Relaxed evolutionary power spectral density functions: A probabilistic approach to model uncertainties of non-stationary stochastic signals

Marius Bittner[a,b,*], Marco Behrendt[a], Michael Beer[a,c,d]

[a] *Institute for Risk and Reliability, Leibniz Universität Hannover, Callinstraße 34, Hannover, 30167, Germany*
[b] *International Research Training Group (IRTG) 2657, Leibniz Universität Hannover, Appelstraße 11/11a, Hannover, 30167, Germany*
[c] *Institute for Risk and Uncertainty, University of Liverpool, Peach Street, Liverpool, L69 7ZF, United Kingdom*
[d] *International Joint Research Center for Engineering Reliability and Stochastic Mechanics, Tongji University, 1239 Siping Road, Shanghai, 200092, China*

Abstract

The identification of patterns and underlying characteristics of natural or engineering time-varying phenomena poses a challenging task, especially in the scope of simulation models and accompanying stochastic models. Because of their complex nature, time-varying processes such as wind speed, seismic ground motion, or vibrations of machinery in the presence of degradation oftentimes lack a closed-form description of their underlying evolutionary power spectral density (EPSD) function. To overcome this issue, a wide range of measurements exist for these types of processes. This opens up the path to a data-driven stochastic representation of EPSD functions. Rather than solely relying on time-frequency transform methods like the familiar short-time Fourier transform or wavelet transform for EPSD estimation, a probabilistic representation of the EPSD can provide valuable insights into the epistemic uncertainty associated with these processes. To address this problem, the evolutionary EPSD function is relaxed based on multiple similar data to account for these uncertainties and to provide a realistic representation of the time data in the time-frequency domain. This results is the so-called relaxed EPSD (REPSD) function, which serves as a modular probabilistic representation of the time-frequency content of stochastic signals. For this purpose, truncated normal distributions and kernel density estimates are used to

determine a probability density function for each time-frequency component. The REPSD function enables the sampling of individual EPSD functions, facilitating their direct application to the simulation model through stochastic simulation techniques like Monte Carlo simulation or other advanced methods. Even though the accuracy is highly dependant on the data available and the time-frequency transformation method used, the REPSD representation offers a stochastic representation of characteristics used to describe stochastic signals and can reduce epistemic uncertainty during the modelling of such time-varying processes. The method is illustrated by numerical examples involving the analysis of dynamic behaviour under random loads. The results show that the method can be successfully employed to account for uncertainties in the estimation of the EPSD function and represent the accuracy of the time-frequency transformation used.

Keywords: Evolutionary power spectral density function, Stochastic processes, Stochastic dynamics, Uncertainty quantification, Stochastic signals, time-frequency transformation

1. Introduction

The description of natural phenomena in the context of simulation models is a challenging problem. These phenomena such as wind and wave movements, seismic activities or climate changes are related to complex, interacting high-dimensional physical models. Also, engineering problems, such as vibrations of a component under changing material behaviour, can often only be modelled with an acceptable level of accuracy using complex models and experiments. Since these phenomena present time-varying properties in engineering they are often referred to as environmental processes or from a mathematical and modelling perspective, stochastic processes. In civil engineering, stochastic dynamics, and structural analysis, stochastic processes play a crucial role [1, 2, 3]. Stochastic dynamics is concerned with the study of probabilistic systems that evolve over time and has applications in structural reliability analysis. Specifically, the treatment of random vibrations is important in this field [4, 5, 6, 7]. To ensure the reliability of a structure, analysts need to consider

*Corresponding author
Email address: bittner@irz.uni-hannover.de (Marius Bittner)

the potential impact of environmental processes such as wind, wave, or seismic loads at the design stage or carry out analyses for existing structures [8]. This often requires complex simulations to accurately predict how the structure will respond to these processes and to ensure that it meets safety requirements.

In addition to the physical models, it is possible to introduce a complementary stochastic model, which includes the formulation of suitable stochastic processes. To approximate stochastic processes mainly three branches of approaches have been established up to now [9]:

- The Karhunen-Lòeve (K-L) expansion, in which for engineering processes orthogonal functions in time and space are combined linearly [10]. To describe the stochastic processes, a formulation of the corresponding covariance functions must be available. K-L can be used to simulate non-Gaussian and non-stationary stochastic processes [11].

- Sampling representations, which are in the classical representation mostly suitable for reproducing a full signal by deterministic samples available, however, methods like the Withaker-Shannon interpolation require a limited bandwidth of the analysed processes [12].

- Spectral representation methods, based on the formulation of a power spectral density (PSD) function [13].

From an engineering perspective, when regarding environmental loads, formulating PSD functions has advantages in vibration analysis, as they provide a method of directly characterising the frequency content of stochastic signals. Artificial stochastic signals can be considered as deterministic realisations of stochastic processes, whereas signals themselves can also be measurements. Via PSD function estimation procedures, a signal can be decomposed into its harmonic components. In particular, the amplitudes and their distribution over the frequencies are determined. However, to accurately calculate the PSD function of a signal, certain mathematical conditions are necessary, such as dealing with continuous

signals and signals of infinite length. Since these requirements cannot be met in practice, estimators are used.

Over time, different spectral density estimators have been developed that offer certain advantages and disadvantages. In the stationary case, common estimators such as the periodogram [7], Welch's method [14], or Bartlett's method [15, 16] are widely employed. For an overview, refer to [17]. These estimators all rely on the discrete Fourier transform (DFT) [18]. Since signals of environmental processes often have a non-stationary character which results in a frequency power change in time, so-called evolutionary PSD (EPSD) functions need to be considered. These take into account the time-varying behaviour of a signal and thus provide a more realistic representation in the resulting time-frequency domain [19, 20, 21]. Different time-frequency transformation methods exist, with certain advantages and disadvantages, particularly in the quality of the transformation as well as the resolution in the time-frequency domain. An EPSD function can be estimated using, among others, the short-time Fourier transform (STFT) [7, 22], wavelet methods [23, 24], or the recently developed multi-taper S-transform (MTST) [25].

Once a suitable estimation of a PSD or EPSD is available, artificial stochastic signals can be generated using the Spectral Representation Method (SRM), either in the stationary case [13] or in the non-stationary case [26]. The latter is of interest in this work. For the generation of stochastic signals, all stochastic simulation methods are applicable, such as the widely used Monte Carlo (MC) simulation [27, 28] but also the usage of advanced techniques such as subset simulation [29], line sampling [30], or directional importance sampling [31] are possible.

However, determining these signals via a transformation method of choice and even further utilise these representations for e.g. structural reliability analysis is challenging due to the presence of uncertainties [32]. Uncertainties can be divided into aleatory and epistemic uncertainties [33], while on the one side, aleatory uncertainties describe irreducible stochastic conditions, epistemic uncertainties are referred to as reducible uncertainties. If both types of uncertainties occur simultaneously and are not separable, they are called hybrid uncertainties [34, 35]. The assessment of uncertainties in simulations, analyses, and

engineering systems is ubiquitous. If uncertainties are incorrectly quantified, consequences can be disastrous. For example, a building under given loads could suffer devastating damage that severely compromises the structural reliability. Different approaches to quantify and propagate uncertainties are available, such as precise probabilistic methods [36, 37], non-probabilistic methods [38] or imprecise probabilities [39].

In general, stochastic signals and specifically real data records are subject to uncertainties. These can result, for example, from poorly calibrated sensors, measurement errors, an insufficient amount of data, while damaged or failed sensors can result in records with missing data. Furthermore, the presence of uncertainties is a direct result of the inherent complexity found in natural environmental processes. In addition, due to the mentioned mathematical requirements on the signal for the time-frequency transformations, only estimates of the EPSD function can be determined, which leaves the epistemic uncertainty of the EPSD functions unidentified. To overcome these issues, a range of PSD and EPSD function estimation techniques involving uncertainty quantification have been proposed. Missing data problems are treated in [40, 41, 42], an interval-valued PSD function from similar data has been derived in [43], while in [44] a set of accelerograms is analysed to derive reliability bounds. An interval-valued signal can be transformed to an interval-valued PSD function using the interval DFT transform [45]. In particular, the steadily growing databases, for instance [46, 47, 48], contribute to a better understanding of environmental processes and the quantification of uncertainties.

In addition to the above mentioned approaches, the authors of this work derived a probabilistic model of a set of similar PSD functions, the relaxed PSD function [49]. However, the proposed methodology is only valid for stationary and Gaussian stochastic signals. The non-stationary case needs to involve the EPSD function estimation for a more accurate representation of environmental processes. Additionally, the spectral representation by SRM and the stochastic simulation need to be reconsidered. In this work a modular framework for the representation of non-stationary stochastic signals via EPSD functions using artificial stochastic signals generated by the SRM is proposed. The EPSD function estimation will be carried out using the recently developed MTST [25]. Once the ensemble is derived, each

spectral density per time-frequency component will be transformed into a probability density function. To underline the modularity of the construction of a REPSD function, simple truncated normal distributions will be used as basis, additionally kernel density estimations are implemented to construct custom probability density functions (PDF) in the time-frequency space and to show, that arbitrary PDF types could be implemented. Both approaches have their advantages and disadvantages, depending on the amount and appearance of the data. The resulting REPSD function is used to perform numerical simulations of the dynamic behaviour of systems subjected to environmental processes by sampling individual EPSD functions applied to the model via MC simulation.

This work is organised as follows: Basic concepts of the SRM and EPSD function establishment, important for the remainder of this work will be explained in Section 2. In Section 3 the methodology of constructing the REPSD function will be elaborated. The obtained REPSD model will be validated by MTST estimations of EPSD functions and compared to the source EPSD of an artificial environmental process. To illustrate the strengths and advantages of the REPSD function, two different numerical examples are presented in Section 4. The final conclusions and a critical discussion of the obtained results are given in Section 5.

2. Preliminaries

In this section necessary methodologies for the REPSD representation are introduced. These mainly include the representation of non-stationary stochastic processes by a spectral representation (Section 2.1), the estimation of EPSD functions from generated stochastic signals by the state of the art MTST method (Section 2.2) and revisiting the kernel density estimation (KDE), which is used to determine PDF representations during the REPSD construction (Section 2.3).

2.1. Representation of non-stationary stochastic processes

A convenient way to generate sample functions $X(t)$ that represent non-stationary stochastic processes in a time-domain is presented in [26]. For a source EPSD $S_X(\omega, t)$ the spectral

representation of non-stationary stochastic processes can be stated as

$$X(t) = \sqrt{2} \sum_{n=0}^{N_\omega - 1} \sqrt{2 S_X(\omega_n, t) \Delta\omega} \cos(\omega_n t + \varphi_n), \tag{1}$$

in which,

$$\begin{aligned} \omega_n &= n\Delta\omega, \quad n = 0, 1, 2, \ldots, N_\omega - 1, \\ \text{and} \quad \Delta\omega &= \frac{\omega_u}{N_\omega}, \end{aligned} \tag{2}$$

$t \in [0, T]$ is the time domain with T being the ultimate time. ω_n are discretised frequency values and $\omega_u = f_s/4\pi$ is the upper cutoff-frequency determined by the time discretisation Δt and sampling ratio $f_s = 1/\Delta t$. φ_n, with $n = 0, 1, \ldots, N_\omega - 1$ are uniformly distributed phase angles in the range $[0, 2\pi]$. N_t is the desired number of time instances, such that $\Delta t = T/N_t$. $S_X(\omega_n, t)$ in Eq. (1) is the source EPSD and could be of arbitrary shape. A non-separable EPSD as in [50] has been chosen as source EPSD:

$$S_X(\omega, t) = \left(\frac{\omega}{5\pi}\right)^2 \cdot \exp\left[-0.15t\right] \cdot t^2 \cdot \exp\left[-\left(\frac{\omega}{5\pi}\right)^2 t\right]. \tag{3}$$

For a discretised time-frequency space this relation is depicted in Fig. 1. As already established in the SRM [13], if $N_\omega \to \infty$ it can be assumed that the sample function in Eq. (1) presents an accurate simulation for a non-stationary stochastic process. Since only a limited number N_ω of approximation terms in Eq. (3) is feasible, the sample functions in $X(t)$ are referred to as stochastic signals or simply signals.

2.2. Evolutionary power spectral density estimation

The generated sample functions in Eq. (1) are regarded as arbitrary stochastic signals but with the same source EPSD. A challenging task remains to find a robust estimator for this source EPSD when considering only a finite number of generated stochastic signals. In [25] the MTST method, which exhibits a significant variance reduction in comparison to other EPSD estimation procedures such as the Priestley method and wavelet-based methods, has been presented. Given a non-stationary stochastic signal $X(t)$ and M time-frequency

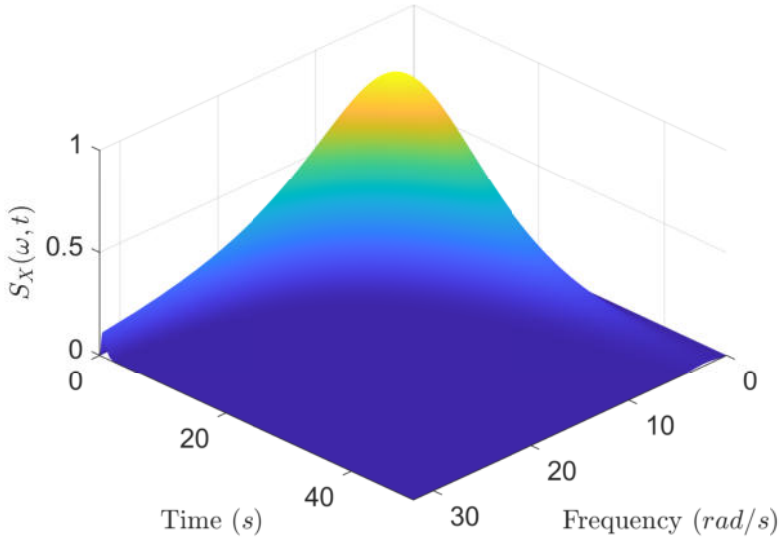


Figure 1: Non-separable EPSD function (Eq. (3)) for $\omega \in [0, 32.17]$ rad/s and $t \in [0, 50]$ s.

Hermite windows, denoted as $\Psi_m(\omega, t)$, $m = 0, 1, ..., M-1$, with zero-padding $X_0(t) = \{0_1, \ldots, 0_{N_t/2}, X(t), 0_1, \ldots 0_{N_t/2-1}\}$, the so-called S-transform is given to be

$$s_m(\omega, t) = \sum_{k=-\underline{c}}^{\bar{c}} \Psi_m(\omega, k\Delta t - t) \cdot X_0(k\Delta t) \cdot \exp\left[-i2\pi\omega k\Delta t\right] \cdot \Delta t \tag{4}$$

in the case that $\underline{c}, \bar{c}, \to \infty$, Eq. (4) has the maximum accuracy dependant only on M. In this study $\underline{c} = \lceil N_t/2 \rceil + 1$ and $\bar{c} = \lfloor N_t/2 \rfloor$. Please note that with the zero-padding a continuous window over the time-domain can be regarded, i.e. $X_0(t)$ is treated as a periodic signal. The Hermite windows are constructed as following

$$\begin{aligned}
\Psi_0(\omega, t) &= \pi^{-1/4} \cdot \sqrt{\mathrm{w}(\omega)} \cdot \exp\left[-1/2 \cdot \mathrm{w}(\omega)^2 \cdot t^2\right], \\
\Psi_1(\omega, t) &= \sqrt{2}\pi^{-1/4} \cdot \mathrm{w}(\omega)^{3/2} \cdot t \cdot \exp\left[-1/2 \cdot \mathrm{w}(\omega)^2 \cdot t^2\right], \\
\Psi_m(\omega, t) &= \sqrt{2/m} \cdot \mathrm{w}(\omega) \cdot t \cdot \Psi_{m-1}(t, \omega) - \sqrt{(m-1)/m} \cdot \Psi_{m-2}(t, \omega).
\end{aligned} \tag{5}$$

With the shape function

$$\mathrm{w}(\omega) = a\left[1 + \frac{b^2 \cdot |\omega/f_s|^{|c|+1}}{|b| \cdot |\omega/f_s|^{|c|+1} + 1}\right], \tag{6}$$

and parameters a, b, c. A detailed study on the choice of these parameters can be found in [51]. With the full definition of the Hermite windows, the estimator for $S_X(\omega, t)$ is established

as

$$\hat{S}_X(\omega, t) = \frac{1}{M} \sum_{m=0}^{M-1} s_m^*(\omega, t) s_m^T(\omega, t). \tag{7}$$

Here $(\cdot)^*$ represents the conjugate operator, and $(\cdot)^T$ the transpose operator.

2.3. Kernel density estimation

The KDE is a method of probability theory and statistics that allows to estimate the PDF of a random variable without assuming a specific distribution. A brief explanation is given here, while the reader is referred to [52, 53, 54] for a detailed explanation.

A PDF is estimated by applying a kernel function to each individual data point. The kernel function serves as a weighting factor to account for the contributions of each data point to the estimated density. Summing up all kernel functions will result in the estimated PDF. The choice of kernel function and its width is crucial for the accuracy of the estimate because it determines the shape of the estimated PDF. Typical kernel functions include the Gaussian distribution or the Epanechinikov function. The KDE is often used to gain an understanding of the distribution of the data, or as a basis for further analysis.

The estimation of the PDF using KDE can be expressed by

$$\hat{f}\,(x) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h} K_h \left(\frac{x - x_i}{h} \right), \tag{8}$$

where $\hat{f}$ is the estimated probability density, K_h is the kernel function, h is the bandwidth of the kernels, x is the point at which the PDF is estimated and x_i are the observations, i.e. the available data points. In this work, a Gaussian kernel is utilised.

The bandwidth has a significant impact on the quality of the resulting estimated PDF. For example, setting the bandwidth too high can lead to an over-smoothed result, while choosing a bandwidth that is too low can overweight individual data points, leading to an under-smoothed PDF characterised by multiple sharp peaks. There are several ways to find an optimal bandwidth for the kernels. Assuming that the PDF to be estimated is Gaussian distributed, a well-known rule is Scott's rule [55], which incorporates the number of data

points n as well as their standard deviation σ

$$h_{opt} = \left(\frac{4\sigma^5}{3n}\right)^{1/5} \approx 1.06\sigma n^{-1/5}. \tag{9}$$

Another rule is Silverman's rule [52], which takes into account the interquartile range IQR in addition to the number of data points n and their standard deviation σ

$$h_{opt} = 0.9 \min\left(\sigma, \frac{IQR}{1.34}\right) n^{-1/5}, \tag{10}$$

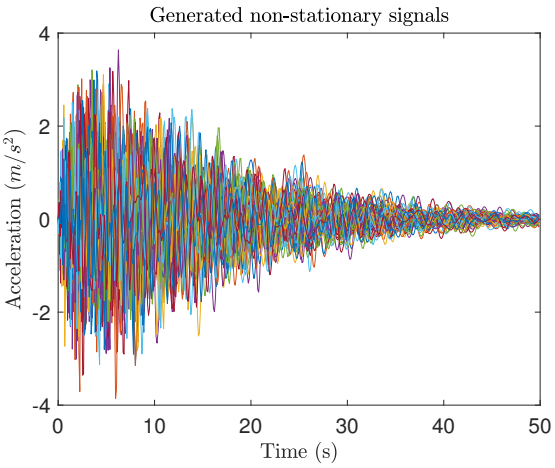
where $IQR = Q_3 - Q_1$ with Q_1 as lower quartile and Q_3 as upper quartile.

3. Methodology

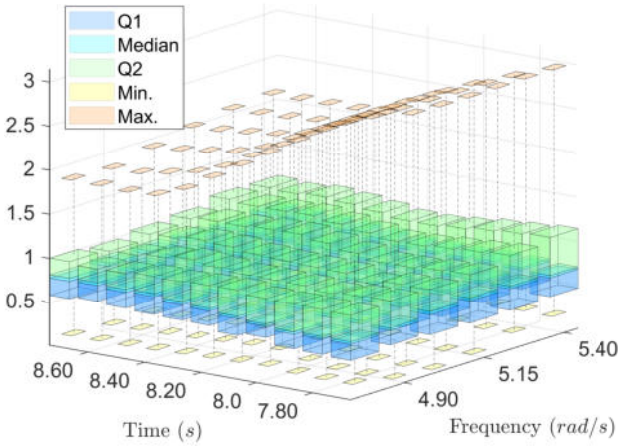
In this section the novel estimation procedure for the REPSD function is presented.

3.1. Relaxed evolutionary power spectral density function

The novel construction of REPSD functions is particularly suitable when many records are available for particular phenomena. The aim is to reduce the epistemic uncertainty when describing stochastic signals. When records of environmental or natural processes such as wind loads, seismic ground motions, or other vibrations are available, it is still not determined which EPSD function is a good approximation for this signal aggregation. For a thorough analysis, it is also interesting to examine whether these phenomena can be considered directly related or not. The latter challenge is not part of this method and would require additional preliminary data analysis. However, from a modelling perspective, a common characteristics formulation of environmental processes is essential to obtain simulation results. Furthermore, a stochastic, relaxed representation of EPSD functions for stochastic signals could lead to a new modelling perspective, as the representation of main features, such as first and second order moments, of natural processes can then be determined by statistical methods. Consider an ensemble that consists of a number of different EPSD estimations from stochastic signals. Fig. 2(a) shall represent signals from a natural process. The estimation procedure can be chosen according to the analyst's needs, throughout this work the presented MTST method in Section 2.2 with the estimator in Eq. (7) is used. This set of



(a) Non-homogeneous stochastic signals with $N_\omega = 256$



(b) Ensemble $\{\hat{S}_{X_i}\}, i = 1, 2, ..., 50$ for $t \in [7.6, 8.8]$ s and $\omega \in [4.75, 5.4]$ rad/s

Figure 2: 50 source signals obtained from Eq. (1)(left), boxplot with full range and quartiles for the ensemble in $\{\hat{S}_{X_i}\}$ (right).

EPSD estimations with cardinality N_e can be stated as an ensemble: $\{\hat{S}_{X_i}\},\ i = 1, 2, ..., N_e$. Here N_e is determined by the number of regarded records available, containing either measurements or artificially generated stochastic signals (as in Fig. 2(a)). Each EPSD function estimation, denoted by the index i, is discretised over the frequency and time domain, i.e. $s_{i,\omega_n,t_k} = \hat{S}_{X_i}(\omega_n, t_k),\ n = 0, 1, \ldots, N_\omega - 1, k = 1, \ldots, N_t$. The ensemble consists of the data-driven input collection of EPSD estimations which lay the foundation of the stochastic input space for the REPSD function described below. In Fig. 2(b) the MTST estimations of the EPSD functions are shown, the label abbreviations of the boxplot for each discretised frequency-time point represent the following: Min.: minimum value, Q1: lower quartile, Median: median value, Q2: upper quartile, Max.: maximum value of the data. It is obvious that the time-frequency transformation from stochastic signals delivers a massive amount of statistical data.

From the ensemble, following statistical moments are derived for each discretised point

of the EPSD ensemble

$$\mu_{\omega_n,t_k} = \frac{1}{N_e}\sum_{i=0}^{N_e-1} s_{i,\omega_n,t_k}, \tag{11a}$$

$$\sigma_{\omega_n,t_k} = \sqrt{\frac{1}{N_e}\sum_{i=0}^{N_e-1}\left(s_{i,\omega_n,t_k}-\mu_{\omega_n,t_k}\right)^2}, \tag{11b}$$

with $\omega_n = n\Delta\omega,\ n = 0,1,\ldots,N_\omega-1$, and $t_k = k\Delta t,\ k = 1,2,\ldots,N_t$. Eq. (11a) and Eq. (11b) need to be calculated for each discretised frequency ω_n and time instance t_k. With this statistical information for each discretised point a probability density function can be constructed. In theory, establishing arbitrary distribution types is possible. For the sake of clarity, first, a simple distribution type is chosen. Assume a truncated normal (TN) distribution for each point. These are then given to be

$$f^{TN}_{\omega_n,t_k}(s;\mu_{\omega_n,t_k},\sigma_{\omega_n,t_k},l_{\omega_n,t_k},u_{\omega_n,t_k}) = \frac{1}{\sigma_{\omega_n,t_k}}\frac{\phi\left(\frac{s-\mu_{\omega_n,t_k}}{\sigma_{\omega_n,t_k}}\right)}{\Phi\left(\frac{u_{\omega_n,t_k}-\mu_{\omega_n,t_k}}{\sigma_{\omega_n,t_k}}\right)-\Phi\left(\frac{l_{\omega_n,t_k}-\mu_{\omega_n,t_k}}{\sigma_{\omega_n,t_k}}\right)}, \tag{12}$$

and $\phi(\eta) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\eta^2\right)$ is the standard normal distribution, $\Phi(\zeta) = \frac{1}{2}\left(1+\mathrm{erf}(\zeta/\sqrt{2})\right)$ is the corresponding cumulative distribution function. The lower and upper truncation bounds are given to be l_{ω_n,t_k} and u_{ω_n,t_k}. The influence of the truncation bounds has been discussed for the frequency domain in [49]. In this study the configuration $l_{\omega_n,t_k} = 0$ and $u_{\omega_n,t_k} = 2\mu_{\omega_n,t_k}$ proved to be more robust. The TN distribution yields a smooth representation of the EPSD function values for each discretised point. Thus, outliers or gaps in the EPSD function value are weighted less.

In some cases, however, it may be useful to represent the data more in its natural appearance. In such a case, the KDE provides a much more inclusive representation of the data. This can result in multiple peaks instead of smooth curve as in the case of the TN distribution. If the data set is to be represented by KDE, Eq. (8) can be reformulated into its KDE-driven REPSD form

$$\hat{f}^{KDE}_{\omega_n,t_k}(x;s_{i,\omega_n,t_k},h_{\omega_n,t_k}) = \frac{1}{N_e}\sum_{i=1}^{N_e}\frac{1}{h_{\omega_n,t_k}}K_{h_{\omega_n,t_k}}\left(\frac{x-s_{i,\omega_n,t_k}}{h_{\omega_n,t_k}}\right). \tag{13}$$

This equation is evaluated for each time-frequency point (ω_n, t_k). The parameter h_{ω_n,t_k} describes an adaptive bandwidth for the KDE, which will be determined according to Eq. (10) and the respective available EPSD values at the corresponding time-frequency point.

For both types of distribution functions it can be argued why they are superior, but this always depends on the individual case, as the form and number of data are important for this assessment. For a high number of data, a truncated normal distribution may be better, as this may well lead to a smooth representation. Also, if a clear interval of the ensemble data can be identified the TN approach would be better because this would make the choice of truncation bounds for the PDF definition easier. If there is less data or multiple PDF peaks are expected, a KDE may provide better results, since gaps and multiple peaks, for instance, can be represented easier. However, here over-fitting could appear, in these cases a careful examination of the sampled REPSDs about the ensemble and the underlying data needs to be carried out. Special consideration to outliers must be given, in particular for KDE approaches, because again here an over representation of outlier data could occur. Note that the distributions presented are only two possibilities that could be used for the REPSD model. Additionally, no correlations or dependencies have been considered so far. To obtain a realisation of a truncated normal REPSD (TN-REPSD) named $TS_X(\omega_n, t_k)$, for each ω_n and t_k a sample is generated from the respective distributions in Eq. (12). To obtain a realisation of a kernel density estimated REPSD (KDE-REPSD) named $KS_X(\omega_n, t_k)$, for each ω_n and t_k a sample is generated from the respective distributions in Eq. (13). In a sense, Eq. (12) and Eq. (13) can be seen as the description of uncorrelated random fields.

3.2. Optimised MTST parameters & error estimates

The MTST estimation for each signal in Fig. 2(a) is dependant on the choice of the parameters a, b, c of the window function in Eq. (6) and the number of Hermite window orders M used. Since the source EPSD for the regarded signals is known as in Eq. (3), it is possible to formulate an objective function. The objective function, also later on used as error estimate, is formulated according to the Frobenius norm of matrices. Respectively for two arbitrary EPSDs a residual matrix of the comparison of the two is stated as $S_{res}(\omega_n, t_k) =$

$S_1(\omega_n, t_k) - S_2(\omega_n, t_k)$ for $n = 0, 1, \ldots, N_\omega - 1, k = 1, \ldots, N_t$. For this residual matrix the Frobenius norm is used to formulate an objective function and error estimate by

$$||S_{res}(\omega_n, t_k)||_F = \sqrt{\sum_{i=0}^{N_\omega - 1} \sum_{j=1}^{N_t} |S_{res}(\omega_i, t_j)|^2}. \tag{14}$$

Following constraints for the MTST parameters are introduced $0.001 \leq a \leq 1$, $0.001 \leq b \leq 30$, $0.001 \leq c \leq 1$, and $M \in 1, 2, \ldots 10$, these constraints were chosen by the authors after assessing the parameter discussion in [51]. For the minimisation of the residual matrix a genetic algorithm optimisation was used as established in [56]. Following objective is minimised once before the REPSD is constructed:

$$\arg\min_{a,b,c,M} \left\{ ||S_{res0}(\omega_n, t_k, a, b, c, M)||_F \right\}, \tag{15}$$

where in this specific case $S_{res0}(\omega_n, t_k) = S_X(\omega_n, t_k) - \overline{\hat{S}_X(\omega_n, t_k)}$, which defines the Frobenius norm between the source EPSD and the mean of one a priori chosen ensemble. The formal calculation of the first and second order moments for the ensemble date can be found in Appendix A.

3.3. Error analysis

In Fig. 3, a single MTST estimation of one arbitrarily chosen signal in Fig. 2(a) is depicted. In Fig. 4, a single realisation of the TN-REPSD is shown. Whilst the MTST estimation is showing gaps for specific parts of the EPSD, the by TN-REPSD generated sample EPSD is resembling the source EPSD in Fig. 1.

The mean of the MTST estimations for all ensemble members is robust and yields a good approximation. But a single TN-REPSD sample, as in Fig. 4, shows in comparison to a single MTST estimation a better resemblance to the source EPSD due to the construction of a non-correlated random field that describes the ensemble's statistics in frequency and time domain. For all error evaluations, the direct absolute differences between two EPSD functions, differences between the ensemble's mean and REPSD sample's mean, or the equivalent difference for the standard deviation are regarded. Also, the errors according

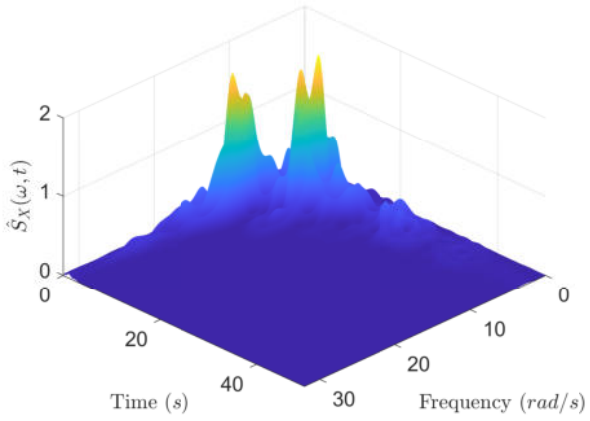


Figure 3: Single MTST estimation for one signal as depicted in Fig. 2(a).

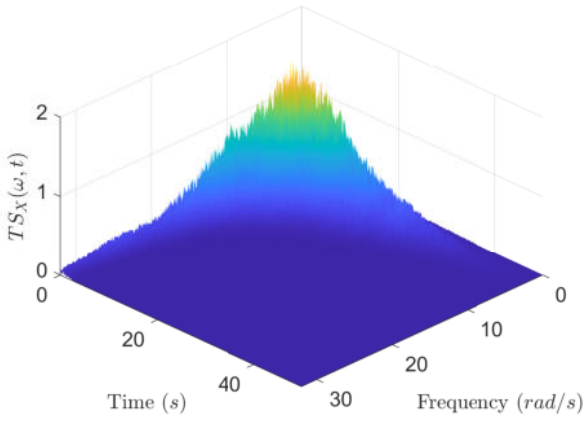


Figure 4: Single TN-REPSD sample generated according to Eq. (12).

to the Frobenius norm introduced in Eq. (14) are shown. The error and convergence analysis for the TN-REPSD representation and sampling has been performed. The results are shown in Fig. 5, here following notations were used: The source EPSD function is described as $S_X(\omega,t)$, the ensemble mean estimated by MTST is $\overline{\hat{S}_X(\omega_n,t_k)}$, the mean of generated TN-REPSD samples is denoted as $\overline{TS_X(\omega,t)}$.

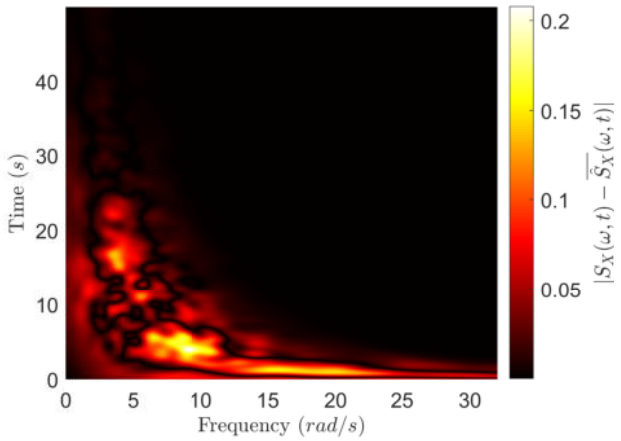
In Fig. 5(a) the error between the source EPSD and the mean of the MTST estimates is shown. The maximum deviation amounts to ≈ 0.2. This error is the reference, since it measures the deviation of the MTST estimation to the source EPSD function. In Fig. 5(b) the error between the mean of $N_{10k} = 10000$ generated TN-REPSD samples and the source EPSD is shown. Here it can be seen that the error is similar to the error of the MTST estimation mean. This leads to the conclusion that the constructed TN-REPSD is mainly dependant on the quality of the EPSD estimation.

TN-REPSD is not adding any errors but offering a fully stochastic description of an estimated EPSD. In Fig. 5(c) the error between the mean of the MTST estimated ensemble is compared to the mean of the N_{10k} generated TN-REPSD samples. From the three results Fig. 5(a) - 5(c), the error map in Fig. 5(c) exhibits the smallest error. This is expected, because the TN-REPSD is constructed out of the $N_e = 50$ ensemble members. In Fig. 5(d)

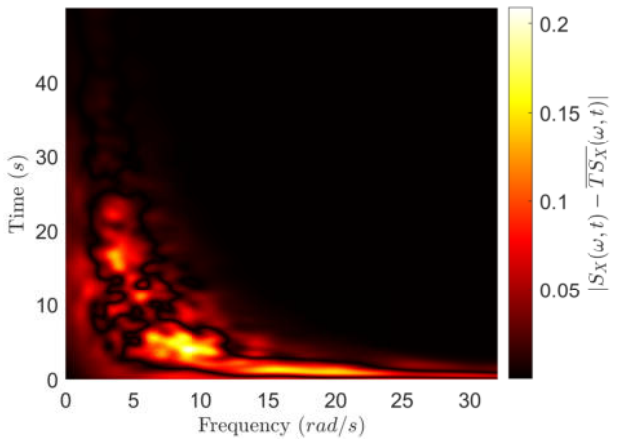
a convergence analysis based on the MC method has been carried out for different number of generated TN-REPSD samples $N \in \{1, 2, 2.5, 3, 4.5, 4, 4.5, 5\} \cdot 10^4$. For larger numbers of N, the error between the TN-REPSD mean and the ensemble mean is decreasing (blue line, blue axis), the error between the source EPSD and the mean of the TN-REPSD is constant (orange line, orange axis). The dashed black line refers to the direct error of the source EPSD function in comparison to the mean of the MTST ensemble in the Frobenius norm. This error serves here as a reference and corresponds to the orange axis. The orange line is below this value. All these results in Fig. 5 indicate that the TN-REPSD does not introduce any additional error and delivers a good representation of the provided input ensemble (data). Additionally, the convergence of the MC simulations follows the law of large numbers, i.e. with a larger sample size the error reduces. Similar considerations are carried out for the KDE-REPSD in Fig. 6.

In Fig. 6, the notations are equivalent to the previous graph, $\overline{KS_X(\omega, t)}$ denotes the mean of the by KDE-REPSD generated EPSD samples. The general trend of these error and convergence results are similar. Which again indicates that the stochastic representation of the EPSD ensemble is mainly reliant on the EPSD estimation procedure. Please note that a different ensemble has been used for the KDE-REPSD analysis, resulting in a slightly different error in Fig. 6(a) compared to Fig. 5(a). In Fig. 6(c) higher local errors are observed. Especially when comparing with the previous results of the TN-REPSD in Fig. 5(c). Also for the KDE-REPSD generated samples it can be observed, that the dashed black line, which is the reference error value for the source EPSD estimation, is always larger than the orange line. This leads to the conclusion that no additional error is introduced by the KDE-REPSD.

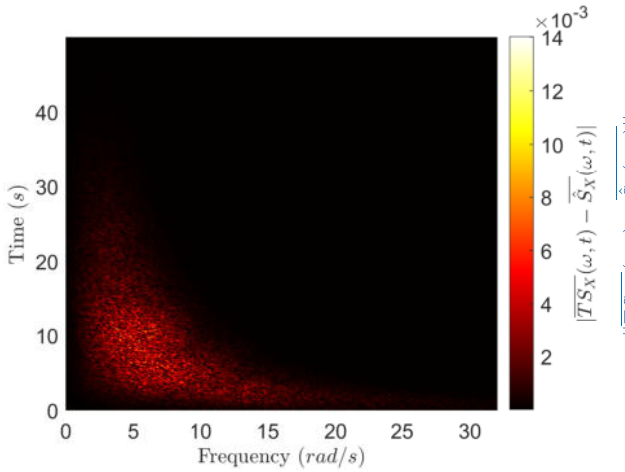
Additionally, for specific time instances $t \in \{0.88\,\mathrm{s}, 9.86\,\mathrm{s}, 19.92\,\mathrm{s}\}$ the EPSD function's frequency space is analysed. The REPSD functions first and second order moment, that are used for the construction of the REPSD (input for the distribution functions) are denoted by $\mu[TS_X(\omega, t)]$ and $\mu[KS_X(\omega, t)]$, the mean parameter and $\sigma(TS_X(\omega, t))$ and $\sigma(KS_X(\omega, t))$, the standard deviation parameter. For the MC sampling of the TN-REPSD and KDE-REPSD the same N_{10k} samples from the previous results were used, the mean of the samples



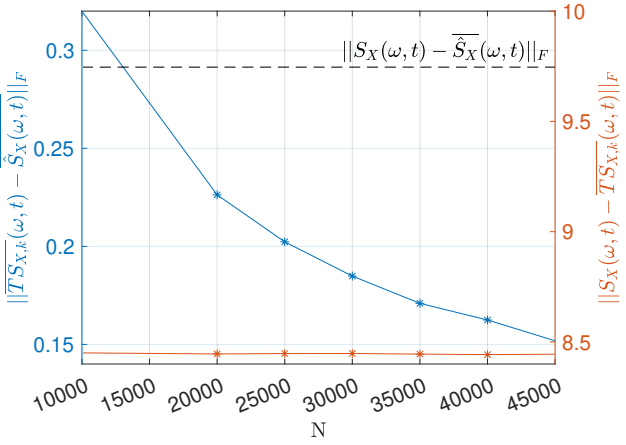
(a) Difference of source EPSD $S_X(\omega, t)$, vs. the mean of the MTST estimations $\overline{\hat{S}_X(\omega_n, t_k)}$.



(b) Difference of source EPSD and the mean of TN-REPSD generated samples $\overline{TS_X(\omega, t)}$.

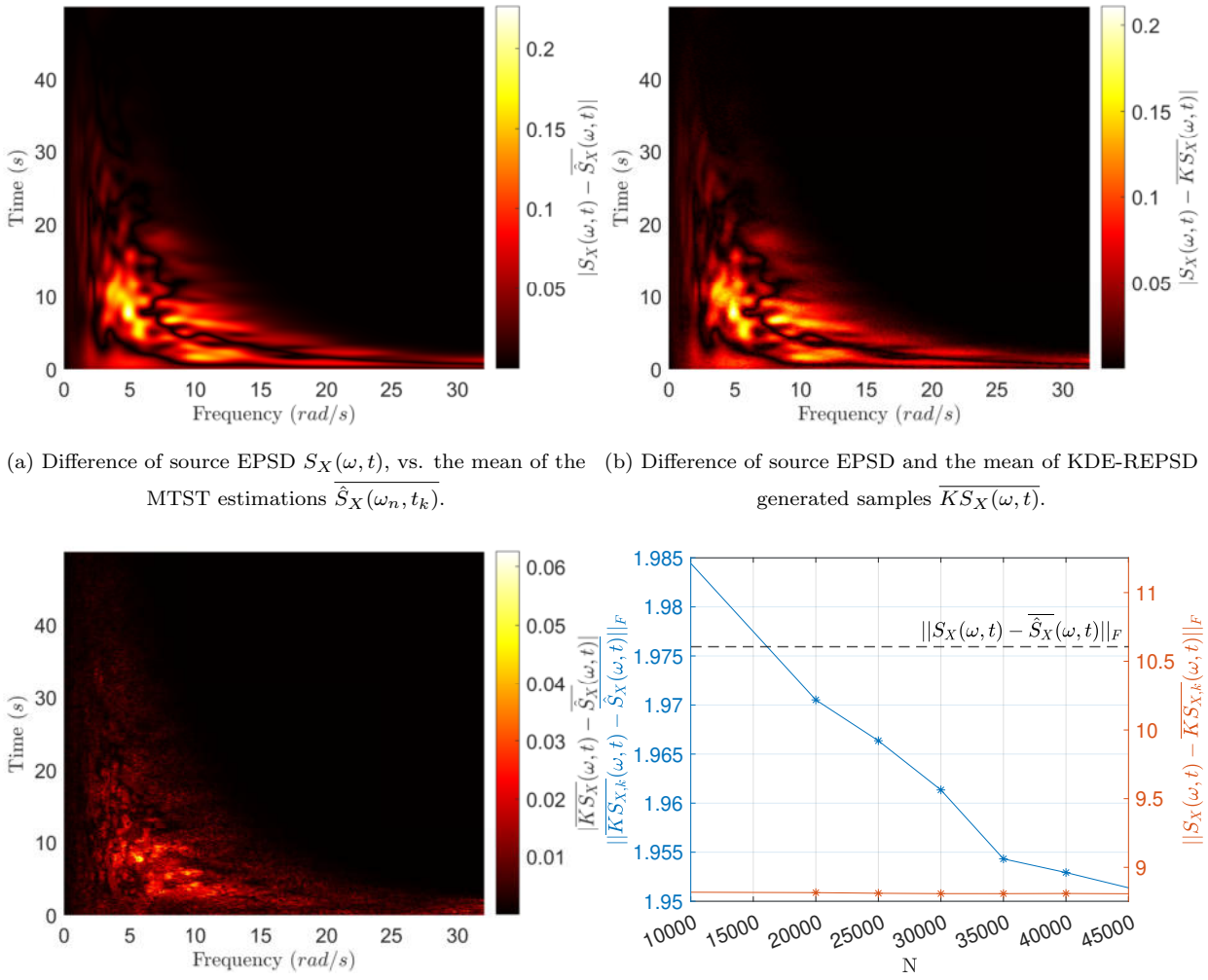


(c) Difference of the mean of the MTST estimation and the mean of TN-REPSD samples.



(d) Three Frobenius norm residuals, the black constant value belongs to the right (orange) axis.

Figure 5: Errors, differences and convergence analysis for the TN-REPSD samples.



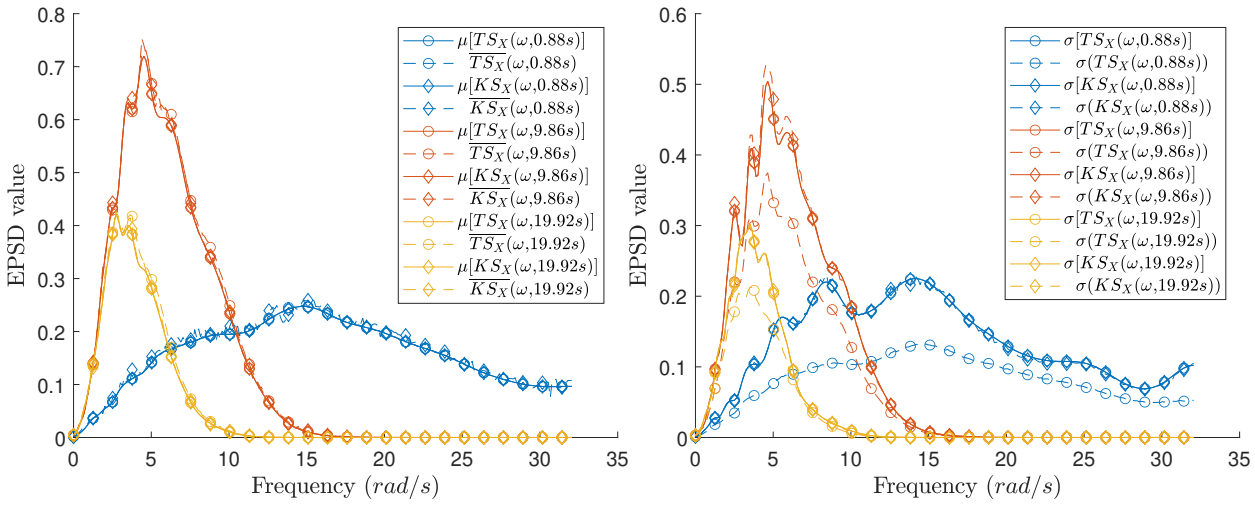
(a) Difference of source EPSD $S_X(\omega,t)$, vs. the mean of the MTST estimations $\overline{\hat{S}_X(\omega_n,t_k)}$.

(b) Difference of source EPSD and the mean of KDE-REPSD generated samples $\overline{KS_X(\omega,t)}$.

(c) Difference of the mean of the MTST estimation and the mean of KDE-REPSD samples.

(d) Three Frobenius norm residuals, the black constant value belongs to the right (orange) axis.

Figure 6: Errors, differences and convergence analysis for the KDE-REPSD samples.



(a) Mean comparison

(b) Standard deviation (std) comparison

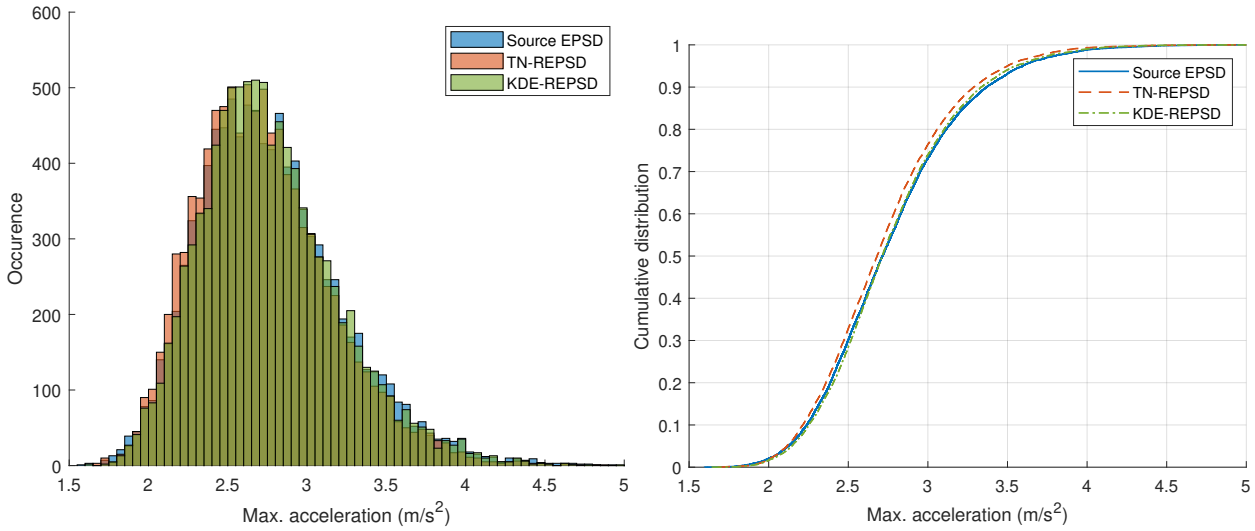
Figure 7: Stochastic modelling parameters in comparison to the MC samples statistics.

is denoted by $\overline{TS_X(\omega,t)}$ and $\overline{KS_X(\omega,t)}$, the standard deviation of the samples is denoted by $\sigma(TS_X(\omega,t))$ and $\sigma(KS_X(\omega,t))$. The mean over the frequency space Fig. 7(a) and the standard deviation over the frequency space Fig. 7(b) were compared.

3.4. Relaxed non-homogeneous spectral representation method

In this section the influence of the REPSD models to the generated time signals is evaluated. For this purpose, samples from both TN-REPSD (Eq. (12)) and the KDE-REPSD (Eq. (13)) are drawn and time signals are generated by using Eq. (1). A total of 10,000 EPSD samples and corresponding signals are compared to each other. As a reference value, time signals generated from the source EPSD given in Eq. (3) are used. The quantity considered for comparison is the absolute maximum value $\max(|\ddot{x}(t)|)$ of the respective time signal $\ddot{x}(t)$, which represents an earthquake ground motion in this case.

The respective results are given in Fig. 8, where the histograms of the three respective cases are given in Fig. 8(a) and the corresponding empirical CDF are depicted in Fig. 8(b). It can be clearly seen that the time signals generated by the three models result in a very similar behviour in terms of maximum acceleration. All histograms show a similar shape and distribution of the maximum values, which consequently is also visible in the empirical



(a) Histogram of the maximum acceleration

(b) Empirical CDF of the maximum acceleration

Figure 8: Histogram and empirical CDF of the maximum acceleration of the generated signals from the sampled EPSD.

CDFs. It should further be notet, that none of the models has some extreme values in any direction, thus the results can be considered of equal quality. However, minor differences can be seen, which may result from the influence of the random variables used in the stochastic process generation.

3.4.1. Energy of generated signals

As a further criterion for comparison, the total energy in the generated signals is considered. This analysis was carried out again for both the TN-REPSD (Eq. (12)) and the KDE-REPSD (Eq. (13)). The total energy of the signal can be determined by the following expression

$$E = \sum_{n=0}^{N_t} |x(n)|^2. \tag{16}$$

The analysis has been carried out with the identical signals generated in the previous section. In Fig. 9 an overview of the energy content of the signals is given. For better comparability, the energy content of the individual signals is arranged in ascending order. It can be clearly seen that all models result in signals with similar energy content and energy distribution.

In addition, Table 1 shows the min, max, mean and median for the 10,000 generated

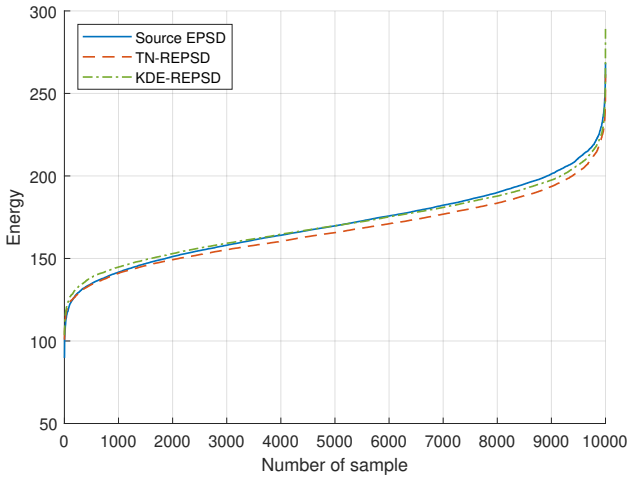


Figure 9: Energy of the generated signals in ascending order of energy content.

Table 1: Comparison of the energy content of the generated signals from the source EPSD, TN-REPSD and KDE-REPSD. Units are given in $[m^2\ s^{-4}]$.

	Source EPSD	TN-REPSD	KDE-REPSD
min	89.7147	100.771	103.5557
max	268.484	261.6281	290.2797
mean	170.8195	166.7248	170.7418
median	169.7479	165.661	169.8419

signals. The min value of the source EPSD seems to be slightly smaller than for both relaxed models and the max value of the KDE-REPSD is also somewhat higher that the other values. However, since these are the extreme values, such a behaviour can be expected. In addition, those outlier values may only be reached by a very small portion of samples. In this comparison, the mean and median are more meaningful as often a high number of samples is applied to the system under investigation. For these two values, it is clear to see that they are in the similar range. Thus, it can be concluded that the signals have an identical energy content, at least in an averaged sense.

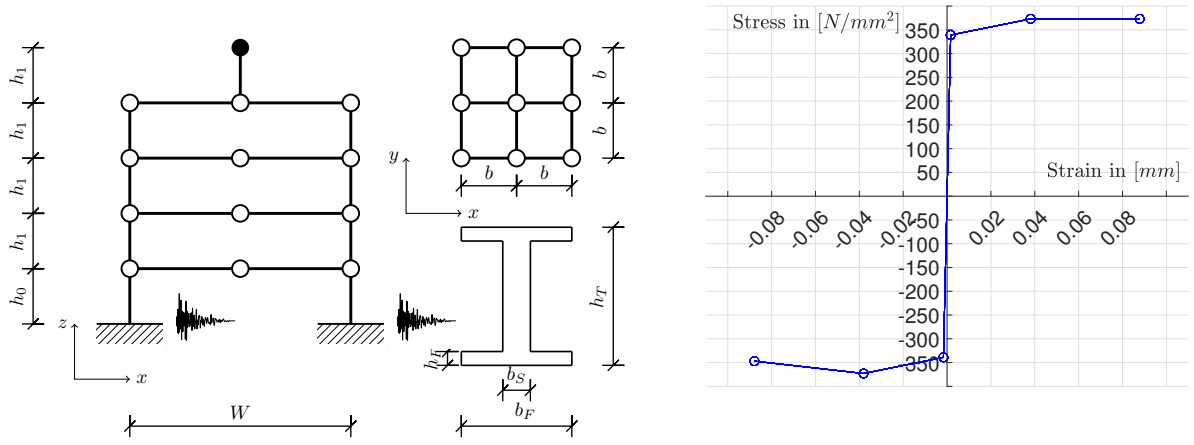
4. Numerical examples

To show possible applications of the REPSD approach, two numerical examples are investigated in this work.

4.1. Finite element model of a steel frame structure with irregular mass

Underlining the versatility of the REPSD approach, the seismic response of a steel moment resisting frame Finite Element Model (FEM) is analysed. The FEM is inspired by the results presented in [57], here a numerical model of a low rise moment steel frame with an irregular mass on the top storey has been presented and validated. In Fig. 10(a) the steel frame's dimensions, nodes and elements are depicted. White dots represent the conjunction nodes, the black dot represents a mass node. The beams and columns are consisting of H-shaped fibre steel material elements which are implemented displacement-based via the Open System for Earthquake Engineering Simulation (OpenSees) [58, 59]. Following length quantities are defined: $h_0 = 680\text{mm}$, $h_1 = 630\text{mm}$, $W = 1600\text{mm}$, $b = 800\text{mm}$, $b_F = 45\text{mm}$, $h_T = 100\text{mm}$, $b_S = 6\text{mm}$, $h_F = 8\text{mm}$. The material properties of the steel is characterised by the skeleton stress-strain curve, which defines compression and tension points, see Fig. 10(b). Within OpenSees the uniaxial material with hysteretic properties has been chosen. The stress-strain points in Fig. 10(b) define the envelope. The model's columns are fixed within the foundation assuming a damping ratio of 0.02.

For the seismic ground motion, the SRM and both REPSD approaches with the relaxed SRM approach have been tested. The source EPSD as in Eq. (3) has been chosen, the SRM and relaxed SRM generated processes are scaled down by a factor of 10. MC simulation results of the SRM generated signals are depicted in Fig. 11(a), the generated signals serve as artificial seismic ground motion signals applied to the FEM. The displacements of the FEM's 4-th storey centre node has been chosen as quantity of interest. Exemplary results for the displacement are shown in Fig. 11(b). Each artificial seismic ground motion again leads to a different EPSD estimation, as respectively depicted in Fig. 12(a). In Fig. 12(b) respectively 3 different response EPSDs of the 50 generated ensemble members are shown.



(a) FEM layout and dimensions of the OpenSEES model.

(b) Skeleton hysteresis curve, stress-strain relation.

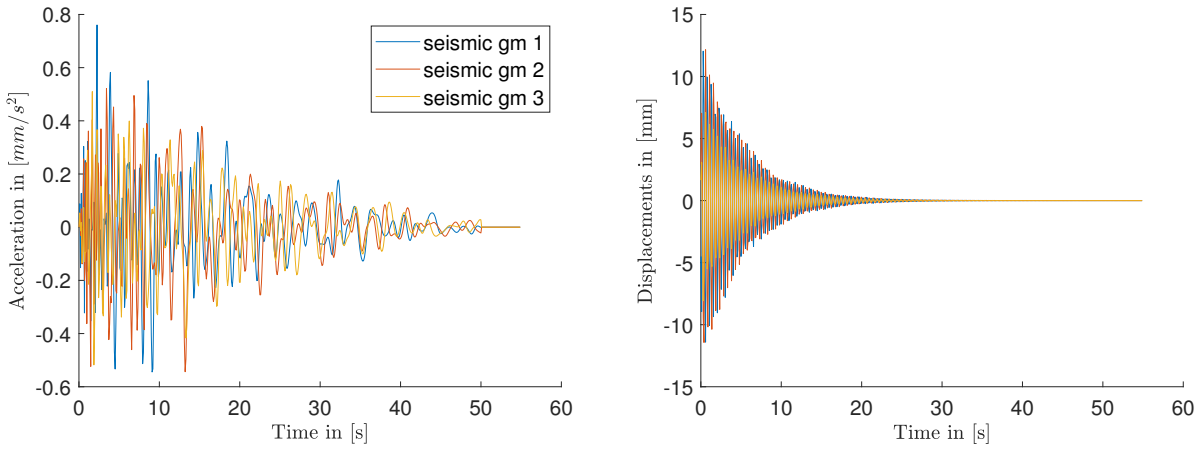
Figure 10: OpenSEES FEM dimensions and stress-strain relations of the material.

Thus the FEM input as well as output are treated as signals from which EPSDs ensembles are constructed.

For the REPSD generation, in each case the same input signal and therefore EPSD ensemble with 50 members has been used. The number of 50 reflects the limited availability of data, in this case artificial ground motion records. The same ensemble for each estimation has been used in order to ensure comparability between the TN-REPSD and KD-REPSD approach. Both REPSD models are then compared with results of 5000 MC simulations. The procedure for the input signal, is already validated in the preliminaries. Only a linear scaling was applied to the signals, which only leads to smaller EPSD function values but no change of the time-frequency components.

First a comparison of the full time history of the FEM response is carried out, the results are shown in Fig. 13(a). A closer comparison of the results for a time interval of $2s$ is shown in Fig. 13(b). In these results it can be seen that the stochastic model of the TN- and KDE-REPSD ground motion data is representing the stochastic properties of the process with satisfying accuracy.

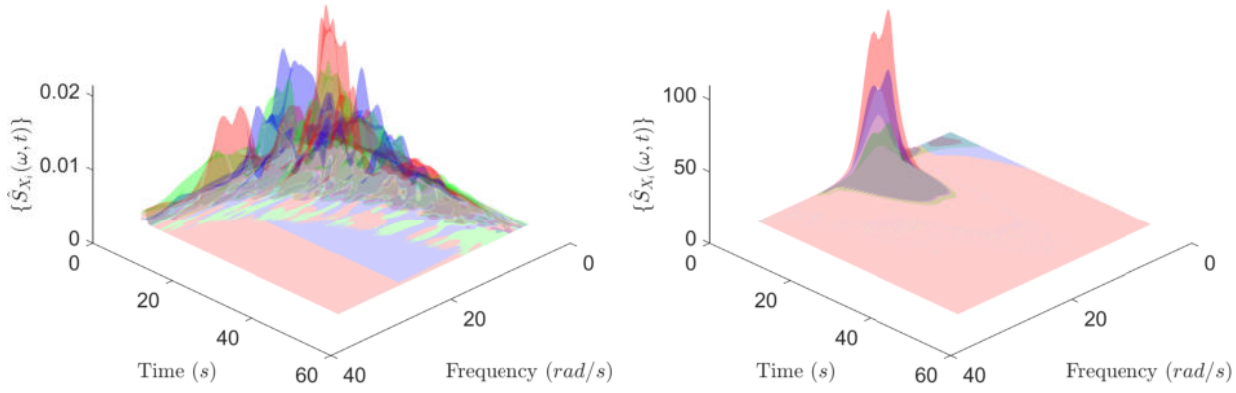
During the analysis of the FEM, no closed-form solution of the response EPSD is available, therefore the comparison with the MC simulation is considered as benchmark. To



(a) Three artificial seismic ground motion signals (obtained by SRM).

(b) Corresponding three response results of the FEM's top centre node.

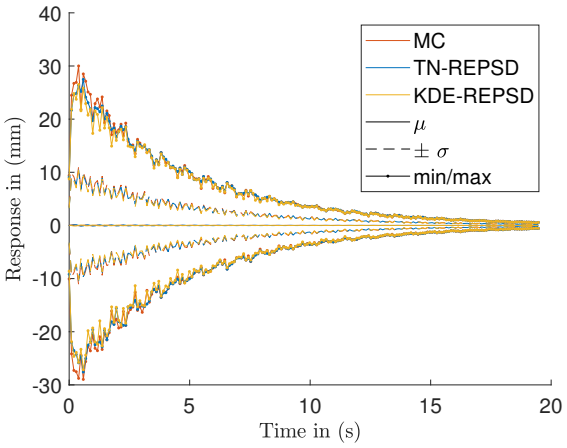
Figure 11: FEM frame structure model input and response. The abbreviation *gm* refers to ground motion.



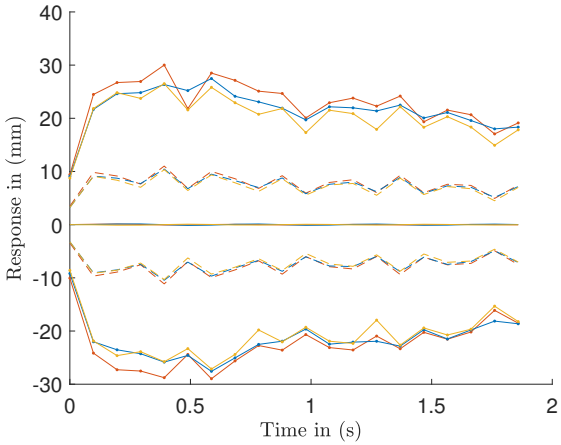
(a) Three deterministic EPSD estimations by MTST for the signals in Fig. 11(a).

(b) Corresponding three response EPSD estimations by MTST for Fig. 11(b).

Figure 12: EPSD estimations by MTST for model input (artificial seismic ground motion) and output (displacement responses).



(a) Full response statistics over time for MC, TN-REPSD and KD-REPSD generated signals

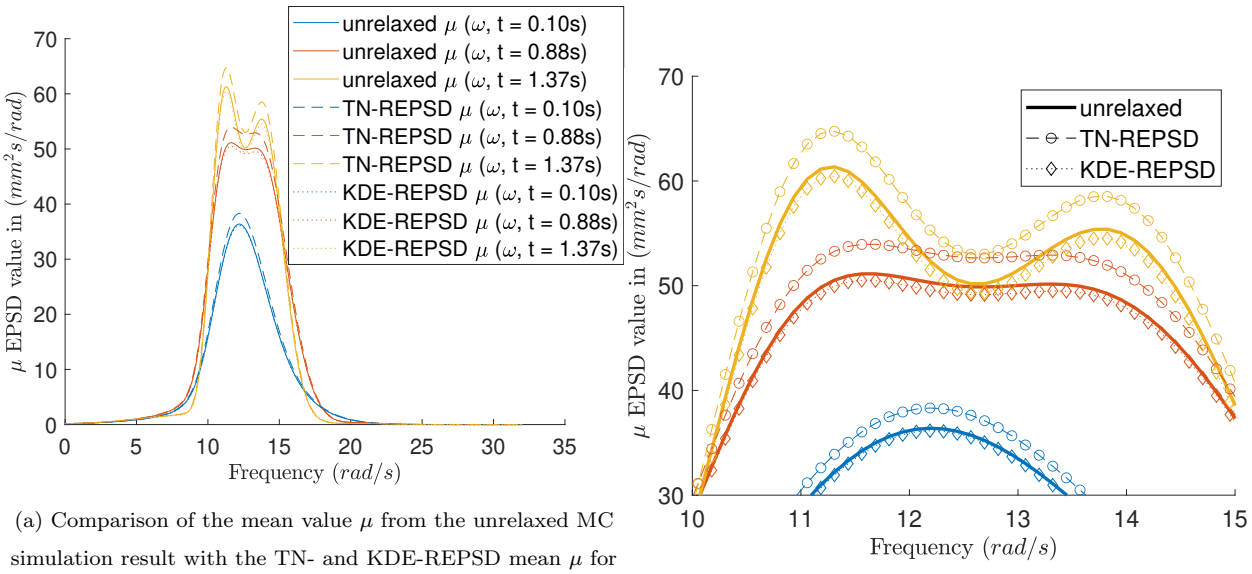


(b) 2 second interval of the response statistics

Figure 13: Full response statistics of each 5000 realised displacement histories of the FEM (both figures share the same legend)

further analyze the accuracy of the REPSD representation, for specific time steps the mean μ in Fig. 14 and standard deviation σ in Fig. 15, for the unrelaxed MC EPSD and the TN- as well as the KDE-REPSD model are regarded. Unrelaxed refers to the fact, that no stochastic model for the EPSD function is considered. Additional stochastic signals are generated via MC with the respective REPSD representation, which would be impossible if only data or records of signals were available.

In Fig. 14(a) the mean value of 5000 MC realisations and their unrelaxed MTST estimation of the response EPSD is compared with the constructed TN- and KDE-REPSD functions. Please remember that the constructed REPSD representations are estimated out of 50 ensemble members. The mean value is compared for different time instants over the full frequency range. The different time instants are depicted in a different colour. From the results in Fig. 14(a) it can be seen that the TN-REPSD function seems to overestimate the benchmark EPSD function. This is specifically true for larger EPSD function values. The KDE-REPSD representation yields a better fit. For a closer comparison the frequency interval $\omega \in [10\,,15\,]$ rad/s is regarded in Fig. 14(b). These results also show that the KDE-REPSD does yield a better representation of the mean. Here in particular for the peak



(a) Comparison of the mean value μ from the unrelaxed MC simulation result with the TN- and KDE-REPSD mean μ for different time instances.

(b) Close up for the peak frequencies mean values.

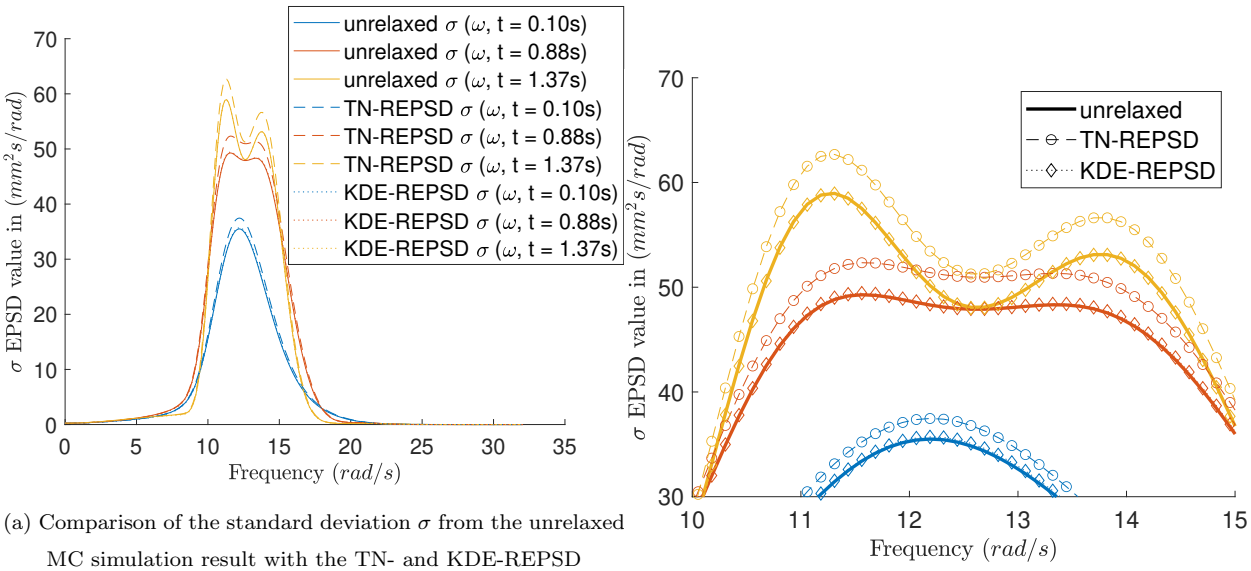
Figure 14: Comparison of mean EPSD values obtained from direct MC simulation (unrelaxed) and TN- and KDE-REPSD representation over the frequency for different time instances of interest.

EPSD values.

Analogous to the analysis of the EPSD functions mean value, the standard deviation σ is analyzed in Fig. 15. Different time instants are regarded and the corresponding EPSD function values over the frequency space are shown. The full frequency space is depicted in Fig. 15(a), where again the TN-REPSD representation is overestimating peak values of the response EPSD function standard deviation. For a closer comparison a smaller frequency interval can be seen in Fig. 15(b). These results indicate that for the standard deviation of the unrelaxed EPSD functions in comparison to the TN- and KDE-REPSD functions, the KDE-REPSD representation can reproduce the second order moment accurately from an ensemble with just 50 members.

4.2. Modulated Davenport's power spectral density function for time dependant fluctuating wind speed simulation

In this example, the empirical Davenport's spectrum for near ground wind velocities as in [60] is regarded and modulated to simulate a time-dependant change of parameters. To achieve this, the basic Davenport's PSD function is modulated by a time dependant term



(a) Comparison of the standard deviation σ from the unrelaxed MC simulation result with the TN- and KDE-REPSD standard deviation σ for different time instances.

(b) Close up for the peak frequencies.

Figure 15: Comparison of the standard deviation σ for different EPSDs obtained from direct MC simulation (unrelaxed) and TN- and KDE-REPSD representation over the frequency for different time steps of interest.

$A(t)$. Similar as in [61] Davenport's power spectral density function can be written as

$$S_X^D(\omega,t) = A(t)v_*^2 \frac{\left(\frac{1200}{2\pi A(t) U_{10}}\omega\right)^2}{|\omega|\left(1+\left(\frac{1200}{2\pi A(t) U_{10}}\omega\right)^2\right)^{4/3}}, \tag{17}$$

where $v_* = 1.691\,\mathrm{m/s}$ is the shear velocity of the wind and $U_{10} = 31.88\,\mathrm{m/s}$ is the $10-\mathrm{min}$ average wind speed in 10 m height. These are empirically estimated parameters.

A crude approach is suggested now. The hypothesis is that these two quantities could change in time. They could also have a correlation if a change in time is assumed. Without respecting the full physics of this complex coupled process at this point and only for the purpose of generating a challenging benchmark problem, we can assume that a time modulation function of these quantities exists. This time modulation is denoted as $A(t)$ and formulated by the following function

$$A(t) = \left|\sin\left(\frac{1}{2}t\right)\right| + \frac{1}{2}. \tag{18}$$

With this simple relation, an oscillatory change of the shear velocity and the the $10-\mathrm{min}$

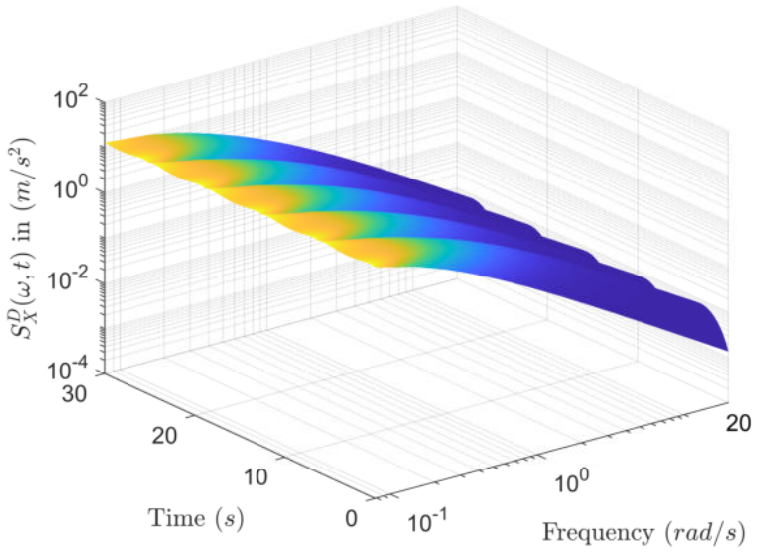


Figure 16: Modulated Davenport's separable EPSD function Eq. (3) for $\omega \in [0\,,20\,]$ rad/s and $t \in [0\,,30\,]$ s

average wind speed in 10m height is modelled. This modulation leads to a so-called separable EPSD function. This modulated Davenport's power spectral density function is denoted as $S_X^D(\omega, t)$ and depicted in Fig. 16. Please note that this is an artificial benchmark problem, this modulation has not been tested or validated on any real processes connected to Davenport's PSD function. The resulting wind speed signals and their magnitude, are of course highly dependant on the modulation function. And in the spirit of the original Davenport's approach, should be validated through experiments empirically.

In this work, since the EPSD function is now fully analytically available, the SRM approach can be used to model stochastic processes that are non-homogeneous in time and frequency. Following parameters are used for the SRM and the domain specification: The total simulation time $T^D = 30\,\text{s}$, the upper cutoff-frequency $\omega_u^D = 20\,\text{rad/s}$, the number of discretised time steps $N_t^D = 512$, resulting in $\Delta t^D = T^D/N_t^D$. SRM yields again the artificial record set of 50 signals, depicted in Fig. 17. From these signals in Fig. 17, respectively the EPSD function ensemble members are estimated via the MTST approach. The same optimisation procedure as presented in Section 3.2 has been performed resulting in following MTST parameters: $M^D = 2$, $a^D = 0.4104$, $b^D = 25.8864$, $c^D = 0.3738$. Three EPSD estimations

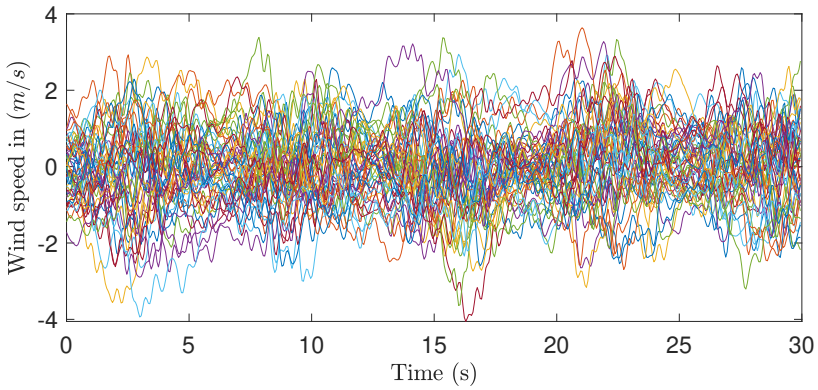
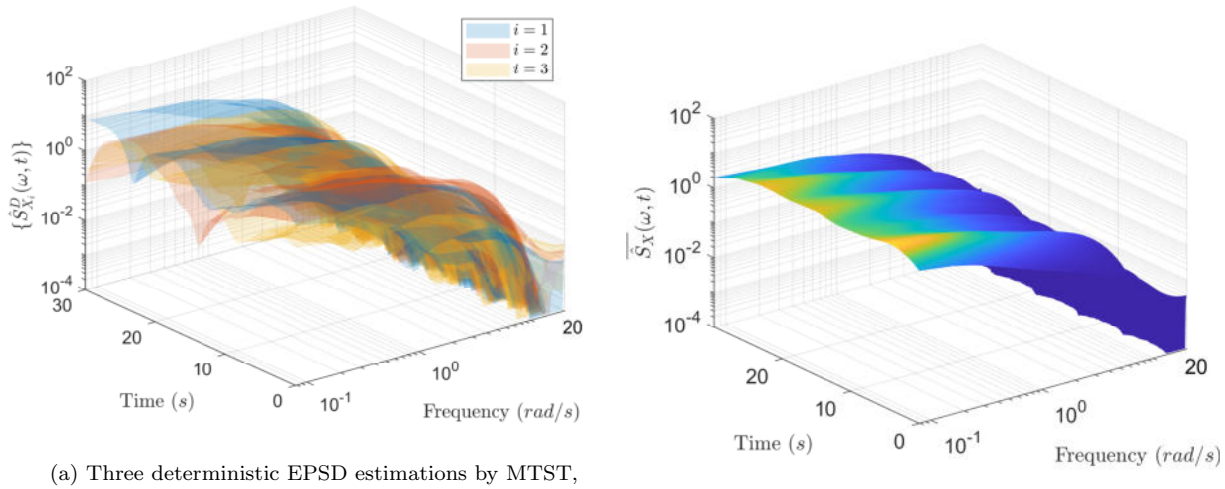


Figure 17: 50 artificial wind speed signals of the modulated Davenport's PSD (obtained by SRM)

by MTST are depicted in Fig. 18(a) and the MTST estimation mean of the ensemble is represented in Fig. 18(b). The deterministic results as well as the statistical mean of the MTST estimations in Fig. 18 suggest that the variance and accuracy of the MTST estimation is limited. The source EPSD function in Eq. (17) represents a highly nonlinear relation, therefore the direct relation of the SRM and the MTST estimation of the EPSD functions can introduce further errors, which are not the research focus of this work. The modulated Davenport's PSD function can pose a challenging benchmark example for signal generation and EPSD function estimation procedures.

In this work, it is assumed that due to a limited amount of data, the estimation by MTST is the only information of the simulated signals in Fig. 17 that is present. However, applying other EPSD estimation procedures like Wavelet, STFT, Hilbert-Huang, or surrogate modelling techniques is beyond the scope of this work. This means, that the information present in Fig. 18 is considered to be the baseline. The hypothesis is, that for some natural processes like earthquakes, wind speed loading and sea wave loading, no source EPSD function is available or known. Only measurements and records are available.

The goal is now to establish an accurate relaxed representation of the EPSD function information within the ensemble, displayed in Fig. 18. This ensemble is hypothetically stemming purely from measurements and records. To achieve the relaxed representation of the EPSD function, first, the two presented REPSD representations TN-REPSD denoted by $T\hat{S}_X^D(\omega, t)$ and KDE-REPSD, denoted by $K\hat{S}_X^D(\omega, t)$ are constructed as discussed in the preliminar-



(a) Three deterministic EPSD estimations by MTST, representative ensemble members

(b) Mean of the MTST EPSD ensemble

Figure 18: The MTST estimation of the EPSD for the by SRM and Eq. (17) generated wind speed signals in Fig. 17

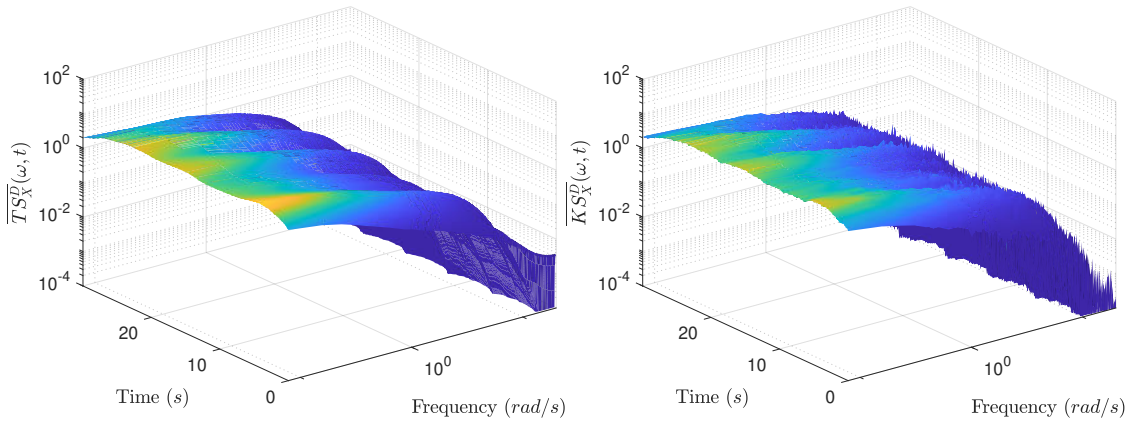
ies. Then the two representations are analysed for the accuracy of the MTST estimations. Therefore, a MC simulation is performed to generate 1000 new EPSD samples from the respective REPSD functions. From these samples, the deviation towards the MTST estimation is analysed. These results are compiled in Fig. 19. Note that for the KDE approach in this example, a manual adjustment of the quartiles, introduced in Eq. (10), was necessary to reduce the modelling error: $Q1 = 35.0646$ and $Q2 = 45.0650$, were chosen. In particular when looking at the mean of the 1000 generated MC samples for the TN-REPSD approach in Fig. 19(a) and the KDE-REPSD approach in Fig. 19(b) it is observable, that the TN-REPSD approach does model the mean of the ensemble better. The aforementioned adjustment of the quartiles was necessary because outliers of the ensemble were weighted too heavily into the establishment of the KDE. For the mean of the generated samples in Fig. 19(b), still some outliers can be identified. From this example it seems that the mean estimation of the TN-REPSD approach is more robust. This is also reflected in the error comparison. Here the mean of the samples for TN-REPSD are compared with the mean of the MTST estimated ensemble in Fig. 19(c), and the KDE-REPSD ensemble mean in Fig. 19(d). The KDE-REPSD representation does exhibit larger errors over the whole do-

main. A contrary result can be observed for the standard deviation of the generated relaxed samples in comparison to the standard deviation of the MTST EPSD function ensemble. The TN-REPSD generated samples exhibit large areas of a larger error of the standard deviation, see Fig. 19(e). The KDE-REPSD generated samples on the other hand, even though having a similar maximum error range, exhibit errors in a smaller area Fig. 19(f). This benchmark example highlights the importance of the choice of the underlying distribution type for the relaxed representation of EPSD functions. Since for the TN-REPSD only the mean value influences the parameter of the standard deviation (Eq. (11b)), no accurate representation of varying data can be achieved. The KDE approach, with regard to an adjusted IQR proofs to be more robust in terms of representing varying data.

All in all it must be pointed out that any relaxed representation of EPSD functions is highly dependant on the EPSD function estimation procedure. For the proposed approaches in this example, the REPSD (of any distribution type) can only be as good as the EPSD estimation procedure. But several techniques could be applied to improve the stochastic model beneath the REPSD representations, such as Bayesian updating procedures, surrogate representations or adaptive sampling approaches.

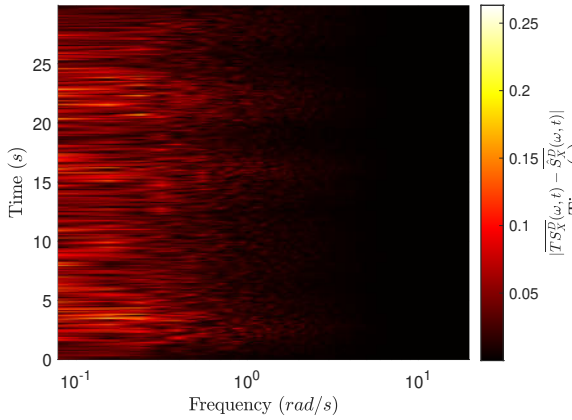
5. Conclusions

In this work, a relaxed representation of EPSD functions for natural phenomena modelled by stochastic processes has been proposed. The scope of this work was built around the assumption, that for a battery of natural processes (e.g. seismic ground motion signals, wind speed signals, wave load signals, or random vibrations in general) no prior information about the source EPSD functions exists. However, data in the form of measurements and records is available. It is possible to estimate EPSD functions deterministically by time-frequency transformation methods, such as wavelet transformation, STFT, HHT, or the recently developed MTST method. These transformations are only estimators and usually are underlying a large variance, especially for complex natural processes. The REPSD function representation allows for the straightforward construction of a stochastic model of the EPSD function, given data for the analysed processes is available. Not only the

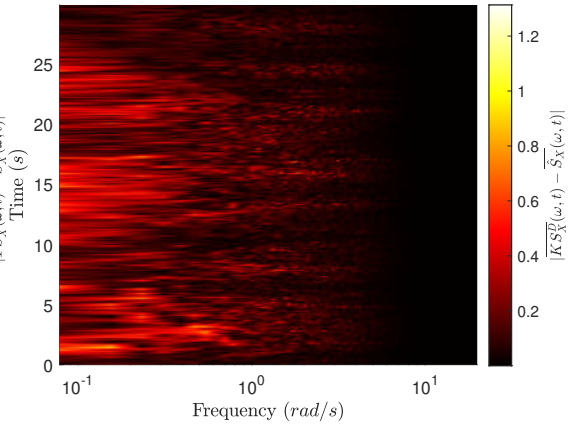


(a) TN-REPSD samples mean

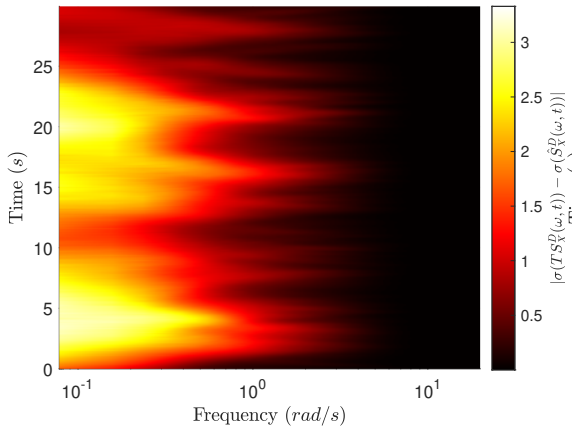
(b) KDE-REPSD samples mean



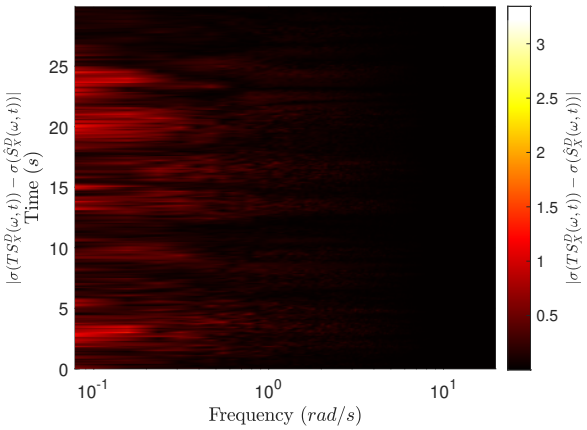
(c) Difference of TN-REPSD samples mean and MTST ensemble mean.



(d) Difference of KDE-REPSD samples mean and MTST ensemble mean.



(e) Difference of TN-REPSD samples standard deviation and MTST ensemble standard deviation σ.



(f) Difference of KDE-REPSD samples standard deviation and MTST ensemble standard deviation σ.

Figure 19: Error analysis of the relaxed representations of the MTST estimated EPSD ensemble for the modulated Davenport's PSD.

choice of time-frequency transformation method is modular, but also the underlying distribution type for the REPSD functions can be adaptively chosen to the specific problem at hand. For two simple approaches, a TN and KDE distribution type, this procedure has been presented in this work. In addition, it has been shown the the presented method is suitable for non-separable as well as separable EPSD functions. Convergence studies have been carried out to validate the results and analyse the accuracy. Additionally, the standard SRM method has been extended to a relaxed SRM approach, to be able to generate new stochastic signals, from the REPSD model. However, at several points in this work, it has been shown that the EPSD construction is not only dependant on the choice of the underlying distribution type, but is also heavily reliant on the accuracy of the time-frequency transformation method. An optimisation procedure for the MTST method has been proposed, which is applicable if further information on the EPSD function space is available. The REPSD formulation is applied to practical benchmark examples, where a FEM of an multiple degree of freedom system is analysed, here the input, as well as the response, are then modelled using the relaxed approach. Additionally, an academic benchmark example of a separable EPSD has been established, based on the empirical Davenport's PSD function, which has been modulated to incorporate temporal features. The relaxed approach is used to construct a model that represents a relaxed stochastic EPSD for signals generated by the modulated PSD. Overall, it can be stated that the REPSD representation's accuracy is as good as its underlying time-frequency transformation method. Nonetheless, the REPSD offers a stochastic representation of a stochastic processes model, already for a limited amount of data and records. The accuracy of these terms should further be validated by experiments and empirical investigations. Additionally, further information such as a correlation between the REPSD sample points could be considered. Also, a Bayesian updating procedure could be incorporated if the data set is of variable size, changing, or if additional parameters e.g. for the correlation are introduced. It is also possible to replace the probability distribution approach with surrogate modelling approaches, such as Gaussian process regression or Neural-Network-Representations. Thus, by the introduction of the REPSD concept, a generalised modular stochastic model for the representation of stochastic

processes is established.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT author statement

Marius Bittner: Methodology, Formal analysis, Visualization, Validation, Software, Writing - Original Draft. **Marco Behrendt:** Conceptualization, Methodology, Visualization, Validation, Software, Writing - Review & Editing. **Michael Beer:** Project administration, Supervision, Conceptualization, Funding acquisition, Writing - Review & Editing.

Acknowledgements

This work was supported by the the German Research Foundation (DFG) within the framework of the International Research Training Group on Computational Mechanics Techniques in High Dimensions GRK 2657, Grant Number 433082294.

References

[1] A. K. Chopra, Dynamics of Structures: Theory and Applications to Earthquake Engineering, Englewood Cliffs, N.J: Prentice Hall, 1995.

[2] Y.-K. Lin, G.-Q. Cai, Probabilistic Structural Dynamics: Advanced Theory and Applications, McGraw-Hill New York, 1995.

[3] J. Li, J. Chen, Stochastic Dynamics of Structures, John Wiley & Sons, 2009.

[4] A. Powell, S. Crandall, Random Vibration, The Technology Press of the Massachusetts Institute of Technology, Cambridge, 1958.

[5] L. D. Lutes, S. Sarkani, Random Vibrations: Analysis of Structural and Mechanical Systems, Butterworth-Heinemann, 2004.

[6] T. Soong, M. Grigoriu, Random Vibration of Mechanical and Structural Systems, PTR Prentice Hall, 1993.

[7] D. E. Newland, An introduction to random vibrations, spectral & wavelet analysis, Courier Corporation, 2012.

[8] G. I. Schuëller, M. Shinozuka, Stochastic Methods in Structural Dynamics, Vol. 10 of Mechanics: Dynamical Systems, Springer Science & Business Media, 2012.

[9] M. Grigoriu, Evaluation of karhunen-loève, spectral, and sampling representations for stochastic processes, Journal of Engineering Mechanics 132 (2) (2006) 179 – 189. doi:10.1061/(ASCE)0733-9399(2006)132:2(179).
URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-30944435311&doi=10.1061%2f%28ASCE%290733-

[10] R. G. Ghanem, P. D. Spanos, Stochastic Finite Elements: A Spectral Approach, Springer-Verlag, Berlin, Heidelberg, 1991.

[11] H. Dai, Z. Zheng, H. Ma, An explicit method for simulating non-gaussian and non-stationary stochastic processes by karhunen-loève and polynomial chaos expansion, Mechanical Systems and Signal Processing 115 (2019) 1–13. doi:https://doi.org/10.1016/j.ymssp.2018.05.026.
URL https://www.sciencedirect.com/science/article/pii/S0888327018302796

[12] M. Broccardo, A. D. Kiureghian, Simulation of stochastic processes by sinc basis functions and application in telm analysis, Journal of Engineering Mechanics 144 (1) (2018) 04017154. arXiv:https://ascelibrary.org/doi/pdf/10.1061/%28ASCE%29EM.1943-7889.0001374, doi:10.1061/(ASCE)EM.1943-7889.0001374.
URL https://ascelibrary.org/doi/abs/10.1061/%28ASCE%29EM.1943-7889.0001374

[13] M. Shinozuka, G. Deodatis, Simulation of stochastic processes by spectral representation, Applied Mechanics Reviews 44 (4) (1991) 191–204. doi:10.1115/1.3119501.

[14] P. Welch, The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms, IEEE Transactions on audio and electroacoustics 15 (2) (1967) 70–73.

[15] M. S. Bartlett, Smoothing periodograms from time-series with continuous spectra, Nature 161 (4096) (1948) 686–687.

[16] M. S. Bartlett, Periodogram analysis and continuous spectra, Biometrika 37 (1-2) (1950) 1–16. arXiv:https://academic.oup.com/biomet/article-pdf/37/1-2/1/486591/37-1-2-1.pdf, doi:10.1093/biomet/37.1-2.1.
URL https://doi.org/10.1093/biomet/37.1-2.1

[17] R. A. Muller, G. J. MacDonald, Ice ages and astronomical causes: data, spectral analysis and mechanisms, Springer Science & Business Media, 2002.

[18] I. Sneddon, Fourier Transforms, Dover Books on Mathematics, Dover Publications, 1995.

[19] M. B. Priestley, Evolutionary spectra and non-stationary processes, Journal of the Royal Statistical

Society: Series B (Methodological) 27 (2) (1965) 204–229.

[20] M. Priestley, Power spectral analysis of non-stationary random processes, Journal of Sound and Vibration 6 (1) (1967) 86–97.

[21] M. Priestley, Spectral Analysis and Time Series, Probability and mathematical statistics : A series of monographs and textbooks, Academic Press, 1982.

[22] E. Sejdić, I. Djurović, J. Jiang, Time–frequency feature representation using energy concentration: An overview of recent advances, Digital Signal Processing 19 (1) (2009) 153–183. doi:https://doi.org/10.1016/j.dsp.2007.12.004.
URL https://www.sciencedirect.com/science/article/pii/S105120040800002X

[23] P. D. Spanos, G. Failla, Wavelets: Theoretical concepts and vibrations related applications, The Shock and vibration digest 37 (5) (2005) 359–376.

[24] P. D. Spanos, J. Tezcan, P. Tratskas, Stochastic processes evolutionary spectrum estimation via harmonic wavelets, Computer Methods in Applied Mechanics and Engineering 194 (12-16) (2005) 1367–1383.

[25] Z. Huang, Y.-L. Xu, T. Tao, Multi-taper S-transform method for evolutionary spectrum estimation, Mechanical Systems and Signal Processing 168 (2022) 108667. doi:https://doi.org/10.1016/j.ymssp.2021.108667.
URL https://www.sciencedirect.com/science/article/pii/S0888327021009912

[26] J. Liang, S. R. Chaudhuri, M. Shinozuka, Simulation of nonstationary stochastic processes by spectral representation, Journal of Engineering Mechanics 133 (6) (2007) 616–627.

[27] E. Zio, The Monte Carlo Simulation Method for System Reliability and Risk Analysis, Springer London, 2013. doi:https://doi.org/10.1007/978-1-4471-4588-2.

[28] G. I. Schuëller, Efficient Monte Carlo simulation procedures in structural uncertainty and reliability analysis - recent advances, Structural Engineering and Mechanics 32 (1) (2009) 1–20.

[29] S.-K. Au, J. L. Beck, Estimation of small failure probabilities in high dimensions by subset simulation, Probabilistic Engineering Mechanics 16 (4) (2001) 263–277. doi:https://doi.org/10.1016/S0266-8920(01)00019-4.
URL https://www.sciencedirect.com/science/article/pii/S0266892001000194

[30] M. de Angelis, E. Patelli, M. Beer, Advanced Line Sampling for efficient robust reliability analysis, Structural Safety 52 (2015) 170–182, engineering Analyses with Vague and Imprecise Information. doi:https://doi.org/10.1016/j.strusafe.2014.10.002.
URL https://www.sciencedirect.com/science/article/pii/S0167473014000927

[31] D. J. Jerez, H. A. Jensen, M. A. Valdebenito, M. A. Misraji, F. Mayorga, M. Beer, On the use of directional importance sampling for reliability-based design and optimum design sen-

sitivity of linear stochastic structures, Probabilistic Engineering Mechanics 70 (2022) 103368. doi:https://doi.org/10.1016/j.probengmech.2022.103368.
URL https://www.sciencedirect.com/science/article/pii/S0266892022001011

[32] G. Schuëller, On the treatment of uncertainties in structural mechanics and analysis, Computers & Structures 85 (5) (2007) 235–243, computational Stochastic Mechanics. doi:https://doi.org/10.1016/j.compstruc.2006.10.009.
URL https://www.sciencedirect.com/science/article/pii/S0045794906003348

[33] A. D. Kiureghian, O. Ditlevsen, Aleatory or epistemic? Does it matter?, Structural Safety 31 (2) (2009) 105–112, risk Acceptance and Risk Communication. doi:https://doi.org/10.1016/j.strusafe.2008.06.020.
URL https://www.sciencedirect.com/science/article/pii/S0167473008000556

[34] M. G. Faes, M. A. Valdebenito, D. Moens, M. Beer, Operator norm theory as an efficient tool to propagate hybrid uncertainties and calculate imprecise probabilities, Mechanical Systems and Signal Processing 152 (2021) 107482. doi:https://doi.org/10.1016/j.ymssp.2020.107482.
URL https://www.sciencedirect.com/science/article/pii/S0888327020308682

[35] C. Dang, P. Wei, M. G. Faes, M. Beer, Bayesian probabilistic propagation of hybrid uncertainties: Estimation of response expectation function, its variable importance and bounds, Computers & Structures 270 (2022) 106860. doi:https://doi.org/10.1016/j.compstruc.2022.106860.
URL https://www.sciencedirect.com/science/article/pii/S0045794922001201

[36] M. Grigoriu, Stochastic Calculus: Applications in Science and Engineering, Springer, 2002. doi:https://doi.org/10.1007/978-0-8176-8228-6.

[37] M. Grigoriu, Stochastic systems: Uncertainty Quantification and Propagation, Springer Science & Business Media, 2012. doi:https://doi.org/10.1007/978-1-4471-2327-9.

[38] M. G. Faes, D. Moens, Recent trends in the modeling and quantification of non-probabilistic uncertainty, Archives of Computational Methods in Engineering 27 (2020) 633–671. doi:https://doi.org/10.1007/s11831-019-09327-x.
URL https://link.springer.com/article/10.1007/s11831-019-09327-x

[39] M. Beer, S. Ferson, V. Kreinovich, Imprecise probabilities in engineering analyses, Mechanical Systems and Signal Processing 37 (1) (2013) 4 – 29. doi:https://doi.org/10.1016/j.ymssp.2013.01.024.
URL http://www.sciencedirect.com/science/article/pii/S0888327013000812

[40] L. Comerford, I. A. Kougioumtzoglou, M. Beer, An artificial neural network approach for stochastic process power spectrum estimation subject to missing data, Structural Safety 52 (2015) 150–160. doi:https://doi.org/10.1016/j.strusafe.2014.10.001.
URL https://www.sciencedirect.com/science/article/pii/S0167473014000915

[41] L. Comerford, I. A. Kougioumtzoglou, M. Beer, On quantifying the uncertainty

of stochastic process power spectrum estimates subject to missing data, International Journal of Sustainable Materials and Structural Systems 2 (1-2) (2015) 185–206. arXiv:https://www.inderscienceonline.com/doi/pdf/10.1504/IJSMSS.2015.078358, doi:10.1504/IJSMSS.2015.078358.
URL https://www.inderscienceonline.com/doi/abs/10.1504/IJSMSS.2015.078358

[42] L. Comerford, I. A. Kougioumtzoglou, M. Beer, Compressive sensing based stochastic process power spectrum estimation subject to missing data, Probabilistic Engineering Mechanics 44 (2016) 66–76. doi:https://doi.org/10.1016/j.probengmech.2015.09.015.
URL https://www.sciencedirect.com/science/article/pii/S0266892015300436

[43] M. Behrendt, M. G. Faes, M. A. Valdebenito, M. Beer, Estimation of an imprecise power spectral density function with optimised bounds from scarce data for epistemic uncertainty quantification, Mechanical Systems and Signal Processing 189 (2023) 110072. doi:https://doi.org/10.1016/j.ymssp.2022.110072.
URL https://www.sciencedirect.com/science/article/pii/S0888327022011402

[44] G. Muscolino, F. Genovese, A. Sofi, Reliability Bounds for Structural Systems Subjected to a Set of Recorded Accelerograms Leading to Imprecise Seismic Power Spectrum, ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering 8 (2) (2022) 04022009. doi:10.1061/AJRUA6.0001215.

[45] M. Behrendt, M. de Angelis, L. Comerford, Y. Zhang, M. Beer, Projecting interval uncertainty through the discrete fourier transform: An application to time signals with poor precision, Mechanical Systems and Signal Processing 172 (2022) 108920. doi:https://doi.org/10.1016/j.ymssp.2022.108920.
URL https://www.sciencedirect.com/science/article/pii/S0888327022001054

[46] T. Kishida, V. Contreras, Y. Bozorgnia, N. A. Abrahamson, Ahdi, T. S.K., Ancheta, D. Boore, K. Campbell, B. Chiou, R. Darragh, N. Gregor, N. Kuehn, D. Kwak, A. Kwok, P. Lin, H. Magistrale, S. Maazoni, S. Muin, S. Midorikawa, H. Si, W. Silva, J. Stewart, K. Wooddell, R. R. Youngs, Nga-sub ground motion database, Proceedings of the Eleventh U.S. National Conference on Earthquake Engineering (2018).

[47] C. A. Goulet, T. Kishida, T. D. Ancheta, C. H. Cramer, R. B. Darragh, W. J. Silva, Y. M. Hashash, J. Harmon, G. A. Parker, J. P. Stewart, R. R. Youngs, Peer nga-east database, Earthquake Spectra 37 (1_suppl) (2021) 1331–1353. arXiv:https://doi.org/10.1177/87552930211015695, doi:10.1177/87552930211015695.
URL https://doi.org/10.1177/87552930211015695

[48] T. Kishida, V. Contreras, Y. Bozorgnia, N. A. Abrahamson, S. K. Ahdi, T. D. Ancheta, D. M. Boore, K. W. Campbell, B. S. Chiou, R. B. Darragh, et al., NGA-sub ground motion database, UCLA (2021).
URL https://escholarship.org/uc/item/3bn528xc

[49] M. Behrendt, M. Bittner, L. Comerford, M. Beer, J. Chen, Relaxed power spectrum estimation from multiple data records utilising subjective probabilities, Mechanical Systems and Signal Processing 165 (2022) 108346. doi:https://doi.org/10.1016/j.ymssp.2021.108346.
URL https://www.sciencedirect.com/science/article/pii/S0888327021007020

[50] P. D. Spanos, I. Kougioumtzoglou, Harmonic wavelets based statistical linearization for response evolutionary power spectrum determination, Probabilistic Engineering Mechanics 27 (1) (2012) 57–68, the IUTAM Symposium on Nonlinear Stochastic Dynamics and Control. doi:https://doi.org/10.1016/j.probengmech.2011.05.008.
URL https://www.sciencedirect.com/science/article/pii/S0266892011000294

[51] Z. Huang, Y.-L. Xu, A multi-taper s-transform method for spectral estimation of stationary processes, IEEE Transactions on Signal Processing 69 (2021) 1452–1467.

[52] B. W. Silverman, Density estimation for statistics and data analysis, Vol. 26, CRC press, 1986.

[53] D. W. Scott, Multivariate density estimation: theory, practice, and visualization, John Wiley & Sons, 2015.

[54] A. W. Bowman, A. Azzalini, Applied smoothing techniques for data analysis: the kernel approach with S-Plus illustrations, Vol. 18, OUP Oxford, 1997.

[55] D. W. Scott, On optimal and data-based histograms, Biometrika 66 (3) (1979) 605–610. doi:10.1093/biomet/66.3.605.
URL https://doi.org/10.1093/biomet/66.3.605

[56] D. E. Goldberg, Genetic Algorithms in Search, Optimization and Machine Learning, 1st Edition, Addison-Wesley Longman Publishing Co., Inc., USA, 1989.

[57] Y. Bai, Y. Li, Z. Tang, M. Bittner, M. Broggi, M. Beer, Earthquake-induced collapse reliability of low-rise steel moment frames with additive-mass based on shaking table test, Bulletin of Earthquake Engineering 19 (2021) 2457–2482.

[58] F. McKenna, M. H. Scott, G. L. Fenves, Nonlinear finite-element analysis software architecture using object composition, Journal of Computing in Civil Engineering 24 (1) (2010) 95–107.

[59] F. McKenna, OpenSees: A Framework for Earthquake Engineering Simulation, Computing in Science & Engineering 13 (4) (2011) 58–66. doi:10.1109/MCSE.2011.66.

[60] A. G. Davenport, The spectrum of horizontal gustiness near the ground in high winds, Quarterly Journal of the Royal Meteorological Society 87 (372) (1961) 194–211. arXiv:https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.49708737208, doi:https://doi.org/10.1002/qj.49708737208.
URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.49708737208

[61] J. Chen, Y. Song, Y. Peng, P. D. Spanos, Simulation of homogeneous fluctuating wind field in two

spatial dimensions via a joint wave number–frequency power spectrum, Journal of Engineering Mechanics 144 (11) (2018) 04018100. arXiv:https://ascelibrary.org/doi/pdf/10.1061/%28ASCE%29EM.1943-7889.0001525, doi:10.1061/(ASCE)EM.1943-7889.0001525.
URL https://ascelibrary.org/doi/abs/10.1061/%28ASCE%29EM.1943-7889.0001525

Appendix A. First and second order moments for matrices

Given the nature of the time-frequency transformation of stochastic processes, the available data of the EPSD functions in the set $\{\hat{S}_{X_i}\}$ appear in matrix form. For the sake of completeness here are the formal calculations of the first and second order moments. The first order moment is given to be

$$\overline{S_X(\omega,t)} = \frac{1}{N_e}\sum_{i=1}^{N_e} S_{X_i}(\omega,t). \tag{A.1}$$

The second order moment is given to be

$$\sigma\left(S_X(\omega,t)\right) = \sqrt{\frac{1}{N_e}\sum_{i=1}^{N_e}\left(S_{X_i}(\omega,t) - \overline{S_X(\omega,t)}\right)}. \tag{A.2}$$