scalars or vectors consisting of real numbers. This enables assessing the sampling results with the aid of established test methods. In this manner, the effectiveness of the model-free sampling may be evaluated.

3.1. BASIC ASPECTS

The critical point of the proposed technique is to formulate an appropriate function for characterizing the statistical dissimilarity $\mathrm{G}^{[.]}$ between the samples $\underline{S}_1^{[.]}$ and $\underline{S}_0$ in each iteration step r (see Step 2 in Section 2). This function $\mathrm{G}^{[.]}$ according to Eq. (6) is required to possess the following four general properties:

1. The measure $\mathrm{G}^{[.]}$ and established statistical test methods (homogeneity tests) must lead to basically analogous propositions regarding the statistical dissimilarity between $\underline{S}_1^{[.]}$ and $\underline{S}_0$. These propositions must be free of contradictions.

2. The mathematical formulation of the dissimilarity measure $\mathrm{G}^{[.]}$ must be extendable to apply for imprecise data in the form of fuzzy-valued samples. That is, the mathematical operations used in the definition of $\mathrm{G}^{[.]}$ for the real-valued case must possess appropriate counterparts in fuzzy arithmetics.

3. $\mathrm{G}^{[.]}$ is required to decrease at least tendentiously with decreasing statistical dissimilarity between $\underline{S}_1^{[.]}$ and $\underline{S}_0$. For samples $\underline{S}_1^{[.]}$ and $\underline{S}_0$ originating from the same population the measure $\mathrm{G}^{[.]}$ should take its global minimum value, see Eq. (7).

4. The mathematical structure of the measure $\mathrm{G}^{[.]}$ should be as simple as possible to ensure a fast numerical evaluation and thus to keep the computational cost reasonably low.

To develop a measure $\mathrm{G}^{[.]}$ that satisfies these requirements the following theoretical experiment is considered.

> *According to statistical estimation theory it is assumed that all available information is contained in the observed sample $\underline{S}_0$. Then, the best description of $\underline{S}_0$ is its empirical distribution function $\mathrm{F}_{\underline{S}_0}^{(\mathrm{e})}(\underline{\mathrm{x}})$, as it is a complete and unique representation of the information in $\underline{S}_0$. Moreover, in inferential statistics, the empirical distribution function is one of the most powerful estimators. If this $\mathrm{F}_{\underline{S}_0}^{(\mathrm{e})}(\underline{\mathrm{x}})$ is taken as the basis for sampling to numerically generate the sample $\underline{S}_1$, and no smoothing is applied, the resulting sample $\underline{S}_1$ and the observed sample $\underline{S}_0$ possess identical empirical distributions (in the limit),*

$$\lim_{\mathrm{n}_1 \to \infty} \mathrm{F}_{\underline{S}_1}^{(\mathrm{e})}(\underline{\mathrm{x}}) = \mathrm{F}_{\underline{S}_0}^{(\mathrm{e})}(\underline{\mathrm{x}})\,. \tag{17}$$

This corresponds to two significant properties of the samples $\underline{S}_0$ and $\underline{S}_1$ with respect to each other. First, the positions of the elements of $\underline{S}_0$ and $\underline{S}_1$ coincide. Second, each element of $\underline{S}_0$ has the same number of uniquely assigned elements from the sampling result $\underline{S}_1$. In the case of an underlying continuous random variable and an accordingly slightly smoothed empirical distribution $\mathrm{F}_{\underline{S}_0}^{(\mathrm{e})}(\underline{\mathrm{x}})$, the elements of the sampling result $\underline{S}_1$ are obtained in a close neighborhood of the elements of $\underline{S}_0$ with the same assignment property. Sampling results generated in this manner are high quality representations of the underlying sample $\underline{S}_0$ as may be shown by applying a variety of two-sample tests of homogeneity.

The measure $\mathrm{G}^{[.]}$ is thus formulated based on the configuration of the sampling result $\underline{S}_1$ from the theoretical experiment. This provides two criteria for monitoring the dissimilarity $\mathrm{G}^{[.]}$ between $\underline{S}_1^{[.]}$ and $\underline{S}_0$, which are defined as an assignment criterion and a distance criterion.

3.2. ASSIGNMENT CRITERION

The assignment criterion evaluates some order in the element configuration in the samples $\underline{S}_1^{[.]}$ and $\underline{S}_0$ with respect to each other. Each element $\underline{\mathrm{s}}_{0,\mathrm{i}}$, $\mathrm{i} = 1, ..., \mathrm{n}_0$ from sample $\underline{S}_0$ is supposed to have the same number $\mathrm{n}_{\mathrm{ass}}\left(\underline{\mathrm{s}}_{0,\mathrm{i}}\right)$ of uniquely assigned elements $\underline{\mathrm{s}}_{1,\mathrm{k}}^{[.]}$, $\mathrm{k} = 1, ..., \mathrm{n}_1$ from sample $\underline{S}_1^{[.]}$. The element assignment is defined on the basis of the Euclidean distance

$$\mathrm{d}\left(\underline{\mathrm{s}}_{0,\mathrm{i}}, \underline{\mathrm{s}}_{1,\mathrm{k}}^{[.]}\right) = \left\| \underline{\mathrm{s}}_{1,\mathrm{k}}^{[.]} - \underline{\mathrm{s}}_{0,\mathrm{i}} \right\| \tag{18}$$

between the respective elements $\underline{\mathrm{s}}_{1,\mathrm{k}}^{[.]}$ and $\underline{\mathrm{s}}_{0,\mathrm{i}}$. For each $\underline{\mathrm{s}}_{1,\mathrm{k}}^{[.]}$ one assigned element $\underline{\mathrm{s}}_{0,\mathrm{i}}\left(\underline{\mathrm{s}}_{1,\mathrm{k}}^{[.]}\right)$ is determined with

$$\underline{\mathrm{s}}_{0,\mathrm{i}}\left(\underline{\mathrm{s}}_{1,\mathrm{k}}^{[.]}\right) = \underline{\mathrm{s}}_{0,\mathrm{i}} \mid \mathrm{d}\left(\underline{\mathrm{s}}_{0,\mathrm{i}}, \underline{\mathrm{s}}_{1,\mathrm{k}}^{[.]}\right) = \min_{\mathrm{i}=1,\ldots,\mathrm{n}_1}\left[\mathrm{d}\left(\underline{\mathrm{s}}_{0,\mathrm{i}}, \underline{\mathrm{s}}_{1,\mathrm{k}}^{[.]}\right)\right], \tag{19}$$

see Figure 1. If Eq. (19) leads to a multiple assignment of elements $\underline{\mathrm{s}}_{0,\mathrm{i}}$ to the same $\underline{\mathrm{s}}_{1,\mathrm{k}}^{[.]}$, which occurs with probability zero in the continuous case but can appear in the numerical procedure due to limited computational precision, the element $\underline{\mathrm{s}}_{0,\mathrm{i}}$ with the smallest index i is selected for the assignment. The number $\mathrm{n}_{\mathrm{ass}}\left(\underline{\mathrm{s}}_{0,\mathrm{i}}\right)$ may then be obtained by means of an indicator function,

$$\mathrm{n}_{\mathrm{ass}}\left(\underline{\mathrm{s}}_{0,\mathrm{i}}\right) = \sum_{\mathrm{k}=1}^{\mathrm{n}_1} \mathrm{I}\left(\underline{\mathrm{s}}_{0,\mathrm{i}}, \underline{\mathrm{s}}_{1,\mathrm{k}}^{[.]}\right), \tag{20}$$

$$\mathrm{I}\left(\underline{\mathrm{s}}_{0,\mathrm{i}}, \underline{\mathrm{s}}_{1,\mathrm{k}}^{[.]}\right) = \begin{cases} 1 & \text{if } \underline{\mathrm{s}}_{0,\mathrm{i}} = \underline{\mathrm{s}}_{0,\mathrm{i}}\left(\underline{\mathrm{s}}_{1,\mathrm{k}}^{[.]}\right) \\ 0 & \text{otherwise} \end{cases}. \tag{21}$$

The target value for the number $\mathrm{n}_{\mathrm{ass}}\left(\underline{\mathrm{s}}_{0,\mathrm{i}}\right)$ is given by the ratio of the sample sizes n_1 and n_0,

$$\mathrm{n}_{\mathrm{ass}}^{\mathrm{target}}\left(\underline{\mathrm{s}}_{0,\mathrm{i}}\right) = \frac{\mathrm{n}_1}{\mathrm{n}_0}. \tag{22}$$

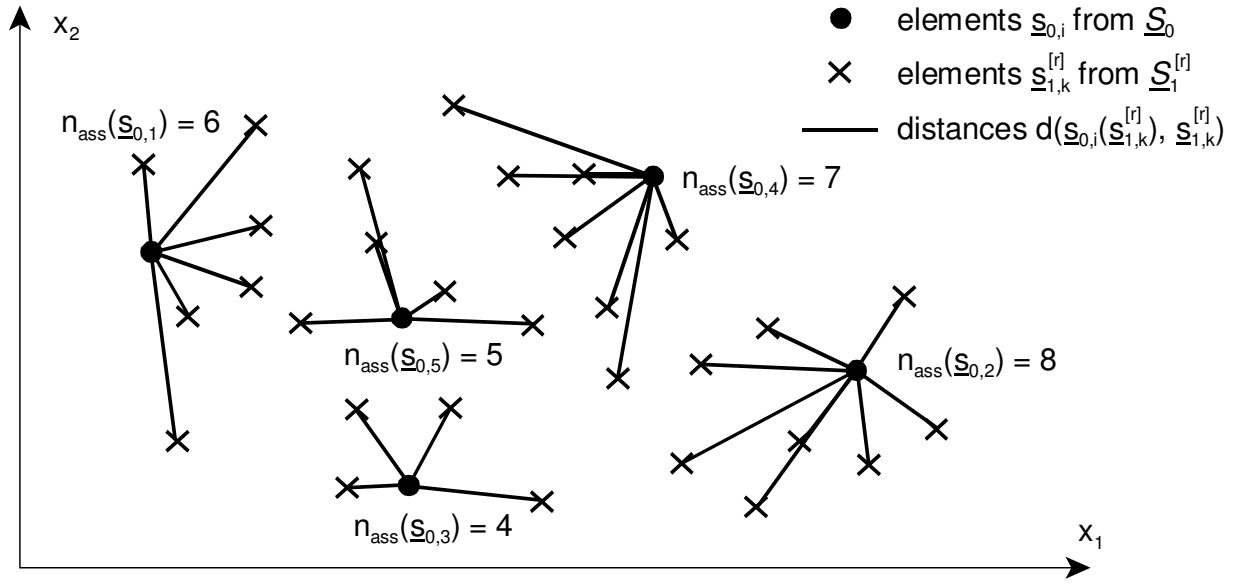

Figure 1. Assignment of sample elements

The assignment criterion is then defined as the total sum of the quadratic differences between the actual numbers $\mathrm{n_{ass}}\left(\underline{\mathrm{s}}_{0,\mathrm{i}}\right)$ and the target value $\mathrm{n_{ass}^{target}}\left(\underline{\mathrm{s}}_{0,\mathrm{i}}\right)$,

$$\mathrm{C_1^{[.]}} = \sum_{\mathrm{i}=1}^{\mathrm{n_0}} \left(\mathrm{n_{ass}}\left(\underline{\mathrm{s}}_{0,\mathrm{i}}\right) - \frac{\mathrm{n_1}}{\mathrm{n_0}} \right)^2 \Rightarrow \mathrm{MIN}\ . \tag{23}$$

The smallest possible value of $\mathrm{C_1^{[.]}}$ depends on the sample sizes $\mathrm{n_1}$ and $\mathrm{n_0}$. With the parameter

$$\mathrm{a} \in \mathbb{N} \mid \mathrm{a} \cdot \mathrm{n_0} \leq \mathrm{n_1} < (\mathrm{a}+1) \cdot \mathrm{n_0} \tag{24}$$

this limit is

$$min\_\mathrm{C_1} = -\frac{1}{\mathrm{n_0}} \left(\mathrm{a} \cdot \mathrm{n_0} - \mathrm{n_1}\right)^2 + \mathrm{n_1} - \mathrm{a} \cdot \mathrm{n_0} \tag{25}$$

In the special case that the size $\mathrm{n_1}$of sample $\underline{S}_1^{[.]}$ is a whole multiple of the size $\mathrm{n_0}$ of $\underline{S}_0$ the value $min\_\mathrm{C_1}$ is equal to zero.

3.3. DISTANCE CRITERION

The distance criterion supplements the assignment criterion by additionally evaluating the particular positions of the sample elements $\underline{\mathrm{s}}_{1,\mathrm{k}}^{[.]}$ and $\underline{\mathrm{s}}_{0,\mathrm{i}}\left(\underline{\mathrm{s}}_{1,\mathrm{k}}^{[.]}\right)$ with respect to each other. The distances between assigned sample elements are supposed to be as small as possible. Specifically,

$$C_2^{[.]} = \sum_{k=1}^{n_1} d\left(\underline{s}_{0,i}\left(\underline{s}_{1,k}^{[.]}\right), \underline{s}_{1,k}^{[.]}\right)^2 \Rightarrow \text{MIN}, \tag{26}$$

with $\underline{s}_{0,i}\left(\underline{s}_{1,k}^{[.]}\right)$ specifying the assignment of $\underline{s}_{1,k}^{[.]}$ to $\underline{s}_{0,i}$ determined with Eq. (19), see Figure 1. The smallest possible value of the distance criterion is zero.

3.4. Composing the dissimilarity measure

To define the dissimilarity measure $G^{[.]}$ for real-valued samples $\underline{S}_1^{[.]}$ and $\underline{S}_0$ the assignment criterion according to Eq. (23) and the distance criterion according to Eq. (26) are combined. As a standard formulation, the quantity

$$G^{[.]} = \sqrt{C_1^{[.]} + C_2^{[.]}} \tag{27}$$

is selected. An extension of Eq. (27) by introducing weighting factors for the criteria $C_1^{[.]}$ and $C_2^{[.]}$ has been investigated in several numerical tests; it has not been found particularly effective for improving the simulation results.

3.5. Assembling the iteration procedure

The dissimilarity measure $G^{[.]}$ in Eq. (27) is implemented into the numerical procedure according to Steps 1 through 4 in Section 2. Moreover, the number m_1 of elements, see Step 3, which are modified in each iteration step, is not held constant during the iteration but varied frequently by a random selection of m_1 from a predefined range of values $[a, b]$ with the aid of a discrete uniform distribution,

$$(X_{m_1} \sim U(a, a+1, ..., b-1, b)) \rightarrow m_1, \; a, b \in \mathbb{N}. \tag{28}$$

As an alternative to the random generation of the m_1 new elements with the aid of a uniform distribution according to Eq. (12), the (slightly smoothed) current empirical distribution $F^{(\bar{e})}_{\underline{S}_1^{[r]}}(\underline{x})$ of the sample $\underline{S}_1^{[r]}$ from the last successful modification r can be used for a kind of bootstrap sampling,

$$\left(\underline{X}_1 \sim F^{(\bar{e})}_{\underline{S}_1^{[r]}}(\underline{x})\right) \rightarrow \underline{S}_1^{[r]+}. \tag{29}$$

In this manner, use is made of the statistical information already gathered in $\underline{S}_1^{[r]}$ during the iteration, which leads to an increase of numerical efficiency. The termination limit in Step 4 is chosen to be 2% and is applied to the moving average of the recent 100 successful iteration steps.

4. Samples of Imprecise Data

4.1. Modeling imprecise data

For dealing with imprecise data, we must select a suitable data model that combines the benefits of the well-established probabilistic approach with an appropriate modeling of non-frequentative uncertainty or

imprecision. From the class of available uncertainty models in this context, the concept of fuzzy random variables originally presented in (Kwakernaak, 1978) is selected for further investigation. This model possesses the advantage of simultaneously covering the models of real-valued random variables, intervals, fuzzy sets, rough sets, random sets, and convex models as special cases.

To define a fuzzy random variable the probability space $[\underline{\mathbf{X}}, \mathfrak{S}, \mathrm{P}]$ is extended by the dimension fuzziness. If the space of the random elementary events, as in probabilistics, is described by Ω, a fuzzy random vector $\underline{\tilde{\mathrm{X}}}$ on the fundamental set $\underline{\mathbf{X}} = \mathbb{R}^{\mathrm{n}}$ may be defined as the fuzzy result of the mapping

$$\Omega \rightarrow \mathbf{F}\left(\mathbb{R}^{\mathrm{n}}\right) \tag{30}$$

where $\mathbf{F}\left(\mathbb{R}^{\mathrm{n}}\right)$ is the set of all fuzzy numbers in $\mathbb{R}^{\mathrm{n}}$. An ordered n-tupel of fuzzy numbers $\tilde{\mathrm{x}}_{\mathrm{i}}$ is assigned to each (crisp) elementary event $\omega \in \Omega$. Every n-tupel $\underline{\tilde{\mathrm{x}}}\left(\omega\right) = \left(\tilde{\mathrm{x}}_1, \ldots, \tilde{\mathrm{x}}_{\mathrm{n}}\right) \subseteq \underline{\mathbf{X}}$ is a realization of the fuzzy random vector $\underline{\tilde{\mathrm{X}}}$. Both objective and subjective information are accounted for simultaneously. The theory of fuzzy random variables permits the modeling of uncertain structural parameters which partly exhibit randomness but which cannot be described using real-valued random variables without an element of doubt. The randomness is "disturbed" by a fuzziness component.

A comprehensive discussion on fuzzy randomness particularly with regard to engineering problems may be found in (Möller and Beer, 2004). In this context the concepts of fuzzy structural analysis, see also (Möller et al., 2000), and fuzzy probabilistic safety assessment, see also (Möller et al., 2003), describe the processing of uncertain structural parameters with the aid of numerical procedures. This basis ensures an appropriate evaluation or further processing of the results from model-free sampling of fuzzy random variables within the framework of structural analysis and safety assessment.

The model-free sampling technique is extended to apply for fuzzy samples by implementing the uncertainty model fuzzy randomness into the basic procedure according to Section 2. Due to the generalized character of this uncertainty model, the capability of processing real-valued samples is hereby preserved as a special case.

4.2. Extension of criteria $\mathrm{C}_1^{[.]}$ and $\mathrm{C}_2^{[.]}$

As the starting point for the extension of the model-free sampling technique to processing imprecise data, these data are described with the aid of Fuzzy Set Theory (Zimmermann, 1992). Each imprecise observation, which represents a sample element as a realization of an underlying fuzzy random vector $\underline{\tilde{\mathrm{X}}}$, is modeled as a normalized fuzzy set or fuzzy vector $\underline{\tilde{\mathrm{s}}} = \underline{\tilde{\mathrm{x}}} \in \mathbf{F}\left(\mathbb{R}^{\mathrm{n}}\right)$ with the membership function $\mu_{\underline{\mathrm{s}}}\left(\underline{\mathrm{s}}\right)$, see Figure 2. The real-valued samples $\underline{S}_0$ and $\underline{S}_1$ from Eqs. (1) and (2) therewith become fuzzy samples,

$$\underline{\tilde{S}}_0 = \left\{\underline{\tilde{\mathrm{s}}}_{0,\mathrm{i}},\ \mathrm{i} = 1,\ \ldots,\ \mathrm{n}_0\right\}, \tag{31}$$

$$\underline{\tilde{S}}_1 = \left\{\underline{\tilde{\mathrm{s}}}_{1,\mathrm{k}},\ \mathrm{k} = 1,\ \ldots,\ \mathrm{n}_1\right\}, \tag{32}$$

with the underlying fuzzy random vector $\underline{\tilde{\mathrm{X}}}_0$ for $\underline{\tilde{S}}_0$.

The processing of the fuzzy samples $\underline{\tilde{S}}_0$ and $\underline{\tilde{S}}_1^{[.]}$ within the procedure according to Steps 1 through 4 in Section 2 requires the extension of the dissimilarity measure $\mathrm{G}^{[.]}$ and thus of the criteria $\mathrm{C}_1^{[.]}$ and $\mathrm{C}_2^{[.]}$ to apply for fuzzy vectors $\underline{\tilde{\mathrm{s}}}_{0,\mathrm{i}}$ and $\underline{\tilde{\mathrm{s}}}_{1,\mathrm{k}}^{[.]}$ as elements of $\underline{\tilde{S}}_0$ and $\underline{\tilde{S}}_1^{[.]}$. As a basis a suitable replacement for the Euclidean

distance $\mathrm{d}\left(\underline{\mathrm{s}}_{0,\mathrm{i}},\, \underline{\mathrm{s}}_{1,\mathrm{k}}^{[.]}\right)$ in Eq. (18) must be introduced as a distance measure between fuzzy vectors $\tilde{\underline{\mathrm{s}}}_{0,\mathrm{i}}$ and $\tilde{\underline{\mathrm{s}}}_{1,\mathrm{k}}^{[.]}$. For this purpose the fuzzy vectors $\tilde{\underline{\mathrm{s}}}_{0,\mathrm{i}}$ and $\tilde{\underline{\mathrm{s}}}_{1,\mathrm{k}}^{[.]}$ are represented with the aid of α-discretization, see Figure 2. For a sufficiently high number of α-levels the fuzzy vectors $\tilde{\underline{\mathrm{s}}}_{0,\mathrm{i}}$ and $\tilde{\underline{\mathrm{s}}}_{1,\mathrm{k}}^{[.]}$ are completely described by the sets of their α-level sets $\underline{\mathrm{s}}_{0,\mathrm{i},\alpha}$ and $\underline{\mathrm{s}}_{1,\mathrm{k},\alpha}^{[.]}$, respectively. Specifically, for real numbers $\alpha \in (0,1]$,

$$\underline{s}_{0,\mathrm{i},\alpha} = \left\{ \underline{\mathrm{s}} \in \mathbb{R}^{\mathbf{n}} \,|\, \mu_{\underline{\mathrm{s}}_{0,\mathrm{i}}}(\underline{\mathrm{s}}) \geq \alpha \right\}, \tag{33}$$

$$\underline{s}_{1,\mathrm{k},\alpha}^{[.]} = \left\{ \underline{\mathrm{s}} \in \mathbb{R}^{\mathbf{n}} \,|\, \mu_{\underline{\mathrm{s}}_{1,\mathrm{k}}^{[.]}}(\underline{\mathrm{s}}) \geq \alpha \right\}, \tag{34}$$

and

$$\tilde{\underline{\mathrm{s}}}_{0,\mathrm{i}} = \left\{ \left(\underline{s}_{0,\mathrm{i},\alpha},\, \mu_{\underline{\mathrm{s}}_{0,\mathrm{i}}}(\underline{s}_{0,\mathrm{i},\alpha}) \right) \,|\, \mu_{\underline{\mathrm{s}}_{0,\mathrm{i}}}(\underline{s}_{0,\mathrm{i},\alpha}) = \alpha \,\forall\, \alpha \in (0,1] \right\}, \tag{35}$$

$$\tilde{\underline{\mathrm{s}}}_{1,\mathrm{k}}^{[.]} = \left\{ \left(\underline{s}_{1,\mathrm{k},\alpha}^{[.]},\, \mu_{\underline{s}_{1,\mathrm{k}}^{[.]}}(\underline{s}_{1,\mathrm{k},\alpha}^{[.]}) \right) \,|\, \mu_{\underline{s}_{1,\mathrm{k}}^{[.]}}(\underline{s}_{1,\mathrm{k},\alpha}^{[.]}) = \alpha \,\forall\, \alpha \in (0,1] \right\}. \tag{36}$$

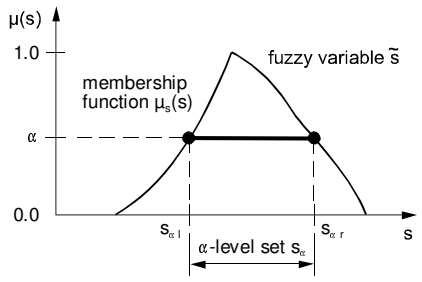

Figure 2. α-discretization of a fuzzy variable

On this basis, the distance $\mathrm{d_F}\left(\tilde{\underline{\mathrm{s}}}_{0,\mathrm{i}},\, \tilde{\underline{\mathrm{s}}}_{1,\mathrm{k}}^{[.]}\right)$ between the fuzzy vectors $\tilde{\underline{\mathrm{s}}}_{0,\mathrm{i}}$ and $\tilde{\underline{\mathrm{s}}}_{1,\mathrm{k}}^{[.]}$ may be defined by recombining the distances $\mathrm{d_H}\left(\underline{\mathrm{s}}_{0,\mathrm{i},\alpha},\, \underline{\mathrm{s}}_{1,\mathrm{k},\alpha}^{[.]}\right)$ between the associated α-level sets $\underline{\mathrm{s}}_{0,\mathrm{i},\alpha}$ and $\underline{\mathrm{s}}_{1,\mathrm{k},\alpha}^{[.]}$ (for the same α-level). Specifically, the metric

$$\mathrm{d_F}\left(\tilde{\underline{\mathrm{s}}}_{0,\mathrm{i}},\, \tilde{\underline{\mathrm{s}}}_{1,\mathrm{k}}^{[.]}\right) = \int_{\alpha=+0}^{\alpha=1} \mathrm{d_H}\left(\underline{\mathrm{s}}_{0,\mathrm{i},\alpha},\, \underline{\mathrm{s}}_{1,\mathrm{k},\alpha}^{[.]}\right) \mathrm{d}\alpha \tag{37}$$

is applied, see (Körner, 1997), which makes use of the Hausdorff metric

$$\mathrm{d_H}\left(\underline{\mathrm{s}}_{0,\mathrm{i},\alpha},\, \underline{\mathrm{s}}_{1,\mathrm{k},\alpha}^{[.]}\right) \;=\; \max\left[\mathrm{d}_{\mathrm{H1,i,k}}^{[.]},\, \mathrm{d}_{\mathrm{H2,i,k}}^{[.]}\right],$$

$$\mathrm{d}_{\mathrm{H1,i,k}}^{[.]}\left(\underline{\mathrm{s}}_{0,\mathrm{i},\alpha},\, \underline{\mathrm{s}}_{1,\mathrm{k},\alpha}^{[.]}\right) = \sup_{\underline{\mathrm{s}}_0 \in \underline{\mathrm{s}}_{0,\mathrm{i},\alpha}} \inf_{\underline{\mathrm{s}}_1 \in \underline{\mathrm{s}}_{1,\mathrm{k},\alpha}^{[.]}} \left[\mathrm{d}\left(\underline{\mathrm{s}}_0,\, \underline{\mathrm{s}}_1\right)\right], \tag{38}$$

$$\mathrm{d}_{\mathrm{H2,i,k}}^{[.]}\left(\underline{\mathrm{s}}_{0,\mathrm{i},\alpha},\, \underline{\mathrm{s}}_{1,\mathrm{k},\alpha}^{[.]}\right) = \sup_{\underline{\mathrm{s}}_1 \in \underline{\mathrm{s}}_{1,\mathrm{k},\alpha}^{[.]}} \inf_{\underline{\mathrm{s}}_0 \in \underline{\mathrm{s}}_{0,\mathrm{i},\alpha}} \left[\mathrm{d}\left(\underline{\mathrm{s}}_0,\, \underline{\mathrm{s}}_1\right)\right],$$

between the associated α-level sets $\underline{\mathrm{s}}_{0,\mathrm{i},\alpha}$ and $\underline{\mathrm{s}}_{1,\mathrm{k},\alpha}^{[.]}$ with $\mathrm{d}\left(\underline{\mathrm{s}}_0,\, \underline{\mathrm{s}}_1\right)$ being the Euclidean distance between crisp elements $\underline{\mathrm{s}}_0$ and $\underline{\mathrm{s}}_1$ from $\underline{\mathrm{s}}_{0,\mathrm{i},\alpha}$ and $\underline{\mathrm{s}}_{1,\mathrm{k},\alpha}^{[.]}$, respectively, see Figure 3. The outcome of Eq. (38) and hence the distance $\mathrm{d}_{\mathrm{F}}\left(\tilde{\underline{\mathrm{s}}}_{0,\mathrm{i}},\, \tilde{\underline{\mathrm{s}}}_{1,\mathrm{k}}^{[.]}\right)$ from Eq. (37) are crisp values, which can be directly applied in Eqs. (19) and (26) to eventually compute criteria $\mathrm{C}_1^{[.]}$ and $\mathrm{C}_2^{[.]}$.

The application of criteria $\mathrm{C}_1^{[.]}$ and $\mathrm{C}_2^{[.]}$ to evaluate the dissimilarity of fuzzy-valued samples enables a consideration of the order in the element configuration and the distance between the respective sample elements. Dissimilarities in the fuzziness of the elements $\tilde{\underline{\mathrm{s}}}_{0,\mathrm{i}}$ and $\tilde{\underline{\mathrm{s}}}_{1,\mathrm{k}}^{[.]}$, however, are taken into account only to a partial degree. In addition to the criteria $\mathrm{C}_1^{[.]}$ and $\mathrm{C}_2^{[.]}$, the fuzziness of the realizations provides a basis for a third dissimilarity criterion.

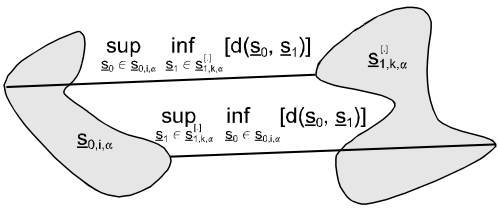

Figure 3. Hausdorff metric applied to α-level sets

4.3. FUZZINESS CRITERION

The fuzziness criterion evaluates the matching in the fuzziness of the respective fuzzy sample elements $\tilde{\underline{\mathrm{s}}}_{0,\mathrm{i}}$ and $\tilde{\underline{\mathrm{s}}}_{1,\mathrm{k}}^{[.]}$. Fuzzy sample elements that are assigned to each other according to the assignment rule Eq (19) are supposed to exhibit the same fuzziness. For this purpose, the fuzziness of the sample elements is computed with an analog to Shannon's entropy applied to the membership functions $\mu(\underline{\mathrm{s}}_{0,\mathrm{i}}) = \mu_{\underline{\mathrm{s}}_{0,\mathrm{i}}}(\underline{\mathrm{s}})$ and $\mu(\underline{\mathrm{s}}_{1,\mathrm{k}}^{[.]}) = \mu_{\underline{\mathrm{s}}_{1,\mathrm{k}}^{[.]}}(\underline{\mathrm{s}})$ of $\tilde{\underline{\mathrm{s}}}_{0,\mathrm{i}}$ and $\tilde{\underline{\mathrm{s}}}_{1,\mathrm{k}}^{[.]}$, respectively. For the fuzzy vector $\tilde{\underline{\mathrm{s}}}$, this uncertainty measure is defined as

$$\mathrm{H_u} = -\mathrm{k} \cdot \int_{\underline{\mathrm{s}}} \mathrm{g}\left(\mu\left(\underline{\mathrm{s}}\right)\right)\, \mathrm{d}\underline{\mathrm{s}},$$

$$\mathrm{g}\left(\mu\left(\underline{\mathrm{s}}\right)\right) = \mu\left(\underline{\mathrm{s}}\right) \cdot \ln\left(\mu\left(\underline{\mathrm{s}}\right)\right) + \left(1 - \mu\left(\underline{\mathrm{s}}\right)\right) \cdot \ln\left(1 - \mu\left(\underline{\mathrm{s}}\right)\right). \tag{39}$$

And the fuzziness criterion is

$$\mathrm{C}_3^{[.]} = \sum_{\mathrm{k}=1}^{\mathrm{n_1}} \left(\mathrm{H_u}\left(\tilde{\underline{\mathrm{s}}}_{0,\mathrm{i}}\left(\tilde{\underline{\mathrm{s}}}_{1,\mathrm{k}}^{[.]}\right)\right) - \mathrm{H_u}\left(\tilde{\underline{\mathrm{s}}}_{1,\mathrm{k}}^{[.]}\right)\right)^2 \Rightarrow \mathrm{MIN}. \tag{40}$$

For a "perfect matching", the fuzziness criterion $C_3^{[.]}$ becomes zero.

4.4. PROCEDURE FEATURES FOR IMPRECISE DATA

The generation and the iterative modification of a fuzzy sample $\underline{\tilde{S}}_1^{[.]}$ require not only determining the position of the sample elements $\underline{\tilde{s}}_{1,k}^{[.]}$ but also specifying their membership functions $\mu(\underline{\tilde{s}}_{1,k}^{[.]})$. New fuzzy realizations $\underline{\tilde{s}}_{1,k}^{[.]}$ are generated in the following three steps.

1. The mean values
$$\underline{s}_{1,k,\mu=1}^{[.]} = \underline{s}_{1,k}^{[.]} \in \underline{\tilde{s}}_{1,k}^{[.]} \mid \mu(\underline{s}_{1,k}^{[.]}) = 1 \tag{41}$$
of the fuzzy sample elements $\underline{\tilde{s}}_{1,k}^{[.]}$ (different from the definition of a statistical mean value, see (Zimmermann, 1992)) are specified analogous to the crisp sample elements $\underline{s}_{1,k}^{[.]}$ of the $\underline{S}_1^{[.]}$ in Eqs. (5), (12), and (29). That is, the initialization is realized with
$$(\underline{X}_1 \sim U(\underline{D})) \rightarrow \left\{\underline{s}_{1,k,\mu=1}^{[0]},\ k = 1,\ ...,\ n_1\right\}, \tag{42}$$
and during the iteration
$$(\underline{X}_1 \sim U(\underline{D})) \rightarrow \left\{\underline{s}_{1,k,\mu=1}^{[r]+},\ k = k_1,\ ...,\ k_{m_1}\right\} \tag{43}$$
or, alternatively,
$$\left(\underline{X}_1 \sim F^{(\bar{e})}_{\underline{S}_{1,\mu=1}^{[r]}}(\underline{x})\right) \rightarrow \left\{\underline{s}_{1,k,\mu=1}^{[r]+},\ k = k_1,\ ...,\ k_{m_1}\right\} \tag{44}$$
are applied, in which $F^{(\bar{e})}_{\underline{S}_{1,\mu=1}^{[r]}}(\underline{x})$ represents the smoothed empirical distribution of the mean values $\underline{s}_{1,k,\mu=1}^{[r]}$ in the fuzzy sample $\underline{\tilde{S}}_1^{[r]}$ in iteration step r.

2. The fuzziness $H_u\left(\underline{\tilde{s}}_{1,k}^{[.]}\right)$ is determined by means of a logarithmic normal distribution $F^{(log)}(H_u)$ estimated from the fuzziness $H_u\left(\underline{\tilde{s}}_{0,i}\right)$ of the fuzzy sample elements $\underline{\tilde{s}}_{0,i}$ in the observed fuzzy sample $\underline{\tilde{S}}_0$,
$$\left\{H_u\left(\underline{\tilde{s}}_{0,i}\right),\ i = 1,\ ...,\ n_0\right\} \rightarrow F^{(log)}(H_u), \tag{45}$$
$$\left(X_H \sim F^{(log)}(H_u)\right) \rightarrow \left\{H_u\left(\underline{\tilde{s}}_{1,k}^{[0]}\right),\ k = 1,\ ...,\ n_1\right\}, \tag{46}$$
$$\left(X_H \sim F^{(log)}(H_u)\right) \rightarrow \left\{H_u\left(\underline{\tilde{s}}_{1,k}^{[r]+}\right),\ k = k_1,\ ...,\ k_{m_1}\right\}. \tag{47}$$

3. The shape of the membership function $\mu(\underline{\tilde{s}}_{1,k}^{[.]})$ is also randomly specified according to the "empirical distribution" of the shape of $\mu(\underline{s}_{0,i})$ in $\underline{\tilde{S}}_0$. This is realized with the aid of a parametric representation of

the fuzzy sample elements $\tilde{\underline{s}}_{0,i}$ and $\tilde{\underline{s}}^{[.]}_{1,k}$ in a zero-mean (in fuzzy terminology, see (Zimmermann, 1992)) form normalized to a unit maximum spread t_r. In general terms,

$$\tilde{\underline{t}} = \tilde{\underline{s}} - \underline{s}_{\mu=1}, \tag{48}$$

$$\tilde{\underline{t}}^{(n)} = \frac{1}{t_r} \cdot \tilde{\underline{t}},\ t_r = \max_{\underline{t}|\mu(\underline{t})>0} \|\underline{t}\|, \tag{49}$$

$$\mu\left(\underline{t}^{(n)}\right) = \mu\left(p_1, ..., p_{n_p}\right). \tag{50}$$

The "empirical distribution" of the shape is then represented by the smoothed joint empirical distribution of the parameters $p_1, ..., p_{n_p}$,

$$\tilde{\underline{t}}_{0,i} = \tilde{\underline{s}}_{0,i} - \underline{s}_{0,i,\mu=1},\ i = 1, ..., n_0, \tag{51}$$

$$\tilde{\underline{t}}^{(n)}_{0,i} = \frac{1}{t_{r,0,i}} \cdot \tilde{\underline{t}}_{0,i},\ t_{r,0,i} = \max_{\underline{t}|\mu\left(\underline{t}_{0,i}\right)>0} \|\underline{t}\|,\ i = 1, ..., n_0, \tag{52}$$

$$\left\{\mu\left(\underline{t}^{(n)}_{0,i}\right),\ i = 1, ..., n_0\right\} \rightarrow \left(F^{(\bar{e})}\left(p_1, ..., p_{n_p}\right)\right). \tag{53}$$

The random shape of new elements $\tilde{\underline{s}}^{[.]}_{1,k}$ is determined according to

$$\left(\underline{X}_p \sim F^{(\bar{e})}\left(p_1, ..., p_{n_p}\right)\right) \rightarrow \left\{\mu\left(\underline{t}^{(n)\,[0]}_{1,k}\right),\ k = 1, ..., n_1\right\} \tag{54}$$

and

$$\left(\underline{X}_p \sim F^{(\bar{e})}\left(p_1, ..., p_{n_p}\right)\right) \rightarrow \left\{\mu\left(\underline{t}^{(n)\,[r]^+}_{1,k}\right),\ k = k_1, ..., k_{m_1}\right\}, \tag{55}$$

respectively. The obtained new normalized fuzzy elements $\underline{t}^{(n)\,[0]}_{1,k}$ and $\underline{t}^{(n)\,[r]^+}_{1,k}$ are backtransformed inverse to Eqs. (48) and (49),

$$\tilde{\underline{s}}^{[0]}_{1,k} = \underline{t}^{(n)\,[0]}_{1,k} \cdot t'_{r,k} + \underline{s}^{[0]}_{1,k,\mu=1},\ k = 1, ..., n_1, \tag{56}$$

$$\tilde{\underline{s}}^{[r]^+}_{1,k} = \underline{t}^{(n)\,[r]^+}_{1,k} \cdot t'_{r,k} + \underline{s}^{[r]^+}_{1,k,\mu=1},\ k = k_1, ..., k_{m_1}, \tag{57}$$

with $\underline{s}^{[0]}_{1,k,\mu=1}$ and $\underline{s}^{[r]^+}_{1,k,\mu=1}$ from Eqs. (42) and (43) or (44). The spread factors $t'_{r,k}$ are obtained implicitely by the fuzziness $H_u\left(\tilde{\underline{s}}^{[0]}_{1,k}\right)$ and $H_u\left(\tilde{\underline{s}}^{[r]^+}_{1,k}\right)$, respectively, specified according to Eqs. (46) and (47).

The consideration of fuzzy samples requires incorporating the criterion $C^{[.]}_3$ into the iterative procedure. Tests have shown that it is effective to perform the iteration for fuzzy samples in two parts. In the first part, only the criteria $C^{[.]}_1$ and $C^{[.]}_2$ are satisfied. Subsequently, the obtained element assignment and the mean value positions are frozen. In the second part, criterion $C^{[.]}_3$ is applied in a separate fuzziness iteration. That is, in the second iteration part, only the $H_u\left(\tilde{\underline{s}}^{[.]}_{1,k}\right)$ and the shape of the membership functions of the fuzzy sample elements $\tilde{\underline{s}}^{[.]}_{1,k}$ are adjusted. The iteration termination criterion is also applied separately in both iteration parts.

5. Application to Structural Engineering Problems

5.1. General application schemes

The general concept of model-free sampling provides a beneficial basis for a coupling with structural engineering computations in various ways. Generally, the following application schemes can be pursued – separately or combined.

5.1.1. *Processing of input data*

Model-free sampling can be applied to crisp or imprecise input data of structural computations, in particular, if the available data are rare, and a probabilistic model is not known to a sufficient degree of confidence. The sampling result $\underline{S}_1{}^{\mathrm{input}}$ or $\underline{\tilde{S}}_1{}^{\mathrm{input}}$ then reflects the stochastic or fuzzy stochastic properties of the input data in the form of a numerically generated data set of crisp or imprecise input vectors for further processing in structural computations. This is equivalent to the result of a Monte Carlo simulation based on a known probabilistic model for the input quantities. The coupling to structural computations can be realized as follows.

- The sampling result $\underline{S}_1^{\mathrm{input}}$ or $\underline{\tilde{S}}_1{}^{\mathrm{input}}$ can be directly used for a subsequent stochastic or fuzzy stochastic structural analysis to compute stochastic or fuzzy stochastic structural responses. It represents the input sample, which contains n_1 input vectors $\underline{\mathrm{s}}_{1,\mathrm{k}}^{\mathrm{input}}$ or $\underline{\tilde{\mathrm{s}}}_{1,\mathrm{k}}^{\mathrm{input}}$ for a subsequent n_1-fold structural analysis. In the case of samples comprising imprecise data, the generalized uncertainty processing algorithms presented in (Möller and Beer, 2004) may be applied. As results fuzzy probabilistic structural responses are obtained, which are characterized by imprecise probability distributions with fuzzy parameters such as a fuzzy mean and a fuzzy variation.

- As a postprocessing attached to a stochastic or fuzzy stochastic structural analysis based on the generated sample $\underline{S}_1^{\mathrm{input}}$ or $\underline{\tilde{S}}_1{}^{\mathrm{input}}$, a safety assessment can be carried out by evaluating limit states in the result space, which may be advantageous if limit state surfaces cannot be specified in the input space for some reason. This is simply realized by counting those sample elements in the stochastic or fuzzy stochastic structural responses, which lead to failure according to the defined limit states. The result is obtained as a failure probability or a fuzzy failure probability. As a difference to traditional methods, this procedure can involve imprecise sample elements, which is explained in Section 5.2.

- In contrast to the latter, a safety assessment can also be performed by evaluating limit state surfaces in the input space. This is particularly useful if the underlying structural analysis to produce structural responses is computationally expensive, and the limit state surfaces can be described in the space of the structural input parameters, for example, within the framework of a response surface method. Again, counting of the elements in the failure domain – with an evaluation of imprecise data according to Section 5.2 – yields a failure probability or a fuzzy failure probability.

5.1.2. *Processing of result data*

The model-free sampling technique can also be used for processing result data from structural computations. This can be instrumental if the set of structural response data is limited and cannot be described with a

probabilistic model on a satisfying confidence level. Also, in the case of imprecise measurements of structural responses this method may be helpful. The stochastic or fuzzy stochastic properties of the structural responses are then described with the aid of the numerical sampling result $\underline{S}_1{}^{\mathrm{result}}$ or $\underline{\tilde{S}}_1{}^{\mathrm{result}}$ of a sufficiently large size n_1 – equivalent to the outcome from a Monte Carlo simulation with a sufficiently high number n_1 of structural analyses. As a prerequisite for obtaining reliable results in this manner, the underlying sample of structural responses must comprise essential information about the properties of the computational model. That is, physical, mechanical, or chemical phenomena that are effective in the underlying structural analysis must be already reflected in the sample of structural responses for being reproduced in a subsequent model-free sampling and thus in the final result result $\underline{S}_1{}^{\mathrm{result}}$ or $\underline{\tilde{S}}_1{}^{\mathrm{result}}$. In correspondence with Section 5.1.1, the following two approaches can be pursued for a coupling to structural computations.

- The uncertain structural responses from stochastic or fuzzy stochastic structural analysis can be introduced into model-free sampling to obtain a sufficiently large sample size n_1 for describing crisp or imprecise probability distributions of the responses empirically instead of performing a weak and ambiguous distribution estimation.
- For a safety assessment, limit states in the result space can be directly evaluated with the aid of the sampling result $\underline{S}_1{}^{\mathrm{result}}$ or $\underline{\tilde{S}}_1{}^{\mathrm{result}}$. For the technique of counting fuzzy sample elements in the failure domain, see Section 5.2.

5.2. Reliability assessment for imprecise data

The application of model-free sampling may be particularly useful in reliability assessment as the available data do usually not cover failure domains. It is thus of great interest in this application field to reproduce the tails of the underlying probability distributions to obtain reliable estimations of failure probabilities.

Structural reliability assessment based on model-free sampling is realized as a straightforward extension to traditional methods. The sampling result $\underline{S}_1$ or $\underline{\tilde{S}}_1$ is directly evaluated with regards to the limit states either in the input space or in the result space, see Section 5.1. That is, the structural reliability is determined by counting the sample elements that lead to failure. For dealing with imprecise data, however, this counting needs to be extended in an appropriate manner, see (Möller and Beer, 2004). Due to their fuzziness, some fuzzy sample elements $\underline{\tilde{\mathrm{s}}}_{1,\mathrm{k}}$ lie only partly in the failure domain $\underline{S}_{\mathrm{f}}$, or, in the case of an underlying computational model that involves model uncertainty as fuzziness (Möller et al., 2003), in the fuzzy failure domain $\underline{\tilde{S}}_{\mathrm{f}}$. This leads to a fuzzy failure probability $\tilde{\mathrm{P}}_{\mathrm{f}}$. For computing $\tilde{\mathrm{P}}_{\mathrm{f}}$ α-discretization is applied again, see Section 4.2. Specifically,

$$\begin{aligned} \tilde{\mathrm{P}}_{\mathrm{f}} &= \{(\mathrm{P}_{\mathrm{f},\alpha},\, \mu\,(\mathrm{P}_{\mathrm{f},\alpha}))\}\,, \\ \mathrm{P}_{\mathrm{f},\alpha} &= [\mathrm{P}_{\mathrm{f},\alpha\,\mathrm{l}},\, \mathrm{P}_{\mathrm{f},\alpha\,\mathrm{r}}]\,, \\ \mu\,(\mathrm{P}_{\mathrm{f},\alpha}) &= \alpha\,\forall\,\alpha \in (0,1]. \end{aligned} \tag{58}$$

The interval bounds $\mathrm{P}_{\mathrm{f},\alpha\,\mathrm{l}}$ and $\mathrm{P}_{\mathrm{f},\alpha\,\mathrm{r}}$ (see Figure 2 for general illustration) are calculated with the aid of indicator functions and particular conditions for evaluating fuzzy realizations, see (Möller and Beer, 2004). Specifically,

$$\mathrm{P}_{\mathrm{f},\alpha\,\mathrm{l}} = \tfrac{1}{\mathrm{n}_1} \cdot \textstyle\sum_{\mathrm{k}=1}^{\mathrm{n}_1} \mathrm{I}_{\alpha\,\mathrm{l}}\left(\underline{\tilde{\mathrm{s}}}_{1,\mathrm{k}}\right),$$

$$\mathrm{I}_{\alpha\,\mathrm{l}}\left(\underline{\tilde{\mathrm{s}}}_{1,\mathrm{k}}\right) = \begin{cases} 1 \text{ if } \underline{\mathrm{s}}_{1,\mathrm{k},\alpha} \subseteq \underline{S}_{\mathrm{f},\alpha} \\ 0 \text{ otherwise} \end{cases}, \tag{59}$$

and

$$\begin{aligned} \mathrm{P}_{\mathrm{f},\alpha\,\mathrm{r}} = & \quad \tfrac{1}{\mathrm{n}_1} \cdot \textstyle\sum_{\mathrm{k}=1}^{\mathrm{n}_1} \mathrm{I}_{\alpha\,\mathrm{r}}\left(\underline{\tilde{\mathrm{s}}}_{1,\mathrm{k}}\right), \\ \mathrm{I}_{\alpha\,\mathrm{r}}\left(\underline{\tilde{\mathrm{s}}}_{1,\mathrm{k}}\right) = & \begin{cases} 1 \text{ if } \underline{\mathrm{s}}_{1,\mathrm{k},\alpha} \cap \underline{S}_{\mathrm{f},\alpha} \neq \emptyset \\ 0 \text{ otherwise} \end{cases}. \end{aligned} \tag{60}$$

6. Examples

6.1. REAL-VALUED DATA

6.1.1. *Sampling*

A one-dimensional real-valued sample S_0 of size $\mathrm{n}_0 = 200$ is taken as the basis, see Figure 4. This is numerically generated from a compound distribution consisting of two extreme value distributions of Ex-Max type I. The extreme values of the sample S_0 are $min\_\mathrm{s}_0 = 5.1$ and $max\_\mathrm{s}_0 = 21.55$.

An initial estimate $S_1^{[0]}$ is numerically generated according to Eq. (5) by uniformly distributing $\mathrm{n}_1 = 10,000$ sample elements $\mathrm{s}_{1,\mathrm{k}}$ over the interval $D = [0,\, 25]$, see Figure 4. Then, the iteration Eq. (15) to improve the generalized estimation $S_1^{[0]}$ is started. The number m_1 of modified elements is randomly selected from the interval $[\mathrm{a},\, \mathrm{b}] = [5,\, 30]$, see Eq. (28), and frequently changed during the iteration. For generating the new elements $\mathrm{s}_{1,\mathrm{k}}^{[.]+}$ the bootstrap-like method of Eq. (29) is applied. After about $\mathrm{r} = 4,000$ iteration steps the average success rate starts decreasing distinctly and attains the termination limit in iteration step $\mathrm{r} = 4,710$, see Figure 4.

Clearly, there is no visible difference between the empirical distribution functions of the samples S_0 and $S_1 = S_1^{[4,710]}$. Homogeneity tests (Kolmogorov-Smirnov, Mann-Whitney, and chi-squared) yield rejection probabilities of P < 0.012 for the H_0-hypothesis that both samples originate from the same population. The tails of the generated sample S_1 run beyond the extreme values of S_0 with $min\_\mathrm{s}_1 = 3.26$ and $max\_\mathrm{s}_1 = 24.01$. A total of 39 elements $\mathrm{s}_{1,\mathrm{k}}$ are smaller than $min\_\mathrm{s}_0 = 5.1$ and, and 48 elements $\mathrm{s}_{1,\mathrm{k}}$ are bigger than $max\_\mathrm{s}_0 = 21.55$. The proportions of S_1 therewith correspond to an extreme value distribution with a thicker tail on the right side than on the left side. Fisher's exact probability test yields a probability of $\mathrm{P} = 0.386$ with which the H_0-hypothesis is not rejected. Further, the sampling result S_1 shows no clumping of the generated sample elements $\mathrm{s}_{1,\mathrm{k}}$ around the original sample elements $\mathrm{s}_{0,\mathrm{i}}$, which has been verified by investigating the distribution of the elements $\mathrm{s}_{1,\mathrm{k}}$ within the "gaps" between the original elements $\mathrm{s}_{0,\mathrm{i}}$.

Results generated via traditionally estimated probability distributions did not attain the quality level of the present sample S_1. Kernel-based estimation methods led to samples showing test results comparable to the present approach. Their tails, however, did not run significantly beyond $min\_\mathrm{s}_0 = 5.1$ and $max\_\mathrm{s}_0 = 21.55$ and were pre-determined in their form depending on the (subjectively) selected kernels. The same applies to even generalized bootstrap methods. In contrast to that, the tails of S_1 from model-free sampling are not influenced by subjectivity and obtained in a form with orientation to the structure of the underlying sample S_0, which possesses significant importance in reliability assessment.

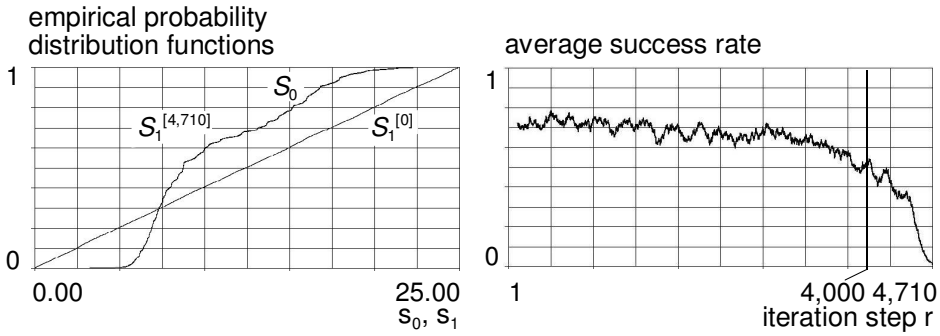

Figure 4. Empirical distribution functions of S_0, $S_1^{[0]}$, and $S_1^{[4,710]}$; average success rate (last 100 steps) during iteration

6.1.2. *Reliability assessment*

The reliability assessment is pursued by directly evaluating the sampling result S_1 with respect to a given limit state surface. Since the related procedures are well-known, these are not highlighted in the example. Herein, it is focused on the dependency of the assessment result on the quality of the sampling result.

For demonstration, the observed sample S_0 is interpreted as a possible record of a live load s resulting from road traffic and acting on a structural member of a road bridge. The sampling result S_1 then represents a statistical loading prognosis for future traffic. For defining a limit state surface, the serviceability requirement s = 22 is defined.

The empirical failure probability obtained from sample S_0 is $\mathrm{P_f} = 0$, whereas the sampling result S_1 yields $\mathrm{P_f} = 3.4 \cdot 10^{-3}$. A compound probability distribution estimated from S_0 without additional prior knowledge leads to $\mathrm{P_f} = 1.7 \cdot 10^{-3}$. According to the underlying extreme value distribution $\mathrm{P_f} = 8.9 \cdot 10^{-3}$ is obtained. These results indicate a good agreement between the prognoses from traditional approaches and from model-free sampling.

6.2. IMPRECISE DATA

6.2.1. *Sampling*

As a starting point the sample S_0 from Section 6.1 is "fuzzified" to represent an uncertain measurement series, for example, of a live load, see Section 6.1.2. That is, the underlying bimodal distribution from Section 6.1.1 is retained for the mean values $\mathrm{s}_{0,\mathrm{i},\mu=1}$. The resulting fuzzy sample $\tilde{S}_0$ consists of $\mathrm{n}_0 = 200$ fuzzy triangular numbers with fluctuating fuzziness $\mathrm{H_u}\left(\tilde{s}_{0,\mathrm{i}}\right)$ over the sample elements $\tilde{s}_{0,\mathrm{i}}$; for relevant concepts and terminology see (Bandemer and Näther, 1992) and (Möller and Beer, 2004). An initial estimate $\tilde{S}_1^{[0]}$ of size $\mathrm{n}_1 = 10,000$ is generated in compliance with Section 4.4 starting from uniformly distributed mean values $\mathrm{s}_{1,\mathrm{k},\mu=1}^{[0]}$ and restricting the fuzzy sample elements completely to $\tilde{\mathrm{s}}_{1,\mathrm{k}}^{[0]} \subseteq [0,\ 25]$, see Figure 5. Again, the iteration in carried out with a randomly selected number $\mathrm{m}_1 \in [5,\ 30]$ of modified elements. First, the dissimilarity measure $\mathrm{G}^{[.]}\left(\mathrm{C}_1^{[.]},\ \mathrm{C}_2^{[.]}\right)$, see Eq. (27) with the extension from Section 4.2, is minimized in 5,990 iteration steps. The empirical fuzzy probability distributions of $\tilde{S}_0$ and $\tilde{S}_1^{[5,990]}$ agree very well. However, there is almost no correspondence between the fuzziness $\mathrm{H_u}\left(\tilde{s}_{0,\mathrm{i}}\right)$ and $\mathrm{H_u}\left(\tilde{\mathrm{s}}_{1,\mathrm{k}}^{[5,990]}\right)$ of the respective

fuzzy sample elements, see Figure 5. The subsequent fuzziness iteration (minimization of criterion $C_3^{[.]}$ up to iteration step r = 16,150) almost does not affect the empirical distribution, but improves considerably the fuzziness agreement, see Figure 5.

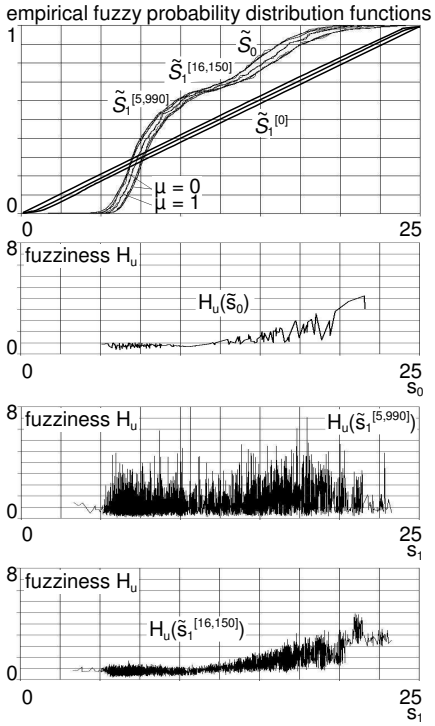

Figure 5. Empirical fuzzy probability distribution functions of $\tilde{S}_0$, $\tilde{S}_1^{[0]}$, $\tilde{S}_1^{[5,990]}$, and $\tilde{S}_1^{[16,150]}$; fuzziness H_u of the associated fuzzy sample elements

6.2.2. *Reliability assessment*

The serviceability requirement s = 22 specified in Section 6.1.2 is evaluated with the fuzzy samples $\tilde{S}_0$ and $\tilde{S}_1 = \tilde{S}_1^{[16,150]}$. The fuzzy failure probability $\tilde{P}_f$ is computed according to Eqs. (58), (59), and (60) with eleven α-levels, see Figure 6. Whereas sample $\tilde{S}_0$ yields an almost useless result with an overestimated fuzziness, sample $\tilde{S}_1$ leads to a more meaningful result. The probability values covered by $\tilde{P}_f$ from $\tilde{S}_1$ again comprise a reasonable range with respect to the results from traditional estimations and from the underlying distribution for the mean values $s_{0,i,\mu=1}$ presented in Sect. 6.1.2.

7. Conclusions

The presented model-free sampling technique may be useful if the data bank comprises, solely, a small sample with uncertain or imprecise elements. It operates free of a probability model, is capable of considering

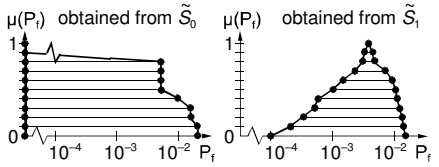

Figure 6. Empirical fuzzy failure probability obtained from $\tilde{S}_0$ and from $\tilde{S}_1$

randomness and non-stochastic uncertainty simultaneously, and can be attached to engineering computations that involve uncertainty in various schemes.

Beyond the demonstrated capabilities in the one-dimensional case, promising experiences have already been made in processing vector valued data including nonlinear stochastic dependencies. A further consideration of multidimensional problems for fuzzy valued data is pursued.

Acknowledgements

The author gratefully acknowledges the support of the German Research Foundation (DFG) and of the Alexander von Humboldt-Foundation (AvH).

References

Bandemer, H. and W. Näther: 1992, *Fuzzy Data Analysis*. Dordrecht: Kluwer Academic Publishers.

Berleant, D. and J. Zhang: 2004, 'Representation and problem solving with Distribution Envelope Determination (DEnv)'. *Reliability Engineering & System Safety* **85**(1–3), 153–168.

Bernardo, J. and A. Smith: 1994, *Bayesian Theory*. Chichester New York Brisbane Toronto Singapore: Wiley.

Colubi, A., C. Fernández-García, and M. A. Gil: 2002, 'Simulation of random fuzzy variables: an empirical approach to statistical/probabilistic studies with fuzzy experimental data'. *IEEE Transactions on Fuzzy Systems* **10**, 384–390.

Deodatis, G., L. Graham-Brady, and R. Micaletti: 2003, 'A hierarchy of upper bounds on the response of stochastic systems with large variation of their properties: random field case'. *Probabilistic Engineering Mechanics* **18**(4), 365–375.

Du, S., B. R. Ellingwood, and J. V. Cox: 2005, 'Initialization strategies in simulation-based SFE eigenvalue analysis'. *Computer-aided Civil and Infrastructure Engineering* **20**(5), 304–315.

Fellin, W., H. Lessmann, M. Oberguggenberger, and R. Vieider (eds.): 2005, 'Analyzing Uncertainty in Civil Engineering'. Berlin Heidelberg New York: Springer.

Fetz, T. and M. Oberguggenberger: 2004, 'Propagation of uncertainty through multivariate functions in the framework of sets of probability measures'. *Reliability Engineering & System Safety* **85**(1-3), 73–87.

Field Jr., R. V. and M. Grigoriu: 2004, 'On the accuracy of the polynomial chaos approximation'. *Probabilistic Engineering Mechanics* **19**(1–2), 65–80.

Ghanem, R. and P. Spanos: 1991, *Stochastic Finite Elements: A Spectral Approach*. New York Berlin Heidelberg: Springer. Revised Edition 2003, Dover Publications, INC., Mineola, New York.

Gutierrez, E. and J. M. Zaldivar: 2000, 'The application of Karhunen-Loeve, or principle component analysis method, to study the non-linear seismic response of structures'. *Earthquake Engineering & Structural Dynamics* **29**(9), 1261–1286.

Hall, J. W. and J. Lawry: 2004, 'Generation, combination and extension of random set approximations to coherent lower and upper probabilities'. *Reliability Engineering & System Safety* **85**(1–3), 89–101.

Helton, J. C. and W. L. Oberkampf: 2004, 'Special Issue on Alternative Representations of Epistemic Uncertainty'. *Reliability Engineering & System Safety* **85**(1–3), 1–369.

Kwakernaak, H.: 1978, 'Fuzzy random variables – I. Definitions and Theorems'. *Information Sciences* **15**, 1–19.

Körner, R.: 1997, 'Linear Models with Random Fuzzy Variables'. Ph.D. thesis, Bergakademie Freiberg, Fakultät für Mathematik und Informatik.

Mood, A., F. Graybill, and D. Boes: 1974, *Introduction to the Theory of Statistics*. New York: McGraw-Hill.

Möller, B. and M. Beer: 2004, *Fuzzy Randomness - Uncertainty in Civil Engineering and Computational Mechanics*. Berlin: Springer.

Möller, B., W. Graf, and M. Beer: 2000, 'Fuzzy structural analysis using alpha-level optimization'. *Computational Mechanics* **26**, 547–565.

Möller, B., W. Graf, and M. Beer: 2003, 'Safety assessment of structures in view of fuzzy randomness'. *Computers and Structures* **81**, 1567–1582.

Papadopoulos, V., G. Deodatis, and M. Papadrakakis: 2005, 'Flexibility-based upper bounds on the response variability of simple beams'. *Computer Methods in Applied Mechanics and Engineering* **194**(1), 1385–1404.

Phoon, K., H. Huang, and S. Quek: 2005, 'Simulation of strongly non-Gaussian processes using Karhunen-Loeve expansion'. *Probabilistic Engineering Mechanics* **20**(2), 188–198.

Rackwitz, R.: 2001, 'Reliability analysis – a review and some perspectives'. *Structural Safety* **23**(4), 365–395.

Red-Horse, J. R. and A. S. Benjamin: 2004, 'A probabilistic approach to uncertainty quantification with limited information'. *Reliability Engineering & System Safety* **85**(1-3), 183–190.

Schenk, C. A., H. J. Pradlwarter, and G. I. Schuëller: 2005, 'Non-stationary response of large, non-linear finite element systems under stochastic loading'. *Computers and Structures* **83**(14), 1086–1102.

Schenk, C. A. and G. I. Schuëller: 2005, *Uncertainty Assessment of Large Finite Element Systems*. Berlin Heidelberg: Springer.

Schuëller, G.: 2001a, 'Computational Stochastic Mechanics - Recent Advances'. *Computers and Structures* **79**(22–25), 2225–2234.

Schuëller, G.: 2001b, 'On Computational Procedures for Processing Uncertainties in Structural Mechanics'. In: Z. Waszczyszyn and J. Pamin (eds.): *2nd Europ. Conf. on Computational Mechanics, Cracow*. pp. 1–24. CD-ROM, Doc 608.

Schuëller, G. and P. Spanos (eds.): 2001, 'Proc. Int. Conf. on Monte Carlo Simulation MCS 2000, Monte Carlo, Monaco, 2000'. Swets & Zeitlinger BV, Lisse, The Netherlands.

Schuëller, G. I., H. J. Pradlwarter, and C. A. Schenk: 2003, 'A computational procedure to estimate the stochastic dynamic response of large non-linear FE-models'. *Computer Methods in Applied Mechanics and Engineering* **192**(7–8), 777–801.

Sickert, J.-U., M. Beer, W. Graf, and B. Möller: 2003, 'Fuzzy probabilistic structural analysis considering fuzzy random functions'. In: A. D. Kiureghian, S. Madanat, and J. Pestana (eds.): *9th Int. Conference on Applications of Statistics and Probability in Civil Engineering*. Berkeley, Rotterdam, pp. 379–386, Millpress.

Spanos, P. D. and R. Ghanem: 1989, 'Stochastic finite element expansion for random media'. *ASCE Journal of the Engineering Mechanics* **115**(5), 1035–1053.

Tonon, F., A. Bernardini, and A. Mammino: 2000, 'Reliability analysis of rock mass response by means of Random Set Theory'. *Reliability Engineering & System Safety* **70**(3), 263–282.

Viertl, R.: 1996, *Statistical Methods for Non-Precise Data*. Boca Raton New York London Tokyo: CRC Press.

Zimmermann, H.-J.: 1992, *Fuzzy set theory and its applications*. Boston London: Kluwer Academic Publishers.