Probabilistic-based approach using Kernel Density Estimation for gap modeling in a statistical tolerance analysis

Edoh GOKA[a,b], Pierre BEAUREPAIRE[b], Lazhar HOMRI[a], Jean-Yves DANTAN[a]

[a] *Arts et Métiers–ParisTech, Université de Lorraine, LCFC, F-57000 Metz, France*
[b] *University of Clermont Auvergne, CNRS, SIGMA-Clermont, Institut Pascal, F-63000 Clermont Ferrand, France*

Abstract

The statistical tolerance analysis has become a key element used in the design stage to reduce the manufacturing cost, the rejection rate and to have high quality products. One of the frequently used methods is the Monte Carlo simulation, employed to compute the non-conformity rate due to its efficiency in handling the tolerance analysis of over-constrained mechanical systems. However, this simulation technique requires excessive numerical efforts. The goal of this paper is to improve this method by proposing a probabilistic model of gaps in fixed and sliding contacts and involved in the tolerance analysis of an assembly. The probabilistic model is carried out on the clearance components of the sliding and fixed contacts for their assembly feasibility considering all the imperfections on the surfaces. The kernel density estimation method is used to deal with the probabilistic model. The proposed method is applied to an over-constrained mechanical system and compared to the classical method regarding their computation time.

1. Introduction

In a competitive industrial area, companies work continuously to improve the quality of their products, reduce the manufacturing costs and decrease the number of defective products. Tolerance analysis is a key tool integrated into the design phase of the manufacturing products, which allows designers to achieve such goals [1, 2]. Tolerance analysis aims to verify if the individual tolerance specified to each part in the product allows the designer to reach the assembly feasibility and the functional requirements. Two methods for the tolerance analysis of mechanical assemblies are usually used: the worst-case and the statistical approaches. The worst-case tolerancing often leads to tight tolerances and high production costs. The statistical tolerancing is widely used because it leads to large tolerances and allows practitioner to get a desire yield and low production cost [3].

The statistical tolerance analysis can be achieved based on several methods [3–5]: the root-sum-squares method, Croft's method, extended Taylor series approximation, Hasofer-Lind index method, approximation by numerical integration or quadrature technique, Taguchi's method and the Monte Carlo simulation.

One of the current challenges in tolerancing is the consideration of the over-constrained mechanical systems with gaps. The assembly of such mechanical systems is possible because of the gaps [6], even though the parts exhibit geometric imperfections.

Consequently, the study of over-constrained mechanical systems is more complex than the iso-constrained mechanical systems. Several methods to study over-constrained mechanical systems have been proposed. The combination of the set point variation, homogeneous transform matrix and Monte Carlo simulation has been proposed in [7]. An alternative strategy consists of identifying the contact configurations of the mechanical assembly [8, 9]. Thus, some system reliability methods are used to compute the non-conformity rate of the over-constrained mechanical assemblies. However, for these system reliability methods, even though the gap is considered, it is up to the designer to define the possible contact configurations of the mechanical assembly, which may be a challenging task. Moreover, these methods are not able to consider parts' form defects and are globally applied on 2D cases. Furthermore, to overcome these previous limitations, a new method based on the Lagrange dual form of the optimization problem and FORM method has been proposed to compute the probability of failure with reduced computational efforts and carrying out a preliminary selective search algorithm to identify the configurations which contribute mostly to this probability of failure [6]. Though this work considers 3D over-constrained mechanical systems with multi-configuration of gaps, it has not considered the form defects nor the different types of contacts in the tolerance analysis procedure.

Therefore, this paper proposes a probabilistic model of gaps in fixed and sliding contacts, applied in a global statistical tolerance analysis method. The statistical tolerance analysis is carried out based on constrained optimization and Monte Carlo Simulation (MCS). The probabilistic model is proposed to determine the gap components in fixed and sliding contacts between non-ideal surfaces, instead of computing these components using the MCS combined with an optimization procedure. The Kernel Density Estimation (KDE) method is used to define the probabilistic model. A first optimization is carried out on the gap components of each sliding or fixed contact. Then the probabilistic model is identified based on these values gathered from this first

optimization to generate new samples in good agreement with the distribution and the correlation of the original gap-values. For the floating contacts a set of constraints is developed to prevent the inferences of the mating surfaces. Then, a general optimization is carried out to determine the values of their gap components. Finally, a Monte Carlo simulation is carried out based on all the previous developments to determine the probabilities of assembly failure and functional failure.

The paper is structured as follows: Section 2 presents the main algorithm currently used to deal with the statistical tolerance analysis and highlights the step at which the probabilistic model of gap is integrated. Section 3 presents the KDE method used to determine the probabilistic model of gaps, Section 4 describes an application of the proposed method to an over-constrained mechanical system, and finally, discussions and conclusion are given at the end.

2. Tolerance analysis of over-constrained mechanical systems

2.1 Classical tolerance analysis approach

The statistical tolerance analysis method aims to calculate the probabilities of failure of the assembly and functioning requirements of a mechanical system. Generally, three main issues are considered in the tolerance analysis procedure: the geometrical deviations modeling, the geometrical behavior modeling of the mechanical systems and the mathematical formulation of the tolerance analysis techniques [10].

The main contribution of this paper concerns the geometrical behavior modeling of assemblies involving form defects, where a probabilistic model is proposed for gaps modeling in fixed and sliding contacts. This section presents a state of art of the statistical tolerance analysis approach.

It is necessary to use a realistic model of the surfaces of a mechanical system (i.e. modeled imperfections in good agreement with the imperfections observed on the manufactured components) to guarantee accurate results of the tolerance analysis. Globally, three models of surfaces are applicable: the nominal surface [11], the substitute surface [11], and the non-ideal surface represented by skin model concept [12]. Both, substitute and non-ideal surface, have an important role in the tolerance analysis method. Several models have been proposed to characterize the geometrical deviations of surfaces [10, 13, 14]. In this paper, the SDT approach [15] is used to model surface deviations and gaps between the mating surfaces.

Thus, each surface is characterized by three components of translation (u, v, w) and three components of rotation (α, β, γ). The set of dimension, position and orientation deviations of the mechanical system is denoted by $\mathbf{X}$ and is randomly generated. González and Sánchez [16] proposed a method for the optimal tolerance allocation using the statistical approach with dependent (e.g. correlated) variables; the structure of dependence can be associated with the manufacturing process. The non-ideal surface is built by adding form defects to the substitute surface. The method used to build the non-ideal surface has been explained in [17, 18]. The vector $\mathbf{F}$ used throughout this paper, represents the set of form defects amplitudes of the surfaces defined in correlation with the discretization of the surfaces.

The geometrical behavior modeling of an over-constrained mechanical system is obtained through the development of a set of constraints [10, 11, 18, 19]: compatibility equations (denoted by $C_c(\mathbf{X},\mathbf{G})=0$), interface constraints (denoted by $C_i(\mathbf{X},\mathbf{F},\mathbf{G})\leq 0$), and the functional conditions (denoted by $C_f(\mathbf{X},\mathbf{F},\mathbf{G})\geq 0$). These constraints have been detailed in [20]

The assembly and functioning problems must be formulated to perform the tolerance analysis of an over-constrained mechanical system. The formulation of the assembly and functioning problems have been detailed in [10, 21] and the mathematical formulations of the probabilities of assembly and functioning failure are given in [20]. To perform the computation of these two probabilities, a Monte Carlo simulation combined with a constrained optimization is employed. This method has been introduced by Qureshi *et al.* [10] considering only size, position and orientation defects; and further extended to integrate parts' form defects [17, 18] and to consider different contact types [20]. The overall tolerance analysis technique of an over-constrained mechanical system is explained in Algorithm 1.

ALGORITHM 1: Computation of the probability of assembly failure (P_{na}) and the probability of functional failure (P_{nf})

Require: Total number of Monte Carlo Samples (N_{mc})
Require: Total number of fixed and sliding contacts in the over-constrained mechanical system (N_{ct})
Ensure: total number of Monte Carlo populations and compute the probabilities P_{na}, P_{nf}

1: Constraints development: $C_c(\mathbf{X},\mathbf{G})$, $C_i(\mathbf{X},\mathbf{G},\mathbf{F})$, $C_f(\mathbf{X},\mathbf{G},\mathbf{F})$

2: **for** each i $\in N_{mc}$ **do**

3:    // Generation: geometrical deviations $\mathbf{X} = \{\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \ldots, \mathbf{X}^{(Nmc)}\}$ and form defects populations $\mathbf{F} = \{\mathbf{F}^{(1)}, \mathbf{F}^{(2)}, \ldots, \mathbf{F}^{(Nmc)}\}$

4:    // Discretization of the mating surfaces in the fixed and sliding contacts
5:    **for** each i $\in N_{ct}$ **do**
6:        min $f(\mathbf{G}^*)$ subject to **Signed distances ($\mathbf{S_1}, \mathbf{S_2}$) $\geq$ 0 ($\mathbf{S_1}, \mathbf{S_2}$: the non-ideal mating surfaces)**
7:    **end for**

8:    // Introduction of $\mathbf{G}^*$ in $C_c(\mathbf{X}, \mathbf{G})$ and $C_f(\mathbf{X}, \mathbf{G}, \mathbf{F})$
9:    // Minimization of $\mathbf{G'}$ and verification of the functional requirements C_f
10:    // min $f(\mathbf{G'})$ subject to $C_c(\mathbf{X}, \mathbf{G})$, $C_i(\mathbf{X}, \mathbf{G'}, \mathbf{F}) \leq 0$, $C_f(\mathbf{X}, \mathbf{G}, \mathbf{F}) \geq 0$
11: **End for**
12: Estimation of P_{na} and P_{nf}, by using the following equations [22]:

$$P_{na} = \frac{1}{N_{mc}} \sum_{i=1}^{N_{mc}} I_{D_{na}}\left(\mathbf{X}^{(i)}, \mathbf{F}^{(i)}\right),$$

$$P_{nf} = \frac{1}{N_{mc}} \sum_{i=1}^{N_{mc}} I_{D_{nf}}\left(\mathbf{X}^{(i)}, \mathbf{F}^{(i)}\right),$$

where $I_D(\mathbf{X}, \mathbf{F})$ represents the indicator function; for assembly conditions, it is represented by:

$$I_D(\mathbf{X}, \mathbf{F}) = \begin{cases} 1 \text{ if no solution is found} \\ 0 \text{ if a solution is found} \end{cases}$$

and for the functional requirements, it is represented by:

$$I_D(\mathbf{X}, \mathbf{F}) = \begin{cases} 1 \text{ if } C_f(\mathbf{X}, \mathbf{G}, \mathbf{F}) \leq 0 \\ 0 \text{ if } C_f(\mathbf{X}, \mathbf{G}, \mathbf{F}) > 0 \end{cases}$$

Algorithm 1 sums up the local simulation to characterize the fixed and sliding contacts and the general optimization carried out on the floating contacts with form defects considered [20]. The steps highlighted in Algorithm 1 (framed steps) represent the focus of this paper, and a novel procedure is proposed. This local simulation on the fixed and sliding contacts is discussed in detail in the upcoming subsection.

2.2   Mathematical formulation of the assembly feasibility of the fixed and sliding contacts of parts with form defects

The goal of the contact study in this section is to determine the clearance components which allow an admissible assembly configuration in case of fixed or sliding contacts.
Several methods to deal with different contacts fitting (fixed, sliding and floating) have been proposed. Some researchers have proposed numerical optimization approaches with several objective functions (convex hull volume, Euclidean distance, weighted unsigned distance, weighted signed distance) [23], experimental and theoretical investigations using the SDT and clearance domain (gap hull) models [24], the Fast Fourier Transform and the Conjugate Gradient Method in case of elastic contact of rough surfaces [25, 26], development of an optimization algorithm and geometric interference relationships [27] and some other methods in [28–30]. Moreover, some others researchers proposed to study the floating contacts by means of a regularized closure function and a constrained optimization problem [31], Virtual Work Principle and a contact model of joints to determine the displacements at each node and relative displacements/rotations of the kinematic pairs [32], constrained registration techniques and difference surface in a sequential positioning problem [33], Tooth Contact Analysis algorithm [2, 34], the Gap Hull estimation [35].
Among all the previous methods, the Signed Distance Approach [17, 18, 20, 23] combined with an optimization technique is selected to characterize the fixed and sliding contacts assembly in the mechanical system. Thus, this subsection presents this local simulation on the gap components of a fixed or sliding contact added to the classical tolerance analysis approach.
Based on the SDT approach [15, 17, 18], the clearance torsor $\mathbf{g}$ of a contact can be written as follows:

$$\mathbf{g} = \begin{Bmatrix} \alpha & u \\ \beta & v \\ \gamma & w \end{Bmatrix}_M \tag{1}$$

where u, v and w represent the translations in the direction of the x-, y- and z- axes, respectively; α, β and γ represent the rotations around the x-, y- and z- axes respectively. Problems involving form defects in the sliding and fixed contacts are more complex than problems without form defects and specific procedures are required to perform the tolerance analysis.

Thus, two cases may be considered when a tolerance analysis problem involves fixed or sliding contacts: (1) form defects are not considered and only size, position and orientation defects are involved; (2) form defects are considered. In the first case, the representation of the clearance torsor depending on the different nature of contact is shown in Fig. 1a. The focus of this paper concerns the second case.
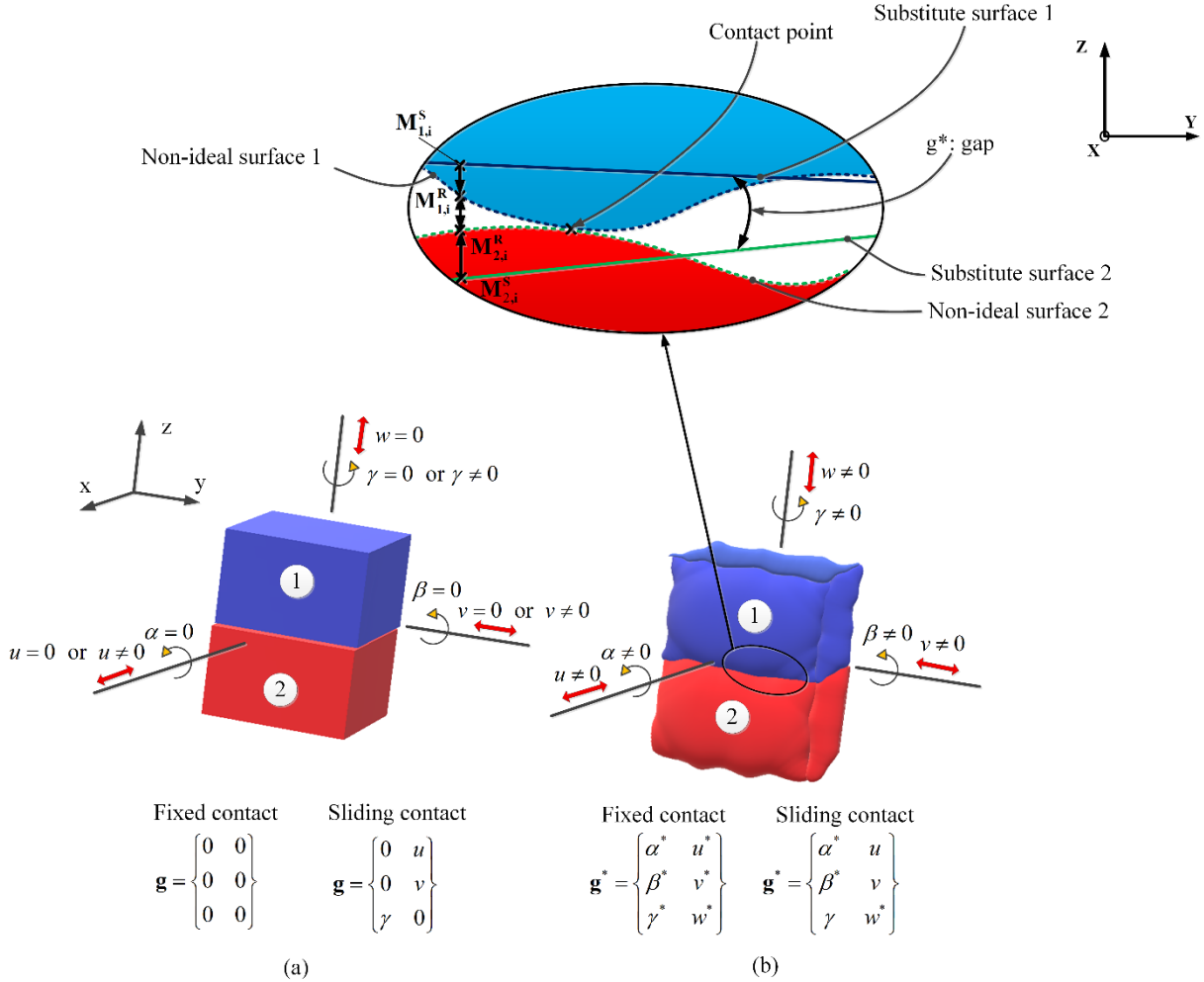

Fig. 1. Gap torsor for: (a) perfect planar contact, (b) contact of non-ideal planar mating surfaces

The second case concerns the contact of planar mating surfaces with form defects (See Fig. 1b). The clearance components depend on the form defects amplitudes and the contact configuration becomes non-trivial. The values of the clearance components are set to a specific value associated with the contact configuration.

For instance $\alpha=\alpha^*$, $\beta=\beta^*$ and $w=w^*$ in the case of a sliding planar contact, where α^*, β^* and w^* are non-zero even though the two surfaces are in contact. The optimization strategy of $\mathbf{g}^*$ consists of minimizing the clearance components such that the signed distances between the two non-ideal surfaces are respected. This optimization technique is well detailed in [17, 18, 20].

For the fixed contact, the six components of the clearance components are determined during the optimization procedure to fully characterize the contact. Regarding the sliding contacts, only three components are identified during the optimization (the translation component w, and the two rotation angles α and β) for a planar contact. The three remaining components of the clearance torsor are not determined at this stage of the procedure. They may be interpreted as "free variables", i.e. these parameters are computed in the general optimization of the tolerance analysis which is associated with the floating contacts (as discussed in section 2.1). The components of $\mathbf{g}^*$ depend on the nature of the contact and they are shown in Fig. 1b. More information about this clearance $\mathbf{g}^*$ and the mathematical formalization of its optimization can be found in [20].

The tolerance analysis algorithm combines a MCS and separate optimization methods for each sample. Both procedures are numerically demanding and combining them may lead to excessive computational efforts. Hence, the probabilistic models are developed to estimate the distributions of the clearance components for fixed and sliding contacts.

3. Probabilistic model for gap modeling in fixed and sliding contacts

In this section, a probabilistic model is proposed for gap components in fixed and sliding contacts to determine their relative positioning without interferences. In the statistical tolerance analysis procedure described above, the optimizations for fixed and sliding contacts are carried out for each sample generated during the Monte Carlo simulation and that is very time-consuming. Therefore, the joint probability density functions associated with the clearance components are identified and subsequently used to generate samples instead of using the optimizations carried out on these contacts. The Kernel Density Estimation (KDE) approach is selected to calculate these joint probability density functions. The main goal of using the KDE approach is to generate new samples from the original gap components optimized in such manner that these new samples have the same distribution and features (mean, standard deviation, etc.) as the original values.

In Section 3.1, the mathematical definition of the KDE approach is given. Section 3.2 presents the method of sampling using the original data from the classical optimizations and the Section 3.3 describes how these samples from the probabilistic model are used in the tolerance analysis.

3.1 Kernel Density Estimation approach.

The Kernel Density Estimation (KDE) is a statistical non-parametric method used to estimate the Probability Density Function (PDF) of a random variable [36]. This method can be used for univariate and multivariate continuous variables. The KDE may be used to estimate a probability density function [37] like some other methods: histograms, naïve estimator, nearest neighbor method, variable kernel estimator, orthogonal series estimators and maximum penalized likelihood estimators. The KDE approach is applied on the realization of random variables of the clearance components to approximate their probability density. The kernel density estimator of a univariate continuous variable y is defined as:

$$f(y) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{y - y_i}{h}\right) \tag{2}$$

where $K(.)$ represents the kernel function centered on the observation y_i, h is the bandwidth or the window width. Several kernels are available in the literature and can be used in this kernel estimation [37]: Epanechnikov kernel, Biweight kernel, Triangular kernel, Gaussian kernel, rectangular kernel. In our work, we chose to use the Gaussian kernel defined as follows:

$$K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} u^2\right) \tag{3}$$

where $u = (y - y_i)/h$.

Here, the clearance components are multivariate continuous variables which can be correlated. Therefore, the Multivariate Kernel Density Estimation (MKDE) is used. Thus, the MKDE is defined in general by the same function [38] as in Eq. (3) (generalized for the multivariate case) but $u = \frac{(y - \mathbf{Y_i})^T \mathbf{S}^{-1} (y - \mathbf{Y_i})}{h^2}$, $K(.)$ is the multivariate Gaussian kernel function, $y = [Y_1, Y_2, ..., Y_d]$ represents the d dimensional random vector whose density is being estimated, $\mathbf{Y_i} = [Y_{1i}, Y_{2i}, ..., Y_{di}]^T$ $i = 1$ to n (n representing the total number of sample available to construct the KDE estimation) represents the sample value of y and $\mathbf{S}$ represents the sample covariance matrix. The multivariate Gaussian kernel function [38] is defined as:

$$K(u) = \frac{1}{(2\pi)^{d/2} \det(\mathbf{S})^{1/2} h^d} \exp(-u/2) \tag{4}$$

An important parameter of the KDE approach is the bandwidth h which has a huge impact on the KDE model. The bandwidth h is used to control the amount of smoothing of the data in the density estimate [38]. There are

several techniques used to determine the best h: the cross-validation method, the plug-in method, the Scott's Rule and the Silverman Rule (Rule of thumb) [39, 40]. The cross-validation method and the plug-in method are based on an optimization procedure to find the best h value, and require considerable numerical efforts [38]. Thus, in this paper, Silverman appropriate h proposed in [37] is used for the Gaussian kernel as the optimal bandwidth. The bandwidth h proposed by Silverman is defined as:

$$h = \left(\left(4 / (2d+1) \right)^{1/(d+4)} \right) . n^{-1/(d+4)} \tag{5}$$

where n represents the number of the data points, d the dimension of the set of random variables. After describing the KDE approach, the resampling method is detailed in the next section.

3.2 Generation of new KDE based variables using the original data.

The KDE approach is introduced in this article to determine firstly, the probability density of the variables and secondly, to generate new samples or realizations based on the KDE in order to reduce the numerical efforts. The process to generate the new samples which are going to be used to replace the repetitive optimization is detailed in this section.

In the classical tolerance analysis method presented above, a sample of variables $\mathbf{X}$ is generated in each iteration of the Monte Carlo simulation. For each sample $\mathbf{X}$, an optimization is carried out on each fixed and sliding contact to determine its corresponding $\mathbf{g*}$ (see Section 2). To introduce the KDE simulation method, a preliminary Monte Carlo simulation is launched on the optimization of the components of $\mathbf{g*}$. Realizing the preliminary Monte Carlo simulation on the optimization of $\mathbf{g*}$ from a fixed or sliding contact, allows one to get several realizations of each component of $\mathbf{g*}$. These realizations are considered as the original data on which the KDE simulation method is applied. Thereafter, the objective is to approximate the joint distribution of these realizations of the components from $\mathbf{g*}$ by using the KDE simulation method.

Several experiments on the optimization of $\mathbf{g*}$ from a fixed or a sliding contacts have shown that, there is a great variability on the marginal distributions of the components from $\mathbf{g*}$ and that the structure of dependence between them is complex. The standard methods of approximation of marginal distributions which on one hand include the use of simple predefined distributions (Gaussian, uniform, beta, gamma, etc.) and on the other hand, the use of usual copulas for the dependence structure (Gaussian copula, Archimedean copula …) are not flexible enough to describe with accuracy the distribution of each component from $\mathbf{g*}$ and their structure of dependence. Hence, it is necessary to find an approach that allows designers to generate new samples with various variables' distribution and which respects also to the dependence structure between the new simulated variables. The KDE simulation method has been chosen because of its simplicity, flexibility and accuracy to approximate the marginal distribution of each component of $\mathbf{g*}$ and to model the structure of dependence between these components of $\mathbf{g*}$.

An algorithm is developed to generate new samples of the variables based on the KDE approach. The method to generate the KDE simulation data is explained as follows. Let us consider N_{mc} as the total number of Monte Carlo simulation, k the number of components of the clearance parameters and N_{KDE} the total number of the preliminary MCS. The local simulation is performed N_{KDE} times and the results are kept in a matrix. After the last iteration, the completed matrix is considered as the original values of the clearance components. This original data is denoted by $\mathbf{D}$ and represents a N_{KDE}-by-k matrix. Then, the probabilistic model based on the Kernel density Estimation is launched by using the matrix $\mathbf{D}$. The sampling method is described by Rajagopalan *et al* [38] and summarized in Algorithm 2.

ALGORITHM 2: Sampling of new KDE-based variables

Step 1: Let us consider $\mathbf{D} = \left[\mathbf{g}_1^*, \mathbf{g}_2^*, ..., \mathbf{g}_{N_{KDE}}^* \right]$ where $\mathbf{g}_i^*$ is of size 1 x k, i=1 to N_{KDE}, obtained from the N_{KDE} time optimization of the clearance parameters of a fixed or a sliding contact. Pick uniformly $\mathbf{g}_i^*$ from $\left[\mathbf{g}_1^*, \mathbf{g}_2^*, ..., \mathbf{g}_{N_{KDE}}^* \right]$ to get a matrix of size N_{mc} x k.

Step 2: Simulate new variables from the kernel i.e. simulate multivariate gaussian variables centered on $\mathbf{g}_i^*$ of the new matrix and have as standard deviation $h^2 * \mathbf{S}$. This simulation $\mathcal{N} (\mathbf{g}_i^*, h^2 \text{ x } \mathbf{S})$ is a matrix with dimension N_{mc} x k. h represents the bandwidth whose calculation is explained in the previous section and $\mathbf{S}$ represents the sample covariance matrix.

Step 3: These new variables obtained, represent the probabilistic model of the clearance components, which are then introduced in the constraints (compatibility equations, interface constraints) developed.

Complete algorithm of the overall method for the probabilistic model of the clearance components is summarized in the following

---

ALGORITHM 3: Probabilistic model of the clearance components of a fixed or a sliding contact

Require: Set the number of Monte Carlo Samples (N_{mc})

Require: Set the total number of the preliminary MCS (N_{KDE})

Require: Ideal mating surfaces of a fixed or sliding contact

Require: An empty matrix $\mathbf{D}$ of size N_{KDE} x k

Ensure: number of the optimization iterations at a contact

1: Generate the form defects amplitudes for each surface of the contact, $\mathbf{F} = \{\mathbf{F}^{(1)}, \mathbf{F}^{(2)}, \ldots, \mathbf{F}^{(N_{KDE})}\}$ where

$\mathbf{F}^{(i)} = [f_1, f_2, \ldots, f_n]$

   n=number of the meshing points on each surface

2: Apply the amplitudes to the perfect discretized mating surfaces to get the real mating surfaces i.e. surfaces with form defects

3: **for** each i $\in N_{KDE}$ **do**

4:      min $f(\mathbf{g}^*)$ subject to **Signed distances ($\mathbf{S_1}, \mathbf{S_2}$)** ≥ 0 ($\mathbf{S_1}, \mathbf{S_2}$: the non-ideal mating surfaces)

5:      // Change the ith row of $\mathbf{D}$ by the components of $\mathbf{g}^*$

6: **end for**

7: Pick uniformly $\mathbf{g}_i^*$ from $\mathbf{D} = \left[\mathbf{g}_1^*, \mathbf{g}_2^*, \ldots, \mathbf{g}_{N_{KDE}}^* \right]$ to get the matrix $\mathbf{M}$ of size N_{mc} x k.

8: Generate a multivariate gaussian variables $\mathbf{N}$ centered on 0 and has h^2xS as the standard deviation, $\mathbf{N}$ is a matrix of size N_{mc} x k

9: Compute $\mathbf{P} = \mathbf{M} + \mathbf{N}$. Each column from $\mathbf{P}$ represents the probabilistic model of each components from the clearance torsor

---

Algorithm 3 is used to get the probabilistic model for each component from the clearance of each fixed and sliding contacts. Therefore, by using this algorithm 3, the distributions of the new samples of the clearance components are in good agreement with the original data as well.

3.3   Integration of the KDE based variables in the statistical tolerance analysis approach

The classical approach for tolerance analysis of over-constrained mechanisms involves a probabilistic model of the geometry of the parts, i.e. the dimension, orientation, shape or form defects are characterized by a set of random variables. Monte Carlo simulation is used to generate samples of the geometry and for each sample, the components of the gap torsor are analyzed. This task is non-trivial and requires to perform (1) an optimization of the gap components for each fixed and sliding contact to compute the value of the limited degrees of freedom; (2) optimizations on the free degrees of freedom of the contacts, which have not been previously identified, in order to determine whether the mechanism can be assembled and whether it is functional. This method allows us to compute the probability of non-functionality and the probability of non-assembly.
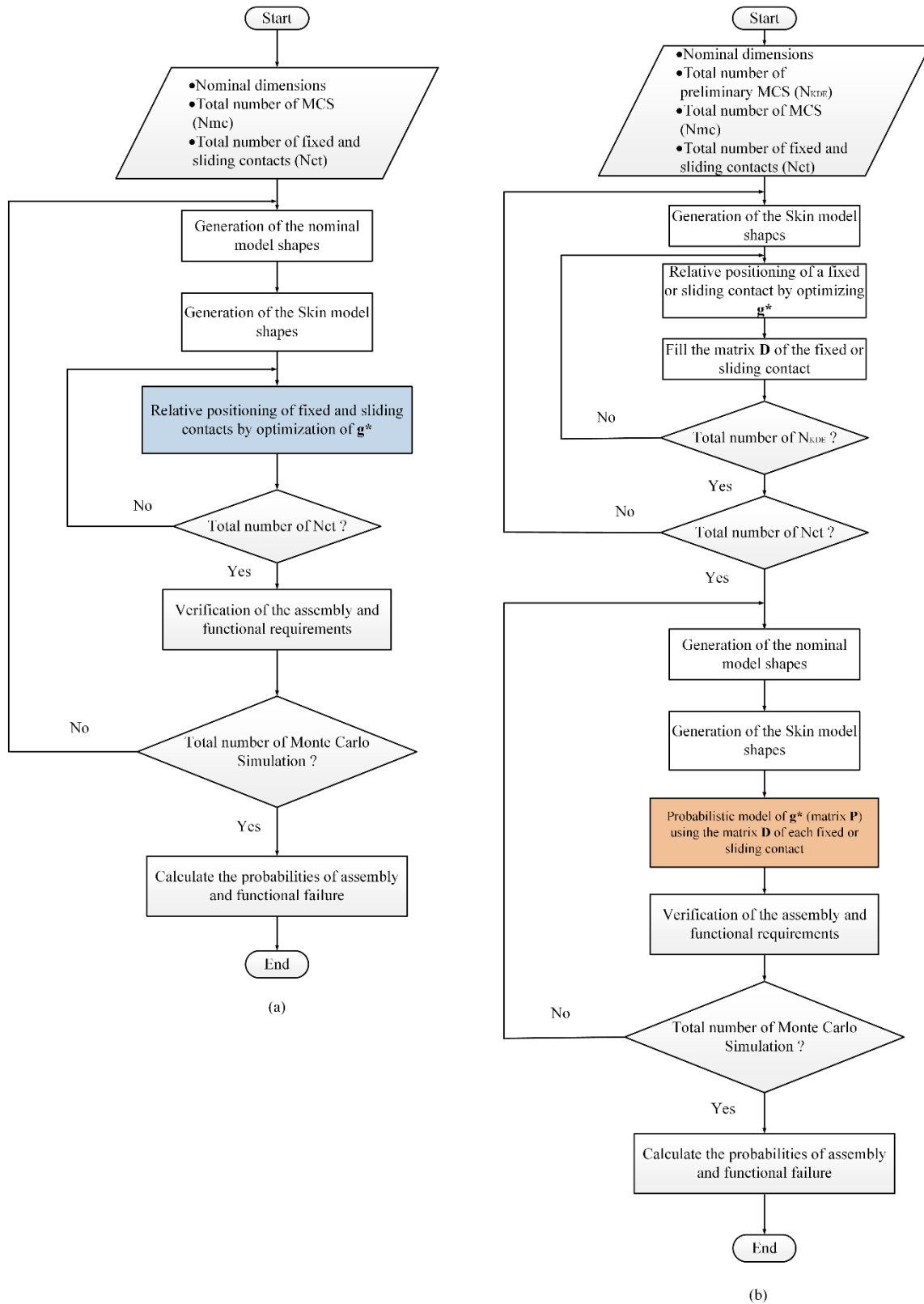
Fig. 2. *Frameworks: (a) classical method, (b) using the proposed method*

The proposed method in this paper is focused on the first step described above. It is suggested to investigate each sliding and fixed contact individually. Samples of the imperfect geometry are generated using Monte Carlo simulation and the corresponding components of the gap torsor are computed. This data is subsequently used to identify the joint probability density function of the limited displacements components of the gap torsor using

the KDE. Once this probability density function is characterized, it is possible to use it in order to directly generate samples of the components of the gap torsor without using any numerically demanding optimization procedure. The interface constraints of the geometrical behavior of a mechanism, involve solely the floating contacts. Therefore, the samples of the proposed KDE based method are not involved in these interface constraints. The compatibility equations and the functional conditions are used in the step (2) detailed in the geometrical behavior modeling. They represent the constraints of a general optimization procedure. The variables in **P** are involved as inputs of the compatibility equations and the functional requirements. The compatibility equations and the functional requirements are concerned about the fixed, the sliding and the floating contacts. Thus, at each iteration of the MCS, a sample of gap components from **P** are introduced in the compatibility equations and in the functional requirements inequalities.

The remaining degrees of freedom of the fixed and sliding contacts which are not limited, are considered as free variables in the step (2) described above, i.e. they are adjustable parameters of the optimization performed at this step.

The whole process for the tolerance analysis with the probabilistic models is depicted in Fig. 2b.

4. Application

In this section, the proposed probabilistic model approach is implemented to model the distribution of the gap components of a sliding contact in an over-constrained mechanical system, chosen as case study. In section 4.1, the mechanical system is presented and then in Section 4.2, the results of the probabilistic modeling are given as well as the results of the tolerance analysis. A comparison is addressed between the classical tolerance analysis method and the proposed method.

4.1 Description of the mechanical system

The case study is a part of an electric engine, see Fig. 3. The mechanical system is composed of a shaft, a housing and a body.
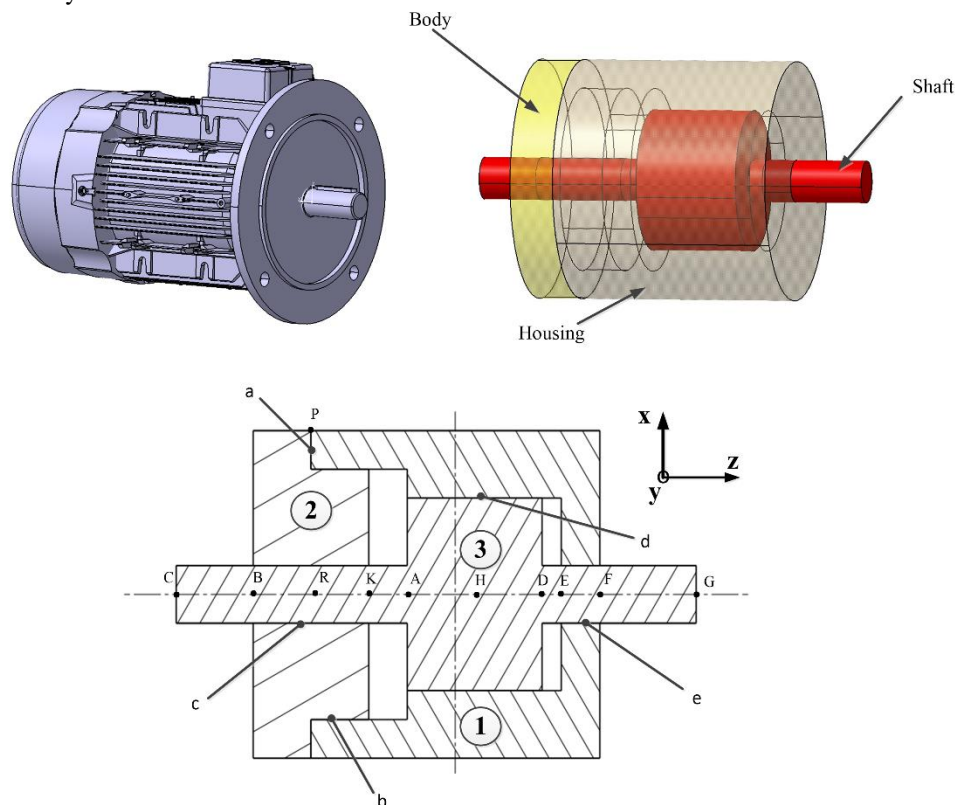

Fig. 3. Studied mechanical system

The mechanical system is composed of 4 floating contacts (1b/2b, 2c/3c, 1d/3d, and 1e/3e) and one sliding contact (1a/2a). The sliding contact is the most important for the approach proposed in this paper because the probabilistic models are established on its clearance parameters. The functional requirement characterizes the floating contact 1d/3d, where a minimum clearance threshold should be respected between the shaft and the housing. The tolerance analysis of this mechanical system is carried out based on the tolerances specifications

taken from [20, 41]. The tolerance specifications of the different parts of the mechanical system used in these studies and more information about the other tolerances used in the overall tolerance analysis method can be found in our previous work [20]. Firstly, a preliminary MCS is launched on the optimization of the gap components in contact 1a/2a. Then, the KDE method is used to get new samples of these realizations, which are after introduced in the overall tolerance analysis method to replace the highlighted steps from algorithm 1.

The contact 1a/2a is composed of two circular planar surfaces. The form defects are generated according to the z axis. The contact 1a/2a is sliding because body is bolted to the shaft and this assembly technique allows slight displacement in the direction of the contact plane. Thus, three clearance parameters are identified as shown in the following clearance **g***:

$$\mathbf{g}^* = \begin{Bmatrix} \alpha & - \\ \beta & - \\ - & w \end{Bmatrix}_M \tag{6}$$

where α, β are components of the rotations around the x- and y- axes, and w is the translation in the z- axes. These components of **g*** are not trivial and are dependent on form defects. An optimization procedure is used to identify them. The probabilistic model method proposed in Section 3 is implemented on these three components. The rest of **g*** components (γ, u, v) are not determined during this optimization or probabilistic model method. They are considered as free variables and are included in the optimization of **G'**. The results for the components w, α, β are presented and interpreted in the next section.

4.2   Results

In this subsection, two categories of simulations have beeen launched to apprehend the different aspects of the use of the KDE method in the tolérance analysis such as the reduction of simulation time, the suitable meshing to modelize corectly the geometrical behavior of the mechanical system.

The first categorie concerns the application of two differents distributions of the mating surfaces of *1a/2a* in order to dertemine the appropriate meshing that modelize properly the geometrical behavior of the mechanical system with steady calculated probabilities. The computing time is not the important aspect of the simulations of this first categorie. The Gaussian distribution and the uniform distribution are considered here to generate the amplitudes of the form defects of the surfaces involved in the sliding contact (contact 1a/2a). These distributions are employed with different standard deviations and supports to test the robutness of the proposed method. The Gaussian distribution is composed of three different standard deviations with the same mean. The uniform distribution is also composed of three different supports. These two distributions are chosen in order to determine whether the distributions of the amplitudes of the form defects have an impact on the distribution of the components of **g*** and if the KDE is capable of approximating the marginal distributions of the clearance components and keeping their structure of dependence.

Three levels of meshing refinement of the mating surfaces are considered: the coarse, the standard and the refined mesh with respectively 100, 2500, and 10000 discretization points. The consideration of multiple levels of refinement allows us to determine whether the mesh is sufficiently fine to capture the behavior of the components of **g***, and to detemine whether the probabilities (of assembly failure and functional failure) have converged towards their reference value. Table 1 shows the different strategies considered to model the form defects.

Table 1: The local simulation parameters

Distribution	Meshing	Standard deviations/ Supports
Gaussian ($\mu=2.5\text{x}10^{-3}$)	100	10^{-6}
	2500	10^{-4}
	10000	10^{-2}
Uniform	100	$[2\text{x}10^{-3}, 3\text{x}10^{-3}]$
	2500	$[1.5\text{x}10^{-3}, 3.5\text{x}10^{-3}]$
	10000	$[-7.5\text{x}10^{-3}, 1\text{x}10^{-2}]$

Based on Table 1, all the combinations possible of parameters shown are studied. Two distributions (Gaussian and uniform) are available to generate the amplitudes of form defects. For each distribution, the three levels of refinement of the mesh are used and for each level of refinement, the three standard deviations (for the Gaussian distribution) or the three supports (for the uniform distribution) are used to generate the form defects amplitudes.

In total, considering the combination between the types of meshing and the standard deviation or support, 18 types of form defects amplitudes are generated (nine for the Gaussian distribution and nine for the uniform distribution). For each distribution of the considered form defects amplitudes the probabilistic model of g^* of the sliding contact is established and the tolerance analysis of the whole over-constrained mechanical system is thereafter carried out. As an example of simulation, the form defects may be generated using 100 discretization points and amplitudes following a Gaussian distribution with a mean equal to 2.5×10^{-3} mm and a standard deviation equal to 10^{-6}.

After every simulation, some comparisons are carried out between the KDE simulation variables and the original variables from the local simulation (i.e. obtained using an optimization procedure). These comparisons aim to verify whether the simulated data from the KDE correspond to the original data and if they have the same features. They involve the components w, α, β of the clearance parameters, their histograms, and the Probability Density Function (PDF). As all the comparison figures of the 18 simulations cannot be shown together, just figures from the Gaussian distribution $\mathcal{N}(2.5 \times 10^{-3}, 10^{-4})$ with 100 discretization points are presented. Nevertheless, the results concerning the probabilities (P_{nf}, P_{na}) and the time of simulations assessed during the tolerance analysis carried out after, are presented for all the 18 simulations (see the Appendix). For each simulation, the matrix $\mathbf{P}$ defined in the Algorithm 3 can be written as:

$$\mathbf{P} = \begin{pmatrix} w_1 & \alpha_1 & \beta_1 \\ w_2 & \alpha_2 & \beta_2 \\ & \mathbf{M}\,\mathbf{M}\,\mathbf{M} & \\ w_{N_{mc}} & \alpha_{N_{mc}} & \beta_{N_{mc}} \end{pmatrix}_{[N_{mc} \times 3]} \tag{7}$$

In summary for each simulation the following mathematical characteristics are assumed:

- N_{mc} random variables following the Gaussian distribution $\mathbf{X} \,\square\, \mathcal{N}(\mu, \sigma_i)$ and the uniform distribution $\mathcal{U}(a_i, b_i)$ where $\mu = 2.5 \times 10^{-3}$, $\sigma_i \in \{10^{-6}, 10^{-4}, 10^{-2}\}$ and $[a_i, b_i]$ represent the different supports of the uniform distribution in Table 1. The total number samples used in the Monte Carlo Simulation is $N_{mc} = 100000$.
- 3 components of g^* (w, α, β) for the first local simulation (optimization) and then after the probabilistic model of the data are determined.
- The number of the optimization constraints (signed distances) depends on the mesh discretization.

The reference simulation involves Gaussian distribution for the amplitudes of form defects, $\mathcal{N}(2.5 \times 10^{-3}, 10^{-4})$ with 100 as meshing type. Based on the simulation from this example, the data from the local simulation (original data) and the data from the KDE simulation are shown respectively in Fig. 4a and Fig. 4b. In these figures, the histogram of each component is shown as well as the repartition between the components. It should be noticed that the original data have the same histograms as the KDE simulation data. Furthermore, the structures of dependence are in good agreement.
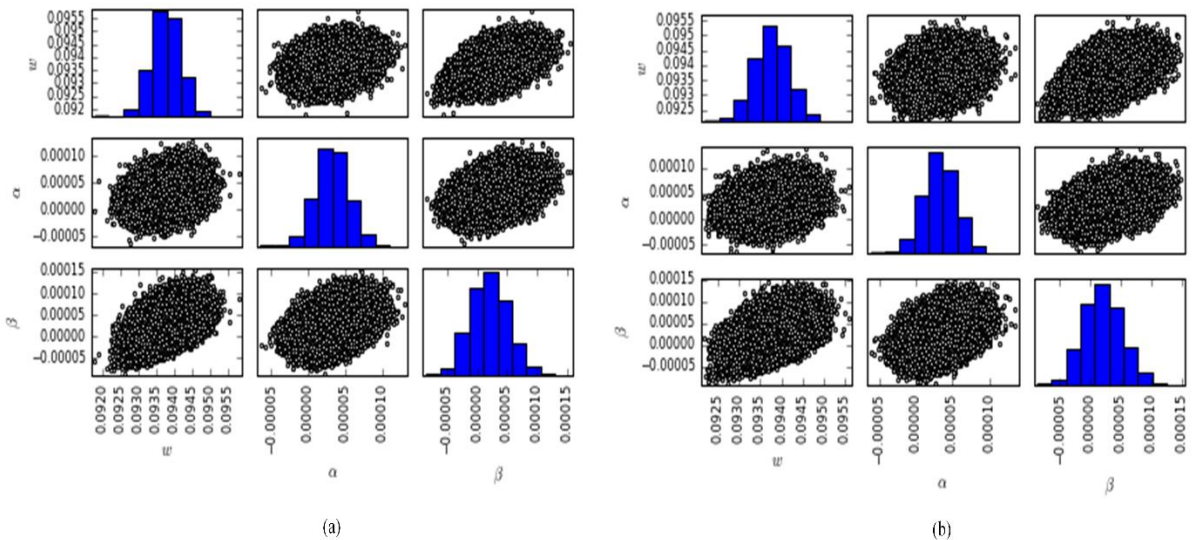

(a)

(b)

Fig. 4 and some other figures presented in the appendix (Fig. 5 and Fig. 6) compare the data obtained using the classical approach (MCS combined with an optimization of the gap torsor components) and the proposed KDE method in the particular case where the amplitude of the form defects exhibits a Gaussian distribution with a mean equal to 2.5×10^{-3}, a standard deviation equal to 10^{-4} and the mesh discretization involves 100 points. The same comparison about the distribution of data points was performed for the other configurations described in Table 1, and we observed that globally, the KDE samples are in good agreement with the Monte Carlo data; however, these comparisons are not shown here and neither in the appendix to keep the paper short and concise.

The KDE simulation data are collected for each simulation and introduced mainly in the compatibility equations already developed. For each kind of distribution of the form defects amplitudes, two tolerance analysis simulations are run. The first tolerance analysis simulation is based on the technique explained in Section 2.1 i.e. based on the direct optimization of the clearance parameters of the sliding contact at each iteration (local simulation) of the Monte Carlo simulation [20]. The second tolerance analysis simulation uses the KDE simulation variables as described in Section 3.2. A comparison is then established between these two tolerance analysis simulations as shown in the Tab. 2 for one type of simulation (F $\square$ $\mathcal{N}(2.5 \times 10^{-3}, 10^{-4})$). For the other cases of simulation their results are presented in the tables presented in the appendix (Table 4-Table 8). It is important to note that the total simulations time presented in these tables in case of tolerance analysis using the KDE simulation samples, includes the time of the preliminary MCS applied to obtain the original data, the time to get the joint PDF of the clearance components and the new samples and finally the time to carry out the whole tolerance analysis simulation. Thus, T_1 represents the addition of the time to carry out the preliminary MCS to get the original data and the time to generate new samples (matrix **P**) of gap components based on the KDE method; T_2 represents the time to carry out the tolerance analysis of the mechanical system using the KDE simulation samples. This particular example does not imply multiple fixed, sliding and floating contacts during simulations. The gain in computation time can be much greater on mechanical systems with more fixed and sliding contacts.

Table 2: Results of the simulation F $\square$ $\mathcal{N}(2.5 \times 10^{-3}, 10^{-4})$, $N_{KDE} = 100000$, $Nmc = 100000$

Simulation	Meshing		
	100	**2500**	**10000**
Classical tolerance analysis	$P_{nf} = 1.72 \times 10^{-2}$ $P_{na} = 1.00 \times 10^{-5}$	$P_{nf} = 1.64 \times 10^{-2}$ $P_{na} = 1.00 \times 10^{-5}$	$P_{nf} = 1.64 \times 10^{-2}$ $P_{na} = 1.00 \times 10^{-5}$
Total time	10h36min	14h06min	46h32min
Tolerance analysis based on KDE simulation	$P_{nf} = 1.73 \times 10^{-2}$ $P_{na} = 2.00 \times 10^{-5}$	$P_{nf} = 1.64 \times 10^{-2}$ $P_{na} = 1.00 \times 10^{-5}$	$P_{nf} = 1.64 \times 10^{-2}$ $P_{na} = 1.00 \times 10^{-5}$
T_1	13min	47min	3h36min
T_2	9h13min	9h39min	9h16min
Total time ($T_1 + T_2$)	9h26min	10h26min	12h52min

The first results shown in Tab. 2 and the others from the appendix imply the reduction of the simulation time for the tolerance analysis based on the KDE simulation data. The speedup obtained using the KDE increases as the mesh is refined. The simulations reduction time is on average 14.50% for the mesh with 100 discretization points, 33.55% for the mesh with 2500 discretization points and 73.24% for the mesh with 10000 discretization points. Consequently, the tolerance analysis method based on the probabilistic models allows a gain of simulation time.
The second finding concerns the probabilities assessed using the different types of meshing. It is observed that the meshes with 2500 and 10000 discretization points have similar values of the probability of assembly failure (P_{na}), whereas its value is significantly different if the mesh with 100 discretization points is used. A similar conclusion can be drawn regarding the probability of functional failure (P_{nf}). Gradually as the mesh refinement is increased, the probabilities converge towards their reference value. The choice of the number of the

discretization points for this tolerance analysis problem is solved in this manner. Based on the results and considering the constraints in terms of simulation time and accuracy of the results, the tolerance analysis using the meshing refinement of 2500 points can be taken as the best compromise solely for these configurations of meshing refinement.

The third finding concerns the impact of the distributions on the probabilities. For both distributions (Gaussian and uniform) chosen to generate the amplitudes of the form defects on the mating surfaces, and for all the types of meshing refinement, the probabilities of assembly failure are increasing gradually as the amplitudes are increasing. This result is in accordance with engineering experiences. Indeed, as the amplitudes of the form defects are increasing, the mating surfaces are more subject to interferences and that can prevent the assembly to occur. It is observed then that the probabilities of assembly failure are very sensitive to the choice of the distributions, whereas the distributions have a reduced impact on the probabilities of functional failure.

The final finding concerns the simulation times, which are increasing gradually as the meshing density increases. Indeed, the total number of constraints involved in the tolerance analysis increases as the mesh is refined, which increases the numerical efforts.

The second category of simulation concerns the use of reduced number of samples during the preliminary MCS in order to gain in computing time. During the first MCS to generate the original data, a small number N_{KDE} of samples can be used and it is considered that this number N_{KDE} of simulation, is sufficient to evaluate a steady joint PDF of the original data. This latter can be used to generate a large number N_{mc} ($N_{mc} > N_{KDE}$ or $N_{mc} \gg N_{KDE}$) of samples which are going to be used in the tolerance analysis. N_{mc} samples are sufficient to allow the tolerance analysis process to evaluate steady and good probabilities. The numerical efforts are reduced because the algorithm of this proposed method and also because a small number of samples is used in the preliminary MCS to generate the original data and identify the joint PDF associated with it. Thus, to show the advantage of the reduction of samples number in the preliminary MCS and the gain in time associated with it, simulations have been launched on the distribution F $\square$ $\mathcal{U}(1.5\text{x}10^{-3}, 3.5\text{x}10^{-3})$ with a total number of MCS N_{mc} = 100000, with discretization type of 2500 and different numbers of N_{KDE}. The results are shown in Table 3, where it is observed that the probabilities can be fairly estimated by using just 10000 samples to generate the KDE data. These results show that it is not necessary to have the number of the preliminary MCS equal to the second MCS where the probabilities are calculated. By doing this, there is a gain in computing time about 6.30% based on the example shown in Table 3, between the simulation in which N_{KDE} is equal to 10000 and the one where it is equal to 100000.

Table 3: Results of the simulation F $\square$ $\mathcal{U}(1.5\text{x}10^{-3}, 3.5\text{x}10^{-3})$, N_{mc}=100000, with different N_{KDE}

Simulations	Results				
	2500				
N_{KDE}	10000	20000	30000	50000	100000
Probabilities	$P_{nf} = 1.63\text{x}10^{-2}$ $P_{na} = 3.00\text{x}10^{-5}$	$P_{nf} = 1.64\text{x}10^{-2}$ $P_{na} = 3.00\text{x}10^{-5}$	$P_{nf} = 1.63\text{x}10^{-2}$ $P_{na} = 3.00\text{x}10^{-5}$	$P_{nf} = 1.63\text{x}10^{-2}$ $P_{na} = 3.00\text{x}10^{-5}$	$P_{nf} = 1.63\text{x}10^{-2}$ $P_{na} = 3.00\text{x}10^{-5}$
T_1	10min	13min	19min	24min	50min
T_2	9h43min	9h44min	9h42min	9h45min	9h43min
Total time ($T_1 + T_2$)	9h53min	9h57min	10h01min	10h09min	10h33min

To sum up all the results, the first simulations based on the variations of the meshing and the distribution of form defects amplitudes, have helped to determine the best meshing choice resulting on steady probabilities of failure. The second batch of simulations on the reduction of the number N_{KDE} have allowed to determine that it is not necessary to have the same number of samples in the preliminary MCS as in the second MCS. And that is very important to reduce the simulation time. Furthermore, based on the results of Tab 3 and Tab 7 (in the Appendix), the simulations based on the computer formatting and development of the algorithm from the proposed method are faster than the simulations of the algorithm from the classical method. This explains the important reduction of time even if the N_{KDE} is equal to N_{mc}.

5. Discussions

The application of the proposed method has shown in the previous section some of its advantages. However, some others can be added.

- Some mechanical systems are composed of several identic contacts. In this case, the KDE method is very suitable because it allows designers to launch just one time, a simulation to determine the original data of a contact. Thereafter, this original data can be used in the evaluation of the joint PDF of its contact and also for the rest of the identic contacts. This helps the designer to save computational time. In fact, the identic contacts have the same elementary defects and the same characteristics. Therefore, they will have the same components of clearance for the best configuration allowing the assembly of the mechanical system. This gain in time is not possible with the classical method because a local simulation should be carried out on each of the identic contacts and at each iteration of the MCS. These repetitive simulations would increase the computational time. Furthermore, additional reduction of computing time can be obtained by using a reduced number of samples (N_{KDE}) during the preliminary MCS as shown in Table 3.

- Moreover, the proposed method allows to study each of the fixed or sliding contacts separately. Therefore, a broad spectrum of these fixed and sliding contacts can be studied in advance to determine their original data. This will be consuming in time but it will help the designers to earn computational time the design starts. During the design, studies must be carried out very fast to be competitive on the market. Thus, the joint PDF of the original data and the new samples will be obtained very fast to proceed to the tolerance analysis of the mechanical system. This computational time gain is possible because the original data of the contacts already exist and also because the KDE method is faster.

- The use of the KDE method enables one to use another method than the MCS to study the tolerance analysis of over-constrained mechanical systems. This other method can be the First Order Reliability Method (FORM) for systems applied in [6] to compute the probabilities of failure. The use of this FORM method in not compatible with the classical statistical tolerance analysis method presented in Section 2.

Moreover, some probabilities calculated in the tables are slightly different. Nonetheless, these differences of probabilities between the local simulation and the proposed method are not sufficiently low to consider that the proposed method converges. Indeed, the probabilities calculated are of the same order of magnitude. It is reminded that the results are obtained using MCS which converges slowly. This method relies on the generation of pseudo-random numbers, and therefore two independent runs of the simulation lead to slightly different results.

6. Conclusion and future work

In this paper, a probabilistic-based model is proposed to deal with the gap components in fixed and sliding contacts. The model is then introduced in the global statistical tolerance analysis method of over-constrained mechanical system with form defects. This statistical tolerance analysis method is based on the Monte Carlo Simulation coupled with multi-scale optimization methods. In this new approach, the contacts (fixed, sliding and floating) are considered differently. The probabilistic models are developed to characterize the clearance parameters from the sliding and fixed contacts. The kernel Density Estimation (KDE) method is used to establish these probabilistic models of the clearance parameters. The main drawback of the Monte Carlo simulation is the considerable numerical efforts associated with it. These numerical efforts are reduced by proposing this new algorithm and by using reduced number of samples in a preliminary MCS which data are employed to get new samples through the KDE simulation method. Thereafter, these new samples are introduced in the general tolerance analysis method. Furthermore, simulations have been proposed to get best choice of the meshing to get steady probabilities of failure. The amplitudes of form defects have an impact on the probability of assembly failure just as the meshing points have an impact on the simulation time.

The proposed method cannot handle yet the deformations which occur on the different parts due the forces applied on them. These deformations can play a significant role in the assembly of the mechanical systems as shown in [42]. A possible outlook can be to see in which manner the studies carried out in [42] can be integrated in the tolerance analysis method of the present work or in the method addressed in [20].

Acknowledgements

Nomenclature

P_{na}	Probability of assembly failure	**g'**	Clearance components of a floating contact
P_{nf}	Probability of functional failure	N_{mc}	Total number of the Monte Carlo simulation
X	Geometrical deviations (random variables)	N_{ct}	Total number of the fixed and sliding contacts in the mechanical system
F	Set of the form defects amplitudes	h	The bandwidth of the Kernel Density Estimation
G	Set of clearance components from the fixed, sliding and floating contacts of the mechanical system	**S**	The covariance matrix of the variables involved in the Kernel Density Estimation method
G*	Set of clearance components from the fixed contacts and the blocked DOFs of the sliding contacts in the mechanical system	**P**	Matrix of the probabilistic model of the clearance components
g*	Clearance components of a fixed or sliding contact	MCS	Monte Carlo Simulation
G'	Set of clearance components from the floating contacts of the mechanical system	**D**	Matrix containing the realizations of the clearance **g*** during the preliminary MCS
PDF	Probability Density Function	N_{KDE}	Total number of the preliminary MCS to get the original data

Appendix

A.1 Figures to show the similarity between the original data and the KDE simulation data

To show the similarity between the original data and the KDE simulation data in more detail, the distributions points between the clearance components are presented in the Fig. 5 for the case of simulation $F \square \mathcal{N}(2.5 \times 10^{-3}, 10^{-4})$. This Fig. 5 shows that the KDE based data globally correspond to the original data.
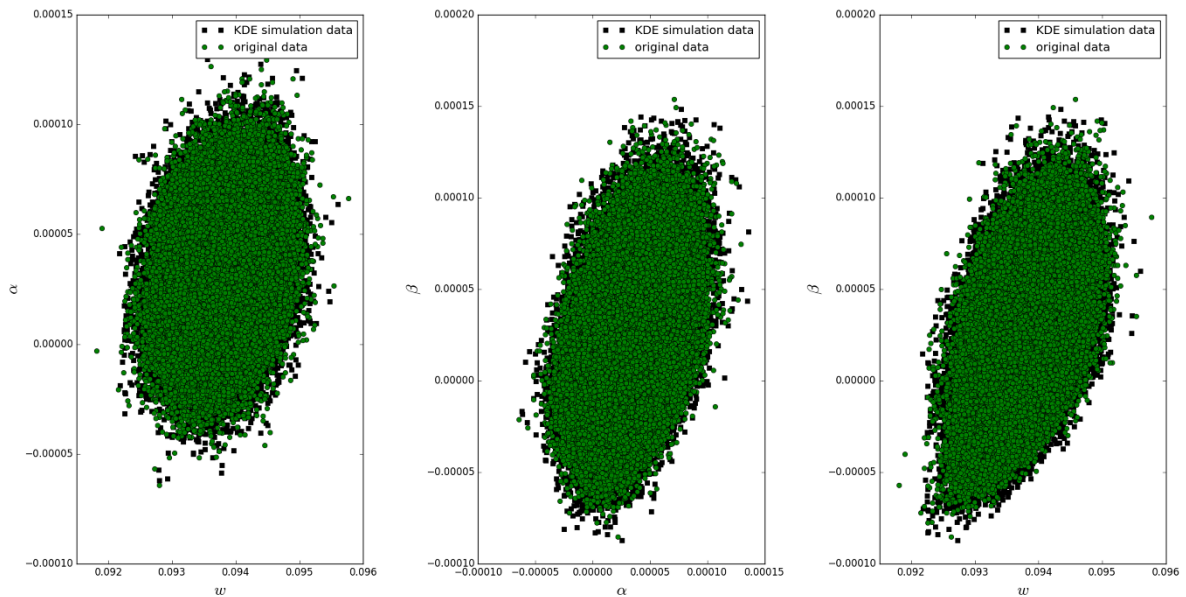

Fig. 5. Comparison of the distribution between the clearance parameters (w, α, β)

Furthermore, the joint Probability Density Function (PDF) between each pair of KDE variables show the match between the original data points and the PDF contour plot. As shown in Fig. 6, for every combination of the clearance parameters, the density of the original data points is corresponding to the color of the PDF contour plot.
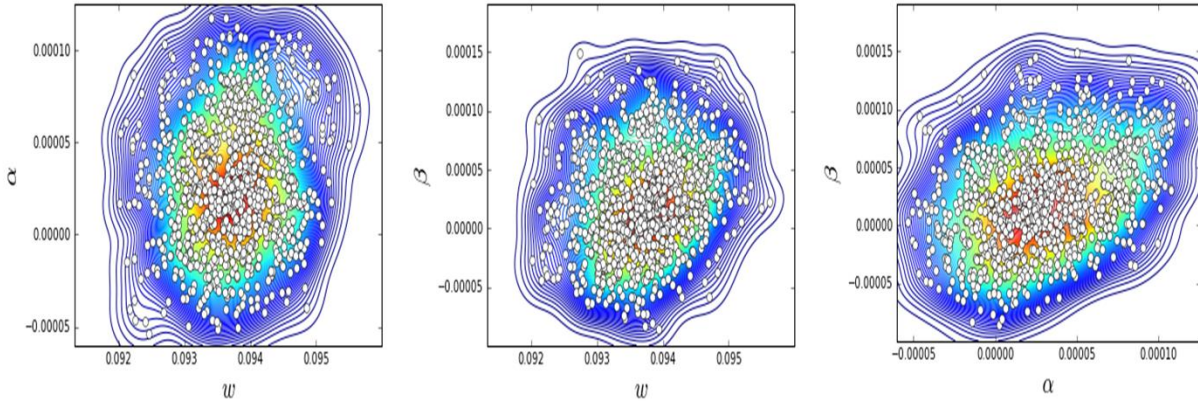

Fig. 6. Contour plot of the Probability Density Function of the KDE data compared to the original data points

A.2 Tables showing the results of the comparisons proposed in Table 1.

Table 4: Results of the simulation $F \sim \mathcal{N}(2.5\text{x}10^{-3}, 10^{-6})$, $N_{KDE} = 100000$, $Nmc = 100000$

Simulation	Meshing		
	100	**2500**	**10000**
Classical tolerance analysis	$P_{nf} = 1.38\text{x}10^{-2}$ $P_{na} = 2.00\text{x}10^{-4}$	$P_{nf} = 1.42\text{x}10^{-2}$ $P_{na} = 1.00\text{x}10^{-5}$	$P_{nf} = 1.41\text{x}10^{-2}$ $P_{na} = 1.00\text{x}10^{-5}$
Total time	13h51min	18h39min	46h42min
Tolerance analysis based on KDE simulation	$P_{nf} = 1.58\text{x}10^{-2}$ $P_{na} = 1.80\text{x}10^{-4}$	$P_{nf} = 1.63\text{x}10^{-2}$ $P_{na} = 1.00\text{x}10^{-5}$	$P_{nf} = 1.62\text{x}10^{-2}$ $P_{na} = 1.00\text{x}10^{-5}$
T_1	15min	46min	3h28min
T_2	9h37min	9h39min	9h53min
Total time ($T_1 + T_2$)	9h52min	10h25min	13h21min

Table 5: Results of the simulation $F \sim \mathcal{N}(2.5\text{x}10^{-3}, 10^{-2})$, $N_{KDE} = 100000$, $Nmc = 100000$

Simulation	Meshing		
	100	**2500**	**10000**
Classical tolerance analysis	$P_{nf} = 4.45\text{x}10^{-3}$ $P_{na} = 8.41\text{x}10^{-1}$	$P_{nf} = 4.39\text{x}10^{-3}$ $P_{na} = 8.81\text{x}10^{-1}$	$P_{nf} = 4.38\text{x}10^{-3}$ $P_{na} = 8.81\text{x}10^{-1}$
Total time	12h36min	21h22min	66h23min
Tolerance analysis based on KDE simulation	$P_{nf} = 4.54\text{x}10^{-3}$ $P_{na} = 8.49\text{x}10^{-1}$	$P_{nf} = 4.39\text{x}10^{-3}$ $P_{na} = 8.80\text{x}10^{-1}$	$P_{nf} = 4.39\text{x}10^{-3}$ $P_{na} = 8.81\text{x}10^{-1}$
T_1	12min	44min	3h30min
T_2	9h20min	9h19min	9h24min

Total time (T_1+ T_2)	9h32min	10h03min	12h54min

Table 6: Results of the simulation $F \sim \mathcal{U}(2\text{x}10^{-3}, 3\text{x}10^{-3})$, <mark>$N_{KDE} = 100000$,</mark> $Nmc = 100000$

Simulation	Meshing		
	100	**2500**	**10000**
Classical tolerance analysis	$P_{nf} = 1.72\text{x}10^{-2}$ $P_{na} = 1.00\text{x}10^{-5}$	$P_{nf} = 1.63\text{x}10^{-2}$ $P_{na} = 1.00\text{x}10^{-5}$	$P_{nf} = 1.63\text{x}10^{-2}$ $P_{na} = 1.00\text{x}10^{-5}$
Total time	10h33min	13h59min	46h42min
Tolerance analysis based on KDE simulation	$P_{nf} = 1.71\text{x}10^{-2}$ $P_{na} = 1.00\text{x}10^{-5}$	$P_{nf} = 1.64\text{x}10^{-2}$ $P_{na} = 1.00\text{x}10^{-5}$	$P_{nf} = 1.64\text{x}10^{-2}$ $P_{na} = 1.00\text{x}10^{-5}$
T_1	13min	39min	3h44min
T_2	9h42min	9h40min	9h42min
Total time (T_1+ T_2)	9h55min	10h19min	13h26min

Table 7: Results of the simulation $F \sim \mathcal{U}(1.5\text{x}10^{-3}, 3.5\text{x}10^{-3})$, <mark>$N_{KDE} = 100000$,</mark> $Nmc = 100000$

Simulation	Meshing		
	100	**2500**	**10000**
Classical tolerance analysis	$P_{nf} = 1.73\text{x}10^{-2}$ $P_{na} = 5.00\text{x}10^{-5}$	$P_{nf} = 1.64\text{x}10^{-2}$ $P_{na} = 1.00\text{x}10^{-5}$	$P_{nf} = 1.64\text{x}10^{-2}$ $P_{na} = 1.00\text{x}10^{-5}$
Total time	10h36min	14h07min	46h41min
Tolerance analysis based on KDE simulation	$P_{nf} = 1.71\text{x}10^{-2}$ $P_{na} = 1.00\text{x}10^{-5}$	$P_{nf} = 1.63\text{x}10^{-2}$ $P_{na} = 3.00\text{x}10^{-5}$	$P_{nf} = 1.64\text{x}10^{-2}$ $P_{na} = 2.00\text{x}10^{-5}$
T_1	12min	50min	3h32min
T_2	9h33min	9h44min	9h36min
Total time (T_1+ T_2)	9h45min	10h34min	13h08min

Table 8: Results of the simulation $F \sim \mathcal{U}(-7.5\text{x}10^{-3}, 1\text{x}10^{-2})$, <mark>$N_{KDE} = 100000$,</mark> $Nmc = 100000$

Simulation	Meshing		
	100	**2500**	**10000**
Classical tolerance analysis	$P_{nf} = 1.71\text{x}10^{-2}$ $P_{na} = 2.00\text{x}10^{-1}$	$P_{nf} = 1.69\text{x}10^{-2}$ $P_{na} = 2.15\text{x}10^{-1}$	$P_{nf} = 1.69\text{x}10^{-2}$ $P_{na} = 2.15\text{x}10^{-1}$
Total time	10h39min	14h03min	46h51min
Tolerance analysis based on KDE simulation	$P_{nf} = 1.71\text{x}10^{-2}$ $P_{na} = 2.09\text{x}10^{-1}$	$P_{nf} = 1.67\text{x}10^{-2}$ $P_{na} = 2.23\text{x}10^{-1}$	$P_{nf} = 1.70\text{x}10^{-2}$ $P_{na} = 2.24\text{x}10^{-1}$
T_1	15min	49min	3h34min
T_2	9h31min	9h47min	9h35min
Total time (T_1+ T_2)	9h46min	10h36min	13h09min

References

[1]   H. Chen, S. Jin, Z. Li, and X. Lai, « A comprehensive study if three dimensional tolerance analysis methods », *Comput. Aided Des.*, vol. 53, p. 1‑13, 2014.

[2]   B. Schleich and S. Wartzack, « A discrete geometry approach for tolerance analysis of mechanism », *Mech. Mach. Theory*, vol. 77, p. 148‑163, juill. 2014.

[3]   Y. S. HONG and T.-C. CHANG, « A comprehensive review of tolerancing research », *International Journal of Production Research*, p. 2425‑2459, 2002.

[4]   S. D. Nigam and J. U. Turner, « Review of statistical approaches of tolerance analysis », *Comput.-Aided Des.*, vol. 27, p. 6‑15, 1995.

[5]   V. J. Skowronski and J. U. Turner, « Using Monte-Carlo variance reduction in statistical tolerance synthesis », *Comput.-Aided Des.*, vol. 29, n° 1, p. 63‑69, janv. 1997.

[6]   A. Dumas, N. Gayton, J.-Y. Dantan, and B. Sudret, « A new system formulation for the tolerance analysis of overconstrained mechanisms », *Probabilistic Eng. Mech.*, vol. 40, p. 66‑74, avr. 2015.

[7]   J. Yang, J. Wang, Z. Wu, and N. Anwer, « Statistical Tolerancing based on Variation of Point-set », *Procedia CIRP*, vol. 10, p. 9‑16, 2013.

[8]   P. Beaucaire, N. Gayton, E. Duc, and J. Y. Dantan, « Statistical tolerance analysis of overconstrained mechanisms with gaps using system reliability methods », *Comput. Aided Des.*, vol. 45, n° 15, p. 47‑55, 2013.

[9]   A. Ballu, J.-Y. Plantec, and L. Mathieu, « Geometrical reliability of overconstrained mechanisms with gaps », *CIRP Ann.-Manuf. Technol.*, vol. 57, n° 1, p. 159–162, 2008.

[10]   A. J. Qureshi, J. Y. Dantan, V. Sabri, P. Beaucaire, and N. Gayton, « A statistical tolerance analysis approach for over-constrained mechanism based on optimisation and Monte-Carlo simulation », *Comput. Aided Des.*, vol. 44, n° 2, p. 132‑142, 2012.

[11]   A. Dumas, J. Y. Dantan, and N. Gayton, « Impact of a behavior model linearization strategy on the tolerance analysis of over-constrained mechanisms », *Comput. Aided Des.*, vol. 62, p. 152‑163, 2015.

[12]   B. Schleich, N. Anwer, L. Mathieu, S. Wartzack, « Skin Model Shapes: A new paradigm shift for geometric variations modelling in mechanical engineering », *Comput. Aided Des.*, p. 1‑15, 0.8 2016.

[13]   J. W. Wittwer, K. W. Chase, and L. L. Howell, « The direct linearization method applied to position error in Kinematic linkages », *Mech. Mach. Theory*, vol. 39, n° 7, p. 681‑693, 2004.

[14]   S. Arroyave-Tobón, D. Teissandier, and V. Delos, « Applying screw theory for summing sets of constraints in geometric tolerancing », *Mech. Mach. Theory*, vol. 112, p. 255‑271, juin 2017.

[15]   P. Bourdet, L. Mathieu, C. Lartigue, and A. Ballu, « The concept of the small displacement torsor in metrology », *Ser. Adv. Math. Appl. Sci. Adv. Math. Tools Metrol. II*, vol. 40, p. 110‑122, 1996.

[16]   I. González and I. Sánchez, « Statistical tolerance synthesis with correlated variables », *Mech. Mach. Theory*, vol. 44, n° 6, p. 1097‑1107, juin 2009.

[17]   L. Homri, E. Goka, G. Levasseur, and J.-Y. Dantan, « Tolerance analysis - form defects modeling and simulation by modal decomposition and optimization », *Comput.-Aided Des.*, juin 2017.

[18]   E. Goka, L. Homri, P. Beaurepaire, and J.-Y. Dantan, « Geometrical variation simulation for assembly with form defects », *ASME 2017 Int. Mech. Eng. Congr. Expo.*, vol. Advanced Manufacturing, n° 58356, p. V002T02A099, 2017.

[19]   J. Y. Dantan and A. J. Qureshi, « Worst case and statistical tolerance analysis based on quantified constraint satisfaction problems and monte carlo simulation », *Comput. Aided Des.*, vol. 41, n° 1, p. 1‑12, 2009.

[20]   E. Goka, L. Homri, P. Beaurepaire, and J.-Y. Dantan, « Statistical tolerance analysis of over-constrained mechanical assemblies with form defects considering contact types », *J. Comput. Inf. Sci. Eng.*, p. 021010-1-021010‑13, nov. 2018.

[21]   J. Y. Dantan, L. Mathieu, A. Ballu, and P. Martin, « Tolerance synthesis: quantifier notion and virtual boundary », *Comput.-Aided Des.*, vol. 37, n° 2, p. 231‑240, 2005.

[22]   A. Dumas, J.-Y. Dantan, N. Gayton, T. Bles, and R. Loebl, « An iterative statistical tolerance analysis procedure to deal with linearized behavior models », *J. Zhejiang Univ.-Sci. A*, vol. 16, n° 5, p. 353‑360, mai 2015.

[23]   B. Schleich, N. Anwer, L. Mathieu, and S. Wartzack, « Contact and mobility simulation for mechanical assemblies based on skin model shapes », *J. Comput. Inf. Sci. Eng.*, vol. 15, n° 2, p. 021009, 2015.

[24]   H. N. Lê, Y. Ledoux, and A. Ballu, « Experimental and theoretical investigations of mechanical joints with form defects », *J. Comput. Inf. Sci. Eng.*, vol. 14, n° 4, 2014.

[25]   Y.-Z. Hu, G. C. Barber, and D. Zhu, « Numerical Analysis for the Elastic Contact of Real Rough Surfaces », *Tribol. Trans.*, vol. 42, n° 3, p. 443‑452, janv. 1999.

[26] W. Wang, « Numerical solution of dry contact problem based on fast fourier transform and conjugate gradient method », *Chin. J. Mech. Eng.*, vol. 42, p. 14, 2006.

[27] S. R. Hummel and C. Chassapis, « Conguration design and optimization of universal joints with manufacturing tolerances », *Mech. Mach. Theory*, p. 14, 2000.

[28] C. Putignano, L. Afferrante, G. Carbone, and G. Demelio, « A new efficient numerical method for contact mechanics of rough surfaces », *Int. J. Solids Struct.*, vol. 49, n° 2, p. 338‑343, janv. 2012.

[29] P. A. Adragna, H. Faverlière, S. Samper, and M. Pillet, « Statiscal assemblies with form errors - a 2D example », in *Micro-Assmbly Technologies and Applications*, (Boston: Springer)., 260 vol., Ratcchev, S., Koelmeijer, S., 2008, p. 23‑33.

[30] J. Grandjean, Y. Ledoux, and S. Samper, « On the role of form defects in assemblies subject to local deformations and mechanical loads », *Int. J. Adv. Manuf. Technol.*, vol. 65, n° 9‑12, p. 1769‑1778, 2013.

[31] J.-F. Rameau, P. Serré, and M. Moinet, « Clearance vs. tolerance for rigid overconstrained assemblies », *Comput.-Aided Des.*, vol. 97, p. 27‑40, avr. 2018.

[32] A. Cammarata, « A novel method to determine position and orientation errors in clearance-affected overconstrained mechanisms », *Mech. Mach. Theory*, vol. 118, p. 247‑264, déc. 2017.

[33] B. Schleich and S. Wartzack, « Approaches for the assembly simulation of skin model shapes », *Comput. Aided Des.*, vol. 65, p. 18‑33, 2015.

[34] K.-S. Lin, K.-Y. Chan, et J.-J. Lee, « Kinematic error analysis and tolerance allocation of cycloidal gear reducers », *Mech. Mach. Theory*, vol. 124, p. 73‑91, juin 2018.

[35] B. Schleich, S. Wartzack, « Gap hull estimation for rigid mechanical joints considering form deviations and multiple pairs of mating surfaces », *Mech. Mach. Theory*, vol. 128, p. 444‑460, oct. 2018.

[36] V. C. Tai and K. Uhlen, « Resampling of Data for Offshore Grid Design Based on Kernel Density Estimation and Genetic Algorithm », *Energy Procedia*, vol. 80, p. 365‑375, 2015.

[37] B. W. Silverman, *Density estimation for statistics and data analysis*, Chapman&Hall/crc. New-York USA, 1986.

[38] B. Rajagopalan, U. Lall, D. G. Tarboton, and D. S. Bowles, « Multivariate nonparametric resampling scheme for generation of daily weather variables », *Stoch. Hydrol. Hydraul.*, vol. 11, n° 1, p. 65‑93, févr. 1997.

[39] D. W. Scott, *Multivariate density estimation: theory, practice, and visualization*. New York: Wiley, 1992.

[40] N.-B. Heidenreich, A. Schindler, and S. Sperlich, « Bandwidth selection for kernel density estimation: a review of fully automatic selectors », *AStA Adv. Stat. Anal.*, vol. 97, n° 4, p. 403‑433, oct. 2013.

[41] B. Anselmetti, « Generation of functional tolerancing based on positioning features », *Comput.-Aided Des.*, vol. 38, p. 902‑919, 2006.

[42] J. Liu, Z. Zhang, X. Ding, and N. Shao, « Integrating form errors and local surface deformations into tolerance analysis based on skin model shapes and a boundary element method », *Comput.-Aided Des.*, vol. 104, p. 45‑59, nov. 2018.